

Sharp phase transition for the random-cluster and Potts models via decision trees

Hugo Duminil-Copin^{*†}, Aran Raoufi[†], Vincent Tassion[‡]

December 27, 2018

Abstract

We prove an inequality on decision trees on monotonic measures which generalizes the OSSS inequality on product spaces. As an application, we use this inequality to prove a number of new results on lattice spin models and their random-cluster representations. More precisely, we prove that

- For the Potts model on transitive graphs, correlations decay exponentially fast for $\beta < \beta_c$.
- For the random-cluster model with cluster weight $q \geq 1$ on transitive graphs, correlations decay exponentially fast in the subcritical regime and the cluster-density satisfies the mean-field lower bound in the supercritical regime.
- For the random-cluster models with cluster weight $q \geq 1$ on planar quasi-transitive graphs \mathbb{G} ,

$$\frac{p_c(\mathbb{G})p_c(\mathbb{G}^*)}{(1 - p_c(\mathbb{G}))(1 - p_c(\mathbb{G}^*))} = q.$$

As a special case, we obtain the value of the critical point for the square, triangular and hexagonal lattices (this provides a short proof of the result of [BD12]).

These results have many applications for the understanding of the subcritical (respectively disordered) phase of all these models. The techniques developed in this paper have potential to be extended to a wide class of models including the Ashkin-Teller model, continuum percolation models such as Voronoi percolation and Boolean percolation, super-level sets of massive Gaussian Free Field, and random-cluster and Potts model with infinite range interactions.

1 Introduction

1.1 OSSS inequality for monotonic measures

In theoretical computer science, determining the computational complexity of tasks is a very difficult problem (think of P against NP). To start with a more tractable problem, computer scientists have studied *decision trees*, which are simpler models of computation. A decision tree associated to a Boolean function f takes $\omega \in \{0,1\}^n$ as an input, and reveals algorithmically the value of ω in different bits one by one. The algorithm stops as soon as the value of f is the same no matter the values of ω on the remaining coordinates. The question is then to determine how many bits of information must be revealed before the algorithm stops. The decision tree can also be taken at random to model random or quantum computers.

The theory of (random) decision trees played a key role in computer science (we refer the reader to the survey [BDW02]), but also found many applications in other fields of mathematics. In particular, random decision trees (sometimes called randomized algorithms) were used in [SS10] to study the noise sensitivity of Boolean functions, for instance in the context of percolation theory.

^{*}Université de Genève

[†]Institut des Hautes Études Scientifiques

[‡]ETH Zurich

The OSSS inequality, originally introduced in [OSSS05] for product measure as a step toward a conjecture of Yao [Yao77], relates the variance of a Boolean function to the influence of the variables and the computational complexity of a random decision tree for this function. The first part of this paper consists in generalizing the OSSS inequality to the context of monotonic measures which are not product measures. A *monotonic* measure is a measure μ on $\{0,1\}^E$ such that for any $e \in E$, any $F \subset E$, and any $\xi, \zeta \in \{0,1\}^F$ satisfying $\xi \leq \zeta$, $\mu[\omega_e = \xi_e, \forall e \in F] > 0$ and $\mu[\omega_e = \zeta_e, \forall e \in F] > 0$,

$$\mu[\omega_e = 1 \mid \omega_e = \xi_e, \forall e \in F] \leq \mu[\omega_e = 1 \mid \omega_e = \zeta_e, \forall e \in F].$$

The motivation to choose such a class of measures comes from the applications to mathematical physics (for example, any positive measure satisfying the FKG-lattice inequality is monotonic, see [Gri06] for more details), but monotonic measures also appear in computer science.

In order to state our theorem, we introduce a few notation. Consider a finite set E of cardinality n . For a n -tuple $e = (e_1, \dots, e_n)$ and $t \leq n$, write $e_{[t]} = (e_1, \dots, e_t)$ and $\omega_{e_{[t]}} = (\omega_{e_1}, \dots, \omega_{e_t})$.

A decision tree encodes an algorithm that takes $\omega \in \{0,1\}^E$ as an input, and then queries the values of ω_e , $e \in E$ one bits after the other. For any input ω , the algorithm always starts from the same fixed $e_1 \in E$ (which corresponds to the root of the decision tree), and queries the value of ω_{e_1} . Then, the second element e_2 examined by the algorithm is prescribed by the decision tree and may depend on the value of ω_{e_1} . After having queried the value of ω_{e_2} , the algorithm continues inductively. At step $t > 1$, $(e_1, \dots, e_{t-1}) \in E^{t-1}$ has been examined, and the values of $\omega_{e_1}, \dots, \omega_{e_{t-1}}$ have been queried. The next element e_t to be examined by the algorithm is a deterministic function of what has been explored in the previous steps:

$$e_t = \phi_t((e_1, \dots, e_{t-1}), \omega_{(e_1, \dots, e_{t-1})}) \in E \setminus \{e_1, \dots, e_{t-1}\}. \quad (1)$$

(ϕ_t should be interpreted as the decision rule at time t : ϕ_t takes the location and the value of the first $t-1$ steps of the induction, and decides of the next bit to query). Formally, we call *decision tree* a pair $T = (e_1, (\phi_t)_{2 \leq t \leq n})$, where $e_1 \in E$, and for each t the function ϕ_t , as above, takes a pair $((e_1, \dots, e_{t-1}), \omega_{(e_1, \dots, e_{t-1})})$ as an input and return an element $e_t \in E \setminus \{e_1, \dots, e_{t-1}\}$.

Let $T = (e_1, (\phi_t)_{2 \leq t \leq n})$ be a decision tree and $f : \{0,1\}^E \rightarrow \mathbb{R}$. Given $\omega \in \{0,1\}^E$ we consider the n -tuple (e_1, \dots, e_n) defined inductively by (1) (this corresponds to the ordering on E that we get when we run the algorithm T starting from the input ω). We define

$$\tau(\omega) = \tau_{f,T}(\omega) := \min \{t \geq 1 : \forall \omega' \in \{0,1\}^E, \quad \omega'_{e_{[t]}} = \omega_{e_{[t]}} \implies f(\omega) = f(\omega')\}. \quad (2)$$

In computer science, a decision tree is usually associated directly to a Boolean function f and defined as a rooted directed tree in which internal nodes are labeled by elements of E , leaves by possible outputs, and edges are in correspondence with the possible values of the bits at vertices (see [OSSS05] for a formal definition). In particular, the decision trees are usually defined up to τ , and not later on. In this paper, we chose the slightly different formalism described above, which is equivalent to the classical one, since it will be more convenient for the proof of the following theorem.

Theorem 1.1 *Fix an increasing function $f : \{0,1\}^E \rightarrow [0,1]$ on a finite set E . For any monotonic measure μ and any decision tree T ,*

$$\text{Var}_\mu(f) \leq \sum_{e \in E} \delta_e(f, T) \text{Cov}_\mu(f, \omega_e), \quad (3)$$

where $\delta_e(f, T) := \mu[\exists t \leq \tau(\omega) : e_t = e]$ is the revealment (of f) for the decision tree T .

A slightly stronger form of this result is stated in Section 2. In this paper, we focus on applications of the previous result to statistical physics but we expect it to have a number of applications in the context of the theory of Boolean functions. The interested reader is encouraged to consult [O'D14] for a detailed introduction to the subject. Theorems regarding Boolean functions have already found several applications in statistical physics, especially in the context of the noise sensitivity. For a review of the relationship between percolation theory and the analysis of Boolean functions we refer the reader to the book of Garban and Steif [GS14].

1.2 Sharpness of the phase transition in statistical physics

We call *lattice* a locally finite (vertex-)transitive infinite graph $\mathbb{G} = (\mathbb{V}, \mathbb{E})$. An (unoriented) edge of the lattice is denoted xy . We also distinguish a vertex $0 \in \mathbb{V}$ and call it the origin. Let $d(\cdot, \cdot)$ denote the graph distance on \mathbb{G} . Introduce a family of *non-negative* coupling constants $J = (J_{xy})_{xy \in \mathbb{E}} \in [0, \infty)^{\mathbb{E}}$ which is non-zero and invariant under a group acting transitively on \mathbb{V} . Notice that the coupling constants are necessarily finite-range (since the graph is locally finite). We call the pair (\mathbb{G}, J) a *weighted lattice*.

Statistical physics models defined on a lattice are useful to describe a large variety of phenomena and objects, ranging from ferro-magnetic materials to lattice gas. They also provide discretizations of Euclidean and Quantum Field Theories and are as such important from the point of view of theoretical physics. While the original motivation came from physics, they appeared as extremely complex and rich mathematical objects, whose study required the development of important new tools that found applications in many other domains of mathematics.

One of the key aspects of these models is that they often undergo order/disorder phase transitions at a certain critical parameter β_c . The regime $\beta < \beta_c$, usually called the disorder regime, exhibits very rapid decay of correlations. While this property is usually simple to derive for very small values of β using perturbative techniques, proving such a statement for the whole range of parameters $\beta < \beta_c$ is a difficult mathematical challenge. Nevertheless, having such a property is the key towards a deep understanding of the disordered regime.

The zoo of lattice models is very diverse: it includes models of spin-glasses, quantum chains, random surfaces, spin systems, percolation models. One of the most famous example of a lattice spin model is provided by the Ising model introduced by Lenz to explain Curie's temperature for ferromagnets. This model has been generalized in many directions to create models exhibiting a wide range of critical phenomena. While the Ising model is very well understood, most of these natural generalizations remain much more difficult to comprehend. In this paper, we prove that the Potts model (one of the most natural of such generalization) undergoes a sharp phase transition, meaning that in the disordered regime, correlations decay exponentially fast. In order to do so, we will study the random-cluster representations of these models, which are often monotonic. The generalized OSSS inequality proved in Theorem 1.1 will play a key role in the proof.

Exponential decay for the subcritical random-cluster model Since random-cluster models were introduced by Fortuin and Kasteleyn in 1969 [FK72], they have become the archetypal example of dependent percolation models and as such have played an important role in the study of phase transitions. The spin correlations of Potts models are rephrased as cluster connectivity properties of their random-cluster representations. This allows the use of geometric techniques, thus leading to several important applications. While the understanding of the model on planar graphs progressed greatly in the past few years [BD12, DST17, DGH⁺16, DRT16], the case of higher dimensions remained poorly understood. We refer to [Gri06, Dum13] for books on the subject and a discussion of existing results.

The model is defined as follows. Consider a finite subgraph $G = (V, E)$ of a weighted lattice (\mathbb{G}, J) and introduce the boundary ∂G of G to be the set of vertices $x \in G$ for which there exists $y \notin G$ with xy an edge of \mathbb{E} . A *percolation configuration* $\omega = (\omega_{xy})_{xy \in E}$ is an element of $\{0, 1\}^E$. A configuration ω can be seen as a subgraph of G with vertex-set V and edge-set given by $\{xy \in E : \omega_{xy} = 1\}$. Let $k_f(\omega)$ (resp. $k_w(\omega)$) be the number of connected components in ω (resp. in the graph obtained from ω by considering all the vertices in ∂G as one single vertex).

Fix $q, \beta > 0$. For $\# \in \{f, w\}$, let $\phi_{G, \beta, q}^\#$ be the measure satisfying, for any $\omega \in \{0, 1\}^E$,

$$\phi_{G, \beta, q}^\#(\omega) = \frac{q^{k_\#(\omega)}}{Z} \prod_{xy \in E} (e^{\beta J_{xy}} - 1)^{\omega_{xy}},$$

where Z is a normalizing constant introduced in such a way that $\phi_{G, \beta, q}^\#$ is a probability measure. The measures $\phi_{G, \beta, q}^f$ and $\phi_{G, \beta, q}^w$ are called the random-cluster measures on G with respectively free and

wired boundary conditions. For $q \geq 1$, the measures $\phi_{G,\beta,q}^\#$ can be extended to \mathbb{G} – the corresponding measure is denoted by $\phi_{\mathbb{G},\beta,q}^\#$ – by taking the weak limit of measures defined in finite volume.

For notational convenience, we set $x \longleftrightarrow y$ if x and y are in the same connected component. We also write $x \longleftrightarrow Y$ if x is connected to a vertex in $Y \subset \mathbb{V}$, and $x \longleftrightarrow \infty$ if the connected component of x is infinite. Finally, let Λ_n be the box of size n around 0 for the graph distance.

For $q \geq 1$, the model undergoes a phase transition: there exists $\beta_c = \beta_c(\mathbb{G}) \in [0, \infty]$ satisfying

$$\theta(\beta) := \phi_{\mathbb{G},\beta,q}^w[0 \longleftrightarrow \infty] = \begin{cases} = 0 & \text{if } \beta < \beta_c, \\ > 0 & \text{if } \beta > \beta_c. \end{cases}$$

The main theorem of this article is the following one.

Theorem 1.2 *Fix $q \geq 1$ and consider the random-cluster model on a weighted lattice (\mathbb{G}, J) . Then,*

- *There exists $c > 0$ such that $\theta(\beta) \geq c(\beta - \beta_c)$ for any $\beta \geq \beta_c$ close enough to β_c .*
- *For any $\beta < \beta_c$ there exists $c_\beta > 0$ such that for every $n \geq 0$,*

$$\phi_{\Lambda_n, \beta, q}^w[0 \longleftrightarrow \partial \Lambda_n] \leq \exp[-c_\beta n].$$

Theorem 1.2 extends to quasi-transitive weighted graphs and to finite range interactions (for the latter, simply interpret finite-range models as nearest-neighbor models on a bigger graph).

For planar graphs, the result was proved for any $q \geq 1$ under some symmetry assumption in [DM16] (see also [MR16] for the case of planar slabs). On \mathbb{Z}^d , the result was restricted to large values of q [LMMS⁺91] and to the special cases of Bernoulli percolation ($q = 1$) [Men86, AB87, DT15] and the FK-Ising model ($q = 2$) [ABF87, DT15].

Numerous results about the subcritical regime have been proved under the assumption of exponential decay, and therefore Theorem 1.2 transform them into unconditional results. To cite but a few, let us mention the Ornstein-Zernike theory of correlations [CIV08], the mixing properties of the model [Ale04], the bounds on the spectral gaps of the associated dynamics [Mar99]. The second item of Theorem 1.2 could be replaced by $\phi_{\mathbb{G}, \beta, q}^w[0 \longleftrightarrow x] \leq \exp[-c_\beta d(0, x)]$, but the stronger statement proved in the theorem is the one useful for these applications.

Applications to computations of critical points for planar graphs Another important application of Theorem 1.2 is the computation of critical points of specific lattices. In this section, we fix coupling constants to be equal to 1 and set $p_c := 1 - e^{-\beta_c}$. In general, the critical parameter p_c is not expected to take any specific value. However, for the square, hexagonal and triangular lattices, the critical values can be predicted using duality. It is proved in [Gri06, Theorem 6.17] that predicted values are indeed the critical ones under the assumption of exponential decay for $p < p_c$. Therefore, our result provides an alternative proof of the following theorem.

Theorem 1.3 *Fix $q \geq 1$. If $y_c := p_c/(1 - p_c)$, we have*

$$\begin{aligned} y_c^2 - q &= 0 && \text{on the square lattice,} \\ y_c^3 + 3y_c^2 - q &= 0 && \text{on the triangular lattice,} \\ y_c^3 - 3qy_c - q^2 &= 0 && \text{on the hexagonal lattice.} \end{aligned}$$

Note that for the square lattice, p_c is equal to $\sqrt{q}/(1 + \sqrt{q})$. This result was originally proved in [BD12], where exponential decay of correlations below p_c is proved using Russo-Seymour-Welsh type arguments and a generalization [GG06] of the KKL result [KKL88, BKK⁺92].

The fact that our proof of exponential decay requires very few conditions on the graphs enables us to study critical points of general planar locally-finite doubly periodic graphs, i.e. embedded planar graphs which are invariant under the action of some lattice $\Lambda \approx \mathbb{Z} \oplus \mathbb{Z}$. Denote the dual of any planar graph \mathbb{G} by \mathbb{G}^* .

Theorem 1.4 *Fix $q \geq 1$ and a planar locally-finite doubly periodic graph \mathbb{G} . We have*

$$\frac{p_c(\mathbb{G})p_c(\mathbb{G}^*)}{(1 - p_c(\mathbb{G}))(1 - p_c(\mathbb{G}^*))} = q. \tag{4}$$

This result should be understood as a generalization of the famous statement $p_c(\mathbb{G}) + p_c(\mathbb{G}^*) = 1$ for Bernoulli percolation. The theorem is a consequence of duality, exponential decay for $p < p_c(\mathbb{G})$ and the following non-coexistence result. For a configuration ω on \mathbb{G} , define a configuration ω^* in \mathbb{G}^* by the formula $\omega_{e^*}^* = 1 - \omega_e$ for every edge e of \mathbb{G} , where e^* is the edge of \mathbb{G}^* between the two vertices of \mathbb{G}^* corresponding to the faces bordered by e .

Theorem 1.5 *There does not exist any translational invariant measure μ on a planar locally-finite doubly periodic graph \mathbb{G} satisfying*

- (FKG) *For any increasing events A and B , $\mu(A \cap B) \geq \mu(A)\mu(B)$.*
- *Almost surely, there exists a unique infinite connected component in ω and in ω^* .*

This result was proved in [She05]. It was also proved for percolation on self-dual polygon configurations in [BR10]. Here, we present a proof which has also the advantage of being quite short.

Applications to the ferromagnetic q -state Potts model The Potts model [Pot52] is one of the most fundamental example of a lattice spin model undergoing an order/disorder phase transition at a critical parameter β_c . It generalizes the Ising model by allowing the spins to take one of q values. In two dimensions, the model has been the object of intense study in the past few years and the behavior is fairly well understood, even at criticality [DST17, DGH⁺16]. In higher dimension, the understanding is limited to the case of the Ising model (i.e. $q = 2$) and of large q [AF86, ADS15, KS82, LMMS⁺91, BC03].

The model is defined as follows. Consider an integer $q \geq 2$. For $G = (V, E)$ a finite subgraph of a weighted lattice (\mathbb{G}, J) , $\nu \in \{1, \dots, q\}$, and $\beta \geq 0$, the q -state Potts measure with boundary condition ν is defined for any $\sigma = (\sigma_x)_{x \in V} \in \{1, \dots, q\}^V$ by

$$\mathbb{P}_{G, \beta, q}^\nu[\sigma] := \frac{\exp(-\beta H_{G, q}^\nu(\sigma))}{\sum_{\sigma' \in \{1, \dots, q\}^V} \exp(-\beta H_{G, q}^\nu(\sigma'))}$$

where

$$H_{G, q}^\nu(\sigma) := - \sum_{xy \in E} J_{xy} \delta_{\sigma_x = \sigma_y} - \sum_{\substack{xy \in E \\ x \in \partial G, y \notin G}} J_{xy} \delta_{\sigma_x = \nu}.$$

The model can be defined in infinite volume by taking the weak limit of measures on a nested sequence of finite graphs. The obtained measure $\mathbb{P}_{\mathbb{G}, \beta, q}^\nu$ is called the Potts measure with boundary conditions ν . The Potts model undergoes a phase transition between absence and existence of long-range order at the so-called critical inverse temperature β_c (which depends on \mathbb{G} and J), see [Gri06] for details.

Theorem 1.6 *Fix an integer $q \geq 2$ and consider the q -state Potts on a weighted lattice (\mathbb{G}, J) . Then, for $\beta < \beta_c$, there exists $c_\beta > 0$ such that for every $x \in \mathbb{V}$,*

$$0 \leq \mathbb{P}_{\Lambda_n, \beta, q}^\nu[\sigma_0 = \nu] - \frac{1}{q} \leq \exp[-c_\beta n].$$

Furthermore, for the nearest-neighbor model on the square lattice, $\beta_c(\mathbb{Z}^2) = \log(1 + \sqrt{q})$.

For the 2-state Potts model, better known as the Ising model, the result goes back to [ABF87] (see also [DT15]). For the q -state Potts model with $q \geq 3$, the result was restricted to either perturbative arguments involving the Pirogov-Sinai theory for $q \gg 1$ or planar arguments (see the discussion on the random-cluster model). The question of deriving this property for $q \geq 3$ and \mathbb{Z}^d with $d \geq 3$ was open. Again, the flexibility in the choice of the lattice \mathbb{G} implies that the result applies to finite range interactions. The statement of Theorem 1.6 is stronger than the statement $\mathbb{P}_{\mathbb{G}, \beta, q}^\nu[\sigma_0 = \nu] - \frac{1}{q} \leq \exp[-c_\beta d(0, x)]$.

The Potts model and the random-cluster models on a weighted lattice (\mathbb{G}, J) can be coupled (see [Gri06, Theorem 1.10] for details) in such a way that

$$\mathbb{P}_{\Lambda_n, \beta, q}^\nu[\sigma_0 = \nu] - \frac{1}{q} = \frac{q-1}{q} \phi_{\Lambda_{n+1}, \beta, q}^w[0 \longleftrightarrow \partial \Lambda_{n+1}],$$

so that Theorem 1.6 is a direct consequence of Theorems 1.2 and 1.3.

Other models. The reasoning above should extend to other lattice spin models for which there exists a random-cluster representation which is monotonic. An archetypal example is provided by the Ashkin-Teller model; see [Bax82] for details. It also extends to continuum percolation models such as Voronoi percolation [DRT17a], occupied and vacant set of Boolean percolation [DRT17b], massive Gaussian free field super-level lines.

Organization The paper is organized as follows. In the next section we prove Theorem 1.1. In the third section, we prove Theorem 1.2 (we tried to isolate a few general statements which may be used for the proof of exponential decay for other models of statistical physics). In the last section, we describe the proof of Theorems 1.4 and 1.5.

Acknowledgments This research was supported by an IDEX grant from Paris-Saclay, a grant from the Swiss FNS, the ERC CriBLaM, and the NCCR SwissMAP. We thank Yvan Velenik, Alain-Sol Sznitman and Ioan Manolescu for many inspiring discussions and insightful comments.

2 Proof of Theorem 1.1

The strategy is a combination of the original proof of the OSSS inequality for product measures (which is an Efron-Stein type reasoning), together with an encoding of monotonic measures in terms of iid random variables. Assume that E is finite and has cardinality n . Let \vec{E} be the set of sequences $e = (e_1, \dots, e_n)$ where each element of E occurs exactly once. Consider a monotonic measure μ on $\{0, 1\}^E$.

We start by a useful lemma explaining how to construct ω with law μ from iid uniform random variables. For $u \in [0, 1]^n$ and $e \in \vec{E}$, define $F_e(u) = x$ inductively for $1 \leq t \leq n$ by

$$x_{e_t} := \begin{cases} 1 & \text{if } u_t \geq \mu[\omega_{e_t} = 0 \mid \omega_{e_{[t-1]}} = x_{e_{[t-1]}},] \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

Lemma 2.1 *Let \mathbf{U} be a iid sequence of uniform $[0, 1]$ random variables, and \mathbf{e} a random variable taking values in \vec{E} . Assume that for every $1 \leq t \leq n$, \mathbf{U}_t is independent of $(\mathbf{e}_{[t]}, \mathbf{U}_{[t-1]})$, then $\mathbf{X} = F_{\mathbf{e}}(\mathbf{U})$ has law μ .*

Proof Let $x \in \{0, 1\}^E$ and $e \in \vec{E}$ such that $\mathbb{P}[\mathbf{X} = x, \mathbf{e} = e] > 0$. The probability $\mathbb{P}[\mathbf{X} = x, \mathbf{e} = e]$ can be written as

$$\prod_{t=1}^n \mathbb{P}[\mathbf{X}_{e_t} = x_{e_t} \mid \mathbf{e}_{[t]} = e_{[t]}, \mathbf{X}_{e_{[t-1]}} = x_{e_{[t-1]}},] \times \prod_{t=1}^n \mathbb{P}[\mathbf{e}_t = e_t \mid \mathbf{e}_{[t-1]} = e_{[t-1]}, \mathbf{X}_{e_{[t-1]}} = x_{e_{[t-1]}},].$$

(All the conditionings are well defined, since we assumed $\mathbb{P}[\mathbf{X} = x, \mathbf{e} = e] > 0$.) Since \mathbf{U}_t is independent of $(\mathbf{e}_{[t]}, U_{[t-1]})$ (and thus $\mathbf{X}_{e_{[t-1]}}$), the definition (5) gives

$$\mathbb{P}[\mathbf{X}_{e_t} = x_{e_t} \mid \mathbf{e}_{[t]} = e_{[t]}, \mathbf{X}_{e_{[t-1]}} = x_{e_{[t-1]}},] = \mu[\omega_{e_t} = x_{e_t} \mid \omega_{e_{[t-1]}} = x_{e_{[t-1]}},]$$

so that the first product is equal to $\mu[\omega = x]$ independently of e . Fixing $x \in \{0, 1\}^E$, and summing on $e \in \vec{E}$ satisfying $\mathbb{P}[\mathbf{X} = x, \mathbf{e} = e] > 0$ gives

$$\begin{aligned} \mathbb{P}[\mathbf{X} = x] &= \sum_e \mathbb{P}[\mathbf{X} = x, \mathbf{e} = e] \\ &= \mu[\omega = x] \sum_e \prod_{t=1}^n \mathbb{P}[\mathbf{e}_t = e_t \mid \mathbf{e}_{[t-1]} = e_{[t-1]}, \mathbf{X}_{e_{[t-1]}} = x_{e_{[t-1]}},] = \mu[\omega = x]. \end{aligned}$$

□

Proof of Theorem 1.1 Our goal is to apply a Lindeberg-type argument on a probability space in which \mathbf{e} and \mathbf{X} (sampled according to μ) are coupled to an independent copy of \mathbf{X} (denoted by \mathbf{Y}^n below). We now present the coupling.

Consider two independent sequences of iid uniform $[0, 1]$ random variables \mathbf{U} and \mathbf{V} . Write \mathbb{P} for the coupling between these variables (and \mathbb{E} for its expectation). Construct $(\mathbf{e}, \mathbf{X}, \tau)$ inductively as follows: set for $t \geq 1$,

$$\mathbf{e}_t = \begin{cases} e_1 & \text{if } t = 1 \\ \phi_t(\mathbf{e}_{[t-1]}, \mathbf{X}_{\mathbf{e}_{[t-1]}}) & \text{if } t > 1 \end{cases} \quad \text{and} \quad \mathbf{X}_{\mathbf{e}_t} = \begin{cases} 1 & \text{if } \mathbf{U}_t \geq \mu(\omega_{\mathbf{e}_t} = 0 | \omega_{\mathbf{e}_{[t-1]}} = \mathbf{X}_{\mathbf{e}_{[t-1]}}, \mathbf{U}_{[t-1]}), \\ 0 & \text{otherwise,} \end{cases}$$

and $\tau := \min \{t \geq 1 : \forall x \in \{0, 1\}^E, x_{\mathbf{e}_{[t]}} = \mathbf{X}_{\mathbf{e}_{[t]}} \Rightarrow f(x) = f(\mathbf{X})\}$. Note that τ is equal to the stopping time defined in (2). Finally, for $0 \leq t \leq n$, define $\mathbf{Y}^t := F_{\mathbf{e}}(\mathbf{W}^t)$, where

$$\mathbf{W}^t := \mathbf{W}^t(\mathbf{U}, \mathbf{V}) = (\mathbf{V}_1, \dots, \mathbf{V}_t, \mathbf{U}_{t+1}, \dots, \mathbf{U}_\tau, \mathbf{V}_{\tau+1}, \dots, \mathbf{V}_n)$$

(in particular \mathbf{W}^t is equal to \mathbf{V} if $t \geq \tau$).

Lemma 2.1 applied to (\mathbf{U}, \mathbf{e}) gives that \mathbf{X} has law μ and is \mathbf{U} -measurable. Lemma 2.1 applied to (\mathbf{V}, \mathbf{e}) implies that \mathbf{Y}^n has law μ and is independent of \mathbf{U} . Therefore, using that f is valued in $[0, 1]$, we deduce that

$$\text{Var}_\mu(f) \leq \frac{1}{2}\mu[|f - \mu[f]|] = \frac{1}{2}\mathbb{E}[|\mathbb{E}[f(\mathbf{X})|\mathbf{U}] - \mathbb{E}[f(\mathbf{Y}^n)|\mathbf{U}]|] \leq \frac{1}{2}\mathbb{E}[|f(\mathbf{X}) - f(\mathbf{Y}^n)|].$$

Since $f(\mathbf{Y}^0) = f(\mathbf{X})$ (the entries of \mathbf{Y}^0 for $t > \tau$ are irrelevant for the value of f by definition of τ), the equation above implies

$$\text{Var}_\mu(f) \leq \frac{1}{2}\mathbb{E}[|f(\mathbf{Y}^0) - f(\mathbf{Y}^n)|].$$

Since $\mathbf{Y}^t = \mathbf{Y}^{t-1}$ for any $t > \tau$, the right-hand side of the previous inequality is less than or equal to

$$\begin{aligned} \sum_{t=1}^n \mathbb{E}[|f(\mathbf{Y}^t) - f(\mathbf{Y}^{t-1})|] &= \sum_{t=1}^n \mathbb{E}[|f(\mathbf{Y}^t) - f(\mathbf{Y}^{t-1})| \cdot \mathbb{1}_{t \leq \tau}] \\ &= \sum_{e \in E} \sum_{t=1}^n \mathbb{E}\left[\mathbb{E}[|f(\mathbf{Y}^t) - f(\mathbf{Y}^{t-1})| \mid \mathbf{U}_{[t-1]}] \mathbb{1}_{t \leq \tau, \mathbf{e}_t = e}\right]. \end{aligned}$$

Recalling that $\sum_{t=1}^n \mathbb{P}[t \leq \tau, \mathbf{e}_t = e] = \delta_e(f, T)$, the proof of the theorem follows from the fact that on $\{t \leq \tau, \mathbf{e}_t = e\}$,

$$\mathbb{E}[|f(\mathbf{Y}^t) - f(\mathbf{Y}^{t-1})| \mid \mathbf{U}_{[t-1]}] \leq 2\text{Cov}_\mu(f, \omega_e). \quad (6)$$

In order to show this, we now restrict ourself to the event $\{t \leq \tau, \mathbf{e}_t = e\}$. First observe that $\mathbf{Y}_e^t = \mathbf{Y}_e^{t-1}$ implies $\mathbf{Y}^t = \mathbf{Y}^{t-1}$, and this together with the fact that f is increasing implies

$$\begin{aligned} |f(\mathbf{Y}^t) - f(\mathbf{Y}^{t-1})| &= (f(\mathbf{Y}^t) - f(\mathbf{Y}^{t-1}))(\mathbf{Y}_e^t - \mathbf{Y}_e^{t-1}) \\ &= f(\mathbf{Y}^{t-1})\mathbf{Y}_e^{t-1} + f(\mathbf{Y}^t)\mathbf{Y}_e^t - f(\mathbf{Y}^{t-1})\mathbf{Y}_e^t - f(\mathbf{Y}^t)\mathbf{Y}_e^{t-1}. \end{aligned} \quad (7)$$

Our goal is to average against $\mathbb{E}[\cdot | \mathbf{U}_{[t-1]}]$. In order to do this, we will use the following claim.

Claim. For any measurable g and $t \leq n$,

$$\mathbb{E}[g(\mathbf{Y}^t) | \mathbf{U}_{[t]}] = \mu[g(\omega)]. \quad (8)$$

Proof Conditioned on $\mathbf{U}_{[t]}$, the random vector \mathbf{W}^t is composed of iid uniform random variables satisfying that \mathbf{W}_i^t is independent of $(\mathbf{e}_1, \dots, \mathbf{e}_i)$ for every $i \leq n$. Therefore, Lemma 2.1 applied to $(\mathbf{e}, \mathbf{W}^t)$ implies that the law of \mathbf{Y}^t conditioned on $\mathbf{U}_{[t]}$ is μ , which gives the claim. \square

Applying (8) to $g(\omega) = f(\omega)\omega_e$ gives that (for the second equality we average on \mathbf{U}_t)

$$\mathbb{E}[f(\mathbf{Y}^{t-1})\mathbf{Y}_e^{t-1} | \mathbf{U}_{[t-1]}] = \mu[f(\omega)\omega_e] = \mathbb{E}[f(\mathbf{Y}^t)\mathbf{Y}_e^t | \mathbf{U}_{[t-1]}]. \quad (9)$$

For fixed $\mathbf{U}_{[n]}$ and s , $\mathbf{Y}^s = F_{\mathbf{e}}(\mathbf{W}^s)$ is an increasing function of \mathbf{V} , by monotonicity of μ . Since f and \mathbf{W}_e are increasing functions of \mathbf{V} , we deduce that $f(\mathbf{Y}^{t-1})$ and $\mathbf{Y}_e^t = F_{\mathbf{e}}(\mathbf{W})_e$ are increasing functions of \mathbf{V} . The FKG inequality applied to the iid random variables \mathbf{V} gives

$$\mathbb{E}[f(\mathbf{Y}^{t-1})\mathbf{Y}_e^t | \mathbf{U}_{[n]}] \geq \mathbb{E}[f(\mathbf{Y}^{t-1}) | \mathbf{U}_{[n]}]\mathbb{E}[\mathbf{Y}_e^t | \mathbf{U}_{[n]}].$$

Taking the expectation with respect to $\mathbb{E}[\cdot | \mathbf{U}_{[t-1]}]$ gives

$$\begin{aligned} \mathbb{E}[f(\mathbf{Y}^{t-1})\mathbf{Y}_e^t | \mathbf{U}_{[t-1]}] &\geq \mathbb{E}[\mathbb{E}[f(\mathbf{Y}^{t-1}) | \mathbf{U}_{[n]}]\mathbb{E}[\mathbf{Y}_e^t | \mathbf{U}_{[n]}] | \mathbf{U}_{[t-1]}] \\ &= \mathbb{E}[f(\mathbf{Y}^{t-1}) | \mathbf{U}_{[t-1]}]\mathbb{E}[\mathbf{Y}_e^t | \mathbf{U}_{[t-1]}] \stackrel{(8)}{=} \mu[f(\omega)]\mu[\omega_e], \end{aligned} \quad (10)$$

where we used that $\mathbb{E}[\mathbf{Y}_e^t | \mathbf{U}_{[n]}]$ is $\mathbf{U}_{[t-1]}$ -measurable (since \mathbf{Y}_e^t depends on $\mathbf{U}_{[t-1]}$ and \mathbf{V} only).

Similarly, $f(\mathbf{Y}^t)$ and \mathbf{Y}_e^{t-1} are increasing functions of \mathbf{V} so that using the FKG inequality and then taking the expectation with respect to $\mathbb{E}[\cdot | \mathbf{U}_{[t]}]$ gives

$$\begin{aligned} \mathbb{E}[f(\mathbf{Y}^t)\mathbf{Y}_e^{t-1} | \mathbf{U}_{[t]}] &\geq \mathbb{E}[\mathbb{E}[f(\mathbf{Y}^t) | \mathbf{U}_{[n]}]\mathbb{E}[\mathbf{Y}_e^{t-1} | \mathbf{U}_{[n]}] | \mathbf{U}_{[t]}] \\ &= \mathbb{E}[f(\mathbf{Y}^t) | \mathbf{U}_{[t]}]\mathbb{E}[\mathbf{Y}_e^{t-1} | \mathbf{U}_{[t]}] \stackrel{(8)}{=} \mu[f(\omega)]\mathbb{E}[\mathbf{Y}_e^{t-1} | \mathbf{U}_{[t]}]. \end{aligned}$$

This time, we used that $\mathbb{E}[\mathbf{Y}_e^{t-1} | \mathbf{U}_{[n]}]$ is $\mathbf{U}_{[t]}$ -measurable. Taking the expectation with respect to $\mathbb{E}[\cdot | \mathbf{U}_{[t-1]}]$ gives

$$\mathbb{E}[f(\mathbf{Y}^t)\mathbf{Y}_e^{t-1} | \mathbf{U}_{[t-1]}] \geq \mu[f(\omega)]\mathbb{E}[\mathbf{Y}_e^{t-1} | \mathbf{U}_{[t-1]}] \stackrel{(8)}{=} \mu[f(\omega)]\mu[\omega_e].$$

This inequality together with (10), (9) and (7) give (6) and therefore concludes the proof. \square

Remark 2.2 For most applications, one may replace covariances in the OSSS inequality by influences $I_e[f] := \mu(f|\omega_e = 1) - \mu(f|\omega_e = 0)$ (we chose not to do so since applications in statistical physics to long-range models would for instance require the statement with covariances). In this case, we do not need to prove (6) anymore and can replace the lengthly end of the proof by the following short argument. Recall the dependency in the measure μ in $F_e(u)$ and write $F_e^\mu(u)$. With this notation, one sees that F is both increasing in u and in μ (for stochastic domination). We deduce that both \mathbf{Y}^{t-1} and \mathbf{Y}^t are sandwiched between $\mathbf{Z} := F_e^{\mu[|\omega_e=0]}(\mathbf{W}^t)$ and $\mathbf{Z}' := F_e^{\mu[|\omega_e=1]}(\mathbf{W}^t)$. Recall that \mathbf{W} is independent of $\mathbf{U}_{[t-1]}$. Lemma 2.1 and the fact that f is increasing give us

$$\mathbb{E}[|f(\mathbf{Y}^t) - f(\mathbf{Y}^{t-1})| | \mathbf{U}_{[t-1]}] \leq \mathbb{E}[f(\mathbf{Z}')] - \mathbb{E}[f(\mathbf{Z})] = \mu[f(\omega)|\omega_e = 1] - \mu[f(\omega)|\omega_e = 0] = I_e[f].$$

Remark 2.3 Note that for the trivial decision tree discovering all the edges, for every edge the revealment is equal to 1. As a consequence, we recover (in a very convoluted way) the discrete Poincaré inequality

$$\text{Var}_\mu(f) \leq \sum_{e \in E} \text{Cov}_\mu(f, \omega_e). \quad (11)$$

Remark 2.4 The proof of the previous statement can be extended in a trivial way as follows. First, we may consider countable sets E by using a very simple martingale argument. Second, we may consider that τ is an arbitrary stopping time (with respect to the filtration $(\mathcal{F}_t = \sigma(\mathbf{e}_{[t]}, \omega_{\mathbf{e}_{[t]}}))_{t \geq 0}$), i.e. that f is not necessarily \mathcal{F}_τ measurable. By simply applying the previous lemma with $g = \mu[f | \mathcal{F}_\tau]$, we obtain the following result, which may be useful in statistical physics.

Theorem 2.5 Fix a countable set E and an increasing function $f : \{0, 1\}^E \rightarrow [-1, 1]$. For any monotonic measure μ on $\{0, 1\}^E$, any decision tree T and any stopping time τ ,

$$\text{Var}_\mu(f) \leq \sum_{e \in E} \delta_e(f, T) \text{Cov}_\mu(f, \omega_e) + \mu[|f - \mu[f | \mathcal{F}_\tau]|]. \quad (12)$$

3 Proof of Theorem 1.2

In order to be able to apply the strategy to other models, we state two useful lemmas.

Lemma 3.1 *Consider a converging sequence of increasing differentiable functions $f_n : [0, \beta_0] \rightarrow [0, M]$ satisfying*

$$f'_n \geq \frac{n}{\Sigma_n} f_n \quad (13)$$

for all $n \geq 1$, where $\Sigma_n = \sum_{k=0}^{n-1} f_k$. Then, there exists $\beta_1 \in [0, \beta_0]$ such that

P1 For any $\beta < \beta_1$, there exists $c_\beta > 0$ such that for any n large enough, $f_n(\beta) \leq \exp(-c_\beta n)$.

P2 For any $\beta > \beta_1$, $f = \lim_{n \rightarrow \infty} f_n$ satisfies $f(\beta) \geq \beta - \beta_1$.

Proof Define

$$\beta_1 := \inf \left\{ \beta : \limsup_{n \rightarrow \infty} \frac{\log \Sigma_n(\beta)}{\log n} \geq 1 \right\}.$$

Assume $\beta < \beta_1$. Fix $\delta > 0$ and set $\beta' = \beta - \delta$ and $\beta'' = \beta - 2\delta$. We will prove that there is exponential decay at β'' in two steps.

First, there exists an integer N and $\alpha > 0$ such that $\Sigma_n(\beta) \leq n^{1-\alpha}$ for all $n \geq N$. For such an integer n , integrating $f'_n \geq n^\alpha f_n$ between β' and β – this differential inequality follows from (13), the monotonicity of the functions f_n (and therefore Σ_n) and the previous bound on $\Sigma_n(\beta)$ – implies that

$$f_n(\beta') \leq M \exp(-\delta n^\alpha), \quad \forall n \geq N.$$

Second, this implies that there exists $\Sigma < \infty$ such that $\Sigma_n(\beta') \leq \Sigma$ for all n . Integrating $f'_n \geq \frac{n}{\Sigma} f_n$ for all n between β'' and β' – this differential inequality is again due to (13), the monotonicity of Σ_n , and the bound on $\Sigma_n(\beta')$ – leads to

$$f_n(\beta'') \leq M \exp\left(-\frac{\delta}{\Sigma} n\right), \quad \forall n \geq 0.$$

Assume $\beta > \beta_1$. For $n \geq 1$, define the function $T_n := \frac{1}{\log n} \sum_{i=1}^n \frac{f_i}{i}$. Differentiating T_n and using (13), we obtain

$$T'_n = \frac{1}{\log n} \sum_{i=1}^n \frac{f'_i}{i} \stackrel{(13)}{\geq} \frac{1}{\log n} \sum_{i=1}^n \frac{f_i}{\Sigma_i} \geq \frac{\log \Sigma_{n+1} - \log \Sigma_1}{\log n},$$

where in the last inequality we used that for every $i \geq 1$,

$$\frac{f_i}{\Sigma_i} \geq \int_{\Sigma_i}^{\Sigma_{i+1}} \frac{dt}{t} = \log \Sigma_{i+1} - \log \Sigma_i.$$

For $\beta' \in (\beta_1, \beta)$, using that $\Sigma_{n+1} \geq \Sigma_n$ is increasing and integrating the previous differential inequality between β' and β gives

$$T_n(\beta) - T_n(\beta') \geq (\beta - \beta') \frac{\log \Sigma_n(\beta') - \log M}{\log n}.$$

Hence, the fact that $T_n(\beta)$ converges to $f(\beta)$ as n tends to infinity implies

$$f(\beta) - f(\beta') \geq (\beta - \beta') \left[\limsup_{n \rightarrow \infty} \frac{\log \Sigma_n(\beta')}{\log n} \right] \geq \beta - \beta'.$$

Letting β' tend to β_1 from above, we obtain $f(\beta) \geq \beta - \beta_1$. \square

We now present an application of Theorem 1.1 to monotonic measures on $\{0, 1\}^E$, where E is the edge set of a finite graph $G = (V, E)$. Let $\Lambda_n(x)$ denote the box of size n around $x \in V$ and write $\Lambda_n = \Lambda_n(0)$. We see elements of $\{0, 1\}^E$ as percolation configurations and use the corresponding notation.

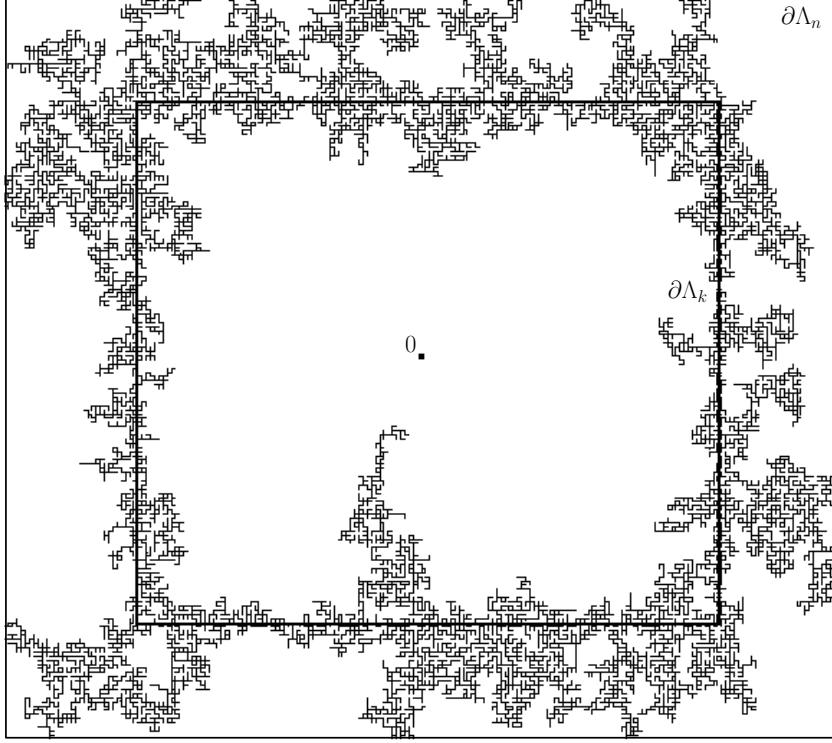


Figure 1: A realization of the clusters intersecting $\partial\Lambda_k$. Every edge having one endpoint in this set is potentially revealed by the decision tree before time τ . Furthermore in this specific case, we know that 0 is not connected to the boundary of Λ_n .

Lemma 3.2 Consider a finite graph $G = (V, E)$ containing 0. For any monotonic measure μ on $\{0, 1\}^E$ and any $n \geq 1$, one has

$$\sum_{xy \in E} \text{Cov}_\mu(\mathbb{1}_{0 \leftrightarrow \partial\Lambda_n}, \omega_e) \geq \frac{n}{4 \max_{x \in \Lambda_n} \sum_{k=0}^{n-1} \mu[x \leftrightarrow \partial\Lambda_k(x)]} \cdot \mu[0 \leftrightarrow \partial\Lambda_n] (1 - \mu[0 \leftrightarrow \partial\Lambda_n]).$$

The proof is based on Theorem 1.1 applied to a well chosen decision tree determining $\mathbb{1}_{0 \leftrightarrow \partial\Lambda_n}$. One may simply choose the trivial algorithm checking every edge of the box Λ_n . Unfortunately, the revealment of this decision tree being 1 for every edge, the OSSS inequality will not bring us more information than the Poincaré inequality (11). A slightly better algorithm would be provided by the decision tree discovering the connected component of the origin “from inside”. Edges far from the origin would then be revealed by the algorithm if (and only if) one of their endpoints is connected to the origin. This provides a good bound for the revealment of edges far from the origin, but edges close to the origin are still revealed with large probability. In order to avoid this last fact, we will rather choose a family of decision trees discovering the connected components of $\partial\Lambda_k$ for $1 \leq k \leq n$ and observe that the average of their revealment for a fixed edge will always be small.

Proof We can assume that $\partial\Lambda_n$ is not empty (otherwise the statement is trivially true). For any $k \in [1, n]$, we wish to construct a decision tree T determining $\mathbb{1}_{0 \leftrightarrow \partial\Lambda_k}$ such that for each $e = uv$,

$$\delta_e(T) \leq \mu[u \leftrightarrow \partial\Lambda_k] + \mu[v \leftrightarrow \partial\Lambda_k]. \quad (14)$$

Note that this would conclude the proof since we obtain the target inequality by applying Theorem 1.1 for each k and then summing on k . As a key, we use that for $u \in \Lambda_n$,

$$\sum_{k=1}^n \mu[u \leftrightarrow \partial\Lambda_k] \leq \sum_{k=1}^n \mu[u \leftrightarrow \partial\Lambda_{|k-d(u,0)|}(u)] \leq 2 \max_{x \in \Lambda_n} \sum_{k=0}^{n-1} \mu[x \leftrightarrow \partial\Lambda_k(x)].$$

We describe the decision tree T , which corresponds first to an exploration of the connected components in Λ_n intersecting $\partial\Lambda_k$ that does not reveal any edge with both endpoints outside these connected components, and then to a simple exploration of the remaining edges.

More formally, we define \mathbf{e} using two growing sequences $\partial\Lambda_k = V_0 \subset V_1 \subset \dots \subset V$ and $\emptyset = F_0 \subset F_1 \subset \dots \subset F$ (where F is the set of edges between two vertices within distance n of the origin) that should be understood as follows: at step t , V_t represents the set of vertices that the decision tree found to be connected to $\partial\Lambda_k$, and F_t is the set of explored edges discovered by the decision tree until time t .

Fix an ordering of the edges in F . Set $V_0 = \partial\Lambda_k$ and $F_0 = \emptyset$. Now, assume that $V_t \subset V$ and $F_t \subset F$ have been constructed and distinguish between two cases:

- If there exists an edge $e = xy \in F \setminus F_t$ with $x \in V_t$ and $y \notin V_t$ (if more than one exists, pick the smallest one for the ordering), then set $\mathbf{e}_{t+1} = e$, $F_{t+1} = F_t \cup \{e\}$ and set

$$V_{t+1} := \begin{cases} V_t \cup \{x\} & \text{if } \omega_e = 1 \\ V_t & \text{otherwise.} \end{cases}$$

- If e does not exist, set \mathbf{e}_{t+1} to be the smallest $e \in F \setminus F_t$ (for the ordering) and set $V_{t+1} = V_t$ and $F_{t+1} = F_t \cup \{e\}$.

As long as we are in the first case, we are still discovering the connected components of $\partial\Lambda_k$, while as soon as we are in the second case, we remain in it. The fact that τ is smaller than or equal to the last time we are in the first case gives us (14). \square

Remark 3.3 Note that τ may a priori be strictly smaller than the last time we are in first case (since the decision tree may discover a path of open edges from 0 to $\partial\Lambda_n$ or a family of closed edges disconnecting the origin from $\partial\Lambda_n$ before discovering the whole connected components of $\partial\Lambda_k$).

We are now in a position to prove Theorem 1.2. We will simply combine a derivative formula for random-cluster models with the previous lemma, and then apply Lemma 3.1.

Proof of Theorem 1.2 Fix $q \geq 1$ and $\beta_0 \geq 0$. For $n \geq 1$ and $\beta \leq \beta_0$, define

$$\mu_n := \phi_{\Lambda_{2n}, \beta, q}^w \quad \theta_k(\beta) := \mu_k[0 \leftrightarrow \partial\Lambda_k] \quad S_n := \sum_{k=0}^{n-1} \theta_k.$$

Now, the comparison between boundary conditions [Gri06, Lemma 4.14] together with the facts that $\Lambda_{2k}(x) \subset \Lambda_{2n}$ and that \mathbb{G} is transitive imply that for $x \in \Lambda_n$,

$$\sum_{k=1}^{n-1} \mu_n[x \leftrightarrow \partial\Lambda_k(x)] \leq 2 \sum_{k \leq n/2} \mu_n[x \leftrightarrow \partial\Lambda_k(x)] \leq 2 \sum_{k \leq n/2} \mu_k[0 \leftrightarrow \partial\Lambda_k] \leq 2S_n.$$

Since μ_n is monotonic [Gri06, Theorem 3.8], Lemma 3.2 (applied to the graph $G = (\Lambda_{2n}, E)$ induced by Λ_{2n}) and the previous bound give

$$\sum_{e \in E} \text{Cov}(\mathbb{1}_{0 \leftrightarrow \partial\Lambda_n}, \omega_e) \geq \frac{n}{8S_n} \cdot \theta_n(1 - \theta_n). \quad (15)$$

Now, a derivative formula for random-cluster models [Gri06, Theorem 3.12] implies that

$$\theta'_n(\beta) = \sum_{e \in E} \frac{J_{xy}}{e^{\beta J_{xy}} - 1} \text{Cov}(\mathbb{1}_{0 \leftrightarrow \partial\Lambda_n}, \omega_e) \geq \min \left\{ \frac{J_{xy}}{e^{\beta_0 J_{xy}} - 1} \right\} \sum_{e \in E} \text{Cov}(\mathbb{1}_{0 \leftrightarrow \partial\Lambda_n}, \omega_e). \quad (16)$$

Notice that the minimum above is positive (since the coupling constants are finite-range and invariant). Inequalities (15) and (16) together lead to

$$\theta'_n \geq c \frac{n}{S_n} \cdot \theta_n, \quad (17)$$

where

$$c = c(\beta_0) := \frac{1 - \theta_1(\beta_0)}{8} \min \left\{ \frac{J_{xy}}{\exp(\beta_0 J_{xy}) - 1} \right\} > 0$$

(we used that $\theta_n \leq \theta_1$ by comparison between boundary conditions and then monotonicity and $\beta \leq \beta_0$). Measurability implies $\limsup \theta_n = \theta$ while the comparison between boundary conditions gives that $\theta_n \geq \theta$ (for all n) so that θ_n converges to θ . Lemma 3.1 applied to $f_n = \theta_n/c$ gives the existence of β_1 such that **P1** and **P2** occur.

Also, for every $n \geq 1$,

$$\phi_{\Lambda_{2n}, \beta, q}^W[0 \longleftrightarrow \partial \Lambda_{2n}] \leq \theta_n(\beta).$$

Overall, the two previous facts combined with **P1** and **P2** implies the theorem readily (note that when $\beta_c < \infty$, $\beta_1 = \beta_c$ as soon as β_0 is chosen larger than β_c). \square

4 Proofs of Theorems 1.4 and 1.5

Without loss of generality, we may assume that \mathbb{G} and \mathbb{G}^* are embedded in such a way that Λ is the set of translations of \mathbb{Z}^2 . We see configurations ω and ω^* as subsets of \mathbb{R}^2 given by the union of the open edges. For three sets $A, B, C \subset \mathbb{R}^2$, denote the event that $\omega \cap C$ contains a continuous path from A to B by $A \xrightarrow{C} B$.

Let us start by explaining how Theorem 1.4 follows from Theorems 1.2 and 1.5.

Proof of Theorem 1.4 If ω has law $\phi_{\mathbb{G}, p, q}^W$ (we write p in the subscript of the measure instead of β) and ω^* is defined by the formula $\omega_{e^*}^* = 1 - \omega_e$, then the duality [Gri06, Theorem 6.13] for random-cluster models states that ω^* has law $\phi_{\mathbb{G}^*, p^*, q}^f$, where

$$\frac{pp^*}{(1-p)(1-p^*)} = q.$$

In particular, we need to prove that $p_c(\mathbb{G})^* = p_c(\mathbb{G}^*)$. The second item of Theorem 1.2, for quasi-transitive graphs, implies that for any $p < p_c(\mathbb{G})$,

$$\sum_{n \geq 0} \phi_{\mathbb{G}, p, q}^W \left[[n, n+1] \times [0, 1] \longleftrightarrow \{0\} \times \mathbb{R} \right] < \infty.$$

The Borel-Cantelli lemma implies that there exist only finite circuits of ω surrounding the origin almost surely. Therefore, by duality, there exists an infinite connected component in ω^* almost surely, which proves that $p^* \geq p_c(\mathbb{G}^*)$. Letting p tend to $p_c(\mathbb{G})$ gives $p_c(\mathbb{G})^* \geq p_c(\mathbb{G}^*)$.

On the other hand, ergodic properties of $\phi_{\mathbb{G}, p, q}^W$ imply that when $p > p_c(\mathbb{G})$, ω contains a unique infinite connected component almost surely (see [Gri06]). Similarly, if p^* was greater than $p_c(\mathbb{G}^*)$, ω^* would contain a unique infinite connected component almost surely (this uses a known fact [Gri06] that, above the critical point, the random-cluster model with free boundary conditions also contains an infinite connected component almost surely). Therefore, Theorem 1.5 shows that $p > p_c(\mathbb{G})$ implies $p^* \leq p_c(\mathbb{G}^*)$. Letting p tend to $p_c(\mathbb{G})$ gives $p_c(\mathbb{G})^* \leq p_c(\mathbb{G}^*)$. \square

We now turn to the proof of Theorem 1.5.

Proof of Theorem 1.5 For $R = [0, n] \times [0, k]$, denote **Top**, **Left**, **Bottom** and **Right** for the top, left, bottom and right sides of the boundary of R . Also, define the crossing probabilities

$$v(n, k) := \mu[\text{Top} \xrightarrow{R} \text{Bottom}] \quad \text{and} \quad h(n, k) := \mu[\text{Left} \xrightarrow{R} \text{Right}].$$

Lemma 4.1 Assume that both ω and ω^* contain a unique infinite connected components almost surely. Then, as $\min\{n, k\}$ tends to infinity,

- $\max\{h(n, k), v(n, k + 1)\}$ tends to 1,
- $\min\{v(n, k), h(n, k)\}$ tends to 0.

Before proving this lemma, let us explain how it implies the theorem. For each n , let k_n be the largest integer for which $v(n, k_n) \geq h(n, k_n)$ (note that by definition $v(n, k_n + 1) < h(n, k_n + 1)$). The uniqueness of the infinite connected component easily implies that k_n tends to infinity as n tends to infinity (for each fixed k , the probability that both the infinite connected component and the dual infinite connected component cross $[0, n] \times [0, k]$ from top to bottom tends to 1 as n tends to infinity).

Now, if both ω and ω^* contain infinite connected components almost surely, the first item of the previous lemma implies that $h(n, k_n)$ or $v(n, k_n + 1)$ tends to 1. This implies that $\min\{v(n, k_n), h(n, k_n)\}$ or $\min\{v(n, k_n + 1), h(n, k_n + 1)\}$ tends to 1, leading to a contradiction with the second item.

Proof of Lemma 4.1 We prove the first item. The second item is implied by the first one (with the roles of ω^* and ω exchanged) since $1 - v(n, k)$ and $1 - h(n, k)$ are the probabilities that R is respectively crossed horizontally and vertically by a path in ω^* .

Fix n, k and s (that should be thought of as satisfying $1 \ll s \ll \min\{n, k\}$). Let S_y be the translate of $S := [0, s]^2$ by $y \in \mathbb{Z}^2$. Define $x = x(R) \in R \cap \mathbb{Z}^2$ such that there exists x' and x'' neighbors of x in \mathbb{Z}^2 satisfying

$$\mu[S_x \xleftrightarrow{R} \text{Bottom}] \geq \mu[S_x \xleftrightarrow{R} \text{Top}], \quad (18)$$

$$\mu[S_x \xleftrightarrow{R} \text{Left}] \geq \mu[S_x \xleftrightarrow{R} \text{Right}], \quad (19)$$

$$\mu[S_{x'} \xleftrightarrow{R} \text{Top}] \geq \mu[S_{x'} \xleftrightarrow{R} \text{Bottom}], \quad (20)$$

$$\mu[S_{x''} \xleftrightarrow{R} \text{Right}] \geq \mu[S_{x''} \xleftrightarrow{R} \text{Left}]. \quad (21)$$

In order to see that this point exists, let X be the set of $x \in \mathbb{Z}^2 \cap R$ such that (18) holds and denote its boundary in R (i.e. the set of points in X with one neighbor in $R \setminus X$) by ∂X . Let Y and ∂Y be defined similarly with (19) instead of (18). (The sets X and Y are illustrated on Fig. 2.) Note that $\partial X \cap \partial Y \neq \emptyset$ since ∂X contains a path of neighboring vertices crossing R from left to right, and ∂Y a path from top to bottom. By definition, any point in $\partial X \cap \partial Y$ satisfies the property above.

Claim. *The distance between $x(R)$ and the boundary of R is tending to infinity as $\min\{n, k\}$ tends to infinity.*

Before proving the claim, let us show how to finish the proof. Let A_R be the event that there is a unique connected component in $\omega \cap R$ going from distance 2 of S_x to the boundary of R .

Assume that $\mu[S_x \xleftrightarrow{R} \text{Bottom}] \geq \mu[S_x \xleftrightarrow{R} \text{Left}]$. The FKG inequality together with (18) and (19) imply that

$$\mu[S_x \xleftrightarrow{R} \text{Bottom}] \leq \mu[S \xleftrightarrow{\infty}]^{1/4}. \quad (22)$$

Now, set $R' = R + (1, 0)$ and Top' for the top side of R' . We find

$$\begin{aligned} \mu[S_{x'+(0,1)} \xleftrightarrow{R'} \text{Top}'] &= \mu[S_{x'} \xleftrightarrow{R} \text{Top}] \\ &\stackrel{(20)}{\geq} \mu[S_{x'} \xleftrightarrow{R} \text{Bottom}] \\ &\geq \mu[\{S_{x'} \xleftrightarrow{\infty}\} \cap \{S_x \xleftrightarrow{R} \text{Bottom}\} \cap A_R] \\ &\stackrel{(22)}{\geq} \mu[S \xleftrightarrow{\infty}] - \mu[S \xleftrightarrow{\infty}]^{1/4} - \mu[A_R^c]. \end{aligned}$$

We deduce that

$$\begin{aligned} v(n, k + 1) &\geq \mu[\{S_x \xleftrightarrow{R} \text{Bottom}\} \cap \{S_{x'+(0,1)} \xleftrightarrow{R'} \text{Top}'\} \cap A_R] \\ &\stackrel{(22)}{\geq} \mu[S \xleftrightarrow{\infty}] - 2\mu[S \xleftrightarrow{\infty}]^{1/4} - 2\mu[A_R^c]. \end{aligned}$$

Assume now that $\mu[S_x \xleftrightarrow{R} \text{Bottom}] < \mu[S_x \xleftrightarrow{R} \text{Left}]$, the same reasoning as above, but using x'' instead of x' and (21) instead of (20), leads to the same bound as above for $h(n, k)$. The uniqueness of the infinite connected component together with the claim imply that $\mu[A_R]$ tends to 1 as $\min\{n, k\}$ tends to infinity. Letting the size s of S tend to infinity finishes the proof of the first item. To conclude the whole proof, we need to prove the claim.

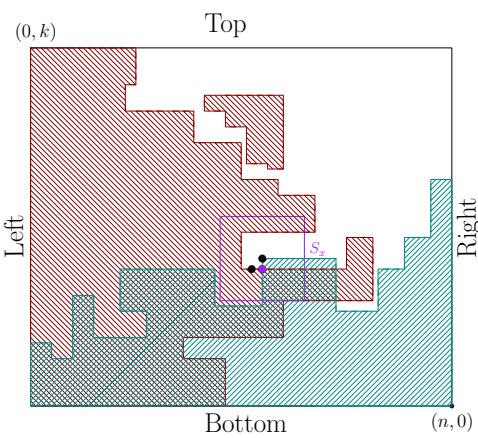


Figure 2: The vertices hatched in green are those such that $\mu[S_x \xleftrightarrow{R} \text{Bottom}] \geq \mu[S_x \xleftrightarrow{R} \text{Top}]$, and the vertices hatched in red are those which satisfy $\mu[S_x \xleftrightarrow{R} \text{Left}] \geq \mu[S_x \xleftrightarrow{R} \text{Right}]$. The point is selected in the intersection of the boundary of the two regions.

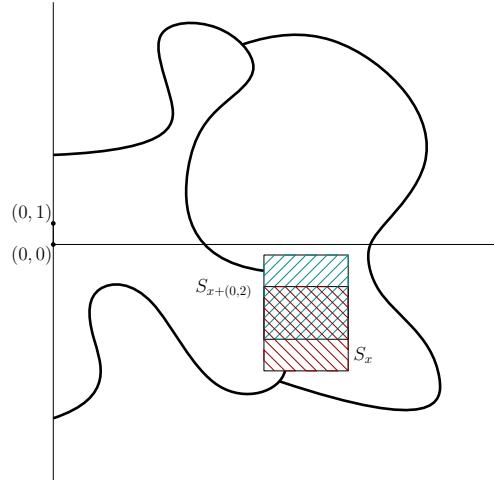


Figure 3: The construction of Claim 1. The black path prevents segment $\{0\} \times [0, 1]$ from belonging to an infinite cluster of the dual.

Proof of the claim We prove that the distance to Left is tending to infinity (the other sides work the same). Note that it is sufficient to prove that $\omega \cap \mathbb{H}$, where $\mathbb{H} = \mathbb{R}_+ \times \mathbb{R}$, does not contain any infinite connected component almost surely. To avoid introducing new notation, we prove the equivalent statement for $\omega^* \cap \mathbb{H}$ instead of $\omega \cap \mathbb{H}$, but the proof is the same. Introduce $\ell_+ := \{0\} \times \mathbb{R}_+$, $\ell_- := \{0\} \times \mathbb{R}_-$ and $\ell = \ell_- \cup \ell_+$.

For an integer $s, m \geq 0$, choose $x = x(m)$ with first coordinate equal to m satisfying

$$\mu[S_x \xleftrightarrow{\mathbb{H}} \ell_-] \geq \mu[S_x \xleftrightarrow{\mathbb{H}} \ell_+] \quad \text{and} \quad \mu[S_{x+(0,1)} \xleftrightarrow{\mathbb{H}} \ell_-] \leq \mu[S_{x+(0,1)} \xleftrightarrow{\mathbb{H}} \ell_+].$$

(This point exists since $\mu[S_x \xleftrightarrow{\mathbb{H}} \ell_\pm]$ increases to $\mu[S_x \xleftrightarrow{\mathbb{H}} \ell]$ as the second coordinate of x tends to $\pm\infty$.) The FKG inequality together with these two inequalities implies that

$$\mu[S_x \xleftrightarrow{\mathbb{H}} \ell_-] \geq 1 - \sqrt{\mu[S_x \not\xleftrightarrow{\ell}]} \quad \text{and} \quad \mu[S_{x+(0,1)} \xleftrightarrow{\mathbb{H}} \ell_+] \geq 1 - \sqrt{\mu[S_{x+(0,1)} \not\xleftrightarrow{\ell}]}.$$

Let A_m be the event that there is a unique connected component in $\omega \cap \mathbb{H}$ going from distance 2 of S_x to ℓ . Let B be the event that $\omega \cap \mathbb{H}$ does not contain an infinite connected component intersecting $\{0\} \times [0, 1]$. We find

$$\begin{aligned} \mu[B] &\geq \mu[\{S_x \xleftrightarrow{\mathbb{H}} \ell_-\} \cap \{S_{x+(0,2)} \xleftrightarrow{\mathbb{H}} \ell_+ + (0, 1)\} \cap A_m] \\ &\geq 1 - \sqrt{\mu[S_x \not\xleftrightarrow{\ell}]} - \sqrt{\mu[S_{x+(0,1)} \not\xleftrightarrow{\ell}]} - \mu[A_m^c]. \end{aligned}$$

(The construction leading to the bound above is illustrated on Fig. 3.) The uniqueness of the infinite connected component in ω^* implies that $\mu[A_m]$ tends to 1 as m tends to infinity, and also that for any $y \in \mathbb{H}$,

$$\mu[S_y \xleftrightarrow{\ell}] \geq \mu[\{S_y \xleftrightarrow{\infty}\} \cap \{S_{-y} \xleftrightarrow{\infty}\}] \stackrel{\text{FKG}}{\geq} \mu[S \xleftrightarrow{\infty}]^2. \quad (23)$$

Letting m tend to infinity and then the size s of S tend to infinity implies that $\mu[B] = 1$. This concludes the proof. \square

References

- [AB87] M. Aizenman and D. J. Barsky. Sharpness of the phase transition in percolation models. *Communications in Mathematical Physics*, 108(3):489–526, 1987.
- [ABF87] M. Aizenman, D. J. Barsky, and R. Fernández. The phase transition in a general class of Ising-type models is sharp. *J. Statist. Phys.*, 47(3-4):343–374, 1987.
- [ADS15] M. Aizenman, H. Duminil-Copin, and V. Sidoravicius. Random currents and continuity of Ising model’s spontaneous magnetization. *Communications in Mathematical Physics*, 334(2):719–742, 2015.
- [AF86] M. Aizenman and R. Fernández. On the critical behavior of the magnetization in high-dimensional Ising models. *J. Statist. Phys.*, 44(3-4):393–454, 1986.
- [Ale04] K. Alexander. Mixing properties and exponential decay for lattice systems in finite volumes. *Annals of probability*, pages 441–487, 2004.
- [Bax82] R. J. Baxter. *Exactly solved models in statistical mechanics*. Elsevier, 1982.
- [BC03] Marek Biskup and Lincoln Chayes. Rigorous analysis of discontinuous phase transitions via mean-field bounds. *Communications in mathematical physics*, 238(1):53–93, 2003.
- [BD12] V. Beffara and H. Duminil-Copin. The self-dual point of the two-dimensional random-cluster model is critical for $q \geq 1$. *Probability Theory and Related Fields*, 153(3-4):511–542, 2012.
- [BDW02] H. Buhrman and R. De Wolf. Complexity measures and decision tree complexity: a survey. *Theoretical Computer Science*, 288(1):21–43, 2002.
- [BKK⁺92] J. Bourgain, J. Kahn, G. Kalai, Y. Katznelson, and N. Linial. The influence of variables in product spaces. *Israel Journal of Mathematics*, 77(1-2):55–64, 1992.
- [BR10] B. Bollobás and O. Riordan. Percolation on self-dual polygon configurations. In *An irregular mind*, pages 131–217. Springer, 2010.
- [CIV08] M. Campanino, D. Ioffe, and Y. Velenik. Fluctuation theory of connectivities for subcritical random cluster models. *The Annals of Probability*, pages 1287–1321, 2008.
- [DGH⁺16] H. Duminil-Copin, M. Gagnebin, M. Harel, I. Manolescu, and V. Tassion. Discontinuity of the phase transition for the planar random-cluster and Potts models with $q > 4$. *arXiv preprint arXiv:1611.09877*, 2016.
- [DM16] H. Duminil-Copin and I. Manolescu. The phase transitions of the planar random-cluster and potts models with $q \geq 1$ are sharp. *Probability Theory and Related Fields*, 164(3-4):865–892, 2016.
- [DRT16] H. Duminil-Copin, A. Raoufi, and V. Tassion. A new computation of the critical point for the planar random-cluster model with $q \geq 1$. *To appear in Annales Institut Henri Poincaré, Probabilités et Statistiques*, 2016.
- [DRT17a] H. Duminil-Copin, A. Raoufi, and V. Tassion. Exponential decay of connection probabilities for subcritical Voronoi percolation in \mathbb{R}^d . *in preparation*, 2017.

- [DRT17b] H. Duminil-Copin, A. Raoufi, and V. Tassion. Subcritical phase of d -dimensional Poisson-boolean percolation and its vacant set. *in preparation*, 2017.
- [DST17] H. Duminil-Copin, V. Sidoravicius, and V. Tassion. Continuity of the phase transition for planar random-cluster and Potts models with $1 \leq q \leq 4$. *Communications in Mathematical Physics*, 349(1):47–107, 2017.
- [DT15] H. Duminil-Copin and V. Tassion. A new proof of the sharpness of the phase transition for Bernoulli percolation and the Ising model. *Communications in Mathematical Physics*, 2015. arXiv:1502.03050.
- [Dum13] H. Duminil-Copin. *Parafermionic observables and their applications to planar statistical physics models*, volume 25 of *Ensaios Matemáticos*. Brazilian Mathematical Society, 2013.
- [FK72] C. M. Fortuin and P. W. Kasteleyn. On the random-cluster model. I. Introduction and relation to other models. *Physica*, 57:536–564, 1972.
- [GG06] B. Graham and G. Grimmett. Influence and sharp-threshold theorems for monotonic measures. *Ann. Probab.*, 34(5):1726–1745, 2006.
- [Gri06] G. R. Grimmett. *The random-cluster model*, volume 333 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2006.
- [GS14] C. Garban and J. Steif. *Noise sensitivity of Boolean functions and percolation*, volume 5. Cambridge University Press, 2014.
- [KKL88] J. Kahn, G. Kalai, and N. Linial. The influence of variables on boolean functions. In *Proceeding of the 29th Annual Symposium on Foundations of Computer Science*, pages 68–80. IEEE, 1988.
- [KS82] R. Kotecký and S. B. Shlosman. First-order phase transitions in large entropy lattice models. *Comm. Math. Phys.*, 83(4):493–515, 1982.
- [LMMS⁺91] L. Laanait, A. Messager, S. Miracle-Solé, J. Ruiz, and S. Shlosman. Interfaces in the potts model i: Pirogov-sinai theory of the fortuin-kasteleyn representation. *Communications in Mathematical Physics*, 140(1):81–91, 1991.
- [Mar99] F. Martinelli. Lectures on glauber dynamics for discrete spin models. In *Lectures on probability theory and statistics*, pages 93–191. Springer, 1999.
- [Men86] M. V. Menshikov. Coincidence of critical-points in the percolation problems. *Doklady akademii nauk sssr*, 288(6):1308–1311, 1986.
- [MR16] I. Manolescu and A. Raoufi. The phase transitions of the random-cluster and potts models on slabs with $q \geq 1$ are sharp. 2016.
- [O'D14] R. O'Donnell. *Analysis of Boolean functions*. Cambridge University Press, New York, 2014.
- [OSSS05] R. O'Donnell, M. Saks, O. Schramm, and R. Servedio. Every decision tree has an influential variable. *FOCS*, 2005.
- [Pot52] R. B. Potts. Some generalized order-disorder transformations. In *Proceedings of the Cambridge Philosophical Society*, volume 48, pages 106–109. Cambridge Univ Press, 1952.
- [She05] S. Sheffield. Random surfaces. *Astérisque*, 2005.
- [SS10] O. Schramm and J. Steif. Quantitative noise sensitivity and exceptional times for percolation. *Annals of Mathematics*, 171:619–672, 2010.

- [Yao77] A. C. Yao. Probabilistic computations: Toward a unified measure of complexity. In *Proceeding of the 18th Annual Symposium on Foundations of Computer Science*, pages 222–227. IEEE, 1977.