J Stat Phys (2008) 130: 983–1009
DOI 10.1007/s10955-007-9459-x

Sharpness of the Phase Transition and Exponential Decay of the Subcritical Cluster Size for Percolation on Quasi-Transitive Graphs

Tonći Antunović · Ivan Veselić

Received: 15 December 2006 / Accepted: 17 October 2007 / Published online: 14 November 2007

Abstract We study homogeneous, independent percolation on general quasi-transitive graphs. We prove that in the disorder regime where all clusters are finite almost surely, in fact the expectation of the cluster size is finite. This extends a well-known theorem by Menshikov and Aizenman & Barsky to all quasi-transitive graphs. Moreover we deduce that in this disorder regime the cluster size distribution decays exponentially, extending a result of Aizenman & Newman. Our results apply to both edge and site percolation, as well as long range (edge) percolation. The proof is based on a modification of the Aizenman & Barsky method.

Keywords Random graphs · Edge percolation · Site percolation · Quasi-transitive graphs · Phase transition

1 Introduction

Percolation theory is devoted to the study of geometric properties of random subgraphs of a given graph. In particular, one wants to understand which disorder regimes exhibit the existence of an infinite cluster, i.e. an infinite component of the subgraph generated by the percolation process.

For percolation models on graphs the low density phase is often defined as the regime of randomness where the expected cluster-size is finite, whereas the high density phase is

T. Antunović (✉) · I. Veselić
Emmy-Noether-Programme of the Deutsche Forschungsgemeinschaft & Fakultät für Mathematik, TU Chemnitz, 09107 Chemnitz, Germany
e-mail: tonci.antunovic@mathematik.tu-chemnitz.de

Present address:
T. Antunović
Department of Mathematics, UC Berkeley, Berkeley, CA 94720, USA

Present address:
I. Veselić
Institut für Angewandte Mathematik, Universität Bonn, 53115 Bonn, Germany

defined as the disorder regime where there exists an infinite cluster almost surely. More specifically, for identically distributed, independent models there is only one scalar disorder parameter (usually denoted by p) which measures the extent of the randomness. If one denotes the supremum of the parameter values which correspond to the low density regime by p_T, and the infimum of the parameter values which correspond to the high density regime by p_H, then the statement $p_T = p_H$ is called *sharpness of the phase transition*. In other words, an intermediate phase between the low and high density regime reduces to (at most) a single value of the parameter p.

This result has been proven by Menshikov in [12] (see also [13] by Menshikov, Molchanov & Sidorenko) and Aizenman & Barsky in [1] for a large class of percolation processes on graphs. More precisely the results of [12, 13] cover independent site percolation on quasi-transitive graphs of subexponential growth. The percolation parameter can be different for the different classes of vertices. In [13], where the proof of [12] is explained in more detail, it is noted in Remark 6.1 that the method of proof works also with a relaxed growth condition on the graph, however, that it is not possible to eliminate it completely. The results of [1] hold for directed and undirected, independent, site and bond, short range percolation models on $\mathbb{Z}^d$. Since the considered graphs are essentially Cayley graphs of $\mathbb{Z}^d$ their volume growth is polynomially bounded. The results of [1] furthermore apply to so called *long range* edge percolation, a model where edges may be present between any pair of vertices, with probability decreasing in the distance between the vertices.

Any edge percolation process can be transformed into a site percolation process by passing to the line graph. If the original graph has a finite number of edge orbits under the automorphism group action, the resulting line graph will be quasi-transitive. Thus the results of [12, 13] apply to edge percolation, too. In contrast to this, if we transform a long range edge percolation process to a site percolation process via the line graph construction we lose quasi-transitivity. More precisely, to avoid triviality for long range edge percolation we need to have an infinite number of different percolation parameters assigned to the edges. Thus the edge set decomposes into an infinite number of classes, which means that the line graph will have infinitely many vertex types, violating the quasi-transitivity property.

We adapt the method of differential inequalities used in [1] for the study of $\mathbb{Z}^d$-models, to show that the sharpness of the phase transition actually holds for all quasi-transitive graphs. Again we can treat short range site and edge percolation, as well as long range edge percolation. On the technical level the differences to [1] are the following: In [1] finite torus graphs are used to approximate the infinite $\mathbb{Z}^d$ graph, which has the advantage that the approximating graph is still homogeneous, i.e. transitive. In the general case this is not possible, thus we work with finite approximation graphs which have a 'boundary'. As a consequence of this and quasi-transitivity rather than transitivity, in comparison to [1] additional finite volume terms appear in the differential inequalities. We control these correction terms to show that the modified inequalities still lead to a proof of the sharpness of the phase transition.

All graphs mentioned so far have a rich algebraic structure which is formulated in terms of transitivity or quasi-transitivity. Looking at the modifications of the Aizenman & Barsky method needed for quasi-transitive graphs one gets the impression that one can adopt the proof to obtain the same results for graphs which have uniform local combinatorial complexity bounds, without having necessarily a large automorphism group. An example would be the Penrose tiling, whose percolation properties were studied by Hof in [7]. For such graphs the finite volume effects can probably be controlled in a similar way as for quasi-transitive graphs. Müller and Richard have recently obtained related results using different techniques [15].

Apart from the fact that the equality $p_T = p_H$ establishes for homogeneous, independent models on $\mathbb{Z}^d$ that percolation has only one phase transition, it has also played a crucial role for the proof of Kesten's theorem [8], namely, that on the two dimensional lattice $\mathbb{Z}^2$ we have $p_T = p_H = 1/2$ for edge percolation.

Closely related to the sharpness of the phase transition are the results [2] by Aizenman & Newman. On the one hand they prove the divergence of the expectation value of the cluster size as the disorder parameter approaches p_T from below, a statement which is used in [1] for the proof of $p_T = p_H$. On the other hand, Aizenman & Newman show the exponential decay of the cluster size distribution for $p < p_T$. We deduce that these results hold actually for all quasi-transitive graphs.

Our interest in the sharpness of the phase transition stems from the study of percolation Hamiltonians, more precisely adjacency and combinatorial Laplacians on percolation subgraphs, and in particular their integrated density of states (IDS). While for the definition of the IDS for general graphs with a free, amenable, quasi-transitive group action, cf. [18, 19], an understanding of the phase transition(s) is not necessary, it seems that for the proof of Lifshitz asymptotics such understanding is crucial.

Lifshitz tails describe the asymptotic behavior of the IDS near the boundaries of the spectrum and have been established for independent site percolation on $\mathbb{Z}^d$ by Biskup & König [5] and independent edge percolation on $\mathbb{Z}^d$ by Kirsch & Müller, resp. Müller & Stollmann in [10, 14]. Before that Klopp and Nakamura [11] derived partial results for the random hopping model, which includes the edge percolation Hamiltonian as a special case. In [3, 4] we study subcritical independent site and edge percolation on amenable Cayley graphs. For the adjacency and the combinatorial Laplace operator we obtain the asymptotics of the IDS at the spectral edge. Our results depend on the decay rate of the cluster size distribution. They apply in particular to Lamplighter graphs which are amenable, but have exponential volume growth. Let us also remark, that the results of [3] cover combinatorial Laplacians on long range percolation graphs in the subcritical phase. This explains our objective to derive the exponential decay of the subcritical cluster size also for graphs with arbitrary growth behavior, and for long range edge percolation.

The structure of this paper is as follows: In the next section we define our percolation model and state the main results. In Sect. 3 we state some basic facts which are common for site and for edge percolation. The two subsequent sections are devoted to (short range) edge percolation. Namely, in Sect. 4 we introduce the (finite volume) order parameters and in Sect. 5 we show that they obey certain differential inequalities. In Sect. 6 we establish the same facts for site percolation. This allows us to complete the proof of our main result for both types of percolation processes in Sect. 7. The last section contains an extension of our main result to long range edge and oriented percolation.

2 Notation and Results

Let $G = (V, E)$ be an infinite, countable, connected graph, with vertex set V and edge set E. The fact that the vertices x and y are adjacent will be denoted by $x \sim y$ and $[x, y]$ will stand for the unoriented edge which connects x and y. By $d\colon V \times V \to \mathbb{R}$ we denote the usual graph distance, that is $d(x, y)$ is the length of a shortest path between two vertices x and y. For a vertex x and a nonnegative integer n, $B(x, n)$ is the ball, with center x, of radius n, in the above metric. For the sphere of radius n around x we shall write $S(x, n) := \{y \in V | d(y, x) = n\}$. The group of graph automorphisms will be denoted by $Aut(G)$. For any subgraph G', $|G'|$ will stand for the number of vertices in G', which may be infinite. When a subgraph G' is given, we will say that two neighboring vertices of G, x and y, are *directly connected in* G', if the edge $[x, y]$ is an edge of the graph G'. For two subgraphs

$G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ we define their intersection $G_1 \cap G_2 := (V_1 \cap V_2, E_1 \cap E_2)$. Notice that the intersection is always well defined. The notation $G' \subset G$ means that G' is either a proper subgraph of G or G itself.

A graph G is called *quasi-transitive*, if there exists a finite set of vertices $\mathcal{F}$ such that for any vertex x there is a $y \in \mathcal{F}$ and $\gamma \in Aut(G)$ such that $\gamma y = x$. In the following we will always assume that the set $\mathcal{F}$ is minimal with respect to inclusion. Such an $\mathcal{F}$ will be called *fundamental domain*. For any graph G we can consider the action of the group $Aut(G)$ on the set of vertices V. Thus, a graph is quasi-transitive, if and only if the set of vertices is decomposed into finitely many orbits, with respect to this action. A fundamental domain is then any set of vertices which intersects each orbit in exactly one vertex. The number of elements in any fundamental domain is the same. If a fundamental domain contains only one element we call the graph *transitive*.

Now we introduce the usual nearest neighbor Bernoulli bond percolation model. We fix some parameter $0 \le p \le 1$. For each edge of the graph G we say that it is open with probability p and closed with probability $1 - p$, independently of all other edges. That is, for each edge $e \in E$ we take a probability space $(\Omega_e, P(\Omega_e), \mathbb{P}_e)$, where $\Omega_e = \{0, 1\}$, $P(\Omega_e)$ is the power set of Ω_e and $\mathbb{P}_e(1) = p$, $\mathbb{P}_e(0) = 1 - p$. The *percolation probability space* $(\Omega, \mathcal{A}, \mathbb{P})$ is defined as the product of these probability spaces, that is $\Omega := \prod_{e \in E} \Omega_e$, $\mathcal{A} := \bigotimes_{e \in E} P(\Omega_e)$ and $\mathbb{P} := \bigotimes_{e \in E} \mathbb{P}_e$. The probability measure $\mathbb{P}$ obviously depends on p. This dependence will sometimes be stressed by writing $\mathbb{P}_p$ instead of $\mathbb{P}$. The same holds for the expectation $\mathbb{E}$.

Elements of Ω will be called *configurations* because each of them uniquely represents some configuration of open and closed edges. For a given configuration ω and a given edge e, the value ω_e will be called the *state of* e. By $G(\omega)$ denote the subgraph of G obtained by deleting all closed edges with respect to the configuration ω, i.e. for the set of vertices of $G(\omega)$ we take the set of vertices of the graph G, while the set of edges of $G(\omega)$ is the set of open edges with respect to the configuration ω. Connected components of $G(\omega)$ are called *clusters*. The cluster containing the vertex x will be denoted by $C_x(\omega)$. The probability measure is invariant under the graph automorphisms and so, in the case of a transitive graph, the probabilistic properties of $C_x(\omega)$ do not depend on the choice of x. Thus, in this case, we will often assume that a certain vertex x is fixed and denote the cluster $C_x(\omega)$ by $C(\omega)$.

The nearest neighbor site percolation model is introduced in an analogous way. For each vertex x we say it is open with some probability $p \in [0, 1]$ and otherwise closed, independently of all other vertices. In other words, we consider the probability space $(\Omega, \mathcal{A}, \mathbb{P}) := \prod_{x \in V} (\Omega_x, P(\Omega_x), \mathbb{P}_x)$, where $(\Omega_x, P(\Omega_x), \mathbb{P}_x)$ is defined in the same way as $(\Omega_e, P(\Omega_e), \mathbb{P}_e)$ before. Now for some given configuration ω the percolation graph $G(\omega)$ is defined simply as the subgraph induced by the set of open vertices with respect to the configuration ω. Clusters are again defined as connected components of $G(\omega)$. We will use the same notation as in the bond model. Note that subgraphs $G(\omega)$ do not have to contain all the vertices of G and thus the event $\{|C_x| = 0\}$ has positive probability (namely equal to $1 - p$).

Remark 1 Everything we mentioned above can be defined for any subgraph $G' = (V', E')$. Of course, the probability space will be different, but every event T in the new probability space can (and will) be identified with the cylinder set $T \times \prod_{e \notin E'} \Omega_e$ ($T \times \prod_{x \notin V'} \Omega_x$ in the case of the site model). In the corresponding probability spaces these events have the same probabilities. Thus we use the same notation for the corresponding probability measures. Since the notion of clusters in G' and G is not the same, we will denote by $C_x^{G'}$ the cluster of x in the graph G'.

Since the statements in the present section hold equally for site and for bond percolation, we will use in this section simply the term percolation. Next we will describe the most basic properties of the percolation process, without giving proofs. In Remark 8 in the next section we briefly sketch how these properties are proven.

An important property of percolation is the existence of a *phase transition* between a percolating and a non-percolating phase. Consider some fixed vertex x and the event $\{|C_x| = \infty\}$. The probability of this event $\mathbb{P}_p(|C_x| = \infty)$ is equal to 0 when $p = 0$ and 1 when $p = 1$. Furthermore, it can be shown that $\mathbb{P}_p(|C_x| = \infty)$ is a non-decreasing function of p. Thus, if we define $p_H := \sup\{p \in [0, 1]; \mathbb{P}_p(|C_x| = \infty) = 0\}$, we see that the probability $\mathbb{P}_p(|C_x| = \infty)$ is equal to 0, if $p < p_H$ and strictly positive, if $p > p_H$. In the case $p < p_H$ there is no infinite cluster almost surely, while in the case $p > p_H$ there exists an infinite cluster almost surely. The value of p_H does not depend on the vertex x. It is often called the *percolation threshold*. The case $p < p_H$ is called *subcritical phase*, the case $p > p_H$ is called *supercritical phase*, while $p = p_H$ is called *critical phase*.

If $p_H < 1$ it is obvious that $\mathbb{E}_p(|C_x|) = \infty$ for all $p > p_H$. The behavior of $\mathbb{E}_p(|C_x|)$ in the subcritical phase is much more interesting. The expectation $\mathbb{E}_p(|C_x|)$ is finite at $p = 0$ and is infinite at $p = 1$. It can be shown that it is a non-decreasing function of p. So, if we define $p_T := \sup\{p \in [0, 1], \mathbb{E}_p(|C_x|) < \infty\}$, we see that $\mathbb{E}_p(|C_x|)$ is finite, if $p < p_T$ and infinite if $p > p_T$. Like the value of p_H, the value of p_T is also independent of the choice of vertex x.

The relation $p_T \leq p_H$ between the critical values is obvious. Our goal is to prove equality of the two values. Our main result is the following.

Theorem 2 *For every quasi-transitive graph G we have $p_T = p_H$.*

As mentioned in the introduction, for general percolation models on the lattice the equality of the two critical points was proven in [1]. The method of proof was the use of differential inequalities for certain (finite volume) order parameters. In [1] one can also find a discussion of the use of such differential inequalities in other models of statistical physics. Using a different method, sharpness of the phase transition for site percolation on quasi-transitive graphs with subexponential growth was proven in [12], see Remark 4 below.

Similarly as in the lattice setting [1], Theorem 2 holds also for long range and oriented percolation on quasi-transitive graphs. To show this, one has only to modify certain steps in the proof of the basic version of Theorem 2. We present and explain these modification in the last section of this paper.

It is well known that, in the subcritical phase on the lattice, the probabilities of the events of the form $\{|C_x| \geq n\}$ decay exponentially in n. The same result holds in the case of quasi-transitive graphs.

Theorem 3 *Let G be a quasi-transitive graph and let $p < p_H$. We can find a constant $\alpha_p > 0$ such that for any positive integer n we have*

$$\mathbb{P}_p(|C_x| \geq n) \leq e^{-\alpha_p n}, \quad \textit{for any vertex } x.$$

In the lattice case, exponential decay was first proven for all p such that $\mathbb{E}_p(|C|) < \infty$. This result follows from Theorem 5.1 in [9]. The same result was proven for more general models on transitive graphs in Proposition 5.1 in [2]. Consequently, the exponential decay in the subcritical phase is just a corollary of the equality of critical points p_T and p_H. The proof of Proposition 5.1 from [2] extends directly from transitive graphs to quasi-transitive ones. Thus Theorem 3 follows directly from Theorem 2.

Remark 4 In [12, 13] Menshikov et al. pursued a different route of argument. They first show that for site percolation with $p < p_H$ on quasi-transitive graphs of subexponential growth the cluster radius distribution decays exponentially. More precisely, for every $p < p_H$ there exists a constant $\tilde{\alpha}_p > 0$ such that for all $x \in V$ and all $n \in \mathbb{N}$

$$\mathbb{P}_p(C_x \cap S(x,n) \neq \emptyset) \leq e^{-\tilde{\alpha}_p n} \tag{1}$$

holds. By the subexponential growth condition on the graph, this implies that the expected cluster size is finite. The key step in the proof of (1) is an estimate on the conditional expectation

$$\mathbb{E}_p(\, |\delta\{C_x \cap S(x,n) \neq \emptyset\}| \mid C_x \cap S(x,n) \neq \emptyset),$$

where $|\delta A|$ denotes the number of pivotal sites for the event A. Note that the estimate (1) on the cluster radius distribution is weaker than the one in Theorem 3 on the cluster size distribution.

In the proof of Theorem 2 we will need the following result.

Proposition 5 *For percolation on a quasi-transitive graph, we have for every vertex x:*

$$\lim_{p \uparrow p_T} \mathbb{E}_p(|C_x|) = \infty. \tag{2}$$

In particular, $\mathbb{E}_{p_T}(|C_x|) = \infty$ *for any* $x \in V$.

In the lattice case the divergence of $\mathbb{E}_{p_T}(|C_x|)$ was proven in Corollary 5.1 in [9]. The stronger statement (2) was then proven for more general percolation processes on transitive graphs in Lemma 3.1 in [2]. The proof of this statement for quasi-transitive graphs is essentially the same.

The versatility of the differential inequalities method as presented in [1] is illustrated by the fact that on the way to prove Theorem 2 one obtains as an aside a bound on the critical exponent δ, cf. (40) for a definition.

Proposition 6 *The critical exponent satisfies* $\delta \geq 2$.

This is a direct consequence of Lemma 22.

3 Basic Facts

Now we shall present some basic definitions and results from percolation theory. To be able to treat both the site and the bond model simultaneously, we shall denote, in the bond case, the edge set of a given graph G by S. In the site case, S will denote the vertex set of G.

Definition 7

(a) We say that the event $A \in \mathcal{A}$ is *increasing*, if

$$\omega_1 \in A, \quad \omega_1 \leq \omega_2 \Rightarrow \omega_2 \in A.$$

Here elements of Ω are ordered as functions from S to $\{0, 1\}$.

(b) We say that a random variable N is *increasing*, if for any two configurations ω_1 and ω_2, such that $\omega_1 \leq \omega_2$ we have $N(\omega_1) \leq N(\omega_2)$.

(c) We say that an event A depends only on finitely many states, if it is contained in some finite dimensional cylinder set in $\mathcal{A}$.

(d) For two increasing events A_1 and A_2, which depend only on finitely many states, we define the event

$$A_1 \circ A_2 := \{\omega \in \Omega; \quad \text{there are disjoint } S_1, S_2 \subset \operatorname{supp} \omega,$$
$$\text{such that for any } \omega_1, \omega_2 \in \Omega \ \omega_i|_{S_i} = 1 \Rightarrow \omega_i \in A_i, \ i = 1, 2\},$$

where $\operatorname{supp} \omega := \{s \in S; \omega_s = 1\}$.

(e) For an increasing event A and $\omega \in \Omega$ we say that $s_0 \in S$ is *pivotal* for A with respect to ω, if $\omega_1 \in A$ and $\omega_0 \notin A$, where ω_0 and ω_1 have the same values as ω on all elements of S except on s_0 where ω_i has value i $(i = 0, 1)$. The set $\{s_0$ is pivotal for the event $A\}$ is obviously an event.

Fundamental Tools

(a) For any increasing event A the function $p \mapsto \mathbb{P}_p(A)$ is non-decreasing.

(b) For any increasing random variable N, the function $p \mapsto \mathbb{E}_p(N)$ is non-decreasing.

(c) **Russo formula**

Suppose A is an increasing event which depends only on states of elements in some finite set S', more precisely, on $\omega|_{S'}$, where $S' \subset S$ is finite. Let $\mathbf{p} = (\mathbf{p}_s)_{s \in S'}$ be a given vector, such that $\mathbf{p}_s \in [0, 1]$, for all $s \in S'$. Let $\mathbb{P}_\mathbf{p}$ be the product probability measure constructed in the same way as the percolation measure before, by declaring an $s \in S'$ to be open with probability $\mathbf{p}_s$. Then the function $\mathbf{p} \mapsto \mathbb{P}_\mathbf{p}$ has all first partial derivatives, which satisfy

$$\frac{d\mathbb{P}_\mathbf{p}(A)}{d\mathbf{p}_s} = \mathbb{P}_\mathbf{p}(s \text{ is pivotal for } A), \quad \text{for any } s \in S'.$$

(d) **FKG inequality**

For any increasing events A_1 and A_2 we have $\mathbb{P}(A_1 \cap A_2) \geq \mathbb{P}(A_1)P(A_2)$.

(e) **BK inequality**

For any increasing events A_1 and A_2, which depend only on finitely many states, we have $\mathbb{P}(A_1 \circ A_2) \leq \mathbb{P}(A_1)\mathbb{P}(A_2)$.

For the proofs of these Fundamental Tools and more background see Chap. 2 in [6].

Remark 8 Having these results one can easily prove some claims from the previous section. Since the event $\{|C_x| = \infty\}$ is increasing, the function $p \mapsto \mathbb{P}_p(|C_x| = \infty)$ is non-decreasing. Similarly, the random variable $|C_x|$ is increasing which implies that the function

$p \mapsto \mathbb{E}_p(|C_x|)$ is non-decreasing. Using the FKG inequality one can easily show that the constants p_T and p_H do not depend on the choice of the vertex x. To see this, first notice

$$\{x \text{ is connected to } y\} \cap \{|C_y| = \infty\} \subseteq \{|C_x| = \infty\}. \tag{3}$$

Now, because G is connected, $\mathbb{P}_p(x \text{ is connected to } y) > 0$ and thus the FKG inequality and (3) imply

$$\mathbb{P}_p(|C_y| = \infty) > 0 \Rightarrow \mathbb{P}_p(|C_x| = \infty) > 0.$$

Because of the symmetry of the role played by x and y, p_H does not depend on x. To prove the claim for p_T, one decomposes $\mathbb{E}_p(|C_x|)$ in the form $\mathbb{E}_p(|C_x|) = \sum_{n=1}^{\infty} \mathbb{P}_p(|C_x| \geq n)$ and uses a relation similar to (3) for the increasing event $\{|C_x| \geq n\}$.

4 The Order Parameter

In this and the following section we shall work exclusively in the nearest neighbor Bernoulli bond percolation model. The site model will be discussed in Sect. 6.

In the remainder of the paper it will be more convenient to work with the parameter $\beta > 0$ such that $p = 1 - e^{-\beta}$, instead with p. Assuming $p_H < 1$, we can define β_T and β_H by $p_T = 1 - e^{-\beta_T}$ and $p_H = 1 - e^{-\beta_H}$. We will prove Theorem 2 in the context of β_T and β_H, but our proof works also in the case $p_H = 1$ (this case corresponds to $\beta_H = \infty$). Also we will abuse notation by writing $\mathbb{P}_\beta$ for the probability measure which corresponds to percolation with parameter $p = 1 - e^{-\beta}$, and use a similar notation for the expectation.

From now on we will assume that we are given a fixed quasi-transitive graph G. Subgraphs of G, which we will consider, will not be required to be quasi-transitive. Moreover, we will assume that some fundamental domain $\mathcal{F}$ is chosen and fixed. For each positive integer l we define a subgraph Λ_l as follows. For the set of vertices of the graph Λ_l take $\bigcup_{x \in \mathcal{F}} B(x, l)$ and connect two vertices, if and only if they are connected in the graph G.

To prove Theorem 2, we will follow the arguments in [1]. The idea of the proof there was to consider a so called *order parameter*, a function of two variables which contains information about both $\mathbb{P}_\beta(|C_x| = \infty)$ and $\mathbb{E}_\beta(|C_x|)$. For any vertex y we define the *order parameter with respect to y* by

$$M_y :]0, \infty[^2 \to [0, 1], \qquad M_y(\beta, h) := 1 - \sum_{n \in \mathbb{N}} \mathbb{P}_\beta(|C_y| = n) e^{-nh}.$$

The *order parameter M* is defined as

$$M :]0, \infty[^2 \to \mathbb{R}, \qquad M(\beta, h) := \sum_{x \in \mathcal{F}} M_x(\beta, h).$$

When a finite subgraph G' is given, we can define an analog function with respect to G'. Namely, for y in G' we define $M_y^{G'}(\beta, h) := 1 - \sum_{n \in \mathbb{N}} \mathbb{P}(|C_y^{G'}| = n) e^{-nh}$. Particularly interesting for our purposes will be the *finite volume order parameter*, defined as $M^{\Lambda_l}(\beta, h) := \sum_{x \in \mathcal{F}} M_x^{\Lambda_l}(\beta, h)$.

In the following Lemmas and Propositions we establish certain basic properties of the order parameter.

Lemma 9 *Let G' be an arbitrary subgraph of G and y an arbitrary vertex in G'. The following formula holds*

$$M_y^{G'}(\beta, h) = \sum_{n \in \mathbb{N}} \mathbb{P}(|C_y^{G'}| \geq n)(e^{-(n-1)h} - e^{-nh}). \tag{4}$$

In particular, (4) *holds in the cases $G' = G$ and $G' = \Lambda_l$.*

Proof The proof is straightforward, using $\mathbb{P}(|C_y^{G'}| \geq n) = 1 - \sum_{k=1}^{n-1} \mathbb{P}(|C_y^{G'}| = k)$. □

Proposition 10 *The order parameter M has the following properties.*

(a) *M is a non-decreasing function in both variables.*
(b) *M has a continuous partial derivative in h, and we have the formula*

$$\frac{\partial M}{\partial h}(\beta, h) = \sum_{x \in \mathcal{F}} \sum_{n \in \mathbb{N}} n \mathbb{P}_\beta(|C_x| = n) e^{-nh}. \tag{5}$$

The analogous claims hold for the finite volume order parameter.

Proof (a) Since the event $\{|C_y| \geq n\}$ is an increasing event, the probability $\mathbb{P}_\beta(|C_y| \geq n)$ is a non-decreasing function of β. From Lemma 9 it is clear that M is non-decreasing in β. On the other hand, from the definition it is clear that M is even strictly increasing in h.

(b) To prove this claim we just have to show that the series of formal partial derivatives $\sum_{n \in \mathbb{N}} n \mathbb{P}_\beta(|C_y| = n) e^{-nh}$ converges locally uniformly. But this is clear since

$$\sum_{n \in \mathbb{N}} n \mathbb{P}_\beta(|C_y| = n) e^{-nh} \leq \sum_{n \in \mathbb{N}} n e^{-nh}$$

and the latter series converges absolutely and locally uniformly. □

The functions M and $\frac{\partial M}{\partial h}$ are positive on $]0, \infty[^2$, $h \mapsto M(\beta, h)$ is non-decreasing and $h \mapsto \frac{\partial M}{\partial h}(\beta, h)$ is non-increasing. The last claim is clear from the formula (5). Thus the limits $\lim_{h \downarrow 0} M(\beta, h)$ and $\lim_{h \downarrow 0} \frac{\partial M}{\partial h}(\beta, h)$ are well-defined with values in $[0, \infty[$, respectively $[0, \infty]$. The next proposition gives the probabilistic interpretation of these limits.

Proposition 11 *For every $\beta \in]0, \infty[$ we have the following*

$$\begin{aligned} \lim_{h \downarrow 0} M(\beta, h) &= \sum_{x \in \mathcal{F}} \mathbb{P}_\beta(|C_x| = \infty), \\ \lim_{h \downarrow 0} \frac{\partial M}{\partial h}(\beta, h) &= \sum_{x \in \mathcal{F}} \mathbb{E}_\beta(|C_x|; |C_x| < \infty). \end{aligned} \tag{6}$$

Proof Since $\lim_{h \downarrow 0} e^{-nh} = 1$, for every $n \in \mathbb{N}$, using the Lebesgue monotone convergence theorem, we get $\lim_{h \downarrow 0} M_x(\beta, h) = 1 - \sum_{n \in \mathbb{N}} \mathbb{P}_\beta(|C_x| = n) = \mathbb{P}_\beta(|C_x| = \infty)$. Now the first equality in (6) follows. The second one can be proved in the same manner, using formula (5). □

Thus we obtain—as indicated earlier—the functions $\beta \mapsto \sum_{x \in \mathcal{F}} \mathbb{P}_\beta(|C_x| = \infty)$ and $\beta \mapsto \sum_{x \in \mathcal{F}} \mathbb{E}_\beta(|C_x|; |C_x| < \infty)$, which we wanted to understand in the first place, as marginals of M and $\frac{\partial M}{\partial h}$. Now we give two lemmas, which will be used repeatedly in the proof of the key inequalities presented in Propositions 17 and 18.

Lemma 12

(a) *Let G_1 be any subgraph of G, and G_2 any finite subgraph of G_1, containing some vertex x. For any nonnegative integer n we have*

$$\mathbb{P}(|C_x^{G_2}| \geq n) \leq \mathbb{P}(|C_x^{G_1}| \geq n). \tag{7}$$

Moreover, $M_x^{G_2}(\beta, h) \leq M_x^{G_1}(\beta, h)$, for all positive β and h.

(b) *Let y be a vertex of Λ_l and x be the unique element of $\mathcal{F}$ in the same orbit as y. For any nonnegative integer n we have*

$$\mathbb{P}(|C_y^{\Lambda_l}| \geq n) \leq \mathbb{P}(|C_y| \geq n) = \mathbb{P}(|C_x| \geq n). \tag{8}$$

(c) *For any vertex $x \in \mathcal{F}$ and $n \leq l$ we have*

$$\mathbb{P}(|C_x^{\Lambda_l}| \geq n) = \mathbb{P}(|C_x| \geq n). \tag{9}$$

Proof (a) Let A be an arbitrary connected subgraph of G_2, containing the vertex x. The identification from Remark 1 implies that the probabilities of the events $\{C_x^{G_2} = A\}$ and $\{A$ is the component of $C_x \cap G_2$ containing $x\}$ are equal. Similarly the probability of the event $\{A$ is the component of $C_x^{G_1} \cap G_2$ containing $x\}$ is equal to probability of $\{A$ is the component of $C_x \cap G_2$ containing $x\}$. So we can write

$$\mathbb{P}(C_x^{G_2} = A) = \mathbb{P}(A \text{ is the component of } C_x^{G_1} \cap G_2 \text{ containing } x).$$

Since the events on the right side are disjoint for different A's, we can write

$$\begin{aligned}\mathbb{P}(|C_x^{G_2}| \geq n) &= \sum_{A;|A|\geq n} \mathbb{P}(C_x^{G_2} = A) \\ &= \sum_{A;|A|\geq n} \mathbb{P}(A \text{ is the component of } C_x^{G_1} \cap G_2 \text{ containing } x) \\ &\leq \mathbb{P}(|C_x^{G_1} \cap G_2| \geq n) \leq \mathbb{P}(|C_x^{G_1}| \geq n).\end{aligned} \tag{10}$$

The last inequality follows from $\{|C_x^{G_1} \cap G_2| \geq n\} \subset \{|C_x^{G_1}| \geq n\}$. The sums in (10) are taken over all connected subgraphs of G_2, which contain x and which are of size greater or equal than n. Since G_2 is a finite graph, these sums are finite. Now the claim $M_x^{G_2}(\beta, h) \leq M_x^{G_1}(\beta, h)$ follows directly from Lemma 9.

(b) The first inequality follows directly from part (a), if we take $G_2 = \Lambda_l$ and $G_1 = G$. The second (in)equality follows from the fact that there is an automorphism τ such that $x = \tau y$ and that the probability measure is invariant under τ.

(c) For any $k < l$ and any connected subgraph A of Λ_l of size k, which contains x, the edge set and the edge boundary of A are contained in Λ_l. So it is clear that for any such A we have $\mathbb{P}(C_x^{\Lambda_l} = A) = \mathbb{P}(C_x = A)$. Taking the sum over all possible A's, when k is fixed, we obtain $\mathbb{P}(|C_x^{\Lambda_l}| = k) = \mathbb{P}(|C_x| = k)$. Taking the sum over $k < n$ and subsequently complements yields the result. □

Lemma 13 *Let y be a vertex of Λ_l and x the unique element of $\mathcal{F}$ in the same orbit as y. Then, for all $(\beta, h) \in]0,\infty[^2$, the following inequality holds*

$$M_y^{\Lambda_l}(\beta, h) \le M_x^{\Lambda_l}(\beta, h) + e^{-lh}. \tag{11}$$

Proof Using Lemma 9 we can write

$$\begin{aligned} M_y^{\Lambda_l}(\beta, h) &= \sum_{n=1}^{l} \mathbb{P}(|C_y^{\Lambda_l}| \ge n)(e^{-(n-1)h} - e^{-nh}) \\ &\quad + \sum_{l+1 \le n < \infty} \mathbb{P}(|C_y^{\Lambda_l}| \ge n)(e^{-(n-1)h} - e^{-nh}). \end{aligned} \tag{12}$$

Using parts (b) and (c) of Lemma 12 we can bound the first summand

$$\begin{aligned} &\sum_{n=1}^{l} \mathbb{P}(|C_y^{\Lambda_l}| \ge n)(e^{-(n-1)h} - e^{-nh}) \\ &\quad \le \sum_{n=1}^{l} \mathbb{P}(|C_x| \ge n)(e^{-(n-1)h} - e^{-nh}) \\ &\quad = \sum_{n=1}^{l} \mathbb{P}(|C_x^{\Lambda_l}| \ge n)(e^{-(n-1)h} - e^{-nh}) \le M_x^{\Lambda_l}(\beta, h). \end{aligned} \tag{13}$$

The second summand can be easily bounded

$$\sum_{l+1 \le n < \infty} \mathbb{P}(|C_y^{\Lambda_l}| \ge n)(e^{-(n-1)h} - e^{-nh}) \le \sum_{l+1 \le n < \infty} (e^{-(n-1)h} - e^{-nh}) = e^{-lh}. \tag{14}$$

Inserting (13) and (14) into (12) we get the result. □

Remark 14 In [1], where percolation on the lattice $\mathbb{Z}^d$ was analyzed, the finite graphs Λ_l where chosen to be tori, i.e. cubes with periodic boundary conditions. This has the advantage that percolation on the finite graphs is still homogeneous under translations. In this situation, (11) simplifies to $M_y^{\Lambda_l}(\beta, h) = M_x^{\Lambda_l}(\beta, h)$. In particular, there is no finite volume correction term e^{-lh}.

Notice that the subgraphs Λ_l exhaust the whole graph G as l goes to ∞. Therefore, in the macroscopic limit $l \to \infty$, we can expect the finite volume order parameters to behave like the order parameter. The proof of this claim is the content of the next proposition.

Proposition 15 *The finite volume order parameter M^{Λ_l} and its partial derivative $\frac{\partial M^{\Lambda_l}}{\partial h}$ converge pointwise to M and $\frac{\partial M}{\partial h}$, respectively. In other words, for all $(\beta, h) \in]0,\infty[^2$ we have*

$$\lim_{l\to\infty} M^{\Lambda_l}(\beta, h) = M(\beta, h), \qquad \lim_{l\to\infty} \frac{\partial M^{\Lambda_l}}{\partial h}(\beta, h) = \frac{\partial M}{\partial h}(\beta, h).$$

Proof Fix some $x \in \mathcal{F}$. Lemma 12(c) implies $\lim_{l\to\infty} \mathbb{P}(|C_x^{\Lambda_l}| \ge n) = \mathbb{P}(|C_x| \ge n)$, for any positive integer n. Using Lemma 9 and the Lebesgue dominated convergence theorem we get

$$\lim_{l\to\infty} M_x^{\Lambda_l}(\beta,h) = M_x(\beta,h), \quad \text{for all } (\beta,h)\in]0,\infty[^2,$$

for any $x \in \mathcal{F}$. Taking the sum over $x \in \mathcal{F}$ we get the desired result for the order parameter. The claim for the partial derivative is obtained in the same way using the formula

$$\frac{\partial M^{\Lambda_l}}{\partial h} = \sum_{x\in\mathcal{F}}\sum_{n\in\mathbb{N}} n\mathbb{P}_\beta(|C_x^{\Lambda_l}| = n)e^{-nh},$$

which is proved in the same way as the formula for $\frac{\partial M}{\partial h}$ in Proposition 10(b). □

Another way of looking at the order parameter is through the idea of "colored sites", which was used in the paper [1]. Fix a positive real $h > 0$. For every vertex y say that it is *blue* with probability $1 - e^{-h}$ independently of all other vertices. The corresponding probability space is defined similarly as the site percolation probability space before. For each vertex y define the probability space $(\Omega_y, P(\Omega_y), \mathbb{P}_y)$, where $\Omega_y := \{0,1\}$, $P(\Omega_y)$ is the power set of Ω_y and $\mathbb{P}_y(1) = 1 - e^{-h}$, $\mathbb{P}_y(0) = e^{-h}$. The probability space $(\Omega', \mathcal{A}', \mathbb{P}'_h)$ is defined as the product of these probability spaces. From now on, we shall actually work on the probability space $(\Omega, \mathcal{A}, \mathbb{P}_\beta) \times (\Omega', \mathcal{A}', \mathbb{P}'_h)$. The probability measure will be denoted by $\mathbb{P}_{\beta,h}$, but again we will often omit the subscript. The random set of blue sites will be denoted by B. Analog functions can be defined on subgraphs of G, and in this case we use the same notation as before.

The event that some vertex y is connected to some blue site with an open path will be denoted by $\{y \leftrightarrow B\}$ or by $\{C_y \cap B \neq \emptyset\}$, while $\{y \not\leftrightarrow B\}$ and $\{C_y \cap B = \emptyset\}$ will stand for the complement of this event. For connectedness with open edges which have end-vertices in some given set A we will use $\{y \leftrightarrow_A B\}$, while for connectedness with open edges which are also edges of some subgraph $G' = (V', E')$ we will write $\{y \leftrightarrow_{G'} B\}$. If V' is a given finite set of vertices and $A : \Omega \to P(V')$ a function from Ω to the power set of V' then $\{y \leftrightarrow_A B\}$ will denote the set of ω's for which there is an open path which connects y with an element in B and contains only vertices in $A(\omega)$. If A is such that for any subset $V'' \subset V'$, the set $\{A(\omega) = V''\}$ is an event, the set $\{y \leftrightarrow_A B\}$ is also an event. Similar notation will be used for random subgraphs.

The blue sites are, in some sense, identified with "infinity". For example, for some fixed vertex x, the event $\{x \leftrightarrow B\}$ is the generalization of $\{|C_x| = \infty\}$, because the open path from x which reaches some vertex in B is considered to have escaped to infinity. Intuitively, if the parameter h decreases to 0, the density of blue sites decreases to 0, and they "move further and further away" from x. So, their effect on the whole picture gets less relevant and in the limit $h \downarrow 0$ we should expect to return to our original percolation setting. Namely, the probability of the event $\{x \leftrightarrow B\}$ should converge to the probability of the event $\{|C_x| = \infty\}$. This is actually a direct consequence of Proposition 11 in view of Proposition 16. The next proposition shows the relationship between the order parameter and blue sites.

Proposition 16 *Let $G' \subset G$ be a subgraph of a quasi-transitive graph G and y some vertex in G'. Using the above notation we have*

$$M_y(\beta,h) = \mathbb{P}_{\beta,h}(y \leftrightarrow B), \qquad M_y^{G'}(\beta,h) = \mathbb{P}_{\beta,h}(y \leftrightarrow_{G'} B). \tag{15}$$

Proof We prove the first equality in (15), while the second one can be proven in the same way. It is enough to show

$$\mathbb{P}(y \nleftrightarrow B) = \sum_{n\in\mathbb{N}} \mathbb{P}(|C_y| = n)e^{-nh}. \tag{16}$$

For any positive integer n we have

$$\begin{aligned}\mathbb{P}(|C_y| = n, y \nleftrightarrow B) &= \sum_{A;|A|=n} \mathbb{P}(C_y = A, A \cap B = \emptyset)\\ &= \sum_{A;|A|=n} \mathbb{P}(C_y = A)\mathbb{P}(A \cap B = \emptyset),\end{aligned}$$

where the last equality is obtained using the independence of bond and site variables and the sums are taken over all connected subgraphs with n vertices containing y. Obviously $\mathbb{P}(A \cap B = \emptyset) = e^{-h|A|}$ and so one obtains,

$$\mathbb{P}(|C_y| = n, y \nleftrightarrow B) = \mathbb{P}(|C_y| = n)e^{-nh}. \tag{17}$$

Now, we are left to estimate $\mathbb{P}(|C_y| = \infty, y \nleftrightarrow B)$. Define the random variable $k_n := \min\{m; |C_y \cap B(y,m)| \geq n\}$ which obviously has only finite values on the event $\{|C_y| = \infty\}$. Next we can write

$$\begin{aligned}\mathbb{P}(k_n = m, y \nleftrightarrow B) &\leq \sum_A \mathbb{P}(C_y \cap B(y,m) = A, A \cap B = \emptyset)\\ &\leq \sum_A \mathbb{P}(C_y \cap B(y,m) = A)e^{-h|A|} \leq e^{-hn}\mathbb{P}(k_n = m),\end{aligned} \tag{18}$$

where the sum is taken over all possible realization of $C_y \cap B(y,m) = A$ such that the condition $k_n = m$ is fulfilled. Since

$$\mathbb{P}(|C_y| = \infty, y \nleftrightarrow B) \leq \mathbb{P}(k_n < \infty, y \nleftrightarrow B) = \sum_{m=0}^{\infty} \mathbb{P}(k_n = m, y \nleftrightarrow B),$$

for all positive integers n, using (18) we obtain $\mathbb{P}(|C_y| = \infty, y \nleftrightarrow B) \leq e^{-hn}$ for all positive integers n and thus

$$\mathbb{P}(|C_y| = \infty, y \nleftrightarrow B) = 0. \tag{19}$$

From (17) and (19) we get (16), and hence the proof is completed. □

5 Differential Inequalities for the Order Parameter

In the following we will prove two differential inequalities involving the order parameter M^{Λ_l}. These inequalities differ from the inequalities (3.1) and (3.2) in [1], by the additional term e^{-hl}. This finite volume correction appears in our situation, since for general automorphism groups we cannot use "periodic" boundary conditions and thus percolation on the finite graph Λ_l is no longer homogeneous. For residually finite automorphism groups it may be possible to use periodic boundary conditions and thus eliminate the correction term e^{-hl}.

The above-mentioned differential inequalities will be crucial for the proof of Theorem 2, because they contain essential information about the behavior of the order parameter $M(\beta, h)$ when h approaches 0. In the proof of Theorem 2 we will forget about the percolation and probability setting, and work with these inequalities instead, using analytic methods. As for the proof of these inequalities, we will use the notion of blue sites extensively.

Proposition 17 *There exists a constant $K > 0$ such that*

$$\frac{\partial M^{\Lambda_l}}{\partial \beta} \le K(M^{\Lambda_l} + e^{-lh}) \frac{\partial M^{\Lambda_l}}{\partial h}, \quad \textit{for all positive } \beta \textit{ and } h. \tag{20}$$

Proof The event $\{x \leftrightarrow_{\Lambda_l} B\}$ is increasing and depends on the states of only finitely many edges. Using the Russo formula for this event, we obtain

$$\frac{\partial M_x^{\Lambda_l}}{\partial \beta} = \sum_{[y,z]\in \Lambda_l} e^{-\beta} \mathbb{P}([y,z] \text{ is pivotal for the event } \{x \leftrightarrow_{\Lambda_l} B\}),$$

where the sum is taken over all edges $[y, z]$ in Λ_l. Since $\mathbb{P}([y, z] \text{ is closed}) = e^{-\beta}$, and the events $\{[y, z] \text{ is closed}\}$ and $\{[y, z] \text{ is pivotal for the event } \{x \leftrightarrow_{\Lambda_l} B\}\}$ are independent, we get

$$\frac{\partial M_x^{\Lambda_l}}{\partial \beta} = \sum_{[y,z]\in \Lambda_l} \mathbb{P}([y,z] \text{ is closed}, [y,z] \text{ is pivotal for the event } \{x \leftrightarrow_{\Lambda_l} B\}).$$

One should notice that

$$\begin{aligned}
&\{[y,z] \text{ is closed}, [y,z] \text{ is pivotal for the event } \{x \leftrightarrow_{\Lambda_l} B\}\} \\
&\quad = \{x \not\leftrightarrow_{\Lambda_l} B, [y,z] \text{ is pivotal for the event } \{x \leftrightarrow_{\Lambda_l} Bt\}\} \\
&\quad = \{C_x^{\Lambda_l} \cap B = \emptyset, y \in C_x^{\Lambda_l}, z \leftrightarrow_{\Lambda_l \backslash C_x^{\Lambda_l}} B\} \cup \{C_x^{\Lambda_l} \cap B = \emptyset, z \in C_x^{\Lambda_l}, y \leftrightarrow_{\Lambda_l \backslash C_x^{\Lambda_l}} B\}.
\end{aligned}$$

Here $\Lambda_l \backslash C_x^{\Lambda_l}$ stands for the graph obtained by deleting the vertices in $C_x^{\Lambda_l}$ and all incident edges of the graph Λ_l. Similar notation will be used often in the rest of the paper. Now we pass from a sum over undirected edges to a sum over directed ones and write

$$\begin{aligned}
\frac{\partial M_x^{\Lambda_l}}{\partial \beta} &= \sum_{\substack{(y,z)\in \Lambda_l^2 \\ y\sim z}} \mathbb{P}(C_x^{\Lambda_l} \cap B = \emptyset, y \in C_x^{\Lambda_l}, z \leftrightarrow_{\Lambda_l \backslash C_x^{\Lambda_l}} B) \\
&= \sum_{\substack{(y,z)\in \Lambda_l^2 \\ y\sim z}} \sum_{A; y\in A} \mathbb{P}(C_x^{\Lambda_l} = A, A \cap B = \emptyset, z \leftrightarrow_{\Lambda_l \backslash A} B),
\end{aligned} \tag{21}$$

where the last sum is taken over all connected subgraphs A of Λ_l containing x and y. The event $\{C_x^{\Lambda_l} = A\}$ depends only on the states of edges which have at least one end-vertex in A. The event $\{z \leftrightarrow_{\Lambda_l \backslash A} B\}$ depends only on the states of edges which do not have end-vertices in A and on the states of vertices outside A. Finally the event $\{A \cap B = \emptyset\}$ depends only on the states of vertices in A. Hence these events are independent. Using this independence and Proposition 16, (21) can be rewritten as

$$\frac{\partial M_x^{\Lambda_l}}{\partial \beta} = \sum_{\substack{(y,z)\in\Lambda_l^2 \\ y\sim z}} \sum_{A;\, y\in A} \mathbb{P}(C_x^{\Lambda_l} = A) M_z^{\Lambda_l\setminus A} e^{-h|A|}. \tag{22}$$

Lemmas 12(a) and 13 imply $M_z^{\Lambda_l\setminus A} \le M_z^{\Lambda_l} \le M^{\Lambda_l} + e^{-lh}$. Inserting this into (22) we obtain

$$\begin{aligned}\frac{\partial M_x^{\Lambda_l}}{\partial \beta} &\le K(M^{\Lambda_l} + e^{-lh}) \sum_A \sum_{y;\, y\in A} \mathbb{P}(C_x^{\Lambda_l} = A) e^{-h|A|} \\ &= K(M^{\Lambda_l} + e^{-lh}) \sum_A |A| \mathbb{P}(C_x^{\Lambda_l} = A) e^{-h|A|}, \end{aligned} \tag{23}$$

where the sum is taken over all possible graphs A for $C_x^{\Lambda_l}$ and K is the maximal vertex degree in the graph G. Now grouping together all A's for which $|A| = n$, we get

$$\sum_A |A| \mathbb{P}(C_x^{\Lambda_l} = A) e^{-h|A|} = \sum_{n\in\mathbb{N}} n \mathbb{P}(|C_x^{\Lambda_l}| = n) e^{-nh} = \frac{\partial M_x^{\Lambda_l}}{\partial h}. \tag{24}$$

Inserting (24) into (23) and taking the sum over $x \in \mathcal{F}$ we have proven the proposition. $\square$

Proposition 18 *The finite volume order parameter satisfies the following inequality*

$$M^{\Lambda_l} \le h \frac{\partial M^{\Lambda_l}}{\partial h} + (M^{\Lambda_l})^2 + \beta (M^{\Lambda_l} + e^{-lh}) \frac{\partial M^{\Lambda_l}}{\partial \beta}, \quad \textit{for all positive } \beta \textit{ and } h. \tag{25}$$

Proof In the proof of this inequality and especially in summations, A will denote vertex sets.

To prove (25), we have to change both our graph and probability space. Let n be an arbitrary, but fixed positive integer. For every pair of adjacent vertices $y \sim z$ in Λ_l we replace the edge $[y,z]$ with n edges which will be denoted by $[y,z]_1, [y,z]_2, \ldots, [y,z]_n$. In this way we obtain a new graph G_n'. We shall consider bond percolation on the graph G_n', and so we define the canonical percolation product probability space in the usual way. A cluster containing some vertex x will be denoted by $C_x^{G_n'}$ and its vertex set by $V_x^{G_n'}$. Notice that, for percolation on Λ_l with percolation parameter $1 - e^{-\beta}$ and for percolation on G_n' with percolation parameter $1 - e^{-\beta/n}$, the probabilities that two adjacent vertices are directly connected in the percolation graph are the same. This implies the fact that the probability of the event $\{V_x^{G_n'} = A\}$ in the new probability space is equal to the probability of the event $\{V_x^{\Lambda_l} = A\}$ in the old probability space, for all possible sets of vertices A, if the parameter is changed from β to β/n. Here $V_x^{\Lambda_l}$ stands for the vertex set of $C_x^{\Lambda_l}$. Next we define a graph G_n, which contains G_n' as a subgraph, by adding to G_n' a new vertex b, which is connected to each of the vertices of G_n' with exactly n edges. The role of the blue sites will be played by the edges incident to b. So, in addition to the percolation on G_n', we have the following rule: We fix $h' > 0$ and for every edge incident to b we say that it is open with probability $1 - e^{-h'}$, independently of the states of all other edges in the graph G_n. Notice that the events that some vertex is blue, in the old probability space, and that some vertex is directly connected to b, in the new probability space, have the same probabilities, if $h' = h/n$ holds for the respective parameters. This implies that the probabilities of the events $\{x \leftrightarrow_{\Lambda_l} B\}$ and $\{x \leftrightarrow_{G_n} b\}$ are the same, if both parameters β and h are divided by n. We will abbreviate the notation for the event $\{x \leftrightarrow_{G_n} b\}$ by writing simply by $\{x \leftrightarrow b\}$. So from now on we will assume that β and h are fixed and we shall work in the new probability space with parameters β/n and h/n, and we will keep in mind that the probabilities of the events mentioned above remain unchanged. The probability measure will still be denoted with $\mathbb{P}$.

We define the sets F_i, $i = 1, 2, 3$, as follows:

$$F_1 := \{\text{There is a unique open edge of } G_n \text{ which connects some vertex of } C_x^{G'_n} \text{ with } b\},$$

$$F_2 := \{x \leftrightarrow b\} \circ \{x \leftrightarrow b\} = \{\text{There are two edge disjoint paths from } x \text{ to } b\},$$

$$F_3 := \bigcup_{\substack{(y,z)\in\Lambda_l^2\\ y\sim z}} \bigcup_{i=1}^{n} \{[y,z]_i \text{ is open and pivotal for } \{x \leftrightarrow b\},$$
$$\{z \leftrightarrow_{G_n\backslash[y,z]_i} b\} \circ \{z \leftrightarrow_{G_n\backslash[y,z]_i} b\}\},$$

where $G_n\backslash[y,z]_i$ denotes the graph obtained from G_n by deleting the edge $[y,z]_i$. It is easy to see that these sets are events. Lemma 3.5 from [1] implies that $\{x \leftrightarrow b\}$ is a disjoint union of the F_i's and so $M_x^{\Lambda_l} = \mathbb{P}(x \leftrightarrow b) = \mathbb{P}(F_1) + \mathbb{P}(F_2) + \mathbb{P}(F_3)$.

The probability $\mathbb{P}(F_1)$ can be calculated as follows

$$\begin{aligned}
\mathbb{P}(F_1) &= \sum_A \mathbb{P}(V_x^{G'_n} = A, A \text{ is directly connected to } b \text{ with a unique open edge})\\
&= \sum_A \mathbb{P}(V_x^{G'_n} = A)\mathbb{P}(A \text{ is directly connected to } b \text{ with a unique open edge})\\
&= \sum_A \mathbb{P}(V_x^{\Lambda_l} = A) n|A|(1-e^{-h/n})e^{-(|A|-1/n)h}\\
&= n(e^{h/n}-1)\sum_A |A|\mathbb{P}(V_x^{\Lambda_l} = A)e^{-|A|h}\\
&= n(e^{h/n}-1)\frac{\partial M_x^{\Lambda_l}}{\partial h}.
\end{aligned} \tag{26}$$

Here the sums are taken over all possible realizations A of the set of vertices $V_x^{G'_n}$. The second equality follows from the independence of the bond variables in G'_n and bond variables which correspond to edges incident to b. The last equality follows from (24).

The probability of the event F_2 is easily bounded from above using the BK-inequality

$$\mathbb{P}(F_2) \le \mathbb{P}(x \leftrightarrow b)^2 = (M_x^{\Lambda_l})^2. \tag{27}$$

Now we bound the probability of the event F_3. Let $C_x^{G'_n\backslash[y,z]_i}$ be the cluster in the graph $G'_n\backslash[y,z]_i$, containing x and $V_x^{G'_n\backslash[y,z]_i}$ its vertex set. The event F_3 can be partitioned with respect to realizations of $V_x^{G'_n\backslash[y,z]_i}$. For a given set of vertices A we write $\{A \overset{d}{\leftrightarrow} b\}$ for the event that an edge between b and some vertex in A is open, and $\{A \overset{d}{\not\leftrightarrow} b\}$ for the complement of this event. We obtain

$$F_3 = \bigcup_{\substack{(y,z)\in\Lambda_l^2 \\ y\sim z}} \bigcup_{i=1}^{n} \bigcup_{A;y\in A} \{V_x^{G'_n\setminus[y,z]_i} = A, [y,z]_i \text{ is open}, A \overset{d}{\not\leftrightarrow} b, \{z \leftrightarrow_{G_n\setminus A} b\} \circ \{z \leftrightarrow_{G_n\setminus A} b\}\}.$$

The union is taken over all possible realizations A of the set of vertices $V_x^{G'_n\setminus[y,z]_i}$. Here $G_n\setminus A$ stands for the set of vertices of the graph G_n which are not elements of A. The event $\{V_x^{G'_n\setminus[y,z]_i} = A\}$ depends only on the states of edges of G'_n which have at least one endpoint in A, but not on $[y,z]_i$. The event $\{A \overset{d}{\leftrightarrow} b\}$ depends on the state of edges between b and the vertices in A. Finally, the event $\{z \leftrightarrow_{G_n\setminus A} b\} \circ \{z \leftrightarrow_{G_n\setminus A} b\}$ depends only on the state of edges of G_n which have no endpoints in A. So we see that these events are independent and also independent of the event $\{[y,z]_i \text{ is open}\}$. Using this independence and the trivial fact $\mathbb{P}([y,z]_i \text{ is open}) = (e^{\beta/n}-1)\mathbb{P}([y,z]_i \text{ is closed})$, we get

$$\begin{aligned}\mathbb{P}(F_3) \le (e^{\beta/n}-1) &\sum_{\substack{(y,z)\in\Lambda_l^2 \\ y\sim z}} \sum_{i=1}^{n} \sum_{A;y\in A} \mathbb{P}(V_x^{G'_n\setminus[y,z]_i} = A)\mathbb{P}([y,z]_i \text{ is closed}) \\ &\times \mathbb{P}(\{z \leftrightarrow_{G_n\setminus A} b\} \circ \{z \leftrightarrow_{G_n\setminus A} b\})\mathbb{P}(A \overset{d}{\not\leftrightarrow} b). \qquad (28)\end{aligned}$$

Notice that the BK inequality, Lemma 12(a) and Lemma 13 imply

$$\begin{aligned}\mathbb{P}(\{z \leftrightarrow_{G_n\setminus A} b\} \circ \{z \leftrightarrow_{G_n\setminus A} b\}) &\le \mathbb{P}(z \leftrightarrow_{G_n\setminus A} b)^2 \\ &\le \mathbb{P}(z \leftrightarrow_{G_n\setminus A} b)(M^{\Lambda_l}(\beta,h) + e^{-lh}). \qquad (29)\end{aligned}$$

Inserting (29) into (28) and using the independence again we obtain

$$\begin{aligned}\mathbb{P}(F_3) &\le (e^{\beta/n}-1)(M^{\Lambda_l} + e^{-lh}) \\ &\quad \times \sum_{(y,z)\in\Lambda_l^2} \sum_{i=1}^{n} \sum_{A;y\in A} \mathbb{P}(V_x^{G'_n\setminus[y,z]_i} = A, [y,z]_i \text{ is closed}) \\ &\quad \times \mathbb{P}(A \overset{d}{\not\leftrightarrow} b)\mathbb{P}(z \leftrightarrow_{G_n\setminus A} b). \qquad (30)\end{aligned}$$

In the last sum there are no contributions from A's which contain z, because, for such A's, the set $\{z \leftrightarrow_{G_n\setminus A} b\}$ is empty. So we can take the sum over all possible realizations A of $V_x^{G'_n\setminus[y,z]_i}$ which contain y but not z. For such A's it is clear that

$$\{V_x^{G'_n\setminus[y,z]_i} = A, [y,z]_i \text{ is closed}\} = \{V_x^{G'_n} = A\}.$$

So (30) can be written in the form

$$\mathbb{P}(F_3) \le n(e^{\beta/n}-1)(M^{\Lambda_l} + e^{-lh}) \sum_{\substack{(y,z)\in\Lambda_l^2 \\ y\sim z}} \sum_{A;y\in A} \mathbb{P}(V_x^{G'_n} = A)\mathbb{P}(A \overset{d}{\not\leftrightarrow} b)\mathbb{P}(z \leftrightarrow_{G_n\setminus A} b).$$

The events $\{V_x^{G'_n} = A\}$, $\{A \overset{d}{\nleftrightarrow} b\}$ and $\{z \leftrightarrow_{G_n \setminus A} b\}$ have the same probabilities as the events $\{V_x^{\Lambda_l} = A\}$, $\{A \cap B = \emptyset\}$ and $\{z \leftrightarrow_{\Lambda_l \setminus A} B\}$, respectively. Since the latter events are independent, we can write

$$\mathbb{P}(F_3) \le n(e^{\beta/n} - 1)(M^{\Lambda_l} + e^{-lh}) \sum_{\substack{(y,z)\in\Lambda_l^2 \\ y\sim z}} \sum_{A;\, y\in A} \mathbb{P}(V_x^{\Lambda_l} = A, A \cap B = \emptyset, z \leftrightarrow_{\Lambda_l \setminus A} B).$$

Using (21), we obtain

$$\mathbb{P}(F_3) \le n(e^{\beta/n} - 1)(M^{\Lambda_l} + e^{-lh}) \frac{\partial M_x^{\Lambda_l}}{\partial \beta}. \tag{31}$$

Summing (26), (27) and (31) we get

$$M_x^{\Lambda_l} \le n(e^{h/n} - 1)\frac{\partial M_x^{\Lambda_l}}{\partial h} + (M_x^{\Lambda_l})^2 + n(e^{\beta/n} - 1)(M^{\Lambda_l} + e^{-lh})\frac{\partial M_x^{\Lambda_l}}{\partial \beta}.$$

For fixed β and h let n go to ∞ and obtain

$$M_x^{\Lambda_l} \le h\frac{\partial M_x^{\Lambda_l}}{\partial h} + (M_x^{\Lambda_l})^2 + \beta(M^{\Lambda_l} + e^{-lh})\frac{\partial M_x^{\Lambda_l}}{\partial \beta}. \tag{32}$$

Now sum (32) over $x \in \mathcal{F}$, use the fact that $\sum_{x\in\mathcal{F}}(M_x^{\Lambda_l})^2 \le (\sum_{x\in\mathcal{F}} M_x^{\Lambda_l})^2$, and (25) is proved. □

6 Site Model

In this part we shall explain how to obtain inequalities similar to those in (20) and (25), in the case of the site model. In the next section we shall use this inequalities to prove Theorem 2. We will follow the arguments from the bond model and explain modifications necessary to proceed in the site case. In the following, just like before, G will always denote a quasi-transitive graph. Throughout the section, $\overline{G'}$ will denote the subgraph induced by the set of vertices which lie in the subgraph G' or have a neighbor in G'.

First we shall slightly change the notion of the cluster. We shall adopt the definitions from Sect. 7 in [1]. Let G be an arbitrary quasi-transitive graph, x an arbitrary element of G and ω an arbitrary site configuration. We define the modified cluster $\widetilde{C}_x(\omega)$ as the subgraph of G induced by the following set of vertices:

$$x \cup \{y;\text{ there is a path from } x \text{ to } y \text{ including only open sites and } x\}.$$

The distribution of the random variable $|\widetilde{C}_x|$ is clearly different from the distribution of $|C_x|$. However, the probabilities $\mathbb{P}(|\widetilde{C}_x| = n)$, $n \in \mathbb{N} \cup \{\infty\}$ will be proportional to $\mathbb{P}(|C_x| = n)$, $n \in \mathbb{N} \cup \{\infty\}$. Therefore the critical values p_H and p_T will remain the same and the property of exponential decay below p_H will be preserved.

For any finite subgraph G', the cluster $\widetilde{C}_x^{G'}(\omega)$ will be defined accordingly. Since the state of the site variable at some vertex x is irrelevant for the properties of the cluster at x, we shall actually define the cluster $\widetilde{C}_x^{G'}(\omega)$ for all vertices x and all subgraphs G' such that $x \in \overline{G'}$. Note that the relation "x lies in the cluster of y" is not symmetric anymore and thus clusters can no longer be represented as connected components of some percolation subgraph.

The functions M_y, M and their finite volume counterparts will be defined in the same way as before, where one just replaces C_x by $\widetilde{C}_x$. Note that for these definitions as well as for Lemma 9, and Propositions 10 and 11 one only needs to define the notion of cluster, while the underlying model is completely irrelevant. This is why these results transfer directly to the site percolation setting. Lemma 12 holds in the site model as well and the proof remains practically the same. This also holds for Lemma 13 and Proposition 15. The notion of blue sites is introduced in the same way as in section 4. Similarly as above, the state of the site variable at some vertex x is irrelevant for the connectedness to any other vertex and so by $x \leftrightarrow_{G'} B$ we will denote the event that x is connected to some blue site by a path in which all vertices, except maybe x, lie in the subgraph G'. Proposition 16 remains unchanged in the site setting.

Now we establish differential inequalities for the site model. In the proofs one has to be careful not to mix two types of site variables, those which correspond to the percolation process and those which correspond to blue sites. In particular, pivotality will refer only to percolation variables. Similarly as before $\Lambda_l \backslash G'$ will denote the subgraph of Λ_l obtained by deleting all vertices in G' and all edges which are incident to some vertex in G'.

The formula in Proposition 17 remains the same in the site percolation setting.

Proposition 19 *There exists a constant $K > 0$ such that*

$$\frac{\partial M^{\Lambda_l}}{\partial \beta} \le K(M^{\Lambda_l} + e^{-lh}) \frac{\partial M^{\Lambda_l}}{\partial h}, \quad \textit{for all positive } \beta \textit{ and } h. \tag{33}$$

Proof The proof is analogous to the proof of Proposition 17. Using the Russo formula one obtains

$$\frac{\partial M_x^{\Lambda_l}}{\partial \beta} = \sum_{y \in \Lambda_l} \mathbb{P}(y \text{ is closed, } y \text{ is pivotal for the event } \{x \leftrightarrow_{\Lambda_l} B\}).$$

We notice that

$$\begin{aligned} &\{y \text{ is closed, } y \text{ is pivotal for the event } \{x \leftrightarrow_{\Lambda_l} B\}\} \\ &\quad = \{\widetilde{C}_x^{\Lambda_l} \cap B = \emptyset, y \in \overline{\widetilde{C}_x^{\Lambda_l}}, y \leftrightarrow_{\Lambda_l \backslash \overline{\widetilde{C}_x^{\Lambda_l}}} B\}. \end{aligned}$$

Now using independence and presenting the above events as unions over possible realizations of $\widetilde{C}_x^{\Lambda_l}$ we obtain by means of the site version of Proposition 16

$$\frac{\partial M_x^{\Lambda_l}}{\partial \beta} = \sum_{y \in \Lambda_l} \sum_{\substack{A \subset \Lambda_l \\ y \in \overline{A} \backslash A}} \mathbb{P}(\widetilde{C}_x^{\Lambda_l} = A) M_y^{\Lambda_l \backslash \overline{A}} e^{-h|A|}. \tag{34}$$

Using the site versions of Lemmas 12(a) and 13 we get

$$\begin{aligned} \frac{\partial M_x^{\Lambda_l}}{\partial \beta} &\le (M^{\Lambda_l} + e^{-lh}) \sum_{A \subset \Lambda_l} |\overline{A} \backslash A| \mathbb{P}(\widetilde{C}_x^{\Lambda_l} = A) e^{-h|A|} \\ &\le K(M^{\Lambda_l} + e^{-lh}) \sum_{A \subset \Lambda_l} |A| \mathbb{P}(\widetilde{C}_x^{\Lambda_l} = A) e^{-h|A|} \end{aligned}$$

$$= K(M^{\Lambda_l} + e^{-lh})\frac{\partial M_x^{\Lambda_l}}{\partial h}, \tag{35}$$

where K is again the maximal vertex degree in G. The last equality in (35) is proven just like its bond analogue (see (23)). □

In Proposition 18 we will modify one term on the left hand side. This is due to the fact that we will not change the graph Λ_l as dramatically as in the proof of Proposition 18.

Proposition 20 *The finite volume order parameter satisfies the following inequality*

$$M^{\Lambda_l} \le h\frac{\partial M^{\Lambda_l}}{\partial h} + (M^{\Lambda_l})^2 + (e^{\beta} - 1)(M^{\Lambda_l} + e^{-lh})\frac{\partial M^{\Lambda_l}}{\partial \beta}, \quad \textit{for all positive } \beta \textit{ and } h. \tag{36}$$

Proof As in the proof of Proposition 18 we will change both our graph and probability space. Rather than considering the probability space of blue sites we shall add a new vertex b to the graph Λ_l. This vertex will be connected to each vertex in Λ_l by exactly n edges. Notice that the new graph, which will be denoted by G_n, contains Λ_l as a subgraph. We shall now consider a mixed site-bond percolation model in which each vertex of Λ_l is open with probability $1 - e^{-\beta}$ and in which each bond incident to b is open with probability $1 - e^{-h/n}$. Again we shall compare this process with the usual percolation model with blue sites which are generated on each vertex with probability $1 - e^{-h}$. For the relation between these two processes see the discussion at the beginning of the proof of Proposition 18. The event $\{x \leftrightarrow_{G_n} b\}$ can be partitioned into following disjoint events:

$$F_1 := \{\text{There is a unique open edge which connects some vertex of } \widetilde{C}_x^{\Lambda_l} \text{ with } b\},$$

$$\begin{aligned} F_2 &:= \{x \leftrightarrow_{G_n} b\} \circ \{x \leftrightarrow_{G_n} b\} \\ &= \{\text{There are two paths from } x \text{ to } b, \text{ which have no common vertices other than } x \\ &\quad \text{and } b\}, \end{aligned}$$

$$F_3 := \bigcup_{y \in \Lambda_l} \{y \in \widetilde{C}_x^{\Lambda_l} \text{ is open and pivotal for } \{x \leftrightarrow_{G_n} b\},\ \{y \leftrightarrow_{G_n} b\} \circ \{y \leftrightarrow_{G_n} b\}\}.$$

The probability of the first event can be calculated in the same way as in the bond case. We obtain:

$$\mathbb{P}(F_1) = n(e^{h/n} - 1)\frac{\partial M_x^{\Lambda_l}}{\partial h}. \tag{37}$$

Since $\mathbb{P}(x \leftrightarrow_{G_n} b) = \mathbb{P}(x \leftrightarrow_{\Lambda_l} B)$, the BK inequality implies

$$\mathbb{P}(F_2) \le \mathbb{P}(x \leftrightarrow_{G_n} b)^2 = (M_x^{\Lambda_l})^2. \tag{38}$$

The last event can be rewritten as

$$F_3 = \bigcup_{y \in \Lambda_l} \bigcup_{A;\, y \in \overline{A} \setminus A} \{\widetilde{C}_x^{\Lambda_l \setminus \{y\}} = A,\ A \overset{d}{\not\leftrightarrow} b,\ y \text{ open},\ \{y \leftrightarrow_{G_n \setminus \overline{A}} b\} \circ \{y \leftrightarrow_{G_n \setminus \overline{A}} b\}\},$$

where the second union is taken over all possible realizations A of $\widetilde{C}_x^{\Lambda_l \setminus \{y\}}$, for which $y \in \overline{A} \setminus A$ (note that $\overline{A}$ is defined as the closure of A in Λ_l). Considering the above formula and using independence we obtain

$$\mathbb{P}(F_3) \le (e^{\beta}-1)\sum_{y\in\Lambda_l}\sum_{A;y\in\overline{A}\setminus A}\mathbb{P}(\widetilde{C}_x^{\Lambda_l\setminus\{y\}}=A, y \text{ is closed})$$

$$\times\,\mathbb{P}(\{y\leftrightarrow_{G_n\setminus\overline{A}} b\}\circ\{y\leftrightarrow_{G_n\setminus\overline{A}} b\})\mathbb{P}(A \overset{d}{\not\leftrightarrow} b)$$

$$\le (e^{\beta}-1)(M^{\Lambda_l}+e^{-lh})\sum_{y\in\Lambda_l}\sum_{A;y\in\overline{A}\setminus A}\mathbb{P}(\widetilde{C}_x^{\Lambda_l}=A)M_y^{\Lambda_l\setminus\overline{A}}e^{-h|A|}$$

$$= (e^{\beta}-1)(M^{\Lambda_l}+e^{-lh})\frac{\partial M_x^{\Lambda_l}}{\partial\beta}. \tag{39}$$

In the second inequality we used the BK inequality and the site versions of Lemma 12(a), Lemma 13 and Proposition 16 in the same way as in the proof of Proposition 18. In the last equality we used (34). Now the result follows after taking the sum of (37), (38) and (39) and letting n tend to ∞. □

7 Completion of the Proof of Theorem 2

In this section we will complete the proof of our main result, using the differential inequalities (20) and (25).

The next result will be useful in the proof of Lemma 22. It is a special case of Lemma 4.1 in [1].

Lemma 21 *Let $M:\mathbb{R}^+\to\mathbb{R}$ be an increasing differentiable function of h obeying*

$$\lim_{h\downarrow 0} M(h)=0,\qquad \lim_{h\downarrow 0}\frac{M(h)}{h}=\infty,$$

$$M\le h\frac{dM}{dh}+M^2+kM^2\frac{dM}{dh},\quad \text{for all } h>0,$$

for some positive constant k. Then there exists a constant $c>0$ such that for all $h>0$ small enough we have

$$M(h)\ge c\sqrt{h}.$$

Lemma 22 and Proposition 23 are the final steps of the proof of Theorem 2. They correspond to Theorem 4.2 and Lemma 5.1 from [1]. We just need some adjustments in the proof of Proposition 23 because of somewhat different differential inequalities. As an aside we obtain an upper bound on the critical exponent defined as

$$\delta := \liminf_{h\downarrow 0}\frac{\ln h}{M(\beta_T,h)}. \tag{40}$$

Lemma 22 *There is a constant $c>0$, such that for $h>0$ small enough*

$$M(\beta_T,h)\ge c\sqrt{h}. \tag{41}$$

In particular, the critical exponent (40) obeys $\delta\ge 2$.

Proof First notice that M satisfies the following differential inequality

$$M(\beta,h) \le h\frac{\partial M}{\partial h}(\beta,h) + M^2(\beta,h) + K\beta M^2(\beta,h)\frac{\partial M}{\partial h}(\beta,h), \tag{42}$$

in the bond case, respectively

$$M(\beta,h) \le h\frac{\partial M}{\partial h}(\beta,h) + M^2(\beta,h) + K(e^{\beta}-1)M^2(\beta,h)\frac{\partial M}{\partial h}(\beta,h), \tag{43}$$

in the site case. These inequalities can be proven by inserting (20) into (25) in the bond case (or inserting (33) into (36) in the site case) and letting l go to ∞. If $\lim_{h\downarrow 0} M(\beta_T,h) > 0$ there is nothing to prove. Suppose $\lim_{h\downarrow 0} M(\beta_T,h) = 0$. This implies by Proposition 11 that $\mathbb{P}_{\beta_T}(|C_x| = \infty) = 0$, for all vertices x and

$$\lim_{h\downarrow 0}\frac{\partial M}{\partial h}(\beta_T,h) = \sum_{x\in\mathcal{F}}\mathbb{E}_{\beta_T}(|C_x|; |C_x|<\infty) = \sum_{x\in\mathcal{F}}\mathbb{E}_{\beta_T}(|C_x|) = \infty,$$

by Proposition 5. Now the Mean Value Theorem implies

$$\lim_{h\downarrow 0}\frac{M(\beta_T,h)}{h} = \infty. \tag{44}$$

In view of (42) (respectively (43) in the site case) and (44), the claim follows directly from Lemma 21. □

Except for having to control the term e^{-hl} in (33) and (36), the proof of the next proposition is the same as the proof of Lemma 5.1 in [1].

Proposition 23 *For any $\beta' > \beta_T$ we can find a positive constant $d > 0$ such that*

$$\lim_{h\downarrow 0} M(\beta,h) \ge d(\beta-\beta_T) \tag{45}$$

holds for every $\beta \in [\beta_T,\beta']$.

Proof Let's consider the bond case first. Change the variables (β,h) to $(\beta,\ln h)$, i.e. define $u := \ln h$ and $\widetilde{M}^{\Lambda_l}(\beta,u) = M^{\Lambda_l}(\beta,h)$. Now $\frac{\partial \ln \widetilde{M}^{\Lambda_l}}{\partial u}(\beta,u) = \frac{h}{M^{\Lambda_l}(\beta,h)}\frac{\partial M^{\Lambda_l}}{\partial h}(\beta,h)$ and so (25) can be rewritten as

$$1 \le \frac{\partial \ln \widetilde{M}^{\Lambda_l}}{\partial u}(\beta,u) + \widetilde{M}^{\Lambda_l}(\beta,u) + \beta\left(1+\frac{e^{-le^u}}{\widetilde{M}^{\Lambda_l}(\beta,u)}\right)\frac{\partial \widetilde{M}^{\Lambda_l}}{\partial \beta}(\beta,u), \tag{46}$$

for every $u\in\mathbb{R}$. Now fix some $0<h_1<h_2$, define $u_1 := \ln h_1$ and $u_2 := \ln h_2$ and integrate (46) over the rectangle $[\beta_T,\beta_1]\times[u_1,u_2]$, where β_1 is an arbitrary real number between β_T and β'. Using the fact that $\widetilde{M}^{\Lambda_l}$ is increasing in both β and u and switching back to h, we get

$$\begin{aligned}(\beta_1-\beta_T)\ln\frac{h_2}{h_1} &\le (\beta_1-\beta_T)\ln\frac{M^{\Lambda_l}(\beta_1,h_2)}{M^{\Lambda_l}(\beta_T,h_1)} + (\beta_1-\beta_T)\ln\frac{h_2}{h_1}M^{\Lambda_l}(\beta_1,h_2)\\ &\quad + \beta'\ln\frac{h_2}{h_1}\left(1+\frac{e^{-lh_1}}{M^{\Lambda_l}(\beta_T,h_1)}\right)(M^{\Lambda_l}(\beta_1,h_2)-M^{\Lambda_l}(\beta_T,h_1)).\end{aligned}$$

Let l go to ∞ and obtain

$$\beta_1 - \beta_T \le (\beta_1 - \beta_T) \frac{\ln \frac{M(\beta_1, h_2)}{M(\beta_T, h_1)}}{\ln \frac{h_2}{h_1}} + (\beta_1 - \beta_T) M(\beta_1, h_2) + \beta'(M(\beta_1, h_2) - M(\beta_T, h_1)). \tag{47}$$

Now notice

$$\frac{\ln \frac{M(\beta_1, h_2)}{M(\beta_T, h_1)}}{\ln \frac{h_2}{h_1}} = \frac{\ln M(\beta_1, h_2) - \ln M(\beta_T, h_1)}{\ln h_2 - \ln h_1} = \frac{\frac{\ln M(\beta_1, h_2)}{\ln h_1} - \frac{\ln M(\beta_T, h_1)}{\ln h_1}}{\frac{\ln h_2}{\ln h_1} - 1}. \tag{48}$$

Using Lemma 22 and (48) we get

$$\limsup_{h_1 \downarrow 0} \frac{\ln \frac{M(\beta_1, h_2)}{M(\beta_T, h_1)}}{\ln \frac{h_2}{h_1}} \le \frac{1}{2}. \tag{49}$$

Inserting (49) to (47) and letting $h_1 \downarrow 0$ leads to

$$\begin{aligned} \frac{1}{2}(\beta_1 - \beta_T) &\le M(\beta_1, h_2)(\beta_1 - \beta_T + \beta') - \beta' \lim_{h_1 \downarrow 0} M(\beta_T, h_1) \\ &\le (2\beta' - \beta_T) M(\beta_1, h_2). \end{aligned}$$

Let $h_2 \downarrow 0$ and the proof is over.

In the site model we start by changing the variables (β, h) to $(p, u) := (1 - e^{-\beta}, \ln h)$. In other words, this time we define $\widetilde{M}^{\Lambda_l} :]0, 1[\times \mathbb{R} \to \mathbb{R}$ such that $\widetilde{M}^{\Lambda_l}(p, u) := M^{\Lambda_l}(\beta, h)$. Now (36) can be rewritten as

$$1 \le \frac{\partial \ln \widetilde{M}^{\Lambda_l}}{\partial u}(p, u) + \widetilde{M}^{\Lambda_l}(p, u) + p\left(1 + \frac{e^{-le^u}}{\widetilde{M}^{\Lambda_l}(p, u)}\right) \frac{\partial \widetilde{M}^{\Lambda_l}}{\partial p}(p, u). \tag{50}$$

This inequality replaces (46) but has the same form. Thus the proof continues the same way as in the bond case after making the transformations $\beta_T \mapsto p_T := 1 - e^{-\beta_T}$ and $\beta' \mapsto p' := 1 - e^{-\beta'}$. □

Proof of Theorem 2 Proposition 23 tells us that $\lim_{h \downarrow 0} M(\beta, h)$ is positive as soon as $\beta > \beta_T$. In the view of Proposition 11 this proves the main result. □

8 Extension of Results to General Bond Models

In this section we will explain how the methods presented above can be applied to more general bond percolation models on quasi-transitive graphs. The model we present here is the *partially oriented long-range model* which was considered in [1] for the lattice case.

Assume that $G = (V, E)$ is again a quasi-transitive graph, with some fixed fundamental domain $\mathcal{F}$. Now make the graph complete, that is connect each pair of vertices $\{x, y\}$ with an unoriented edge $[x, y]$. Moreover, connect x and y with two oriented edges $[x, y\rangle$ (oriented from x to y) and $[y, x\rangle$ (oriented from y to x). The distance function on the vertices is the one inherited from the graph G. Thus it makes sense to define the length of an edge as the distance (in G) between its endvertices. Paths in our graph can contain both oriented and unoriented edges, but the orientation of oriented edges must be in accordance with the orientation of the considered path.

On the complete graph the usual nearest neighbor bond percolation is uninteresting, because any parameter $p > 0$ will correspond to the supercritical phase. To avoid this triviality, one has to introduce certain damping of the probabilities that x and y are connected, as the distance between x and y goes to infinity. This is done by introducing for each pair of vertices (x, y) two positive parameters $J_{[x,y]}$ and $J_{[x,y\rangle}$. The unoriented edge $[x, y]$ will be open with probability $1 - e^{-\beta J_{[x,y]}}$ and the oriented edge $[x, y\rangle$ will be open with probability $1 - e^{-\beta J_{[x,y\rangle}}$. Of course, we assume that all these events are mutually independent and thus the product probability space can be constructed similarly as before. The structure of the quasi-transitive graph G is reflected through the invariance of the parameters J: we assume that the parameters J are invariant under the automorphisms of the graph G. In other words, $J_{[\gamma x,\gamma y]} = J_{[x,y]}$ and $J_{[\gamma x,\gamma y\rangle} = J_{[x,y\rangle}$, for all $\gamma \in Aut(G)$ and all vertices x and y. Next we define $J_x := \sum_{y\in V}(J_{[x,y]} + J_{[x,y\rangle})$. To avoid the triviality mentioned above, we will assume that

$$J_0 := \sup_{x\in V} J_x = \max_{x\in\mathcal{F}} J_x < \infty. \tag{51}$$

Without this assumption, some vertices would be directly connected with infinitely many other vertices almost surely.

The subgraphs Λ_l are also defined similarly as before, using the distance function of the original graph G. The vertex set remains unchanged, but for the set of edges we take all possible oriented and unoriented edges between pairs of vertices contained in Λ_l. Percolation on the graph Λ_l inherits the probabilities for edges to be open from the percolation on the whole graph.

Since the graph contains oriented edges, the relation "being connected in a percolation subgraph" defined on the set of vertices is not symmetric any more and thus the notion of the connected components is now meaningless. However, the percolation cluster containing some vertex x can be defined in a natural way, as the graph $C_x(\omega)$ for which the vertex set is the set of all vertices which can be reached from x by an open path. The edge set is defined as the set of all open edges between vertices of $C_x(\omega)$. A percolation cluster $C_x^{\Lambda_l}(\omega)$ in Λ_l is defined similarly. Using this new definition of clusters, the order parameter M and the finite volume order parameter M^{Λ_l} can be defined in the same way as before. The probabilistic interpretation with colored sites is also applicable just as before, since Proposition 16 is true in this setting, too.

The critical parameters β_T and β_H are defined in the same way as before. In the nearest neighbor model, the fact that these values are well defined relied on the Fundamental Tools presented in Sect. 3. Both Definition 7 and the Fundamental Tools can be generalized to the present model in a natural way. These results are scattered in the literature. For example, a general Russo inequality can be found in [16] and a general BK inequality can be found in [17]. For more explanations, one can also look at the arguments in [1] regarding this general model. Just as before, these generalizations imply the fact that the critical parameters are well defined. The generalizations of the Fundamental Tools also ensure that the basic properties of the order parameter remain valid in the new model. Lemma 9 and Propositions 10 and 11 are still valid in the new setting. One can easily convince oneself that this is also true for parts (a) and (b) of Lemma 12. However, we need to be more careful with part (c). Rather than the equality stated in this part, in the general model we obtain an inequality formulated in Lemma 24 below. This inequality will be used in Lemma 25 which replaces

Lemma 13. Under the same assumptions as in Lemma 13, Lemma 25 gives the following inequality

$$M_y^{\Lambda_l}(\beta,h) \le M_x^{\Lambda_l}(\beta,h) + f_l(\beta,h), \tag{52}$$

where $(f_l)_{l\in\mathbb{N}}$ is some sequence of positive continuous functions which converges to zero locally uniformly for $l \to \infty$. Notice that this bound is sufficient to prove differential inequalities similar to those in (20) and (25). Namely, one obtains the following differential inequalities

$$\frac{\partial M^{\Lambda_l}}{\partial \beta} \le K(M^{\Lambda_l} + f_l)\frac{\partial M^{\Lambda_l}}{\partial h}, \quad \text{and} \tag{53}$$

$$M^{\Lambda_l} \le h\frac{\partial M^{\Lambda_l}}{\partial h} + (M^{\Lambda_l})^2 + \beta(M^{\Lambda_l} + f_l)\frac{\partial M^{\Lambda_l}}{\partial \beta}. \tag{54}$$

These inequalities are sufficient to conclude the equality of the critical values $\beta_H = \beta_T$. More precisely, the proof of Lemma 22 extends to our new setting. For this we use the fact that Proposition 5 is also true in this general model (this also follows from the proof of Lemma 3.1 in [2]). Proposition 23 still gives the main result, since its proof does not require any special form of the functions (f_l), but only the fact, that they decay locally uniformly for $l \to \infty$.

Now we state the mentioned inequality which replaces part (c) of Lemma 12.

Lemma 24 *There exists a nondecreasing sequence of positive integers $(n_l)_{l\in\mathbb{N}}$, which converges to infinity and a sequence of positive continuous functions $(g_l)_{l\in\mathbb{N}}$, $g_l\colon]0,\infty[\to \mathbb{R}$ which converges to zero locally uniformly for $l \to \infty$, such that the following inequality holds for any $x \in \mathcal{F}$, any $l \in \mathbb{N}$, and any positive integer $1 \le k \le n_l$*

$$\mathbb{P}(|C_x| \ge k) \le \mathbb{P}(|C_x^{\Lambda_l}| \ge k) + g_l(\beta), \quad \textit{for all } \beta \in]0,\infty[. \tag{55}$$

Proof We follow the arguments in the proof of Lemma A.3 from [1]. For any positive real r define

$$J_r := \max_{x\in\mathcal{F}} \sum_{\substack{y\in V\\ d(x,y)\ge r}} (J_{[x,y]} + J_{[x,y\rangle}).$$

Since J_0 is finite, $\lim_{r\to\infty} J_r = 0$. In the following we will use the identification from Remark 1 in our new setting. We have to estimate $\mathbb{P}(|C_x| \ge k, |C_x^{\Lambda_l}| < k)$. For any $\omega \in \{|C_x| \ge k, |C_x^{\Lambda_l}| < k\}$ there exists a path consisting of at most k edges which connects some x with some vertex outside Λ_l. This path connects two vertices which are at distance greater or equal to l. Thus there has to be an edge in this path which has length greater or equal to l/k. To reach this edge we have to make j steps in the path, for some j such that $0 \le j \le k-1$. The probability that there exists an open edge of length greater or equal than l/k which can be reached from x by an open path of length j, can be bounded above by $\beta J_{l/k}(\beta J_0)^j$. Here we used the inequality $1 - e^{-t} \le t$ for any positive t. So from the arguments above we deduce that

$$\mathbb{P}(|C_x| \ge k, |C_x^{\Lambda_l}| < k) \le \beta J_{l/k} \sum_{j=1}^{k} (\beta J_0)^{j-1}. \tag{56}$$

For any positive integer n define $K_n := n\sum_{k=1}^{n}(nJ_0)^{k-1}$. Since $\lim_{r\to\infty} J_r = 0$, we can find an increasing sequence of positive integers L_n such that

$$\lim_{n\to\infty} J_{L_n/n}K_n = 0. \tag{57}$$

Now define $n_l := \max\{n; L_n \le l\}$ and $g_l(\beta) := \beta J_{l/n_l}\sum_{k=1}^{n_l}(\beta J_0)^{k-1}$. From (57) it is clear that $\lim_{l\to\infty} g_l(\beta) = 0$ locally uniformly on $\mathbb{R}^+$. The other claimed properties of the sequences $(n_l)_l$ and $(g_l)_l$ are obvious. Using (56) and the fact that $r \mapsto J_r$ is a non-increasing function we obtain $\mathbb{P}(|C_x| \ge k, |C_x^{\Lambda_l}| < k) \le g_l(\beta)$. This proves the lemma. □

Using the previous result and part (a) of Lemma 12 which, as we said, still holds in our new setting, one easily obtains $\lim_{l\to\infty}\mathbb{P}(|C_x^{\Lambda_l}| \ge k) = \mathbb{P}(|C_x| \ge k)$. This can be used to prove the pointwise convergence of the finite volume order parameter to the order parameter and the same claim for the partial derivative in h, that is Proposition 15.

Using Lemma 24 one can easily obtain an inequality as in (52).

Lemma 25 *Let y be a vertex of Λ_l and x be the unique element of $\mathcal{F}$ in the same orbit as y. Then there exists a sequence of positive continuous functions $(f_l)_{l\in\mathbb{N}}$, converging locally uniformly to 0 for $l \to \infty$, such that the following inequality holds*

$$M_y^{\Lambda_l}(\beta,h) \le M_x^{\Lambda_l}(\beta,h) + f_l(\beta,h), \quad \textit{for all } (\beta,h) \in]0,\infty[^2. \tag{58}$$

Proof Using a similar decomposition as in the proof of Lemma 13, and then Lemma 24 and Lemma 9 we get

$$\begin{aligned} M_y^{\Lambda_l}(\beta,h) &= \sum_{k=1}^{n_l}\mathbb{P}(|C_y^{\Lambda_l}| \ge k)(e^{-(k-1)h} - e^{-kh}) + \sum_{n_l+1\le k<\infty}\mathbb{P}(|C_y^{\Lambda_l}| \ge k)(e^{-(k-1)h} - e^{-kh}) \\ &\le \sum_{k=1}^{n_l}\mathbb{P}(|C_x^{\Lambda_l}| \ge k)(e^{-(k-1)h} - e^{-kh}) + \sum_{k=1}^{n_l} g_l(\beta)(e^{-(k-1)h} - e^{-kh}) + e^{-n_l h} \\ &\le M_x^{\Lambda_l}(\beta,h) + g_l(\beta) + e^{-n_l h}. \end{aligned} \tag{59}$$

Since $(n_l)_l$ converges to infinity, the claim of the lemma is proven. □

References

1. Aizenman, M., Barsky, D.J.: Sharpness of the phase transition in percolation models. Commun. Math. Phys. **108**(3), 489–526 (1987)
2. Aizenman, M., Newman, C.M.: Tree graph inequalities and critical behavior in percolation models. J. Stat. Phys. **36**(1-2), 107–143 (1984)
3. Antunović, T., Veselić, I.: Equality of Lifshitz and van Hove exponents on amenable Cayley graphs. http://www.arxiv.org/abs/0706.2844
4. Antunović, T., Veselić, I.: Spectral asymptotics of percolation Hamiltonians on amenable Cayley graphs. In: Proceedings of OTAMP 2006. Operator Theory: Advances and Applications (2007, in press)
5. Biskup, M., König, W.: Long-time tails in the parabolic Anderson model with bounded potential. Ann. Probab. **29**(2), 636–682 (2001)
6. Grimmett, G.: Percolation, Grundlehren der Mathematischen Wissenschaften, vol. 321. Springer, Berlin (1999)

7. Hof, A.: Percolation on Penrose tilings. Can. Math. Bull. **41**(2), 166–177 (1998)
8. Kesten, H.: The critical probability of bond percolation on the square lattice equals $1/2$. Commun. Math. Phys. **74**(1), 41–59 (1980)
9. Kesten, H.: Percolation Theory for Mathematicians. Progress in Probability and Statistics, vol. 2. Birkhäuser, Boston (1982)
10. Kirsch, W., Müller, P.: Spectral properties of the Laplacian on bond-percolation graphs. Math. Z. **252**(4), 899–916 (2006). http://www.arXiv.org/abs/math-ph/0407047
11. Klopp, F., Nakamura, S.: A note on Anderson localization for the random hopping model. J. Math. Phys. **44**(11), 4975–4980 (2003)
12. Men'shikov, M.: Coincidence of critical points in percolation problems. Sov. Math. Dokl. **33**, 856–859 (1986)
13. Men'shikov, M.V., Molchanov, S.A., Sidorenko, A.F.: Percolation theory and some applications. In: Probability Theory. Mathematical Statistics. Theoretical Cybernetics (Russian), Itogi Nauki i Tekhniki, vol. 24, pp. 53–110. Akad. Nauk SSSR Vsesoyuz. Inst. Nauchn. i Tekhn. Inform., Moscow (1986). Translated in J. Sov. Math. **42**(4) 1766–1810. http://dx.doi.org/10.1007/BF01095508
14. Müller, P., Stollmann, P.: Spectral asymptotics of the Laplacian on supercritical bond-percolation graphs. http://www.arxiv.org/math-ph/0506053
15. Müller, P., Richard, C.: Random colourings of aperiodic graphs: ergodic and spectral properties. http://www.arxiv.org/abs/0709.0821
16. Russo, L.: On the critical percolation probabilities. Z. Wahrscheinlichkeitstheor. Verw. Geb. **56**(2), 229–237 (1981)
17. van den Berg, J., Kesten, H.: Inequalities with applications to percolation and reliability. J. Appl. Probab. **22**(3), 556–569 (1985)
18. Veselić, I.: Quantum site percolation on amenable graphs. In: Proceedings of the Conference on Applied Mathematics and Scientific Computing, pp. 317–328. Springer, Dordrecht (2005). http://arXiv.org/math-ph/0308041
19. Veselić, I.: Spectral analysis of percolation Hamiltonians. Math. Ann. **331**(4), 841–865 (2005). http://arXiv.org/math-ph/0405006