WEIGHTED DEPENDENCY GRAPHS AND THE ISING MODEL

JEHANNE DOUSSE AND VALENTIN FÉRAY

Abstract. Weighted dependency graphs have been recently introduced by the second author, as a toolbox to prove central limit theorems. In this paper, we prove that spins in the d-dimensional Ising model display such a weighted dependency structure. We use this to obtain various central limit theorems for the number of occurrences of local and global patterns in a growing box.

1. Introduction and statement of results

1.1. **Cumulants in the Ising model.** The Ising model is a mathematical model of ferro-magnetism in statistical physics. It was introduced in 1920 by Wilhelm Lenz who gave it as a problem to his Ph.D. student Ernst Ising [19]. It can be defined on any finite graph, but we restrict ourselves to finite subsets Λ of $\mathbb{Z}^d$. For any lattice site $i \in \Lambda$, there is a random variable σ_i which is equal to either 1 or -1 and represents the *spin* at site i. A *spin configuration* $\omega = (\sigma_i(\omega))_{i\in\Lambda}$ is an assignment of spins to every site of Λ.

The distribution of spins depends on the magnetic field h and the inverse temperature β in a way that is detailed in Section 2.1. In particular, spins corresponding to neighbour sites i and j are more likely to be equal. The bigger β is, i.e. the lower the temperature is, the more important is this phenomenon.

In his Ph.D. thesis [19], Ising solved the model for the one-dimensional case $d = 1$, and showed that there is no phase transition. But in 1936, Peierls [33] showed that, in dimensions 2 and 3, when $h = 0$, the Ising model undergoes a phase transition at a critical inverse temperature β_c. He used a combinatorial argument now known as Peierls' argument. The two-dimensional model for $h = 0$ was then exactly solved by Onsager [32] in 1944, using analytic techniques and the transfer matrix method. It turns out that in higher dimensions, there is also a phase transition for $h = 0$ (see [33], or [13] for a more modern treatment). However, there is no phase transition when there is a magnetic field $h \neq 0$ [25, 36].

The Ising model is *a priori* defined on a finite subset $\Lambda \subset \mathbb{Z}^d$, but it is well-known that we can take the *thermodynamic limit* $\Lambda \uparrow \mathbb{Z}^d$ (see eg. [13]). This defines, for each pair of parameters (β, h), a measure $\mu_{\beta,h}$ on the set $\{-1,1\}^{\mathbb{Z}^d}$ of spin configurations on the whole d-dimensional lattice $\mathbb{Z}^d$. In low temperature without magnetic field, i.e. β large and $h = 0$, this measure is not unique; we will consider the one corresponding to + boundary conditions, see Section 2.1 for details.

The Ising model has been studied in thousands of research articles, under various aspects. Among many others, a subject of interest has been the decay of joint cumulants of the spins (also called *truncated k-point functions* or *Ursell functions* in the physics literature). Consider random variables $X_1, \dots, X_r$ with finite moments defined on the same probability space Their *joint cumulant* of order r is defined as

$$\kappa(X_1, \dots, X_r) = [t_1 \dots t_r] \log \big\langle \exp(t_1 X_1 + \dots + t_r X_r) \big\rangle,$$

where $\langle Y \rangle$ denotes the expectation of Y and the notation $[t_1 \dots t_r]F$ stands for the coefficient of $t_1 \dots t_r$ in the series expansion of F. The finite moments assumption ensures that this series expansion exists, at least formally. The joint cumulant of order 2 is simply the covariance. If all random variables

2010 *Mathematics Subject Classification.* 82B20,60F05.

Key words and phrases. Ising model, cumulants, central limit theorems, cluster expansions, weighted dependency graphs.

Both authors are partially supported by grant SNF-149461 from Swiss National Science Fundation.

$X_1, \cdots, X_r$ are equal to the same variable X, then $\kappa_r(X) := \kappa(X, \dots, X)$ is the usual *cumulant* of a single random variable.

Joint cumulants have a long history in statistics and theoretical physics, see *e.g.* [35]. In the case where the X_i's are indicator functions of the presence of particles or + spins for example, they are often referred to as *truncated correlation functions* or *Ursell functions* in the statistical physics literature. In this paper, we will denote by $\kappa_{\beta,h}(\sigma_{i_1}, \dots, \sigma_{i_r})$ the joint cumulant of order r of spins $\sigma_{i_1}, \dots, \sigma_{i_r}$, with inverse temperature β and magnetic field h.

Bounds on cumulants in the physics literature are often called *cluster properties*. There is in fact a hierarchy of cluster properties, corresponding to sharper or weaker bounds on cumulants; we refer to [5] or [26, Chapter 6, §1] for definitions of various kinds of cluster properties.

In the case of the Ising model, a first bound on cumulants was obtained by Martin-Löf [27, Eq. (20)] — see also [24, Section 1] —: he proved that the joint cumulant $\kappa_{\beta,h}(\sigma_i; i \in A)$ decreases exponentially in $\mathrm{diam}(A)/r$, where $\mathrm{diam}(A)$ is the the diameter of A and r the order of the cumulants. In [6], Duneau, Iagolnitzer and Souillard sharpened this bound in presence of a magnetic field ($h \neq 0$), or for $h = 0$ and very high temperature: $\kappa_{\beta,h}(\sigma_i; i \in A)$ decays exponentially in $\ell_T(A)$. Here $\ell_T(A)$ denotes the tree-length of A, i.e. the minimal size of a connected set of edges of the lattice $\mathbb{Z}^d$ such that each vertex of A is incident to at least one edge in the set. In [26], Malyshev and Minlos have a similar result in the case $h = 0$ and very low temperature. Both their approaches use cluster expansion, a powerful tool introduced by Mayer and Montroll [28], which is now standard in the study of the Ising model. Both proofs use additional ingredients of different nature: Duneau, Iagolnitzer and Souillard use the Lee-Yang circle theorem and complex analysis arguments, while Malyshev's and Minlos' approach relies on combinatorial developments and bounds on joint cumulants for *contours* as an intermediate step.

These bounds on cumulants will be our starting point to prove central limit theorems for patterns in the Ising model. In order to make the article more self-contained, we give a simpler and more unified approach of the decays of joint cumulants in several regimes where the cluster expansion converges. To do so, for $\beta_1, \beta_2, h_1 > 0$, let us introduce the three different regimes

$$\begin{aligned}
R^{HT}(\beta_1, h_1) &:= \{(\beta, h) : 0 \leq \beta < \beta_1, |h| \leq h_1\} \quad \textit{(very high temperature)},\\
R^{LT}(\beta_2) &:= \{(\beta, h) : \beta > \beta_2, h = 0\} \quad \textit{(very low temperature)},\\
R^{SF}(h_1) &:= \{(\beta, h) : |h| > h_1\} \quad \textit{(strong magnetic field)},
\end{aligned}$$

and their union

$$R(\beta_1, \beta_2, h_1) = R^{HT}(\beta_1, h_1) \cup R^{LT}(\beta_2) \cup R^{SF}(h_1).$$

The result is stated as follows.

Theorem 1.1. *We consider the Ising model on $\mathbb{Z}^d$ with parameters (β, h). There exist positive constants $\varepsilon(d) < 1$, $\beta_1 = \beta_1(d)$, $\beta_2 = \beta_2(d)$ and $h_1 = h_1(d)$ depending on the dimension d, and for any integer $r \geq 1$ a constant D_r such that for all $A = \{i_1, \dots, i_r\} \subset \mathbb{Z}^d$, we have*

$$\left|\kappa_{\beta,h}(\sigma_{i_1}, \dots, \sigma_{i_r})\right| \leq D_r \varepsilon(d)^{\ell_T(A)},$$

for any $(\beta, h) \in R(\beta_1, \beta_2, h_1)$.

The notion of tree-length $\ell_T(A)$ is important to make the connection with weighted dependency graphs, which we discuss now.

Remark 1.2. In the very high temperature and strong magnetic field regimes, one can chose $\varepsilon(d)$ independent of the dimension and $D_r = 1$ for all $r \geq 1$ (see the end of Section 3.1). This choice of D_r is important to study the speed of convergence in the central limit theorem for the magnetization (see [11, Section 5.3]), but is irrelevant for the purpose of the present article.

1.2. **Weighted dependency graphs.** The theory of weighted dependency graphs, recently introduced by the second author in [10], is a toolbox to prove central limit theorems. It extends the well-known concept of dependency graphs; see [1, 20].

Throughout the article, a weighted graph is a graph such that a weight w_e in $[0,1]$ is associated to each edge e, where an edge of weight 0 is the same as no edge.

A *spanning tree* of a graph $G = (V, E)$ is a subset E' of E such that (V, E') is a tree. If G is an edge-weighted graph, we define the weight $w(T)$ of a spanning tree T of G as the *product* of the weights of the edges in T. The maximum weight of a spanning tree of G is denoted $\mathcal{M}\big(G\big)$. By convention, if G is disconnected, we set $\mathcal{M}\big(G\big) = 0$.

We are now ready to define weighted dependency graphs:

Definition 1.3. Let $\{Y_\alpha, \alpha \in A\}$ be a family of random variables with finite moments, defined on the same probability space; and let $\boldsymbol{C} = (C_1, C_2, \cdots)$ be a sequence of positive real numbers.

A weighted graph G is a $\boldsymbol{C}$*-weighted dependency graph* for $\{Y_\alpha, \alpha \in A\}$ if, for any multiset $B = \{\alpha_1, \dots, \alpha_r\}$ of elements of A, one has

$$\left|\kappa\big(Y_\alpha; \alpha \in B\big)\right| \leq C_r\, \mathcal{M}\big(G[B]\big), \tag{1.1}$$

where $G[B]$ denotes the graph induced by G on the vertex set B.

Remark 1.4. This is actually a simplified definition, sufficient for the purpose of this paper. It corresponds to the case $\Psi \equiv 1$ of the general definition given in [10].

Informally, that a family of random variables $\{Y_a, a \in A\}$ admits a weighted graph G as weighted dependency graph means the following:

- G has vertex-set A, i.e. we have one vertex in G per variable in $\{Y_a, a \in A\}$.
- The smaller the weight of an edge $\{a, b\}$ is, the *closer to independent* Y_a and Y_b should be. In particular, an edge of weight 0 – or equivalently no edge – between a and b means that Y_a and Y_b are independent. This closeness to independence is not only measured, as one could expect, by a bound on the covariance, but also involves bounds on higher order cumulants.

As we will see in Section 1.3, weighted dependency graphs allows one to easily obtain central limit theorems. Another nice feature of weighted dependency graphs is the following stability property: a weighted dependency graph for a family $\{Y_a, a \in A\}$ automatically gives a weighted dependency graph for monomials $Y_I = \prod_{a \in I} Y_a$ in the Y_a's with a fixed bound m on the degree (*i.e.* I runs over multiset of elements of A of size at most m). As a consequence, we can potentially prove central limit theorems for sums of such monomials. We refer the reader to [10] for a detailed presentation of the theory of weighted dependency graphs.

Let us come back to the Ising model. The bounds on joint cumulants of Theorem 1.1 can be naturally translated in terms of weighted dependency graphs for the random variables $\{\sigma_i : i \in \mathbb{Z}^d\}$. In the next statement, and throughout the paper, we let $\|i - j\|_1$ denote the graph distance in $\mathbb{Z}^d$ between two points i and j.

Theorem 1.5. *Let $\omega = (\sigma_i(\omega))_{i \in \mathbb{Z}^d}$ be a spin configuration distributed according to $\mu_{\beta,h}$, where $(\beta, h) \in R(\beta_1, \beta_2, h_1)$. Let G be the complete weighted graph with vertex set $\mathbb{Z}^d$, such that every edge $e = \{i, j\}$ has weight $w_e = \varepsilon(d)^{\frac{\|i-j\|_1}{2}}$, where $\varepsilon(d)$ comes from Theorem 1.1.*

Then G is a $\mathbf{C}$-weighted dependency graph for the family $\{\sigma_i : i \in \mathbb{Z}^d\}$, for some sequence $\mathbf{C} = (C_r)_{r \geq 1}$, depending only on d.

Theorem 1.5 is proved in Section 4.1.1. The proof uses Theorem 1.1, some general results of [10] and elementary considerations. As explained above, this automatically yields a weighted dependency graph for products of a finite number of spins, which will be presented in Theorem 4.6 below.

We conclude this subsection with the motivation behind Theorem 1.5. The Ising model is the prototypical example of a Markov random field. Recall that a Markov random field on a graph G

with vertex set A is a family of random variables $\{Y_a, a \in A\}$ such that, for subsets $B, B', B'' \subset A$, $\{Y_a, a \in B\}$ and $\{Y_a, a \in B'\}$ are independent conditionally on $\{Y_a, a \in B''\}$ as soon as every path going from B to B' in G goes through B''; this is also sometimes called *global Markov property* [23].

Informally, in a Markov random field, a variable interacts directly only with its neighbours. We can thus expect that the dependency between variables is weaker when their distance in the graph G increases. Indeed, such variables only interact through all variables lying between them in the graph. In other terms, we expect to have a weighted dependency graph which is complete (because there is no reason to have unconditionally independent variables), but whose weights decrease with the graph distance. This was observed in the case of Markov chains, which are one-dimensional Markov random fields, in [10, Section 10]. The present paper gives such a statement for the d-dimensional Ising model. In both cases, weights decrease exponentially with the graph distance.

1.3. **Central limit theorems.** Central limit theorems (CLTs) play a key role in probability theory and have also been a subject of interest in the study of the Ising model. We refer to the second edition of Georgii's classical book [14, Bibliographic Notes on Section 8.2, p.469] for an overview of the different methods used to obtain such results. The goal of this paper is to prove new central limit theorems, using the weighted dependency graph technique. To this end, we shall use the following normality criterion, which is a slightly modified version of the main theorem in [10].

Theorem 1.6. *Suppose that, for each n, $\{Y_{n,i}, 1 \le i \le N_n\}$ is a family of random variables with finite moments defined on the same probability space. Let $\mathbf{C} = (C_r)_{r\ge 1}$ be a fixed sequence that does not depend on n.*

Assume that, for each n, one has a $\mathbf{C}$-weighted dependency graph G_n for $\{Y_{n,i}, 1 \le j \le N_n\}$ and denote $\Delta_n - 1$ its maximal weighted degree.

Let $X_n = \sum_{i=1}^{N_n} Y_{n,i}$ and $v_n^2 = Var(X_n)$. Assume that there exists a sequence (a_n), an integer $s \ge 3$ and a real number v such that

$$(1)\ \frac{v_n^2}{a_n^2} \xrightarrow[n\to\infty]{} v^2, \quad (2)\ \text{for all } n,\ a_n^2 \le C_2 N_n \Delta_n, \quad (3)\ \left(\frac{N_n}{\Delta_n}\right)^{\frac{1}{s}} \frac{\Delta_n}{a_n} \xrightarrow[n\to\infty]{} 0.$$

Then in distribution,

$$\frac{X_n - \mathbb{E}(X_n)}{a_n} \xrightarrow[n\to\infty]{d} \mathcal{N}(0, v^2).$$

Remark 1.7. The proof of Theorem 1.6 is almost identical to the proof of the normality criterion in [10, Section 4.3] replacing σ_n by a_n. Indeed, as noticed in [10, Section 4.3], in the special case $\Psi \equiv 1$ to which we restrict ourselves in this article, the quantities R_n and Q_n defined there can be replaced respectively by N_n (the number of vertices) and Δ_n (the maximal weighted degree plus one).

Recall from the previous section that spins, and therefore products of spins, admit a weighted dependency graph. The normality criterion above can thus be used to find CLTs for polynomials of spins in a growing box $\Lambda_n := [-n, n]^d$. To illustrate this, we consider number of occurrences of two kinds of spin patterns: *local* and *global* patterns.

We define a *local pattern* $\mathcal{P}$ to be a pair $(\mathcal{D}, \mathfrak{s})$, where $\mathcal{D}$ is a finite subset of $\mathbb{Z}^d$ containing 0 and $\mathfrak{s}$ is a function $\mathcal{D} \longrightarrow \{+,-\}$. The cardinality of $\mathcal{D}$ is called the *size* of the pattern $\mathcal{P}$. An example of local pattern is a positive spin surrounded by negative ones. In that case the subset is $\mathcal{D} = \{j \in \mathbb{Z}^d : \|j\|_1 \le 1\}$, while the sign function is given by $\mathfrak{s}(0) = +$ and $\mathfrak{s}(j) = -$ for all $j \in \mathcal{D} \setminus \{0\}$. This pattern has size $2d+1$. An *occurrence* of a local pattern $\mathcal{P} = (\mathcal{D}, \mathfrak{s})$ is a set $\{(i+j, \mathfrak{s}(j)) : j \in \mathcal{D}\}$, where $i \in \mathbb{Z}^d$ is the *position* of the occurrence.

While in local patterns we consider spins that are at a fixed distance from one another, in global patterns they can be as far as we want, as long as they have a certain global shape. Formally, we define a *global pattern* $\tilde{\mathcal{P}}$ of size m to be a pair $(\mathcal{O}, \mathfrak{s})$, where $\mathcal{O} = (\le_1, \dots, \le_d)$ is a d-tuple of total orders over $\{1, \dots, m\}$, and $\mathfrak{s}$ is a function $\{1, \dots, m\} \longrightarrow \{+,-\}$. An *occurrence* of $\tilde{\mathcal{P}}$ in a spin configuration ω is a set $\{x^{(1)}, \dots, x^{(m)}\}$ of m elements of $\mathbb{Z}^d$ such that there exists some ordering $(x^{(1)}, \dots, x^{(m)})$ of these elements such that

(1) for all $i \in \{1, \dots, m\}$, $\sigma_{x^{(i)}}(\omega) = \mathfrak{s}(i)$,
(2) for all $i, j \in \{1, \dots, m\}$, for all $k \in \{1, \dots d\}$, $x_k^{(i)} \leq x_k^{(j)}$ if and only if $i \leq_k j$.

For example, if $d = 2$, $\leq_1, \leq_2$ are both the natural ordering and $\mathfrak{s}(i) = +$ for all i, then an occurrence of the global pattern $(\mathcal{O}, \mathfrak{s})$ is a set of m positive spins such that each of them is located to the North-East of the previous one.

CLTs for local and global patterns in other structures than the Ising model have attracted attention in the literature. We mention Markov chains (see [34, 12, 10] and references therein), patterns in random permutations (see [2, 21] for global patterns and [16, 2, 7] for local patterns) and arc configurations in random set-partitions (CLTs for the number of arcs of size 1, which is a local pattern, and the number of crossings, which is a global pattern, were given in [4]). Note that Markov chains are (discrete) one-dimensional Markov random fields, while the random permutation model is a non-Markovian two-dimensional model (when considering patterns, we think of permutations as permutation matrices). Finding such CLT results in Markov random fields of dimension two or more, and in particular in the Ising model, is therefore a natural problem.

We first prove a CLT for local patterns. Let $S_{n,\mathcal{P}}$ denote the number of occurrences of a given local pattern $\mathcal{P}$ in Λ_n.

Theorem 1.8. *Consider the Ising model on $\mathbb{Z}^d$, with inverse temperature β and magnetic field h, such that $(\beta, h) \in R(\beta_1, \beta_2, h_1)$. Let $\mathcal{P}$ be a local pattern. Then*

$$\frac{S_{n,\mathcal{P}} - \mathbb{E}(S_{n,\mathcal{P}})}{\sqrt{|\Lambda_n|}} \xrightarrow[n\to\infty]{d} \mathcal{N}(0, v_{\mathcal{P}}^2),$$

where

$$v_{\mathcal{P}}^2 = \lim_{n\to\infty} \frac{\mathrm{Var}(S_{n,\mathcal{P}})}{|\Lambda_n|}.$$

Similarly, if $S_{n,\tilde{\mathcal{P}}}$ denotes the number of occurrences of a global pattern $\tilde{\mathcal{P}}$ in Λ_n, we have the following result.

Theorem 1.9. *Consider the Ising model on $\mathbb{Z}^d$, with inverse temperature β and magnetic field h, such that $(\beta, h) \in R(\beta_1, \beta_2, h_1)$. Let $\tilde{\mathcal{P}}$ be a global pattern of size m. We assume that, for some positive constants A and η*

$$\mathrm{Var}(S_{n,\tilde{\mathcal{P}}}) \geq A|\Lambda_n|^{2m-2+\eta}. \tag{1.2}$$

Then

$$\frac{S_{n,\tilde{\mathcal{P}}} - \mathbb{E}(S_{n,\tilde{\mathcal{P}}})}{\sqrt{\mathrm{Var}(S_{n,\tilde{\mathcal{P}}})}} \xrightarrow[n\to\infty]{d} \mathcal{N}(0, 1).$$

We do not have in general an estimate for the variance $\mathrm{Var}(S_{n,\tilde{\mathcal{P}}})$. However, when the pattern consists of positive spins only, we can prove that (1.2) is satisfied (with $\eta = 1$) — see Proposition 4.10 below. The reverse inequality $\mathrm{Var}(S_{n,\tilde{\mathcal{P}}}) \leq B|\Lambda_n|^{2m-1}$ is always fulfilled (see the proof of Theorem 1.9).

We finish this introduction with a comparison with other methods. Standard methods to get CLT in random fields are mixing properties [29] or FKG inequalities [30, 31]. It seems that the CLT for local patterns can be easily obtained with these methods. Indeed, in an exponentially mixing field such as the Ising model, if we denote $Z_i^{\mathcal{P}}$ the characteristic function of the occurrence of $\mathcal{P}$ in position i (see (4.3)), then the field $(Z_i^{\mathcal{P}})_{i\in\mathbb{Z}^d}$ is also exponentially mixing and we can use the criterion given by Neaderhouser [29, Section 3]. CLT for functions of neighbouring spins are also accessible with methods based on FKG inequalities, see [31] for a general result in this direction.

On the contrary, we do not know how to adapt these methods to global patterns. Obtaining a CLT for subword occurrences in Markov chains (the analogue problem in one-dimensional Markov field) is already a difficult problem, see [3, 12] for some history on this problem. The technique of dependency graphs gives access to CLTs for such global patterns, as shown in Theorem 1.9.

In principle, it would also be feasible to mix local and global conditions (as in vincular patterns for permutations [18]) or to consider more generally polynomials in $X_i^+ = \frac{1+\sigma_i}{2}$ and $X_i^- = \frac{1-\sigma_i}{2}$, in the spirit of [15]; a major difficulty is then to get general estimates for the variance.

1.4. **Outline of the paper.** The remainder of the paper is organised as follows. In Section 2, we give some preliminary definitions and basic results about the Ising model and tree lengths. In Section 3, we discuss the cluster expansion for the Ising model in the three different regimes we consider (very high temperature, very low temperature, high magnetic field), and deduce bounds on joint cumulants. In Section 4, we use the theory of weighted dependency graphs to prove our central limit theorems.

Note: all constants throughout the paper depend on the dimension d of the space and we shall not make it explicit from now on.

2. Preliminaries

2.1. **The Ising model.** We consider the Ising model on a finite subset Λ of $\mathbb{Z}^d$. We use the notation of [13], that we present now.

Let $\mathcal{E}_\Lambda := \{\{i,j\} \subset \Lambda : \|i-j\|_1 = 1\}$ be the set of nearest neighbour pairs in Λ. To each spin configuration ω, we associate its Hamiltonian

$$H_{\Lambda;\beta,h}(\omega) := -\beta \sum_{\{i,j\}\in\mathcal{E}_\Lambda} \sigma_i(\omega)\sigma_j(\omega) - h \sum_{i\in\Lambda} \sigma_i(\omega),$$

where $\beta \geq 0$ and h are two real parameters, respectively called *inverse temperature* and *magnetic field.*

The probability of a spin configuration ω is given by the Gibbs distribution

$$\mu_{\Lambda;\beta,h}(\omega) := \frac{1}{Z_{\Lambda;\beta,h}} e^{-H_{\Lambda;\beta,h}(\omega)},$$

where

$$Z_{\Lambda;\beta,h} := \sum_{\omega\in\{-1,1\}^\Lambda} e^{-H_{\Lambda;\beta,h}(\omega)}$$

is called the partition function.

The quantities defined so far are with "free boundary conditions", which means that the value of the spins outside of Λ is not taken into consideration. We can also define the same quantities with boundary condition, by considering the Ising model on the full lattice $\mathbb{Z}^d$, but where the values of the spins outside of Λ are fixed. Fixing a spin configuration $\eta \in \{-1,1\}^{\mathbb{Z}^d}$, we define a spin configuration in Λ with boundary condition η as an element of the set

$$\Omega_\Lambda^\eta := \left\{\omega \in \{-1,1\}^{\mathbb{Z}^d} : \omega_i = \eta_i, \forall i \notin \Lambda\right\}.$$

We now define the Hamiltonian as

$$H^\eta_{\Lambda;\beta,h}(\omega) := -\beta \sum_{\{i,j\}\in\mathcal{E}^b_\Lambda} \sigma_i(\omega)\sigma_j(\omega) - h \sum_{i\in\Lambda} \sigma_i(\omega),$$

where $\mathcal{E}^b_\Lambda := \left\{\{i,j\} \subset \mathbb{Z}^d : \|i-j\|_1 = 1 \text{ and } \{i,j\}\cap\Lambda \neq \emptyset\right\}.$

The Gibbs distribution of the Ising model in Λ with boundary condition η and parameters β and h is the probability distribution defined on Ω^η_Λ by

$$\mu^\eta_{\Lambda;\beta,h}(\omega) := \frac{1}{Z^\eta_{\Lambda;\beta,h}} e^{-H^\eta_{\Lambda;\beta,h}(\omega)},$$

where

$$Z^\eta_{\Lambda;\beta,h} := \sum_{\omega\in\Omega^\eta_\Lambda} e^{-H^\eta_{\Lambda;\beta,h}(\omega)}$$

is the partition function with boundary condition η.

The most classical boundary conditions are the $+$ boundary condition, where $\eta_i = +1$ for all $i \in \mathbb{Z}^d$, and the $-$ boundary condition, where $\eta_i = -1$ for all $i \in \mathbb{Z}^d$. When considering quantities with $+$ (resp. $-$) boundary condition, we write them with superscript $+$ (resp. $-$), e.g. $\mu^+_{\Lambda;\beta,h}(\omega)$.

We now take an increasing sequence Λ_n of finite subsets of $\mathbb{Z}^d$ with $\bigcup_{n\geq 1} \Lambda_n = \mathbb{Z}^d$. It is a well-known fact (see, *e.g.*, [13, Chapter 3]) that the sequence of measures $\mu^+_{\Lambda_n;\beta,h}$ converges in the weak sense towards a measure denoted $\mu^+_{\beta,h}$ as $n \to \infty$. In the high temperature case ($\beta < \beta_c(d)$, $h = 0$) or in the presence of a magnetic field ($h \neq 0$), the limiting measure is independent of the choice of boundary conditions. At low temperature ($\beta > \beta_c(d)$, $h = 0$), the limiting measure depends on the boundary conditions; in this article, we restrict ourselves to $+$ boundary conditions to have a well-defined limiting measure in all cases. Also, we drop the superscript $+$ and denote the limiting measure by $\mu_{\beta,h}$.

In this article, we work with this limiting measure $\mu_{\beta,h}$ and prove our central limit theorem under this measure. In comparison with the measure $\mu^+_{\Lambda_n;\beta,h}$, it has the advantage to be translation invariant, which simplifies in particular the variance estimates.

2.2. **Spanning trees of maximal weight and tree lengths.** We recall from the introduction that, if G is an edge-weighted graph, we denote by $\mathcal{M}\big(G\big)$ the maximum weight of a spanning tree of G.

We will be mainly interested in the case where V is a finite subset A of $\mathbb{Z}^d$, E consists of all pairs of vertices of A (i.e. we have a complete graph), and the weights are of the form $w(i,j) = \varepsilon^{\|i-j\|_1}$ for some positive constant $\varepsilon < 1$. We denote this weighted graph by $G[A]$. Then, for a spanning tree T of $G[A]$,

$$w(T) = \varepsilon^{\sum_{(i,j)\in T} \|i-j\|_1},$$

and the maximal such weight $\mathcal{M}\big(G[A]\big)$ is obtained by *minimizing* the quantity $\sum_{(i,j)\in T} \|i-j\|_1$. Therefore we define

$$\ell'_T(A) = \min_T \sum_{(i,j)\in T} \|i-j\|_1,$$

where the minimum is taken over all spanning trees T of $G[A]$, *i.e.* of the complete graph on A. Then we have $\mathcal{M}\big(G[A]\big) = \varepsilon^{\ell'_T(A)}$.

The quantity $\ell'_T(A)$ is sometimes referred to as the *tree-length* of A. There is another closely related notion of tree-length, used in the introduction, which is defined as $\ell_T(A) = \min_B \ell'_T(A \cup B)$, where the minimum is taken over all finite subsets B of $\mathbb{Z}^d$. In other words, this is the minimum length of a tree connecting vertices of A and possibly other vertices of $\mathbb{Z}^d$. Equivalently, this is the minimal size of a connected set of edges of the lattice $\mathbb{Z}^d$ such that each vertex of A is incident to at least one edge in the set. These two notions of tree-length are illustrated on Fig. 1.

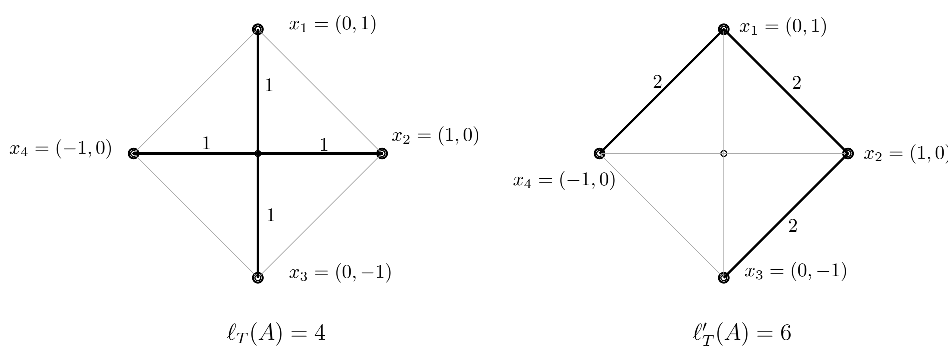

Figure 1. The two notions of tree-length on an example

In [5, page 197], Duneau, Iagolnitzer and Souillard proved the following bound, which will be useful later in our computations.

Proposition 2.1. *For all $A = \{x_1, \dots, x_n\}$ finite subset of $\mathbb{Z}^d$, we have*

$$\ell_T(A) \leq \ell'_T(A) \leq 2\ell_T(A).$$

3. Cluster expansions and bounds on joint cumulants

The cluster expansion is a powerful tool in statistical mechanics, which consists in studying a system in terms of macroscopic geometrical objects instead of considering its original microscopic components. It was introduced in a work of Mayer and Montroll [28] studying molecular distribution and has since been used in several other topics; for the Ising model, see for example [9] or Chapter 5 of [13]. In this section, we will use the cluster expansion in three different regimes of the Ising model to prove the bounds on joint cumulants of Theorem 1.1. This will later be useful to apply the theory of weighted dependency graphs. Theorem 1.1 is proved in Sections 3.1.2, 3.2.2 and 3.3, depending on the considered regime.

Remark 3.1. In each section below, we use some classical notation for cluster expansion, such as Ξ, W, X, ... Note however that these quantities may have different meanings in different regimes. Since they are only used for the proof of Theorem 1.1 and since the proofs in the different regimes are independent from each other, this should not create any difficulty.

3.1. At very high temperature, with a weak magnetic field.

3.1.1. *The cluster expansion of the (multivariate) moment generating function.* Fix $h_1 > 0$. For Theorem 1.1, we will specialize h_1 to the threshold value of the strong magnetic field regime, but, in this section, we work with an arbitrary value of h_1. We consider the regime where $|h| \leq h_1$ and β is sufficiently small (very high temperature and weak magnetic field).

We fix a finite domain $\Lambda \subset \mathbb{Z}^d$ and let $A = \{x_1, \dots, x_r\}$ be a set of points in Λ. We consider the (multivariate) moment generating function

$$\left\langle \exp\left(\sum_{j=1}^{r} t_j \sigma_{x_j}\right)\right\rangle_{\Lambda;\beta,h} = \frac{\sum_{\omega\in\Omega_\Lambda} \exp\left(\sum_{j=1}^{r} t_j \sigma_{x_j}(\omega)\right) e^{-H_{\Lambda;\beta,h}(\omega)}}{Z_{\Lambda;\beta,h}}.$$

To state the high temperature representation of this moment generating function, we need some notation. Let us call $Z^A_{\Lambda;\beta,h}$ the numerator of the right-hand side. The denominator $Z_{\Lambda;\beta,h}$ is then exactly $Z^{\emptyset}_{\Lambda;\beta,h}$. For a subset $E \subseteq \mathcal{E}_{\mathbb{Z}^d}$, we denote $V_o(E)$ the set of vertices in $\mathbb{Z}^d$ that are incident to an odd number of edges in E. As usual, we use Δ for the symmetric difference operator on sets. Finally, if $E \subseteq \mathcal{E}_{\mathbb{Z}^d}$ and $B \subseteq \mathbb{Z}^d$ fulfill $B\Delta V_o(E) \subseteq A$, we define

$$\mathrm{W}(E,B) = (\tanh\beta)^{|E|}(\tanh\beta h)^{|B|} \prod_{x_j\in B\Delta V_o(E)} (\tanh t_j).$$

Lemma 3.2 (high temperature representation)**.** *We have*

$$Z^A_{\Lambda;\beta,h} = (\cosh\beta)^{|\mathcal{E}_\Lambda|}(2\cosh\beta h)^{|\Lambda|}\left(\prod_{j=1}^{r}\cosh(t_j)\right)\Xi^A_{\Lambda;\beta,h}, \tag{3.1}$$

where

$$\Xi^A_{\Lambda;\beta,h} = \sum_{\substack{E\subseteq\mathcal{E}_\Lambda, B\subseteq\Lambda:\\ B\Delta V_o(E)\subseteq A}} \mathrm{W}(E,B).$$

Proof: This proof is a straight-forward extension of the case $A = \emptyset$ and $h = 0$, see e.g. [13, Eq. (5.40)]. We write in short σ_i for $\sigma_i(\omega)$. Since every σ_i lies in $\{-1,+1\}$, we can write

$$\exp(t_j\sigma_{x_j}) = \cosh(t_j) + \sigma_{x_j}\sinh(t_j) = \cosh(t_j)\big(1+\sigma_{x_j}\tanh(t_j)\big);$$
$$\exp(\beta h\sigma_i) = \cosh(\beta h)\big(1+\sigma_i\tanh(\beta h)\big);$$
$$\exp(\beta\sigma_i\sigma_j) = \cosh(\beta)\big(1+\sigma_i\sigma_j\tanh(\beta)\big).$$

This gives the following expression for $Z^A_{\Lambda;\beta,h}$:

$$Z^A_{\Lambda;\beta,h} = (\cosh\,\beta)^{|\mathcal{E}_\Lambda|}(\cosh\,\beta h)^{|\Lambda|}\left(\prod_{j=1}^{r}\cosh(t_j)\right)$$
$$\cdot \sum_{\omega\in\Omega_\Lambda}\left(\left[\sum_{E\subseteq\mathcal{E}_\Lambda}\prod_{\{i,j\}\in E}(\sigma_i\sigma_j\tanh(\beta))\right]\cdot\left[\sum_{B\subseteq\Lambda}\prod_{i\in B}(\sigma_i\tanh(\beta h))\right]\cdot\left[\sum_{C\subseteq A}\prod_{x_j\in C}(\sigma_{x_j}\tanh(t_j))\right]\right).$$

Changing the order of summation we get

$$Z^A_{\Lambda;\beta,h} = (\cosh\,\beta)^{|\mathcal{E}_\Lambda|}(\cosh\,\beta h)^{|\Lambda|}\left(\prod_{j=1}^{r}\cosh(t_j)\right)$$
$$\cdot\left(\sum_{E,B,C}\tanh(\beta)^{|E|}\tanh(\beta h)^{|B|}\left(\prod_{x_j\in C}\tanh(t_j)\right)\left[\sum_{\omega\in\Omega_\Lambda}\left(\prod_{\{i,j\}\in E}\sigma_i\sigma_j\prod_{i\in B}\sigma_i\prod_{x_j\in C}\sigma_{x_j}\right)\right]\right),$$

where, as above, we sum over $E\subseteq\mathcal{E}_\Lambda$, $B\subseteq\Lambda$ and $C\subseteq A$. By an easy symmetry argument, the sum in square brackets is zero unless all σ_i's appear an even number of times, which corresponds to the condition $B\Delta V_o(E)=C$. In this case, the sum is the number of spin configurations $|\Omega_\Lambda|=2^{|\Lambda|}$. This ends the proof of the high temperature expansion. $\square$

Pairs (E,B) with $E\subseteq\mathcal{E}_\Lambda$ and $B\subseteq\Lambda$ can be considered as subgraphs of Λ, where the vertex set $V(E,B)$ consists of B and of vertices incident to an edge of E and the edge-set is precisely E. This graph has a unique decomposition (up to reordering) into s connected components, each again being the graph of some (E_i,B_i) (for $1\le i\le s$). Clearly, if $B\Delta V_o(E)\subseteq A$, then each (E_i,B_i) satisfies $B_i\Delta V_o(E_i)\subseteq A$. Moreover, the weight function is multiplicative with respect to connected components, *i.e.* $\mathrm{W}(E,B)=\prod_i \mathrm{W}(E_i,B_i)$. Therefore using notation of [13],

$$\Xi^A_{\Lambda;\beta,h} = 1+\sum_{s\ge1}\frac{1}{s!}\sum_{\substack{(E_1,B_1),\dots,(E_s,B_s)\\ \text{connected}}}\prod_{i=1}^{s}\mathrm{W}(E_i,B_i)\prod_{1\le i<j\le s}\mathbf{1}\big[V(E_i,B_i)\cap V(E_j,B_j)=\emptyset\big],$$

where $\mathbf{1}$[event] is the indicator function of the corresponding event. The last product in the above display encodes the fact that connected components should not intersect. We set

$$\zeta((E,B),(E',B')) = \mathbf{1}\big[V(E_i,B_i)\cap V(E_j,B_j)=\emptyset\big]-1,$$

as usual in cluster expansions.

To compute cumulants, we need an expansion of $\log\left(\left\langle\exp\left(\sum_{j=1}^r t_j\sigma_{x_j}\right)\right\rangle_{\Lambda;\beta,h}\right)$ and thus of $\log(\Xi^A_{\Lambda;\beta,h})$. Such an expansion will be given by the theory of cluster expansions, but we should first check some conditions ensuring convergence, *e.g.* the ones given in [13, Section 5.4]. For this, we define

$$\overline{\mathrm{W}}(E,B) = (\tanh\,\beta)^{|E|}|\tanh\,\beta h|^{|B|},$$

which dominates all functions $\mathrm{W}(E,B)$ when the t_j's are complex parameters of moduli at most $\tanh^{-1}(1)$.

Lemma 3.3. *There exists a constant $\beta^{ht}_{ce}(d,h_1)$ such that the following holds for $|h|\le h_1$ and $\beta<\beta^{ht}_{ce}(d,h_1)$. For each fixed pair $(E_\star,B_\star)$ where $E_\star$ and $B_\star$ are finite subsets of $\mathcal{E}_{\mathbb{Z}^d}$ and $\mathbb{Z}^d$ respectively, one has*

$$S_{(E_\star,B_\star)} := \sum_{\substack{E\subseteq\mathcal{E}_{\mathbb{Z}^d},B\subseteq\mathbb{Z}^d:\\ (E,B)\text{ connected}}}\overline{\mathrm{W}}(E,B)\,e^{|V(E,B)|}\,\big|\zeta\big[(E,B),(E_\star,B_\star)\big]\big|\le|V(E_\star,B_\star)|.$$

Remark 3.4. To prove the convergence of cluster expansion, it is actually enough to prove a weaker version of this lemma, where $E_\star$ and $B_\star$ fulfil $B_\star \Delta V_o(E_\star) \subseteq A$ and where the sum only runs over pairs (E, B) with the additional conditions $E \subseteq \mathcal{E}_\Lambda$, $B \subseteq \Lambda$ and $B\Delta V_o(E) \subseteq A$. The stronger version stated here will be useful in the proof of Lemma 3.6 below.

Proof: By definition, $\zeta\big[(E,B),(E_\star,B_\star)\big] = -1$ if (E, B) and $(E_\star, B_\star)$ share a vertex and 0 otherwise. Thus

$$S_{(E_\star,B_\star)} \leq \sum_{v \in V(E_\star,B_\star)} \left[\sum_{\substack{E \subset \mathcal{E}_{\mathbb{Z}^d}, B \subset \mathbb{Z}^d \text{ s.t. } (E,B) \text{ connected} \\ \text{and } v \in V(E,B)}} \overline{\mathrm{W}}(E,B)\, e^{|V(E,B)|} \right].$$

A simple translation argument shows that the quantity in the brackets is independent of v, so that

$$S_{(E_\star,B_\star)} \leq |V(E_\star,B_\star)| \left[\sum_{\substack{E \subset \mathcal{E}_{\mathbb{Z}^d}, B \subset \mathbb{Z}^d \text{ s.t. } (E,B) \text{ connected} \\ \text{and } 0 \in V(E,B)}} \overline{\mathrm{W}}(E,B)\, e^{|V(E,B)|} \right].$$

Note that (E, B) connected implies in particular that B is included in the vertex set of the graph associated to E. For such pair (E, B), we therefore have $\overline{\mathrm{W}}(E,B) \leq (\tanh\beta)^{|E|}(\tanh\beta h)^{|V(E)|}$. Moreover, for a given E, the number of corresponding sets B is $2^{|V(E)|}$. Finally, connectedness implies $|V(E,B)| = |V(E)| \leq |E|+1$. Thus we get:

$$S_{(E_\star,B_\star)} \leq |V(E_\star,B_\star)| \left[\sum_{\substack{E \subset \mathcal{E}_{\mathbb{Z}^d} \text{ connected} \\ \text{s.t. } 0 \in V(E)}} 2^{|E|+1} (\tanh\beta)^{|E|}\, |\tanh\beta h|^{|E|+1}\, e^{|E|+1} \right].$$

The summand depends only on the size k of E. From [13, Lemma 3.59], the number of connected sets $E \subset \mathcal{E}_{\mathbb{Z}^d}$ containing 0 of size k is bounded from above by $(2d)^{2k}$, so that

$$S_{(E_\star,B_\star)} \leq |V(E_\star,B_\star)| \left[\sum_{k\geq 1} (2d)^{2k}\, (\tanh\beta)^k\, (2e|\tan\beta h|)^{k+1} \right].$$

For β small enough, uniformly on h with $|h| \leq h_1$, say $\beta < \beta_{\mathrm{ce}}^{\mathrm{ht}}(d,h_1)$, the sum is smaller than 1, which proves the lemma. □

We can now state the *cluster expansion* of $\log(\Xi^A_{\Lambda;\beta,h})$. In the remaining part of Section 3.1, we write X to represent a list $((E_i,B_i))_{1\leq i\leq s}$ of pairs of subsets of $\mathcal{E}_{\mathbb{Z}^d}$ and $\mathbb{Z}_d$, where each (E_i, B_i) should correspond to a *connected* graph and satisfy $B_i\Delta V_o(E_i) \subseteq A$. Such a list is called an (ordered) *cluster*. We also write $|X| = s$ for its length, $\mathrm{W}(X) = \prod_{i=1}^s \mathrm{W}(E_i,B_i)$ for its weight and lastly $\overline{X} = \bigcup_{i=1}^s V(E_i,B_i)$ for its support.

Proposition 3.5. *For $\beta < \beta_{ce}^{ht}(d,h_1)$ and $|h| \leq h_1$, we have the following expansion:*

$$\log(\Xi^A_{\Lambda;\beta,h}) = \sum_X \varphi(X)\ \mathrm{W}(X), \tag{3.2}$$

where the sum runs over clusters X of all lengths $s \geq 1$,

$$\varphi(X) = \varphi((E_1,B_1),\dots,(E_s,B_s)) = \frac{1}{s!} \sum_{G \subseteq G_s \text{ connected}} \left(\prod_{\{i,j\}\in G} \zeta\big[(E_i,B_i),(E_j,B_j)\big] \right),$$

and G_s denotes the complete graph on s vertices. The convergence of the series in Eq. (3.2) *holds in the sense of locally uniform convergence of analytic functions in the complex parameters $t_1, \dots, t_r$ for $|t_1|, \dots, |t_r| \leq \tanh^{-1}(1)$.*

Proof: This follows from Lemma 3.3 and the general theory of cluster expansions, see e.g. [13, Chapter 5]. For the analyticity in the parameters, see specifically [13, Section 5.5]. □

3.1.2. *Bounds on joint cumulants.* Recall that $A = \{x_1, \dots, x_r\}$ is a set of points in the finite domain Λ. The joint cumulant $\kappa_{\Lambda;\beta,h}(\sigma_{x_1}, \dots, \sigma_{x_r})$ is the coefficient of $t_1 \dots t_r$ in

$$\log \left\langle \exp\left(\sum_{j=1}^{r} t_j \sigma_{x_j}\right)\right\rangle_{\Lambda,\beta,h} = \log Z^A_{\Lambda;\beta,h} - \log Z^{\emptyset}_{\Lambda;\beta,h}$$

$$= |\Lambda| \, \log(2 \, \cosh \, \beta h) + |\mathcal{E}_\Lambda| \, \log(\cosh \, \beta) + \sum_{j=1}^{r} \log(\cos \, t_j) + \log \Xi^A_{\Lambda;\beta,h} - \log Z^{\emptyset}_{\Lambda;\beta,h}.$$

Only the summand $\log \Xi^A_{\Lambda;\beta,h}$ contributes to the coefficient of $t_1 \dots t_r$. Using Proposition 3.5, we have

$$\kappa_{\Lambda;\beta,h}(\sigma_{x_1}, \dots, \sigma_{x_r}) = [t_1 \dots t_r] \sum_{X : \overline{X} \subseteq \Lambda} \varphi(X) \, \mathrm{W}(X) = \sum_{X : \overline{X} \subseteq \Lambda} \varphi(X) \, [t_1 \dots t_r] \, \mathrm{W}(X). \tag{3.3}$$

The exchange of infinite sum and coefficient extraction is valid since we have uniform convergence of analytic functions on a neighborhood of 0. A cluster X contributes to the coefficient of $t_1 \dots t_r$ only if

$$(B_1 \Delta V_o(E_1)) \uplus \dots \uplus (B_s \Delta V_o(E_s)) = A. \tag{3.4}$$

In particular, we should have $A \subseteq \overline{X}$. When (3.4) is fulfilled, since $[t_i] \tanh(t_i) = 1$, we have

$$[t_1 \dots t_r] \, \mathrm{W}(X) = (\tanh \beta)^{e(X)} (\tanh \beta h)^{b(X)},$$

where $e(X) = |E_1| + \dots + |E_s|$ and $b(X) = |B_1| + \dots + |B_s|$. Back to Eq. (3.3), we get

$$\left|\kappa_{\Lambda;\beta,h}(\sigma_{x_1}, \dots, \sigma_{x_r})\right| \le \sum_{\substack{X: \\ A \subseteq \overline{X} \subseteq \Lambda}} |\varphi(X)| \, (\tanh \beta)^{e(X)} |\tanh \beta h|^{b(X)}.$$

Taking the limit $\Lambda \uparrow \mathbb{Z}^d$, we get a similar upper bound for the cumulant $\kappa_{\beta,h}(\sigma_{x_1}, \dots, \sigma_{x_r})$ under the probability measure $\mu_{\beta,h}$ corresponding to the whole lattice $\mathbb{Z}^d$:

$$\left|\kappa_{\beta,h}(\sigma_{x_1}, \dots, \sigma_{x_r})\right| \le \sum_{\substack{X: \\ A \subseteq \overline{X}}} |\varphi(X)| \, (\tanh \beta)^{e(X)} |\tanh \beta h|^{b(X)}.$$

But, by definition, $\varphi(X) = 0$ unless the graph corresponding to $(\bigcup_{i=1}^{s} E_i, \bigcup_{i=1}^{s} V_i)$ is connected. Together with the condition $A \subseteq \overline{X}$, this forces $e(X) \ge \ell_T(A)$. Hence, we can write

$$\left|\kappa_{\beta,h}(\sigma_{x_1}, \dots, \sigma_{x_r})\right| \le \sum_{\substack{X: \\ x_1 \in \overline{X}, \, e(X) \ge \ell_T(A)}} |\varphi(X)| \, (\tanh \beta)^{e(X)} |\tanh \beta h|^{b(X)}. \tag{3.5}$$

We now bound the right-hand side in the following lemma, whose proof is inspired by the end of the proof of Theorem 5.16 in [13].

Lemma 3.6. *Fix $h_1 > 0$. Then there exist constants $\beta^{ht}_{jc}(d, h_1) > 0$ and $\varepsilon > 0$ such that, for (β, h) with $\beta \le \beta^{ht}_{jc}(d)$ and $|h| \le h_1$, we have the following inequality:*

$$\sum_{X: \; x_1 \in \overline{X} \, and \; e(X) \ge R} |\varphi(X)| \, (\tanh \beta)^{e(X)} |\tanh \beta h|^{b(X)} \le \varepsilon^R.$$

Proof: The proof involves different values of the parameter (β, h) so that we will here make explicit the dependency of the weight in (β, h): we write $\overline{\mathrm{W}}_{(\beta,h)}(E, B)$ instead of $\overline{\mathrm{W}}(E, B)$. We first prove the following inequality: for $\beta' < \beta^{\mathrm{ht}}_{\mathrm{ce}}(d, h_1)$, we have

$$\sum_{X: \; x_1 \in \overline{X}} |\varphi(X)| \, (\tanh \beta')^{e(X)} (\tanh \beta' h_1)^{b(X)} = \sum_{X: \; x_1 \in \overline{X}} |\varphi(X)| \prod_{i=1}^{r} \overline{\mathrm{W}}_{(\beta',h_1)}(E_i, B_i) \le 1. \tag{3.6}$$

This uses the same argument as in [13, Eq. (5.29)]:

$$\sum_{X:\ x_1\in\overline{X}} |\varphi(X)| \prod_{i=1}^{r} \overline{\mathrm{W}}_{(\beta',h_1)}(E_i,B_i) \leq \sum_{r\geq 1} r \sum_{\substack{(E_1,B_1):\\ x_1\in V(E_1,B_1)}} \sum_{(E_2,B_2),\dots,(E_r,B_r)} \big|\varphi((E_i,B_i)_{i\leq s})\big| \prod_{i=1}^{r} \overline{\mathrm{W}}_{(\beta',h_1)}(E_i,B_i)$$
$$\leq \sum_{\substack{(E_1,B_1):\\ x_1\in V(E_1,B_1)}} \overline{\mathrm{W}}_{(\beta',h_1)}(E_1,B_1)\, e^{|V(E_1,B_1)|} \leq |V(\emptyset,\{x_1\})| = 1,$$

where we used Lemma 3.3 and [13, Theorem 5.4]. This proves (3.6).

Let us fix a value β' as above. There exists a constant $\varepsilon<1$ such that for β small enough, we have $\tanh\beta < \varepsilon \tanh\beta'$. We can now write, for β small enough and $|h|\leq h_1$,

$$\sum_{X:\ x_1\in\overline{X} \text{and } e(X)\geq R} |\varphi(X)|\,(\tanh\beta)^{e(X)} |\tanh\beta h|^{b(X)}$$
$$\leq \varepsilon^R \sum_{X:\ x_1\in\overline{X} \text{and } e(X)\geq R} |\varphi(X)|\,(\tanh\beta')^{e(X)} (\tanh\beta' h_1)^{b(X)} \leq \varepsilon^R,$$

where the last inequality uses (3.6). This ends the proof of the lemma. □

Combining Eq. (3.5) and Lemma 3.6, we get the desired bound: for $\beta\leq\beta^{\mathrm{ht}}_{\mathrm{jc}}(d)$,

$$|\kappa_{\beta,h}(\sigma_{x_1},\dots,\sigma_{x_r})| \leq \varepsilon^{\ell_T(A)}.$$

3.2. **At very low temperature, without magnetic field.**

3.2.1. *The cluster expansion of the partition function.* We now turn to the regime without magnetic field ($h=0$) and very low temperature (β large). Intuitively, in that case, the spin configurations with fewer pairs of neighbours having opposite spins appear with higher probability. To emphasize the role of these pairs, we rewrite the Hamiltonian as follows:

$$H^{\eta}_{\Lambda;\beta,0}(\omega) = -\beta|E^{\eta}_{\Lambda}| - \beta \sum_{\{i,j\}\in\mathcal{E}^{\eta}_{\Lambda}} (\sigma_i(\omega)\sigma_j(\omega)-1).$$

The only non-zero terms in the sum are those where two neighbours i and j have opposite spins. Let us consider a finite subset $\Lambda\subset\mathbb{Z}^d$ with + boundary condition. A typical spin configuration will then look as a sea of +'s with some islands of −'s. Therefore the interesting macroscopic components for the cluster expansion in that case are the frontiers between the areas of +'s and those of −'s, which are called *contours*. Let us define them more rigorously.

Given $\omega\in\Omega^+_\Lambda$, let $\Lambda^-(\omega)$ denote the set of lattice points i where $\sigma_i(\omega)=-1$. For each $i\in\mathbb{Z}^d$ we define $\mathcal{S}_i := i+[\frac{-1}{2},\frac{1}{2}]^d$ to be the unit cube of $\mathbb{R}^d$ centred at i. Now let

$$\mathcal{U}(\omega) := \bigcup_{i\in\Lambda^-(\omega)} \mathcal{S}_i,$$

and consider the set of maximal connected components of the boundary of $\mathcal{U}(\omega)$, which we denote

$$\Gamma'(\omega) = \{\gamma_1,\dots,\gamma_r\}.$$

Each of the γ_i's is a *contour* of ω. Contours are connected sets of closed $(d-1)$-dimensional faces of the cubes $\mathcal{S}_i$. We denote by $|\gamma_i|$ the number of such faces in γ_i. Let $\Gamma_\Lambda := \{\gamma\in\Gamma'(\omega) : \omega\in\Omega^+_\Lambda\}$ denote the set of all possible contours in Λ. Finally, a collection of contours $\Gamma'\subset\Gamma_\Lambda$ is said to be *admissible* if there exists a spin configuration $\omega\in\Omega^+_\Lambda$ such that $\Gamma'(\omega)=\Gamma'$.

We say that Λ is *c-connected* if $\mathbb{R}^d\setminus\bigcup_{i\in\Lambda}\mathcal{S}_i$ is connected, which we will assume from now on in this paper. Then, according to [13, eq (5.42)], the partition function can be rewritten as

$$Z^+_{\Lambda;\beta,0} = e^{\beta|\mathcal{E}^+_\Lambda|}\,\Xi^+_{\Lambda;\beta,0},$$

where

$$\Xi^+_{\Lambda;\beta,0} := \sum_{\Gamma' \subset \Gamma_\Lambda \text{admissible}} \prod_{\gamma \in \Gamma'} e^{-2\beta|\gamma|}.$$

The cluster expansion is an expression of $\log \Xi^+_{\Lambda;\beta,0}$ as an absolutely convergent series. In this case, an (ordered) cluster is a list $X = (\gamma_1, \dots, \gamma_s)$ of contours. We denote by $\overline{X}$ the support of X, ie $\overline{X} = \cup_{\gamma \in X} \gamma$. In the following we write $\overline{X} \subseteq \Lambda$ to say that $\overline{X} \subseteq \cup_{i\in\Lambda} \mathcal{S}_i$ as subsets of $\mathbb{R}^d$.

In can be shown (see e.g. [13, Chapter 5]) that the cluster expansion converges for β large enough.

Proposition 3.7. *There exists $\beta^{lt}_{ce}(d)$ such that for all $\beta > \beta^{lt}_{ce}(d)$,*

$$\log \Xi^+_{\Lambda;\beta,0} = \sum_{X : \overline{X} \subseteq \Lambda} \varphi(X)\, e^{-2\beta \sum_{i=1}^s |\gamma_i|},$$

where

$$\varphi(\gamma_1, \dots, \gamma_s) = \frac{1}{s!} \sum_{G \subseteq G_s \, connected} \prod_{\{i,j\} \in G} \zeta(\gamma_i, \gamma_j),$$

$$\zeta(\gamma_i, \gamma_j) := \begin{cases} 0 & \text{if } \gamma_i \cap \gamma_j = \emptyset, \\ -1 & \text{otherwise,} \end{cases}$$

and G_s denotes as above the complete graph on s vertices.

3.2.2. *Bounds on joint cumulants.* This cluster expansion can be used to compute expectations and therefore deduce some bounds on joint cumulants.

Let $A \subseteq \Lambda$ and let us define $\sigma_A := \prod_{i\in A} \sigma_i$. Its expectation is given by

$$\langle \sigma_A \rangle^+_{\Lambda;\beta,0} = \sum_{\omega \in \Omega^+_\Lambda} \sigma_A(\omega) \frac{e^{-H_{\Lambda;\beta,0}(\omega)}}{Z^+_{\Lambda;\beta,0}}.$$

For any spin configuration $\omega \in \Omega^+_\Lambda$ and any contour $\gamma \in \Gamma'(\omega)$, let us define the *interior* of γ (written $\mathrm{Int}(\gamma)$) as the set of points of Λ which would have spin -1 if γ was the only contour of ω. We also write $\mathrm{Int}(X) := \bigcup_{\gamma \in X} \mathrm{Int}(\gamma)$ for any collection X of contours. Thus for any $\omega \in \Omega^+_\Lambda$ and any $i \in \Lambda$,

$$\sigma_i(\omega) = (-1)^{|\{\gamma \in \Gamma'(\omega) : i \in \mathrm{Int}(\gamma)\}|},$$

and thus

$$\begin{aligned} \sigma_A(\omega) &= (-1)^{\sum_{i\in A} |\{\gamma \in \Gamma'(\omega) : i \in \mathrm{Int}(\gamma)\}|} \\ &= (-1)^{\sum_{\gamma \in \Gamma'(\omega)} |\{i \in A : i \in \mathrm{Int}(\gamma)\}|}. \end{aligned}$$

Therefore one can write

$$\langle \sigma_A \rangle^+_{\Lambda;\beta,0} = \frac{\Xi^{+,A}_{\Lambda;\beta,0}}{\Xi^+_{\Lambda;\beta,0}},$$

where

$$\Xi^{+,A}_{\Lambda;\beta,0} := \sum_{\Gamma' \subseteq \Gamma_\Lambda \text{admissible}} \prod_{\gamma \in \Gamma'} (-1)^{|\{i \in A : i \in \mathrm{Int}(\gamma)\}|} e^{-2\beta|\gamma|}.$$

The cluster expansion of $\log \Xi^{+,A}_{\Lambda;\beta,0}$ converges, which means that we have an analogue of Proposition 3.7 for $\Xi^{+,A}_{\Lambda;\beta,0}$, replacing $e^{-2\beta|\gamma|}$ by $(-1)^{|\{i\in A : i \in \mathrm{Int}(\gamma)\}|} e^{-2\beta|\gamma|}$. Thus $\langle \sigma_A \rangle^+_{\Lambda;\beta,0}$ can be expressed as

$$\langle \sigma_A \rangle^+_{\Lambda;\beta,0} = \exp\left(\sum_{X : \overline{X} \subseteq \Lambda} \Psi^A_\beta(X) - \sum_{X : \overline{X} \subseteq \Lambda} \Psi^\emptyset_\beta(X) \right),$$

where for a cluster $X = (\gamma_1, \dots, \gamma_s)$,

$$\Psi^A_\beta(X) := \varphi(X)\, (-1)^{\sum_{j=1}^r |\{i \in A : i \in \mathrm{Int}(\gamma_j)\}|} e^{-2\beta \sum_{j=1}^r |\gamma_j|}.$$

In particular, in the domain of convergence of the cluster expansion, joint moments are nonzero. If a cluster X has no vertex of A in the interior of any of its contours, then $\Psi_\beta^A(X) = \Psi_\beta^\emptyset(X)$. Such clusters do not contribute to $\langle \sigma_A \rangle^+_{\Lambda;\beta,0}$. Therefore,

$$\langle \sigma_A \rangle^+_{\Lambda;\beta,0} = \exp\left(\sum_{X \sim A : \overline{X} \subseteq \Lambda} (\Psi_\beta^A(X) - \Psi_\beta^\emptyset(X)) \right),$$

where $X \sim A$ means that X contains at least one contour γ such that a point of A is in the interior of γ. The series is absolutely convergent and we can let $\Lambda \uparrow \mathbb{Z}^d$, obtaining the following proposition.

Proposition 3.8 (Equation (5.47) in [13])**.** *For β large enough,*

$$\langle \sigma_A \rangle^+_{\beta,0} = \exp\left(\sum_{X \sim A} (\Psi_\beta^A(X) - \Psi_\beta(X)) \right).$$

We now want to find estimates on the joint cumulants of the variables $\{\sigma_i : i \in A\}$, for all $A \subset \mathbb{Z}^d$ finite. But in this case, it is easier to estimate first another quantity related to cumulants. We define, for some set B and random variables $(Y_i)_{i \in B}$ defined on the same probability space with nonzero joint moments,

$$Q\left(Y_j; j \in B\right) := \prod_{\substack{\delta \subseteq B \\ \delta \neq \emptyset}} \left\langle \prod_{j \in \delta} Yj \right\rangle^{(-1)^{|\delta|}}.$$

For example,

$$Q_{\beta,0}(\sigma_1, \sigma_2) = \frac{\langle \sigma_1 \sigma_2 \rangle^+_{\beta,0}}{\langle \sigma_1 \rangle^+_{\beta,0} \langle \sigma_2 \rangle^+_{\beta,0}}, \qquad Q_{\beta,0}(\sigma_1, \sigma_2, \sigma_3) = \frac{\langle \sigma_1 \sigma_2 \sigma_3 \rangle^+_{\beta,0} \langle \sigma_1 \rangle^+_{\beta,0} \langle \sigma_2 \rangle^+_{\beta,0} \langle \sigma_3 \rangle^+_{\beta,0}}{\langle \sigma_1 \sigma_2 \rangle^+_{\beta,0} \langle \sigma_1 \sigma_3 \rangle^+_{\beta,0} \langle \sigma_2 \sigma_3 \rangle^+_{\beta,0}}.$$

We show a bound on the quantities $Q_{\beta,0}\left(\sigma_j; j \in A\right)$ for all finite $A \subset \mathbb{Z}^d$.

Lemma 3.9. *Let A be a finite subset of $\mathbb{Z}^d$ of size r. Then for β large enough,*

$$\left| Q_{\beta,0}\left(\sigma_j; j \in A\right) - 1 \right| \leq C_r e^{-c\beta \ell_T(A)},$$

where $c = c(d)$ and C_r are positive constants depending respectively on d and r.

Proof: Using Proposition 3.8, we have

$$\log Q\left(\sigma_j; j \in A\right) = \sum_{\delta \subseteq A, \delta \neq \emptyset} (-1)^{|\delta|} \sum_{X \sim \delta} (\Psi_\beta^\delta(X) - \Psi_\beta(X)).$$

Recall that $X \sim \delta$ means that at least one point of δ is in the interior of a contour in X. We split the second sum depending on the exact subset $I \subseteq \delta$ of points that are in the interior of a contour in X. By definition of Ψ, observe that, if I is as above, then $\Psi_\beta^\delta(X) = \Psi_\beta^I(X)$. Therefore

$$\begin{aligned}
\log Q\left(\sigma_j; j \in A\right) &= \sum_{\delta \subseteq A, \delta \neq \emptyset} (-1)^{|\delta|} \sum_{\substack{I \subseteq \delta: \\ I \neq \emptyset}} \sum_{\substack{X: \\ \mathrm{Int}(X) \cap \delta = I}} (\Psi_\beta^I(X) - \Psi_\beta(X)) \\
&= \sum_{\substack{I \subseteq A: \\ I \neq \emptyset}} \sum_{\substack{X: \\ I \subseteq \mathrm{Int}(X)}} \left[(\Psi_\beta^I(X) - \Psi_\beta(X)) \sum_{\substack{\delta: \\ I \subseteq \delta \subseteq (I \cup (A \setminus \mathrm{Int}(X)))}} (-1)^{|\delta|} \right].
\end{aligned}$$

But the last sum is equal to 0 unless A is contained in $\mathrm{Int}(X)$, in which case it is $(-1)^{|I|}$. Therefore we obtain

$$\log Q\left(\sigma_j; j \in A\right) = \sum_{\substack{X: \\ A \subseteq \mathrm{Int}(X)}} \sum_{\substack{I \subseteq A: \\ I \neq \emptyset}} (-1)^{|I|} (\Psi_\beta^I(X) - \Psi_\beta(X)).$$

Finally, there are $2^r - 1$ non-empty subsets of A, and $|\Psi^I_\beta(X)| \leq \overline{\Psi}_\beta(X) := |\varphi(X)|e^{-2\beta\sum_{j=1}^r |\gamma_j|}$ for all X and I, thus

$$|\log Q\,(\sigma_j; j \in A)| \leq \sum_{\substack{X:\\ A\subseteq \mathrm{Int}(X)}} 2(2^r-1)\overline{\Psi}_\beta(X). \tag{3.7}$$

We conclude by using a trick similar to Lemma 3.6. By [13, Equation (5.29)], if $\beta \geq \beta^{\mathrm{lt}}_{\mathrm{ce}}(d)$, for any face f of any $\mathcal{S}_i$ ($i \in \mathbb{Z}^d$), we have the bound

$$\sum_{X:\overline{X}\ni f} \overline{\Psi}_\beta(X) \leq 1.$$

Thus if $\beta \geq 2\beta^{\mathrm{lt}}_{\mathrm{ce}}(d)$,

$$\sum_{X:\overline{X}\ni f} \overline{\Psi}_\beta(X)e^{\beta|\overline{X}|} \leq \sum_{X:\overline{X}\ni f} \overline{\Psi}_{\frac{\beta}{2}}(X) \leq 1.$$

So for any positive integer R,

$$\sum_{\substack{X:\overline{X}\ni f\\ |\overline{X}|\geq R}} \overline{\Psi}_\beta(X) \leq e^{-\beta R} \sum_{X:\overline{X}\ni f} \overline{\Psi}_\beta(X)e^{\beta|\overline{X}|} \leq e^{-\beta R}. \tag{3.8}$$

Let us now turn back to Eq. (3.7). Every cluster X such that $\varphi(X) \neq 0$ and $A \subseteq \mathrm{Int}(X)$ satisfies $|\overline{X}| \geq 2\ell_T(A)$. Moreover, if a cluster of size R has $j_1 \in A$ in its interior, then it contains at least a face f which is at distance at most R of j_1. There are at most CR^d such points, for some constant $C = C(d)$, therefore

$$\begin{aligned}\sum_{\substack{X:\\ A\subseteq \mathrm{Int}(X)}} \overline{\Psi}_\beta(X) &\leq \sum_{R\geq 2\ell_T(A)} CR^d \left[\sum_{\substack{X:\overline{X}\ni f\\ |\overline{X}|=R}} \overline{\Psi}_\beta(X)\right] \\ &\leq C \sum_{R\geq 2\ell_T(A)} R^d e^{-\beta R} \leq C \\ &\leq C' e^{-c\beta\ell_T(A)},\end{aligned}$$

for β large enough, where C' and c are some positive constants depending on the dimension d of the ambient space. Thus by (3.7),

$$|\log Q\,(\sigma_j; j \in A)| \leq C'_r e^{-c\beta\ell_T(A)},$$

for some positive constant C'_r depending on r and d. Exponentiating completes the proof. □

Now we can convert this estimate into the desired bound on joint cumulants.

Proposition 3.10. *Let A be a finite subset of $\mathbb{Z}^d$ of size r. Then, for β large enough,*

$$|\kappa_{\beta,0}\,(\sigma_j; j \in A)\,| \leq D_r e^{\frac{-c\beta\ell_T(A)}{2}},$$

where $c = c(d)$ is given by Lemma 3.9 and D_r is a positive constant depending on r.

Proof: By Lemma 3.9,

$$Q_{\beta,0}\,(\sigma_j; j \in A) = 1 + \mathcal{O}(\mathcal{M}\big(G[A]\big)),$$

where G is the weighted graph defined on $\mathbb{Z}^d$ such that for each $e = (i,j)$, one has $w_e = e^{\frac{-c\beta\|i-j\|_1}{2}}$. Indeed in that case $\mathcal{M}\big(G[A]\big) = e^{\frac{-c\beta\ell'_T(A)}{2}} \geq e^{-c\beta\ell_T(A)}$; see the discussion in Section 2.2.

Then using [10, Proposition 5.8], we deduce that

$$|\kappa_{\beta,0}\,(\sigma_j; j \in A)\,| = \prod_{j\in A}\langle\sigma_j\rangle^+_{\beta,0} \times \mathcal{O}(\mathcal{M}\big(G[A]\big)).$$

The first factor is trivially bounded by 1, while the second is bounded using Proposition 2.1:

$$\mathcal{M}\big(G[A]\big) = e^{\frac{-c_d\beta\ell'_T(A)}{2}} \leq e^{\frac{-c_d\beta\ell_T(A)}{2}}.$$

Thus $|\kappa_{\beta,0}\left(\sigma_j; j \in A\right)| \leq D_r e^{\frac{-c_d \beta \ell_T(A)}{2}}$, as claimed. $\square$

3.3. **With a strong magnetic field.** The last regime we consider is the Ising model with a strong magnetic field, *i.e.* h is bigger than some value $h_1 > 0$ (h_1 is to be determined later). The case of negative h (smaller than $-h_1 < 0$) is obviously symmetric.

In this regime, there is also a well-known cluster expansion for the partition function [13, Section 5.7]. Let us present it briefly.

Fix $\Lambda \subset \mathbb{Z}^d$ and consider the Ising model on Λ with + boundary conditions. We first write its partition function in a suitable form. For a subset Λ^- of Λ, we denote

$$\delta_e \Lambda^- = \{\{i,j\} \in \mathcal{E}^b_\Lambda,\ i \in \Lambda^-, j \notin \Lambda^-\}.$$

Define also $\mathrm{W}(\Lambda^-) = \exp(-2\beta\,|\delta_e \Lambda^-| - 2h|\Lambda^-|)$. Then we have:

Lemma 3.11 (strong magnetic field representation)**.** *With the above notation,*

$$Z^+_{\Lambda,\beta,h} = \exp\left(\beta|\mathcal{E}^b_\Lambda| + h|\Lambda|\right)\left(\sum_{\Lambda^- \subseteq \Lambda} \mathrm{W}(\Lambda^-)\right).$$

Proof: The proof is not difficult and can be found, *e.g.*, in [13, Section 5.7]. It is important to note that the sum over $\Lambda^- \subseteq \Lambda$ corresponds to the sum over spin configurations in the definition of the partition function: the correspondence simply associates with a spin configuration the set Λ^- of positions of its minus spins. $\square$

Let A be a subset of Λ. It is straightforward to modify the argument to find a similar expression for the numerator of $\langle \sigma_A \rangle$ (as in Section 3.2) or of $\left\langle \exp\left(\sum_{i \in A} t_i \sigma_i\right)\right\rangle$ (as in Section 3.1).

$$\sum_{\omega \in \Omega_\Lambda} \sigma_A(\omega) \exp\left(-H^+_{\Lambda;\beta,h}(\omega)\right) = \exp\left(\beta|\mathcal{E}^b_\Lambda| + h|\Lambda|\right)\left(\sum_{\Lambda^- \subseteq \Lambda} (-1)^{|A \cap \Lambda^-|}\, \mathrm{W}(\Lambda^-)\right). \tag{3.9}$$

$$\sum_{\omega \in \Omega_\Lambda} \exp\left(-H^+_{\Lambda;\beta,h}(\omega) + \sum_{i \in A} t_i \sigma_i\right) = \exp\left(\beta|\mathcal{E}^b_\Lambda| + h|\Lambda| + \sum_{i \in A} t_i\right)\left(\sum_{\Lambda^- \subseteq \Lambda}\left[\prod_{i \in A \cap \Lambda^-} \exp(-2t_i)\right] \mathrm{W}(\Lambda^-)\right). \tag{3.10}$$

A set $\Lambda^- \subseteq \Lambda$ can be seen as a subgraph of the lattice $\mathbb{Z}^d$. As such, it admits a unique decomposition as disjoint union of its connected components $\Lambda^- = S_1 \sqcup S_2 \sqcup \cdots \sqcup S_r$. The weight W behaves multiplicatively with respect to this decomposition $\mathrm{W}(\Lambda^-) = \prod_{i=1}^r \mathrm{W}(S_i)$; the same is true for the modified weights $(-1)^{|A \cap \Lambda^-|}\,\mathrm{W}(\Lambda^-)$ and $\left[\prod_{i \in A \cap \Lambda^-} \exp(-2t_i)\right] \mathrm{W}(\Lambda^-)$ which appear in Eqs. (3.9) and (3.10) above. This enables us to use the technique of cluster expansion.

The convergence of this cluster expansion is proved for the partition function in [13, Section 5.7]. The argument can be directly adapted to get a cluster expansion of the expression in Eqs. (3.9) and (3.10) above. The same reasoning as in Section 3.1 or in Section 3.2 leads to similar bounds on joint cumulants, which proves Theorem 1.1 in the strong magnetic field regime.

4. Weighted dependency graphs and central limit theorems

We will now use the bounds on cumulants obtained in the previous section to show that the family of random variables $\{\sigma_i : i \in \mathbb{Z}^d\}$ has a weighted dependency graph, and we will use this fact to deduce central limit theorems. We consider any of the regimes studied in the previous section: strong magnetic field, very high temperature, or very low temperature with + boundary condition. To have uniform notation, we omit from now on the notation of the boundary condition in low temperature.

4.1. Weighted dependency graph for the σ_i's and central limit theorem for the magnetization.

4.1.1. *The weighted dependency graph.* We start by proving Theorem 1.5, which gives a weighted dependency graph for $\{\sigma_i : i \in \mathbb{Z}^d\}$.

Proof of Theorem 1.5: Let $B = \{i_1, ..., i_r\}$ be a multiset of elements of $\mathbb{Z}^d$ and consider the induced subgraph $G[B]$. Then the maximum weight $\mathcal{M}(G[B])$ of a spanning tree in $G[B]$ satisfies

$$\mathcal{M}(G[B]) = \varepsilon^{\frac{\ell'_T(B)}{2}}.$$

Thus by Proposition 2.1,

$$\varepsilon^{\ell_T(B)} \leq \mathcal{M}(G[B]) \leq \varepsilon^{\frac{\ell_T(B)}{2}}. \tag{4.1}$$

By Proposition 5.2 of [10], it is sufficient to show that

$$\left|\kappa_{\beta,h}\left(\prod_{\alpha\in B_1}\sigma_\alpha, \dots, \prod_{\alpha\in B_k}\sigma_\alpha\right)\right| \leq D_r\mathcal{M}(G[B]),$$

for some sequence $\mathbf{D} = (D_r)_{r\geq 1}$, where $B_1, \dots, B_k$ are the vertex-sets of the connected components of $G_1[B]$, which is the graph induced by edges of weight 1 of G on B.

The vertices i and j are connected in G_1 if and only if $i = j$, because of the definition of the weights w_e in G. Consequently, each multiset in the list $B_1, \dots, B_k$ consists of a single spin σ_i, possibly repeated several times. The random variables σ_i taking value $+1$ or -1, we have

$$\sigma_i^j = \begin{cases} \sigma_i \text{ if } j \text{ odd,} \\ 1 \text{ if } j \text{ even.} \end{cases}$$

Therefore it is sufficient to prove that for any set B' of distinct $i_1, \dots, i_r$,

$$|\kappa_{\beta,h}(\sigma_{i_1}, \dots, \sigma_{i_r})| \leq D_r\mathcal{M}(G[B']). \tag{4.2}$$

But by Theorem 1.1,

$$|\kappa_{\beta,h}(\sigma_{i_1}, \dots, \sigma_{i_r})| \leq D_r\varepsilon^{\ell_T(B')},$$

for some sequence $\mathbf{D}$ depending only on r. Thus using (4.1), Equation (4.2) is proved, which completes the proof of the theorem. □

4.1.2. *The central limit theorem for the magnetization.* As a first application, we use the weighted dependency graph from last section to obtain the well-known central limit theorem for the magnetization (for an early reference, see [30]), in the regime $(\beta, h) \in R(\beta_1, \beta_2, h_1)$.

We consider the Ising model on $\mathbb{Z}^d$, with inverse temperature β and magnetic field h. For any positive integer n, we define $\Lambda_n := [-n, n]^d$ the d-dimensional cube centred at 0 of side $2n$. We define the *magnetization*

$$S_n := \sum_{i\in\Lambda_n}\sigma_i,$$

and let v_n^2 denote the variance of S_n. Let us further define the covariance

$$\langle\sigma_i;\sigma_j\rangle_{\beta,h} := \langle\sigma_i\sigma_j\rangle_{\beta,h} - \langle\sigma_i\rangle_{\beta,h}\langle\sigma_j\rangle_{\beta,h}.$$

Theorem 4.1. *Consider the Ising model on $\mathbb{Z}^d$, with inverse temperature β and magnetic field h, such that $(\beta, h) \in R(\beta_1, \beta_2, h_1)$. Then there exists $v = v(\beta, h, d)$ such that*

$$\frac{S_n - \mathbb{E}(S_n)}{\sqrt{|\Lambda_n|}} \xrightarrow[n\to\infty]{d} \mathcal{N}(0, v^2).$$

Moreover $v^2 > 0$ so the Gaussian law is non-degenerate.

We start by a lemma of [8] on the asymptotics of the variance of S_n.

Lemma 4.2. [8, Lemma V.7.1] *The limit*

$$v^2 := \lim_{n\to\infty} \frac{v_n^2}{|\Lambda_n|}$$

exists as an extended real valued number and

$$v^2 = \sum_{i\in\mathbb{Z}^d} \langle \sigma_0; \sigma_i \rangle_{\beta,h}.$$

But by Theorem 1.1, in the regimes we consider, the cumulants (so in particular the covariance) are exponentially small, so the sum is absolutely convergent and we actually have the stronger statement:

Corollary 4.3. *Suppose that* $(\beta, h) \in R(\beta_1, \beta_2, h_1)$. *The limit*

$$v^2 := \lim_{n\to\infty} \frac{v_n^2}{|\Lambda_n|} = \sum_{i\in\mathbb{Z}^d} \langle \sigma_0; \sigma_i \rangle_{\beta,h}$$

is finite.

Proof of Theorem 4.1: We will use Theorem 1.6. Let G be the weighted dependency graph defined in Theorem 1.5. Then for all n, $G[\Lambda_n]$ is a $\mathbf{C}$-weighted dependency graph for $\{\sigma_i; i \in \Lambda_n\}$. The number of vertices of $G[\Lambda_n]$ is

$$N_n = |\Lambda_n| = (2n+1)^d,$$

and its maximal weighted degree is

$$\Delta_n - 1 = \max_{i\in\Lambda_n} \sum_{j\in\Lambda_n} \varepsilon^{\frac{\|i-j\|_1}{2}}.$$

There are $2^d\binom{d+y-1}{d-1}$ points at distance y of 0 in $\mathbb{Z}^d$. Indeed such a point has coordinates $(y_1, \dots, y_d)$ such that $|y_1| + \cdots + |y_d| = y$. There are $\binom{d+y-1}{d-1}$ choices for the values of $|y_1|, \dots, |y_d|$, and each y_i can be either positive or negative, which multiplies the number of choices by 2^d. Thus there are at most $2^d\binom{d+y-1}{d-1}$ points at distance y of any point x in Λ_n, and

$$\Delta_n - 1 \le \sum_{y=0}^{2dn} \varepsilon^{\frac{y}{2}} 2^d \binom{d+y-1}{d-1} \le C,$$

for some constant C because the infinite series is absolutely convergent.

We now have to find a sequence (a_n) and integers s and v such that conditions (1)-(3) of Theorem 1.6 are satisfied. We set for all n, $a_n = \sqrt{|\Lambda_n|} = (2n+1)^{\frac{d}{2}}$, $v = \sqrt{\sum_{i\in\mathbb{Z}^d} \langle \sigma_0; \sigma_i \rangle_{\beta,h}}$ as in Lemma 4.2 and we can choose s to be any integer ≥ 3.

Now condition (1) is satisfied because of Lemma 4.2, as

$$\frac{v_n^2}{|\Lambda_n|} = \frac{v_n^2}{a_n^2} \xrightarrow[n\to\infty]{} v^2.$$

Condition (2) is also satisfied as $a_n^2 = (2n+1)^d = N_n$.

Finally, for some constant C',

$$\left(\frac{N_n}{\Delta_n}\right)^{\frac{1}{s}} \frac{\Delta_n}{a_n} \le C' \frac{(2n+1)^{\frac{d}{s}}}{(2n+1)^{\frac{d}{2}}},$$

and the right-hand side tends to 0 as n tends to infinity for $s \ge 3$. So (3) is satisfied too.

The central limit theorem is proved.

Moreover, whatever the values of β and h are, the spin at 0 is not constant, thus $\langle \sigma_0, \sigma_0 \rangle_{\beta,h} > 0$. On the other hand, because of the GKS inequalities [17, 22], for all $i \in \mathbb{Z}^d$, $\langle \sigma_0; \sigma_i \rangle_{\beta,h} \ge 0$ (see e.g. [13]). Therefore,

$$v^2 = \langle \sigma_0, \sigma_0 \rangle_{\beta,h} + \sum_{i\in Z^d\setminus\{0\}} \langle \sigma_0; \sigma_i \rangle_{\beta,h} > 0,$$

which ends the proof of the theorem. □

4.2. **Central limit theorem for occurrences of given patterns.**

4.2.1. *Power of weighted dependency graphs.* A major advantage of the theory of weighted dependency graphs is that this structure is stable by taking *powers*.

Definition 4.4. Let G be an edge-weighted graph with vertex set A and weight function w; we also consider a positive integer m. We denote by $\mathrm{MSet}_{\leq m}(A)$ the set of multisets of elements of A with cardinality at most m. Then the m-th power G^m of G is by definition the graph with vertex-set $\mathrm{MSet}_{\leq m}(A)$ and where the weight between I and J is given by $w_m(I,J) = \max_{i\in I, j\in J} w(i,j)$. As usual, edges not in the graph should be seen as edges of weight 0.

This definition is justified by the following property, proved in [10, Section 5.3].

Proposition 4.5. *Let $\{Y_a, a \in A\}$ be a family of random variables with a $\boldsymbol{C}$-weighted dependency graph G. Then G^m is a $\boldsymbol{D}$-weighted dependency graph for the family $\{Y_I, I \in \mathrm{MSet_{\leq m}(A)}\}$, where $Y_I = \prod_{a\in I} Y_a$, and $\boldsymbol{D}$ depends only on m and $\boldsymbol{C}$.*

Instead of applying this to the variables σ_i, we will rather work with the variables $X_{(i,+)} := X_i = \frac{1+\sigma_i}{2}$ and $X_{(i,-)} = 1 - X_i$. Start with the following observation. For all $A = \{(i_1,s_1),\dots,(i_r,s_r)\} \subseteq \mathbb{Z}^d \times \{+,-\}$, we have, for $r \geq 2$,

$$\left|\kappa_{\beta,h}(X_{(i_1,s_1)},\dots,X_{(i_r,s_r)})\right| = \tfrac{1}{2^r}\left|\kappa_{\beta,h}(\sigma_{i_1},\dots,\sigma_{i_r})\right|,$$

Mimicking the proof of Theorem 1.5, we obtain the following. Let G_s be the complete weighted graph with vertex set $\mathbb{Z}^d \times \{+,-\}$, such that for all $i,j \in \mathbb{Z}^d$,

$$w'((i,+),(j,+)) = w'((i,+),(j,-)) = \varepsilon^{\frac{1}{2}\|i-j\|_1}.$$

In other words, we ignore the sign and use the weight function w from the previous section. Then G_s is a $\mathbf{C}$-weighted dependency graph for the family $\{X_{(i,s)}; i \in \mathbb{Z}^d, s \in \{+,-\}\}$, for some sequence $\mathbf{C} = (C_r)_{r\geq 1}$.

By considering the powers of G_s and using Proposition 4.5, we obtain weighted dependency graphs for the products of $X_{(i,+)}$'s and $X_{(i,-)}$'s with a bounded number of terms.

Theorem 4.6. *Consider the Ising model on $\mathbb{Z}^d$, with inverse temperature β and magnetic field h, such that $(\beta,h) \in R(\beta_1,\beta_2,h_1)$. Let m be a fixed positive integer; for multisets I of elements of $\mathbb{Z}^d \times \{+,-\}$, we define $Z_I := \prod_{i\in I} X_i$. Then G_s^m is a $\mathbf{D}_m$-weighted dependency graph for the family of random variables $\{Z_I; I \in \mathrm{MSet_{\leq m}}\left(\mathbb{Z}^d \times \{+,-\}\right)\}$, for some sequence $\mathbf{D}_m$ depending only on m.*

4.2.2. *Local patterns.* In this section, we prove Theorem 1.8, that is the CLT for the number of occurrences of a given local pattern of spins (for example, the number of isolated + spins).

To find a weighted dependency graph for the potential occurrences of a pattern $\mathcal{P} = (\mathcal{D}, \mathfrak{s})$ of size m, we consider $G_{\mathcal{P}}$, the restriction of G_s^m to the Z_I's of the form

$$Z_i^{\mathcal{P}} = \prod_{j\in\mathcal{D}} X_{(i+j,\mathfrak{s}(j))}. \tag{4.3}$$

Note that vertices of $G_{\mathcal{P}}$ are canonically indexed by $i \in \mathbb{Z}^d$ so that we will think of $G_{\mathcal{P}}$ as a graph with vertex set $\mathbb{Z}^d$. The weight of the edge between i_1 and i_2 is then

$$w_{\mathcal{P}}(i_1,i_2) = \max_{j_1\in\mathcal{D}, j_2\in\mathcal{D}} w(i_1+j_1, i_2+j_2) \leq \varepsilon^{\frac{1}{2}(\|i_1-i_2\|_1 - \max_{\alpha,\beta\in\mathcal{D}}\|\alpha-\beta\|_1)}. \tag{4.4}$$

The graph $G_{\mathcal{P}}$ is a $\mathbf{D}$-weighted dependency graph for $\{Z_i, i \in \mathbb{Z}^d\}$, for some sequence $\mathbf{D}$ depending only on $\mathcal{P}$. Indeed, it is a restriction of the weighted dependency graph given in Theorem 4.6.

We define

$$S_{n,\mathcal{P}} := \sum_{i\in\Lambda_n} Z_i^{\mathcal{P}},$$

the number of occurrences of $\mathcal{P}$ whose position is in Λ_n. In the example of isolated + spins, we have

$$S_{n,\mathcal{P}} := \sum_{i \in \Lambda_n} \left(X_i \prod_{j : \| i-j \|_1 = 1} (1 - X_j) \right).$$

It is also easy to encode in this framework the number of + connected components of any given shape.

Let $v_{n,\mathcal{P}}^2$ denote the variance of $S_{n,\mathcal{P}}$. We have a lemma analogous to Lemma 4.2.

Lemma 4.7. *As n tends to infinity, the quantity $\frac{v_{n,\mathcal{P}}^2}{|\Lambda_n|}$ tends to $v_{\mathcal{P}}^2 := \sum_{k \in \mathbb{Z}^d} \langle Z_0^{\mathcal{P}}; Z_k^{\mathcal{P}} \rangle_{\beta,h} < \infty$.*

Proof: That $G_{\mathcal{P}}$ is a weighted dependency graph for the family $Z_i^{\mathcal{P}}$ implies that

$$\langle Z_0^{\mathcal{P}}; Z_k^{\mathcal{P}} \rangle_{\beta,h} \leq D_2 \varepsilon^{\frac{1}{2}(k - \max_{\alpha,\beta \in \mathcal{D}} \| \alpha - \beta \|_1)}.$$

This proves that $v_{\mathcal{P}}^2$ is finite as claimed.

Let $\varepsilon > 0$ be fixed. We want to show that for n large enough,

$$\left| \frac{v_{n,\mathcal{P}}^2}{|\Lambda_n|} - \sum_{k \in \mathbb{Z}^d} \langle Z_0^{\mathcal{P}}; Z_k^{\mathcal{P}} \rangle_{\beta,h} \right| \leq \varepsilon.$$

We have

$$\begin{aligned} \frac{v_{n,\mathcal{P}}^2}{|\Lambda_n|} &= \frac{1}{|\Lambda_n|} \sum_{i \in \Lambda_n} \sum_{j \in \Lambda_n} \langle Z_i^{\mathcal{P}}; Z_j^{\mathcal{P}} \rangle_{\beta,h} \\ &= \frac{1}{|\Lambda_n|} \sum_{i \in \Lambda_n} \sum_{j \in \mathbb{Z}^d} \langle Z_i^{\mathcal{P}}; Z_j^{\mathcal{P}} \rangle_{\beta,h} - \frac{1}{|\Lambda_n|} \sum_{i \in \Lambda_n} \sum_{j \in \mathbb{Z}^d \setminus \Lambda_n} \langle Z_i^{\mathcal{P}}; Z_j^{\mathcal{P}} \rangle_{\beta,h}. \end{aligned}$$

By translation invariance of $\langle Z_i^{\mathcal{P}}; Z_j^{\mathcal{P}} \rangle_{\beta,h}$, the first sum equals $\sum_{k \in \mathbb{Z}^d} \langle Z_0^{\mathcal{P}}; Z_k^{\mathcal{P}} \rangle_{\beta,h}$. Thus the only thing left to do is to show that for n large enough, the absolute value of the second term is bounded by ε. We cut the sum on i into two parts : the points that are far from the boundary of Λ_n and those which are not. Recall that the boundary $\partial \Lambda_n$ consists of points j not in Λ, which have a neighbour in Λ_n. We denote $\delta(i, \partial \Lambda_n) = \min_{j \in \partial \Lambda_n} \| i - j \|_1$, which is the distance between i and $\partial \Lambda_n$. For R a positive integer, let us consider the points $i \in \Lambda_n$ at distance more than R from the boundary of Λ_n. We have, again by translation invariance

$$\frac{1}{|\Lambda_n|} \sum_{\substack{i \in \Lambda_n \\ \delta(i, \partial \Lambda_n) > R}} \sum_{j \in \mathbb{Z}^d \setminus \Lambda_n} |\langle Z_i^{\mathcal{P}}; Z_j^{\mathcal{P}} \rangle_{\beta,h}| \leq \sum_{\substack{k \in Z^d \\ |k| > R}} |\langle Z_0^{\mathcal{P}}; Z_k^{\mathcal{P}} \rangle_{\beta,h}|.$$

But the series $\sum_{k \in \mathbb{Z}^d} |\langle Z_0^{\mathcal{P}}; Z_k^{\mathcal{P}} \rangle_{\beta,h}|$ is absolutely convergent so the sum above tends to 0 as R tends to infinity. Therefore, there exists some integer R_0 such that

$$\frac{1}{|\Lambda_n|} \sum_{\substack{i \in \Lambda_n \\ \delta(i, \partial \Lambda_n) > R_0}} \sum_{j \in \mathbb{Z}^d \setminus \Lambda_n} |\langle Z_i^{\mathcal{P}}; Z_j^{\mathcal{P}} \rangle_{\beta,h}| \leq \frac{\varepsilon}{2}. \tag{4.5}$$

Now let us consider the points of Λ_n that are at distance at most R_0 of $\partial \Lambda_n$. There are at most $C |\partial \Lambda_n| R_0^d$ such points. Therefore

$$\frac{1}{|\Lambda_n|} \sum_{\substack{i \in \Lambda_n \\ \delta(i, \partial \Lambda_n) \leq R_0}} \sum_{j \in \mathbb{Z}^d \setminus \Lambda_n} |\langle Z_i^{\mathcal{P}}; Z_j^{\mathcal{P}} \rangle_{\beta,h}| \leq \frac{1}{|\Lambda_n|} \sum_{\substack{i \in \Lambda_n \\ \delta(i, \partial \Lambda_n) \leq R_0}} \sum_{k \in \mathbb{Z}^d} |\langle Z_0^{\mathcal{P}}; Z_k^{\mathcal{P}} \rangle_{\beta,h}| \leq C' \frac{|\partial \Lambda_n|}{|\Lambda_n|} R_0^d.$$

But as n tends to ∞, $\frac{|\partial \Lambda_n|}{|\Lambda_n|}$ tends to 0. Therefore for n large enough,

$$\frac{1}{|\Lambda_n|} \sum_{\substack{i \in \Lambda_n \\ \delta(i, \partial \Lambda_n) \leq R_0}} \sum_{j \in \mathbb{Z}^d \setminus \Lambda_n} |\langle Z_i^{\mathcal{P}}; Z_j^{\mathcal{P}} \rangle_{\beta,h}| \leq \frac{\varepsilon}{2}. \tag{4.6}$$

Adding (4.5) and (4.6) completes the proof. □

We are now ready to prove the central limit theorem.

Proof of Theorem 1.8: We proceed as in the proof of Theorem 4.1. We consider $G_{\mathcal{P}}[\Lambda_n]$.

The number of vertices is $N_n = |\Lambda'_n| = |\Lambda_n|$ and, from (4.4), its maximal weighted degree $\Delta_n - 1$ is bounded as follows:

$$\Delta_n - 1 \leq \max_{i\in\Lambda_n} \sum_{j\in\Lambda_n} \varepsilon^{\frac{1}{2}(\|i-j\|_1 - \max_{\alpha,\beta\in\mathcal{D}}\|\alpha-\beta\|_1)} \leq \max_{i\in\Lambda_n} \sum_{j\in\Lambda_n} C_{\mathcal{P}}\varepsilon^{\frac{1}{2}\|i-j\|_1},$$

where $C_{\mathcal{P}}$ is a positive constant depending only on the pattern $\mathcal{P}$. Thus by the same argument as in the proof of Theorem 4.1, $\Delta_n - 1 \leq C'_{\mathcal{P}}$, for some other constant $C'_{\mathcal{P}}$.

Again we set for all n, $a_n = \sqrt{|\Lambda_n|}$. We also set $v = v_{\mathcal{P}}$ as in Lemma 4.7 and we can choose s to be any integer ≥ 3.

Conditions (1) to (3) of Theorem 1.6 are satisfied again and the theorem is proved. □

Remark 4.8. The variance $v_{\mathcal{P}}$ appearing in Theorem 1.8 might be equal to 0 for some patterns $\mathcal{P}$, in which case the central limit theorem is degenerate. If the pattern has only plus spins, the same proof as before gives $v_{\mathcal{P}} > 0$.

4.2.3. *Global patterns.* In this final section, we establish Theorem 1.9, the central limit theorem for the number of occurrences of a global pattern of spins.

To find a weighted dependency graph for the potential occurrences of $\tilde{\mathcal{P}}$ of size m, we consider $G_{\tilde{\mathcal{P}}}$, the restriction of G_s^m to the Z_I's of the form

$$Z^{\tilde{\mathcal{P}}}_{\{x^{(1)},\ldots,x^{(m)}\}} = \prod_{i=1}^{m} X_{(x^{(i)},\mathfrak{s}(i))}. \tag{4.7}$$

In $G_{\tilde{\mathcal{P}}}$, the weight of the edge between $\{x^{(1)},\ldots,x^{(m)}\}$ and $\{y^{(1)},\ldots,y^{(m)}\}$ is given by

$$w_{\tilde{\mathcal{P}}}(\{x^{(1)},\ldots,x^{(m)}\},\{y^{(1)},\ldots,y^{(m)}\}) = \max_{i,j\in\{1,\ldots,m\}} w(x^{(i)},y^{(j)}) = \varepsilon^{\frac{1}{2}\min_{i,j\in\{1,\ldots,m\}}\|x^{(i)}-y^{(j)}\|_1}.$$

Again, the graph $G_{\tilde{\mathcal{P}}}$ is a $\mathbf{D}$-weighted dependency graph for $\{Z^{\tilde{\mathcal{P}}}_{\{x^{(1)},\ldots,x^{(m)}\}}, \{x^{(1)},\ldots,x^{(m)}\} \subset \mathbb{Z}^d\}$, for some sequence $\mathbf{D}$ depending only on $\mathcal{P}$ as it is a restriction of the weighted dependency graph given in Theorem 4.6.

Now define

$$S_{n,\tilde{\mathcal{P}}} := \sum_{\{x^{(1)},\ldots,x^{(m)}\}\subset\Lambda_n} Z^{\tilde{\mathcal{P}}}_{\{x^{(1)},\ldots,x^{(m)}\}},$$

the number of occurrences of $\tilde{\mathcal{P}}$ in Λ_n. Let $v^2_{n,\tilde{\mathcal{P}}}$ denote the variance of $S_{n,\tilde{\mathcal{P}}}$.

Proof of Theorem 1.9: Consider the weighted dependency graph $G_{\tilde{\mathcal{P}}}[\Lambda_n]$. Its number of vertices is $N_n^m = |\Lambda_n|^m$. Let us now bound its maximal weighted degree $\Delta_n - 1$. Fix $\{x^{(1)},\ldots,x^{(m)}\} \subset \Lambda_n$. We have

$$\begin{aligned}\sum_{\{y^{(1)},\ldots,y^{(m)}\}\subset\Lambda_n} \varepsilon^{\frac{1}{2}\min_{i,j\in\{1,\ldots,m\}}\|x^{(i)}-y^{(j)}\|_1} &\leq \sum_{y^{(1)},\ldots,y^{(m)}\in\Lambda_n} \sum_{i=1}^{m}\sum_{j=1}^{m} \varepsilon^{\frac{1}{2}\|x^{(i)}-y^{(j)}\|_1} \\ &\leq \sum_{i=1}^{m} m|\Lambda_n|^{m-1} \sum_{y\in\Lambda_n} \varepsilon^{\frac{1}{2}\|x^{(i)}-y\|_1}.\end{aligned}$$

By the proof of Theorem 4.1, the last sum is bounded by a certain constant C. Thus

$$\Delta_n - 1 = \max_{\{x^{(1)},\ldots,x^{(m)}\}\subset\Lambda_n} \sum_{y^{(1)},\ldots,y^{(m)}\in\Lambda_n} \varepsilon^{\frac{1}{2}\min_{i,j\in\{1,\ldots,m\}}\|x^{(i)}-y^{(j)}\|_1} \leq m^2|\Lambda_n|^{m-1}C.$$

We want to apply Theorem 1.6 and set $a_n = \sqrt{v^2_{n,\tilde{\mathcal{P}}}}$. Condition (1) is trivial, while (2) holds for all weighted dependency graphs when a_n is the standard deviation of X_n (see [10, Lemma 4.10]).

Condition (3) is fulfilled since, using (1.2) and the above inequality for Δ_n,

$$\left(\frac{N_n}{\Delta_n}\right)^{\frac{1}{s}} \frac{\Delta_n}{a_n} \leq \left(\frac{|\Lambda_n|^m}{m^2|\Lambda_n|^{m-1}C}\right)^{\frac{1}{s}} \frac{m^2|\Lambda_n|^{m-1}C}{\sqrt{A|\Lambda_n|^{2m-2+\eta}}} \leq C'|\Lambda_n|^{1/s-\eta/2}$$

for some constant C', and, for s large enough, the right-hand side tends to 0 as n tends to infinity. □

We now show a simple sufficient condition – the pattern consisting of positive spins only – so that the bound (1.2) on the variance is satisfied. We start with a lemma.

Lemma 4.9. *Fix $m \geq 2$. There exist some constants $R > 0$ and $B > 0$ such that the following holds. For any lists $(x^{(1)}, \dots, x^{(m)})$ and $(y^{(1)}, \dots, y^{(m)})$ such that $x^{(1)} = y^{(1)}$ but no two elements in the set $\{x^{(1)}, \dots, x^{(m)}, y^{(2)}, \dots, y^{(m)}\}$ are at distance less than R from each other, we have*

$$\mathrm{Cov}\left(\prod_{i=1}^m X_{(x^{(i)},+)}, \prod_{i=1}^m X_{(x^{(i)},+)}\right) \geq B.$$

Proof: By definition, and using that $X_{(x^{(1)},+)}X_{(y^{(1)},+)} = X^2_{(x^{(1)},+)} = X_{(x^{(1)},+)}$, we have

$$\mathrm{Cov}\left(\prod_{i=1}^m X_{(x^{(i)},+)}, \prod_{i=1}^m X_{(x^{(i)},+)}\right) = \mathbb{E}\big[X_{(x^{(1)},+)} \cdots X_{(x^{(m)},+)} X_{(y^{(2)},+)} \cdots X_{(y^{(m)},+)}\big]$$
$$- \mathbb{E}\big[X_{(x^{(1)},+)} \cdots X_{(x^{(m)},+)}\big] \mathbb{E}\big[X_{(x^{(1)},+)} X_{(y^{(2)},+)} \cdots X_{(y^{(m)},+)}\big].$$

Using the expression of joint moments in terms of cumulants – see, e.g. [10, Eq. (3)] – and the bound for cumulants of spins (Theorem 1.1), we have that there exists a constant C_m such that

$$\big|\mathbb{E}\big[X_{(x^{(1)},+)} \cdots X_{(x^{(m)},+)}\big] - \mathbb{E}\big[X_{(x^{(1)},+)}\big] \cdots \mathbb{E}\big[X_{(x^{(m)},+)}\big]\big| \leq C_m \varepsilon^R,$$

whenever the $x^{(i)}$ all lie at distance at least R from each other. The same holds for the other joint moments in the above expression for the covariance and we get

$$\mathrm{Cov}\left(\prod_{i=1}^m X_{(x^{(i)},+)}, \prod_{i=1}^m X_{(x^{(i)},+)}\right) = \left(\mathbb{E}\big[X_{(x^{(1)},+)}\big] - \mathbb{E}\big[X_{(x^{(1)},+)}\big]^2\right)$$
$$\cdot\, \mathbb{E}\big[X_{(x^{(2)},+)}\big] \cdots \mathbb{E}\big[X_{(x^{(m)},+)}\big] \mathbb{E}\big[X_{(y^{(2)},+)}\big] \cdots \mathbb{E}\big[X_{(y^{(m)},+)}\big] + \text{error},$$

where the error is uniformly bounded by $C_m \varepsilon^R$. The main term in the above equation is positive (as a product of positive terms) and independent from the $x^{(i)}$ and the $y^{(i)}$ (by translation invariance), while the error can be made as small as wanted by making R tend to infinity. This proves the lemma. □

Proposition 4.10. *Let $\tilde{\mathcal{P}}$ be a global pattern of size m and assume that the function $\mathfrak{s}$ defining $\tilde{\mathcal{P}}$ takes only value $+1$. Then there exists a constant A such that $\mathrm{Var}(S_{n,\tilde{\mathcal{P}}}) \geq A|\Lambda_n|^{2m-1}$.*

Proof: We expand the variance as

$$\mathrm{Var}(S_{n,\tilde{\mathcal{P}}}) = \sum_{\substack{\{x^{(1)},\dots,x^{(m)}\} \subset \Lambda_n \\ \{y^{(1)},\dots,y^{(m)}\} \subset \Lambda_n}} \mathrm{Cov}\left(Z^{\tilde{\mathcal{P}}}_{\{x^{(1)},\dots,x^{(m)}\}}, Z^{\tilde{\mathcal{P}}}_{\{y^{(1)},\dots,y^{(m)}\}}\right).$$

When $\tilde{\mathcal{P}}$ involves only positive spins, the FKG inequalities ensure that all summands are positive. Restricting the sum to sets with an ordering that fulfils the hypothesis of Lemma 4.9 gives a lower bound. Therefore $\mathrm{Var}(S_{n,\tilde{\mathcal{P}}}) \geq B \cdot N_1$, where N_1 is the number of pairs of sets $(\{x^{(1)}, \dots, x^{(m)}\}, \{y^{(1)}, \dots, y^{(m)}\})$ as in Lemma 4.9. For fixed $R > 0$, this number is clearly of order $|\Lambda_n|^{2m-1}$, finishing the proof of the proposition. □

Acknowledgements

The authors are grateful to H. Duminil-Copin, R. Kotecký and D. Ueltschi for discussions on cluster expansions and bounds on cumulants in the Ising model.

They are also indebted to anonymous referees whose suggestions improved significantly the presentation of the paper. One of the referees also suggested the extension of the very high temperature regime from $h = 0$ (as it was in a previous version) to $|h| \leq h_1$ (as it is now), making our results slightly more general.

Both authors have been partially supported by the Swiss National Science Fundation (SNSF), grant number 200021-149461.

References

[1] P. Baldi and Y. Rinott, *On normal approximations of distributions in terms of dependency graphs*, Ann. Prob., (1989), pp. 1646–1650.
[2] M. Bóna, *On three different notions of monotone subsequences*, in Permutation Patterns, vol. 376 of London Math. Soc. Lecture Note Series, Cambridge University Press, 2010, pp. 89–113.
[3] J. Bourdon and B. Vallée, *Generalized pattern matching statistics*, in Mathematics and Computer Science II, Springer, 2002, pp. 249–265.
[4] B. Chern, P. Diaconis, D. Kane, and R. Rhoades, *Central limit theorems for some set partition statistics*, Adv. Appl. Math., 70 (2015), pp. 92–105.
[5] M. Duneau, D. Iagolnitzer, and B. Souillard, *Decrease properties of truncated correlation functions and analyticity properties for classical lattices and continuous systems*, Comm. Math. Phys., 31 (1973), pp. 191–208.
[6] ———, *Strong cluster properties for classical systems with finite range interaction*, Comm. Math. Phys., 35 (1974), pp. 307–320.
[7] S. Elizalde, H. Crane, and S. DeSalvo, *The probability of avoiding consecutive patterns in the Mallows distribution.* preprint arXiv:1609.01370, 2016.
[8] R. Ellis, *Entropy, Large Deviations, and Statistical Mechanics*, Springer, 1985.
[9] R. A. Farrell, T. Morita, and P. H. E. Meijer, *Cluster expansion for the Ising model*, J. Chem. Phys., 45 (1966), pp. 349–363.
[10] V. Féray, *Weighted dependency graphs.* arXiv preprint 1605.03836, 2016.
[11] V. Féray, P.-L. Méliot, and A. Nikeghbali, *Mod-ϕ convergence II: Estimates on the speed of convergence.* To appear in *Séminaire de Probabilités*, arXiv:1705.10485, 2017.
[12] P. Flajolet, W. Szpankowski, and B. Vallée, *Hidden word statistics*, J. ACM, 53 (2006), pp. 147–183.
[13] S. Friedli and Y. Velenik, *Statistical Mechanics of Lattice Systems: a Concrete Mathematical Introduction*, Cambridge University press, 2017.
[14] H.-O. Georgii, *Gibbs measures and phase transitions*, vol. 9 of De Gruyter Studies in Mathematics, De Gruyter, 2011. 2nd edn.
[15] R. Gheissari, E. Lubetzky, and Y. Peres, *Concentration inequalities for polynomials of contracting Ising models.* preprint arXiv:1706.00121, 2017.
[16] L. Goldstein, *Berry-Esseen bounds for combinatorial central limit theorems and pattern occurrences, using zero and size biasing*, J. Appl. Prob., 42 (2005), pp. 661–683.
[17] R. Griffiths, *Correlation in Ising ferromagnets I, II*, J. Math. Phys., 8 (1967), pp. 478–489.
[18] L. Hofer, *A central limit theorem for vincular permutation patterns.* preprint arXiv:1704.00650, 2017.
[19] E. Ising, *Contribution to the theory of ferromagnetism*, Z. Phys., 31 (1925), pp. 253–258.
[20] S. Janson, *Normal convergence by higher semiinvariants with applications to sums of dependent random variables and random graphs*, Ann. Prob., 16 (1988), pp. 305–312.
[21] S. Janson, B. Nakamura, and D. Zeilberger, *On the asymptotic statistics of the number of occurrences of multiple permutation patterns*, J. Comb., 6 (2015), pp. 117–143.
[22] D. Kelly and S. Sherman, *General Griffiths' inequalities on correlations in Ising ferromagnets*, J. Math. Phys., 9 (1968), pp. 466–484.
[23] R. Kindermann and L. Snell, *Markov random fields and their applications*, vol. 1 of Contemporary Mathematics, American Mathematical Society, Providence, R.I., 1980.
[24] J. L. Lebowitz, *Bounds on the correlations and analyticity properties of ferromagnetic Ising spin systems*, Comm. Math. Phys., 28 (1972), pp. 313–321.
[25] T. Lee and C. Yang, *Statistical theory of equations of state and phase transitions. II. Lattice gas and Ising model*, Phys. Rev., 87 (1952), pp. 410–419.
[26] V. Malyshev and R. Minlos, *Gibbs random fields*, Springer, 1991.
[27] A. Martin-Löf, *Mixing properties, differentiability of the free energy and the central limit theorem for a pure phase in the Ising model at low temperature*, Comm. Math. Phys., 32 (1973), pp. 75–92.
[28] J. E. Mayer and E. Montroll, *Molecular distribution*, J. Chem. Phys., 9 (1941), pp. 2–16.

[29] C. Neaderhouser, *Limit theorems for multiply indexed mixing random variables, with application to Gibbs random fields*, Ann. Probab., 6 (1978), pp. 207–215.
[30] C. Newman, *Normal fluctuations and the FKG inequalities*, Comm. Math. Phys., 74 (1980), pp. 119–128.
[31] ———, *A general central limit theorem for FKG systems*, Comm. Math. Phys., 91 (1983), pp. 75–80.
[32] L. Onsager, *Crystal statistics, I. A two-dimensional model with an order-disorder transition*, Phys. Rev., 65 (1944), pp. 117–149.
[33] R. Peierls, *On Ising's ferromagnet model*, Proc. Camb. Phil. Soc., 32 (1936), pp. 477–481.
[34] M. Régnier and W. Szpankowski, *On pattern frequency occurrences in a Markovian sequence*, Algorithmica, 22 (1998), pp. 631–649.
[35] H. D. Ursell, *The evaluation of Gibbs' phase-integral for imperfect gases*, Math. Proc. Cambridge Phil. Soc., 23 (1927), pp. 685–697.
[36] C. Yang, *The spontaneous magnetization of a two-dimensional Ising model*, Phys. Rev. (2), 85 (1952), pp. 808–816.

Institut für Mathematik, Universität Zürich, Winterthurerstr. 190, CH-8032 Zürich, Switzerland
Email address: `jehanne.dousse@math.uzh.ch, valentin.feray@math.uzh.ch`