1

PERCOLATION AND RANDOM GRAPHS

Remco van der Hofstad
Department of Mathematics and Computer Science
Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven, The Netherlands.

Abstract

In this section, we define percolation and random graph models, and survey the features of these models.

1.1 Introduction and notation

In this section, we discuss random networks. In Section 1.2, we study *percolation*, which is obtained by independently removing vertices or edges from a graph. Percolation is a model of a porous medium, and is a paradigm model of statistical physics. Think of the bonds in an infinite graph that are not removed as indicating whether water can flow through this part of the medium. Then, the interesting question is whether water can *percolate*, or, alternatively, whether there is an infinite connected component of bonds that are kept? As it turns out, the answer to this question depends sensitively on the fraction of bonds that are kept. When we keep most bonds, then the kept or occupied bonds form most of the original graph. In particular, an infinite connected component may exist, and if this happens, we say that the system **percolates**. On the other hand, when most bonds are removed or vacant, then the connected components tend to be small and insignificant. Thus, percolation admits a *phase transition*. Despite the simplicity of the model, the results obtained up to date are far from complete, and many aspects of percolation, particularly of its critical behavior, are ill understood. In Section 1.2 we shall discuss the basics of percolation, and highlight some important open questions. The key challenge in percolation is to uncover the relation between the percolation critical behavior and the properties of the underlying graph from which we obtain percolation by removing edges.

In Section 1.3, we discuss *random graphs*. While in percolation, the random network considered naturally lives on an *infinite graph*, in random graph theory one considers random *finite* graphs. Thus, all random graphs are obtained by removing edges from the complete graph, or by adding edges to an empty graph. An important example of a random graph is obtained by *independently* removing bonds from a finite graph, which makes it clear that there is a strong link to percolation. However, also other mechanisms are possible to generate a random graph. We shall discuss some of the basics of random graph theory, focussing

on the phase transition of the largest connected component and the distances in random graphs. The random graph models studied here are inspired by applications, and we shall highlight real-world networks that these random graphs aim to model to some extent.

The fields that this contribution covers, percolation and random graph theory, have attracted tremendous attention in the past decades, and enormous progress has been made. It is impossible to cover all material appearing in the literature, and we believe that one should not aim to do so. Thus, we have strived to cover the main results which have been proved, as well as recent results in which we expect that more progress shall be made in the (near?) future, and we list open problems which we find of interest. We have tried to give particular attention to results that are of interest to the stochastic geometry community, thus giving detailed accounts of the recent progress on two-dimensional percolation and percolation on tesselations, as well as on continuum percolation and random geometric graphs and its relations to telecommunications. We hope that this contribution gives an idea of the breadth and depth of these fields, as well as on the challenges for the future.

We now start by introducing some notation. Let $\mathbb{G} = (\mathbb{V}, \mathbb{E})$ be a graph, where $\mathbb{V}$ is the vertex set and $\mathbb{E} \subseteq \mathbb{V} \times \mathbb{V}$ is the edge set. For percolation, the number of vertices, denoted by $|\mathbb{V}|$, is naturally infinite, while for random graphs, $|\mathbb{V}|$ is naturally finite. A random network is obtained by a certain rule that determines which subset of the edges $\mathbb{E}$ is occupied, the remaining edges being vacant. Let $v \in \mathbb{V}$, and denote by $C(v)$ the set of vertices which can be reached from v by occupied edges. More precisely, for $u \in \mathbb{V}$, we say that $u \longleftrightarrow v$ when there exists a path of occupied edges that connects u and v, and we write

$$C(v) = \{u \in \mathbb{V} : u \longleftrightarrow v\}. \tag{1.1}$$

The central question in the study of random networks involves the cluster size distributions, i.e., for percolation whether there exists an infinite connected component, and for random graphs what the distribution is of the largest connected component.

1.2 Percolation

In this section, we discuss percolation. For more details, see the books (Grimmett, 1999), (Hughes, 1996), (Kesten, 1982) and (Bollobás and Riordan, 2006*b*). For an expository account of recent progress with a long list of open problems, we refer to (Kesten, 2002). There is a nice account of the history of percolation in (Hughes, 1996, Section 1.1.1), the foundation of percolation as a mathematical discipline being generally ascribed to (Broadbent and Hammersley, 1957).

Introduction of the model.

In this section, $\mathbb{G} = (\mathbb{V}, \mathbb{E})$ shall denote an infinite graph. We shall assume that $\mathbb{G} = (\mathbb{V}, \mathbb{E})$ is *transitive*, i.e., the neighborhoods of all points are the same. More precisely, transitivity means that for every $x, y \in \mathbb{V}$, there exists a bijection

$\phi\colon \mathbb{V} \to \mathbb{V}$ for which $\phi(x) = y$ and $\{\phi(u), \phi(v)\} \in \mathbb{E}$ precisely when $\{u, v\} \in \mathbb{E}$. Such a bijection $\phi\colon \mathbb{V} \to \mathbb{V}$ is called an *automorphism*. In particular, transitivity of a graph $\mathbb{G}$ implies that each vertex has the same degree. We shall denote the degree of $\mathbb{G}$ by r. We sometimes assume the weaker condition of *quasi-transitivity*, which means that there is a *finite* set of vertices such that for each vertex v, there is a graph automorphism of the graph which maps v to one of the vertices in the finite set. We note that if a graph is quasi-transitive and each vertex has a bounded degree, then the degrees are uniformly bounded.

Bond percolation is obtained by independently removing each of the bonds or edges with a fixed probability $1 - p$. Thus, each edge is *occupied* with probability p, and *vacant* otherwise, and the edge statuses are independent. We shall restrict to the setting where the probability that a edge is occupied is fixed. In the literature, also the setting is studied where $\mathbb{E} = \mathbb{V} \times \mathbb{V}$ and, for $b \in \mathbb{E}$, the probability that b is occupied depends on b in a translation invariant way. For simplicity, we refrain from this generality. The resulting probability measure is denoted by $\mathbb{P}_p$, and $\mathbb{E}_p$ denotes expectation w.r.t. $\mathbb{P}_p$.

We define the *percolation function $p \mapsto \theta(p)$* by

$$\theta(p) = \mathbb{P}_p(|C(v)| = \infty), \tag{1.2}$$

where $v \in \mathbb{V}$ is an arbitrary vertex. By transitivity, the above probability does not depend on the choice of v. We shall therefore often write $C = C(o)$ where $o \in \mathbb{V}$ is an appropriately chosen origin. When $\mathbb{G}$ is quasi-transitive, then $\theta^v(p) = \mathbb{P}_p(|C(v)| = \infty)$ possibly depends on $v \in \mathbb{V}$, but these is a finite connection of vertices V_0 such that for every v there exists a $v_0 \in V_0$ such that $\theta^v(p) = \theta^{v_0}(p)$.

When $\theta(p) = 0$, then the probability that o is inside an infinite connected component is 0, so that there will be no infinite connected component a.s. When $\theta(p) > 0$, on the other hand, then, by ergodicity, the proportion of vertices in infinite connected components equals $\theta(p) > 0$, and we say that the system *percolates*. For quasi-transitive graphs, if $\theta^v(p) = 0$ for *some* v, then, in fact, $\theta^v(p) = 0$ for *all* v.

We define the *critical value* by

$$p_c = p_c(\mathbb{G}) = \inf\{p : \theta(p) > 0\}. \tag{1.3}$$

For quasi-transitive graphs, it might appear that this definition depends on the choice of v for which $\theta(p) = \theta^v(p)$. However, since $\theta^v(p) = 0$ for some v implies that $\theta^v(p) = 0$ for all v, in fact, $p_c(\mathbb{G})$ is independent of the choice of v. The above critical value is sometimes written as $p_c(\mathbb{G}) = p_H(\mathbb{G})$ in honor of Hammersley, who defined it in (Hammersley, 1961).

An important question is whether the critical value is non-trivial, i.e., whether $p_c \in (0, 1)$. We shall study this question in detail below. When $\theta(p) > 0$, then a natural question is *how many* infinite connected components there can be. Denote this number of infinite connected components by N. We shall now prove that, for any infinite graph, $N \in \{0, 1, \infty\}$ a.s., and that N is constant a.s.

Indeed, by ergodicity arguments and the fact that N is a *translation invariant* random variable, $N = k$ almost surely for some $k \in \{0, 1, 2, \ldots\} \cup \{\infty\}$. Moreover, since N is a.s. constant, it is not changed when we change the status of a *finite* number of edges. Indeed, for $B \subseteq \mathbb{E}$, let $N_B(0)$ denote the number of infinite components when all edges in B are declared vacant, and $N_B(1)$ the number of infinite components when all edges in B are declared occupied. Then $N_B(0) = N_B(1) = k$ a.s. When $k < \infty$, $N_B(0) = N_B(1) = k$ only when B intersects *at most one* infinite connected component, and we conclude that the number of infinite connected components which intersect with B is at most 1 a.s. When $B \uparrow \mathbb{V}$, this number increases to N, and we arrive at the claim that if $k < \infty$, then $k \leq 1$. This completes the proof that $N = k$ a.s., for some $k \in \{0, 1, \infty\}$.

Instead of considering bond percolation, one can also study *site percolation*, for which we independently and with fixed probability $1-p$ remove the vertices in $\mathbb{V}$, and we are only allowed to make use of edges for which *both* endpoints are kept. In the literature (see e.g. (Grimmett, 1999)), the main focus has been on bond percolation, despite the fact that, as we shall show now, site percolation is more general. Indeed, we shall show that for each bond percolation model indicated by $\mathbb{G}$, there exists a site percolation model, denoted by $\mathbb{G}_s$, which is equivalent to bond percolation on $\mathbb{G}$. Indeed, we take $\mathbb{V}_s$ to contain the edges in $\mathbb{E}$, and say that $\{a_s, b_s\} \in \mathbb{E}_s$ precisely when, in $\mathbb{G}$, the edges to which a_s and b_s are identified share a common endpoint. In this setting, bond percolation on $\mathbb{G}$ becomes site percolation on $\mathbb{G}_s$, and the connected component of $v \in \mathbb{V}$ is infinite precisely when there exists an edge $b \in \mathbb{E}$ such that the connected component of $b \in \mathbb{V}_s$ is infinite. The reverse is not true, i.e., not every site percolation model is equivalent to a bond percolation model. In this paper, we shall restrict to bond percolation, having in mind that almost all arguments can be straightforwardly adapted to the site percolation setting, possibly except for certain *duality* arguments which play an important role in two dimensions, and are described in more detail in Section 1.2.4. Interestingly, already in (Hammersley, 1961) it was shown that the critical value for bond percolation is never larger than the one for site percolation.

It is useful to note that obviously percolation is *monotone* in the parameter p. To make this notion precise, we say that an event E is *increasing*, when for every $\omega \in E$ and $\eta \geq \omega$, where $\eta \geq \omega$ when $\eta(e) \geq \omega(e)$ for every edge $e \in \mathbb{E}$, we have that also $\eta \in E$. For example, the event $\{|C(v)| = \infty\}$ is increasing. Then, we can couple percolation for all probabilities p simultaneously as follows. We let $\{U_e\}_{e \in \mathbb{E}}$ be i.i.d. uniform random variables, and note that percolation with percolation probability p is obtained by declaring an edge e occupied precisely when $U_e \leq p$ (see (Hammersley, 1963)). This implies that when $p_1 \leq p_2$ and E an arbitrary increasing event, we have that

$$\mathbb{P}_{p_1}(E) \leq \mathbb{P}_{p_2}(E). \tag{1.4}$$

In particular, we obtain that $\theta(p_1) \leq \theta(p_2)$, i.e., $p \mapsto \theta(p)$ is non-decreasing. We also say that a random variable X is increasing when $\{X \geq x\}$ is increasing for all $x \in \mathbb{R}$.

We mention two inequalities that play a profound role in percolation theory, namely the FKG and BK-inequalities. The FKG-inequality in the context of percolation is called the *Harris' inequality* and was first proved in (Harris, 1960). The more general FKG inequality, which, for example, also applies to the Ising model, was derived in (Fortuin, Kasteleyn and Ginibre, 1971). The Harris' inequality states that for two increasing events E and F,

$$\mathbb{P}_p(E \cap F) \geq \mathbb{P}_p(E)\mathbb{P}_p(F), \tag{1.5}$$

the FKG-inequality gives the same conclusion under weaker assumptions on the measure involved. In words, for increasing events E and F, the occurrence of E makes the simultaneous occurrence of F more likely. For example, the FKG-inequality yields that, for every $x, y, u, v, \in \mathbb{V}$, we have $\mathbb{P}_p(x \longleftrightarrow y, u \longleftrightarrow v) \geq \mathbb{P}_p(x \longleftrightarrow y)\mathbb{P}_p(u \longleftrightarrow v)$. The intuition for the FKG-inequality is that if the increasing event E holds, then this makes it more likely for edges to be occupied, and, therefore, it becomes more likely that the increasing event F also holds. Thus, $\mathbb{P}_p(F|E) \geq \mathbb{P}_p(F)$, which is equivalent to (1.5). See (Häggström, 2007) for a Markov chain proof of the FKG-inequality.

The BK-inequality gives, in a certain sense, an opposite inequality. We shall only state it in the case of increasing events, for which it was proved in (van den Berg and Kesten, 1985). The most general version is proved in (Reimer, 2000). For $K \subseteq \mathbb{E}$ and $\omega \in \{0,1\}^{\mathbb{E}}$, we write $\omega_K(e) = \omega(e)$ for $e \in K$, and $\omega_K(e) = 0$ otherwise. Let E and F again be increasing events, and write

$$E \circ F = \{\omega : \exists K \subseteq \mathbb{E} \text{ such that } \omega_K \in E, \omega_{K^c} \in F\}. \tag{1.6}$$

Then, the van den Berg-Kesten (BK) inequality states that

$$\mathbb{P}_p(E \circ F) \leq \mathbb{P}_p(E)\mathbb{P}_p(F). \tag{1.7}$$

For example, the event $\{x \longleftrightarrow y\} \circ \{u \longleftrightarrow v\}$ is the event that there are edge-disjoint occupied paths from x to y and from u to v, and (1.7) implies that $\mathbb{P}_p(\{x \longleftrightarrow y\} \circ \{u \longleftrightarrow v\}) \leq \mathbb{P}_p(x \longleftrightarrow y)\mathbb{P}_p(u \longleftrightarrow v)$. Intuitively, this can be understood by noting that, if $x \longleftrightarrow y$ and $u \longleftrightarrow v$ must occur disjointly, then we can first fix an occupied path connecting x and y in a certain arbitrary manner, and remove the occupied edges used in this path. Then $\{x \longleftrightarrow y\} \circ \{u \longleftrightarrow v\}$ occurs when in the configuration with the edges removed, we still have that $u \longleftrightarrow v$. Since we have removed the edges in the occupied path from x to y, this event now has smaller probability than $\mathbb{P}_p(u \longleftrightarrow v)$.

Many objects we shall study are increasing or decreasing. For example, $|C(v)|$ is obviously increasing. The number of infinite connected components N is an example of a random variable that is neither increasing nor decreasing, and we shall see the complications of this fact later on.

We next discuss an important tool to study probabilities which goes under the name of Russo's formula (Russo, 1981). Let E be an increasing event. Then we say that the bond (u, v) is *pivotal* for the event E when E occurs when the

status of (u, v) in the (possibly modified) configuration where (u, v) is turned occupied, while E does not occur in the (possibly modified) configuration where (u, v) is turned vacant. Thus, the bond (u, v) is essential for the occurrence of the event E. The set of pivotal bonds for an event is *random*, as it depends on which other bonds are occupied and vacant in the configuration. Russo's Formula states that for every increasing event E which depends on a *finite number of bonds*,

$$\frac{d}{dp}\mathbb{P}_p(E) = \sum_{(u,v)\in\mathbb{E}} \mathbb{P}_p\big((u, v) \text{ is pivotal for } E\big). \tag{1.8}$$

Russo's Formula allows us to study how the probability of an event changes as p varies. The fact that (1.8) is only valid for events that only depend on a finite number of bonds is a nuisance, and there are many settings in which Russo's Formula can be extended to events depending on infinitely many bonds by an appropriate cutting procedure.

We shall be interested in several key functions that describe the connections in bond percolation. The *susceptibility* $p \mapsto \chi(p)$ is the expected cluster size

$$\chi(p) = \mathbb{E}_p[|C(v)|]. \tag{1.9}$$

Clearly, we have that $\chi(p) = \infty$ for $p > p_c$, since then, with probability $\theta(p) > 0$, we have that $|C(v)| = \infty$. Further, $p \mapsto \chi(p)$ is clearly increasing. Define the critical value $p_T = p_T(\mathbb{G})$ by

$$p_T = \sup\{p : \chi(p) < \infty\}. \tag{1.10}$$

The subscript T in $p_T(\mathbb{G})$ is in honor of Temperley.

A natural question is whether $p_T(\mathbb{G}) = p_c(\mathbb{G})$, i.e., is $\chi(p) < \infty$ for every $p < p_c$? The latter indeed turns out to be true by the celebrated results independently proved by (Menshikov, 1986) and (Aizenman and Barsky, 1987), as we shall discuss in more detail below.

For $p \in [0, 1]$, let

$$\chi^{\mathrm{f}}(p) = \mathbb{E}_p[|C(v)|\mathbb{1}_{\{|C(v)|<\infty\}}] \tag{1.11}$$

denote the mean finite cluster size. Clearly, for $p < p_T$, we have that $\chi^{\mathrm{f}}(p) = \chi(p)$, but for $p > p_T$, this may not be true. We define the *two-point function* $\tau_p \colon \mathbb{V} \times \mathbb{V} \to [0, 1]$ by

$$\tau_p^{\mathrm{f}}(x, y) = \mathbb{P}_p(x \longleftrightarrow y, |C(x)| < \infty). \tag{1.12}$$

On $\mathbb{Z}^d$, when the model is translation invariant, we have that $\tau_p^{\mathrm{f}}(x, y) = \tau_p^{\mathrm{f}}(y - x, 0) \equiv \tau_p^{\mathrm{f}}(y - x)$. Also for transitive $\mathbb{G}$, the two-point function is characterized by $\tau_p^{\mathrm{f}}(x) = \tau_p^{\mathrm{f}}(o, x)$, where o is an appropriately chosen origin in the model. In terms of $\tau_p^{\mathrm{f}}(x)$, we can identify $\chi^{\mathrm{f}}(p)$ as

$$\chi^{\mathrm{f}}(p) = \sum_{x\in\mathbb{V}} \tau_p^{\mathrm{f}}(x). \tag{1.13}$$

We shall also be interested in the *mean number of clusters per vertex* $\kappa(p)$, which is defined as

$$\kappa(p) = \mathbb{E}_p[1/|C(v)|]. \tag{1.14}$$

The significance of $\kappa(p)$ is that it measures the average number of connected components in large volumes. Indeed, let $B(n) = \{x \in \mathbb{V} : d(o,x) \leq n\}$ denote a ball of radius n around o, where $d(x,y)$ denotes an appropriately chosen distance function on $\mathbb{G}$. Then, with C_n the number of different connected components obtained when making every edges not entirely in $B(n)$ vacant, and when $\partial B(n) = o(B(n))$, $C_n/|B(n)| \to \kappa(p)$ a.s.

An important measure of the *spatial extent* of clusters is the *correlation length* $\xi(p)$ defined by

$$\xi(p) = \sqrt{\frac{1}{\chi^{\mathrm{f}}(p)} \sum_{x \in \mathbb{V}} |x|^2 \tau_p^{\mathrm{f}}(x)}, \tag{1.15}$$

where we write $|x| = d(o,x)$ for the distance of x to an appropriately chosen origin $o \in \mathbb{V}$. For many graphs $\mathbb{G}$, there are several ways of defining the correlation length, many of them being equivalent in the sense that they are bounded above and below by finite and positive constants times $\xi(p)$ defined in (1.15). The correlation length measures the *dependence* between finite clusters at a given distance. If $d(x,y) \gg \xi(p)$ and, for $p > p_c$, x and y are in different finite clusters, then we can think of $C(x)$ and $C(y)$ as being close to independent, while if $d(x,y) \ll \xi(p)$, then $C(x)$ and $C(y)$ are quite dependent.

1.2.1 *Critical behavior.*

The behavior of percolation models is most interesting and richest for p values which are close to the critical value. Clearly, the precise value of $p_c(\mathbb{G})$ depends sensitively on the nature of the underlying graph $\mathbb{G}$. By drawing an analogy to physical systems, physicists predict that the behavior of percolative systems close to criticality is rather insensitive to the precise details of the model, and it is only characterized by the macroscopic behavior. Thus, percolation is expected to behave in a *universal* manner. For example, it is predicted that the critical nature of finite-range percolation systems on $\mathbb{Z}^d$, under suitable symmetry conditions, is similar in the sense that the nature of the critical behavior are similar. While this prediction is far from rigorous, it does offer us a way of summarizing percolation models by only looking at their simplest examples. It is one of the key challenges in percolation of making the paradigm of universality rigorous.

We shall now make the notion of universality more tangible, by discussing *critical exponents*. The critical nature of many physical systems is believed to be characterized by the validity of *power laws*, the exponent of which is a robust or universal measure of the underlying critical behavior. We start by giving an example of a critical exponent. It is predicted that

$$\theta(p) \sim (p - p_c)^\beta, \qquad \text{as} \qquad p \downarrow p_c, \tag{1.16}$$

for some $\beta > 0$. The value of β is expected to be different for different $\mathbb{G}$, but (1.16) remains valid. The symbol $\sim$ in (1.16) can have several meanings, which we now elaborate on. We say that the critical exponent β exists in the *logarithmic form* if

$$\lim_{p \downarrow p_c} \frac{\log \theta(p)}{\log (p - p_c)} = \beta, \tag{1.17}$$

while β exists in the *bounded-ratios form* if there exist $0 < c_1 < c_2 < \infty$ such that, uniformly for $p \geq p_c$,

$$c_1(p - p_c)^\beta \leq \theta(p) \leq c_2(p - p_c)^\beta, \tag{1.18}$$

Finally, we say that β exists in the *asymptotic form* if, as $p \downarrow p_c$, there exists a $c > 0$ such that

$$\theta(p) = c(p - p_c)^\beta (1 + o(1)). \tag{1.19}$$

The existence of a critical exponent is *a priori* unclear, and needs a mathematical proof. Unfortunately, in general such a proof is missing, and we can only give proofs of the existence in special cases, which we shall discuss below in quite some detail. Indeed, the existence of the critical exponent $\beta > 0$ is stronger than continuity of $p \mapsto \theta(p)$, which is unknown in general, and is arguably the holy grail of percolation theory. Indeed, $p \mapsto \theta(p)$ is clearly continuous on $[0, p_c)$, and it is also continuous (and even infinitely differentiable) on $(p_c, 1]$ by the results of (van den Berg and Keane, 1984) (for infinite differentiability of $p \mapsto \theta(p)$ for $p \in (p_c, 1]$, see (Russo, 1978)). Thus, continuity of $p \mapsto \theta(p)$ is equivalent to the statement that $\theta(p_c(\mathbb{G})) = 0$.

We now introduce several more critical exponents. The critical exponent γ for the expected cluster size is given by

$$\chi^{\mathrm{f}}(p) \sim |p - p_c|^{-\gamma}, \qquad p \to p_c. \tag{1.20}$$

More precisely, we can think of (1.20) as defining the critical exponents $\gamma, \gamma' > 0$ defined by

$$\chi(p) \sim (p - p_c)^{-\gamma}, \quad p \uparrow p_c, \qquad \chi^{\mathrm{f}}(p) \sim (p - p_c)^{-\gamma'}, \quad p \downarrow p_c, \tag{1.21}$$

with the predicted equality $\gamma = \gamma'$. For (1.20) and (1.21), we are implicitly assuming that $p_T(\mathbb{G}) = p_c(\mathbb{G})$, this equality shall be discussed in more detail below.

Further, ν, ν' are defined by

$$\xi(p) \sim (p - p_c)^{-\nu}, \quad p \uparrow p_c, \qquad \xi(p) \sim (p - p_c)^{-\nu'}, \quad p \downarrow p_c, \tag{1.22}$$

again with the prediction that $\nu = \nu'$. The exponent $-1 \leq \alpha < 0$ is defined by the blow up of the third derivative of $p \mapsto \kappa(p)$ at $p = p_c$, i.e.,

$$\kappa'''(p) \sim |p - p_c|^{-1-\alpha}, \tag{1.23}$$

while the *gap exponent* $\Delta > 0$ is defined by, for $k \geq 1$,

$$\frac{\mathbb{E}_p\big[|C|^{k+1}\mathbb{1}_{\{|C|<\infty\}}\big]}{\mathbb{E}_p\big[|C|^{k+1}\mathbb{1}_{\{|C|<\infty\}}\big]} \sim |p - p_c|^{-\Delta}. \tag{1.24}$$

Also α and Δ can be defined, similarly to (1.21), as an exponent for $p \uparrow p_c$ and one for $p \downarrow p_c$, the values being equal, and we shall always use the convention that α and Δ denote the $p \uparrow p_c$ versions, while α' and Δ denote the $p \downarrow p_c$ versions.

As mentioned before, it is highly unclear that these critical exponents are well-defined, and that the value of $\Delta > 0$ does not depend on k. However, there are good physical reasons why these exponents are defined as they are. The exponents $\beta, \gamma, \nu, \alpha, \Delta$ can be thought of as *approach exponents* which measure the blow-up of various aspects of the cluster size as p approaches the critical value $p = p_c(\mathbb{G})$.

We finally define two critical exponents at criticality. The exponent $\delta \geq 1$ measures the power-law exponent of the critical cluster tail, i.e.,

$$\mathbb{P}_{p_c}(|C| \geq n) \sim n^{-1/\delta}, \qquad n \to \infty, \tag{1.25}$$

the assumption that $\delta \geq 1$ following from the prediction that $\chi(p_c) = \infty$. Further, we define the exponent $\rho > 0$ by

$$\mathbb{P}_{p_c}(o \longleftrightarrow \partial B(n)) \sim n^{-1/\rho}, \qquad n \to \infty. \tag{1.26}$$

Finally, η is defined by

$$\mathbb{E}_{p_c}\big[|C \cap B(n)|\big] = \sum_{x:|x|\leq n} \tau_{p_c}(x) \sim n^{2-\eta}, \qquad n \to \infty, \tag{1.27}$$

where we recall that $|x| = d(o, x)$ is the distance of x to o in $\mathbb{G}$.

The above definitions give rise to eight critical values that each describe a certain aspect of the (near-)critical behavior of the percolation system. Several relations between these critical exponents are predicted by (non-rigorous) physics argument, and go under the name of *scaling relations*. These scaling relations assert that

$$\gamma + 2\beta = \beta(\delta + 1) = 2 - \alpha = \Delta + \beta, \qquad \gamma = \nu(2 - \eta). \tag{1.28}$$

The validity of the scaling relations in (1.28) is widely accepted, but few proofs exist, in particular since the existence of the critical exponents is in general unknown. We can intuitively understand the scaling relations in (1.28) by assuming the existence of certain*scaling functions*, which describe certain percolation

quantities close to criticality. An example is to assume that there exist scaling functions f_+ and f_-, as well as some exponents $\sigma, \tau > 0$, such that

$$\mathbb{P}_p(|C| = n) \sim \begin{cases} n^{-\sigma} f_-(n/\xi(p)^\tau) & \text{for } p \leq p_c \\ n^{-\sigma} f_+(n/\xi(p)^\tau) & \text{for } p \geq p_c, \end{cases} \tag{1.29}$$

for some sufficiently smooth and integrable functions f_+ and f_-. When working out the consequences of (1.29), we can see that it must imply the first three scaling relations in (1.28), and when also assuming that a scaling function exists for $x \mapsto \tau_p^{\mathrm{f}}(x)$, the relation $\gamma = \nu(2-\eta)$ follows. The existence of scaling functions is, as far as we know, unknown except for percolation on a tree. Nice discussions of scaling functions can be found in (Grimmett, 1999, Sections 9.1-9.2) or (Hughes, 1996, Section 4.2). The most complete reference to results on critical exponents until 1996 is (Hughes, 1996, Chapter 4).

Percolation is a paradigm model in statistical physics. A central notion in this field is *universality*, a notion strongly rooted in theoretical physics which states that the near-critical behavior of a system does is rather insensitive to the microscopic details of the system. An example in the setting of percolation is that any finite-range percolation model on a vertex space $\mathbb{V}$ has the *same* critical behavior. In particular, it has the same critical exponents. While universality is quite plausible when describing real physical systems from the viewpoint of statistical physics, and while universality is a very useful notion since it allows us to study only the simplest finite-range model available, there are very few examples where universality can be rigorously proved. We shall discuss a few universality results below.

So far, we have discussed percolation in full generality. We shall now treat examples of percolation models. In Section 1.2.2, we shall describe percolation on regular trees, in Section 1.2.3 we discuss percolation on $\mathbb{Z}^d$ for general d, in Section 1.2.4 we specialize to the two-dimensional setting, in Section 1.2.5, we study the high-dimensional case for which $d > 6$ and in Section 1.2.6, we study oriented or directed percolation. Finally, in Section 1.2.7, we study the case of percolation on non-amenable graphs and we close this section in Section 1.2.8 by discussing continuum percolation and its applications.

1.2.2   *Percolation on the regular tree.*

In this section, we study percolation on the regular tree. Let $\mathbb{T}_r$ denote the r-regular tree of degree r. The advantage of trees is that they do not contain cycles, which make explicit computations possible. In order to compute the critical exponents ν and η, we first identify

$$|x| = \sqrt{h(x)}, \tag{1.30}$$

where $h(x)$ is the *height* of the vertex x in the tree, i.e., the length of the shortest path linking o and x, so that (1.30) is the *Euclidean distance* in the tree. We shall first prove that the critical exponents for percolation on a regular tree exist and identify their values in the following theorem:

Theorem 1.1. (Critical behavior on the r-regular tree) *On the r-regular tree $\mathbb{T}_r$, $p_c = p_T = 1/(r-1)$, and $\beta = \gamma = \gamma' = 1$, $\delta = \Delta = \Delta' = \rho = 2$, $\nu = \nu' = 1/2$ and $\alpha = \alpha' = -1$ in the asymptotic sense.*

Proof   We shall make substantial use of the fact that percolation on a tree can be described in terms of *branching processes*. Recall that o is the root of the tree. For $x \neq o$, we write $C_{\mathrm{BP}}(x)$ for the forward cluster of x, i.e., the vertices $y \in \mathbb{T}_r$ which are connected to x and for which the unique path from x to y only moves away from the root o. Then, clearly,

$$|C(o)| = 1 + \sum_{e \sim o} I_{(o,e)} |C_{\mathrm{BP}}(e)|, \tag{1.31}$$

where the sum is over all neighbors e of o, $I_{(o,e)}$ is the indicator that the edge (o, e) is occupied, and $C_{\mathrm{BP}}(e)$ is the forward cluster of e. The random vector $\{I_{(o,e)}\}_{e \sim o}$ forms a collection of r independent Bernoulli random variables with success probability p, and $\{|C_{\mathrm{BP}}(e)|\}_{e \sim o}$ is an i.i.d. sequence independent of $\{I_{(o,e)}\}_{e \sim o}$. Equation (1.31) allows us to deduce all information concerning $|C(o)|$ from the information of $|C_{\mathrm{BP}}(e)|$. Also, for each e, $|C_{\mathrm{BP}}(x)|$ satisfies the formula

$$|C_{\mathrm{BP}}(x)| = 1 + \sum_{e \sim x : h(e) > h(x)} I_{(x,e)} |C_{\mathrm{BP}}(e)|, \tag{1.32}$$

where $h(x)$ is the height of x in $\mathbb{T}_r$, and $\{|C_{\mathrm{BP}}(e)|\}_{e \sim x : h(e) > h(x)}$ is a set of $r-1$ independent copies of $|C_{\mathrm{BP}}(x)|$. Thus, $|C_{\mathrm{BP}}(x)|$ is the *total progeny* of a branching process. As a result,

$$\chi_{\mathrm{BP}}(p) = \mathbb{E}[|C_{\mathrm{BP}}(x)|] = \frac{1}{1 - (r-1)p}. \tag{1.33}$$

From (1.31), we then obtain that, for $p < 1/(r-1)$,

$$\chi(p) = 1 + rp\chi_{\mathrm{BP}}(p) = 1 + \frac{rp}{1 - (r-1)p} = \frac{1+p}{1 - (r-1)p}, \tag{1.34}$$

while, for $p > 1/(r-1)$, $\chi(p) = \infty$. In particular, $p_T = 1/(r-1)$, and $\gamma = 1$ in the asymptotic sense. The computation of $\chi(p)$ can also be performed without the use of (1.32), by noting that, for $p \in [0,1]$,

$$\tau_p(x) = p^{h(x)}, \tag{1.35}$$

and the fact that, for $n \geq 1$, there are $r(r-1)^{n-1}$ vertices in $\mathbb{T}_r$ at height n, so that, for $p < 1/(r-1)$,

$$\chi(p) = 1 + \sum_{n=1}^{\infty} r(r-1)^{n-1} p^n = \frac{r+1 - (r-1)p}{1 - (r-1)p}. \tag{1.36}$$

However, for related results for percolation on a tree, the connection to branching processes in (1.32) is vital. We defer the proof that $\gamma' = 1$ to later.

Let $\theta_{\mathrm{BP}}(p) = \mathbb{P}_p(|C_{\mathrm{BP}}(x)| = \infty)$. Then $\theta_{\mathrm{BP}}(p)$ is the survival probability of a branching process with a binomial offspring distribution with parameters $r - 1$ and p. Thus, $\theta_{\mathrm{BP}}(p)$ satisfies the equation

$$\theta_{\mathrm{BP}}(p) = 1 - \big(1 - p + p(1 - \theta_{\mathrm{BP}}(p))\big)^{r-1} = 1 - \big(1 - p\theta_{\mathrm{BP}}(p)\big)^{r-1}. \qquad (1.37)$$

To compute $\theta_{\mathrm{BP}}(p)$, it is more convenient to work with the extinction probability $\zeta_{\mathrm{BP}}(p) = 1 - \theta_{\mathrm{BP}}(p)$, which is the probability that the branching process dies out. The extinction probability $\zeta_{\mathrm{BP}}(p)$ satisfies

$$\zeta_{\mathrm{BP}}(p) = \big(1 - p + p\zeta_{\mathrm{BP}}(p)\big)^{r-1}. \qquad (1.38)$$

This equation can be solved explicitly when $r = 2$, when the unique solution is $\theta_{\mathrm{BP}}(p) = 0$ for $p \in [0, 1)$ and $\theta_{\mathrm{BP}}(1) = 1$, so that $p_c = 1$. When $r = 3$, we obtain that

$$p^2 \zeta_{\mathrm{BP}}(p)^2 + (2p(1-p) - 1)\zeta_{\mathrm{BP}}(p) + (1-p)^2 = 0, \qquad (1.39)$$

so that

$$\zeta_{\mathrm{BP}}(p) = \frac{1 - 2p(1-p) \pm |2p - 1|}{2p^2}. \qquad (1.40)$$

Since $\zeta_{\mathrm{BP}}(0) = 1, \zeta_{\mathrm{BP}}(1) = 0$, we must have that

$$\zeta_{\mathrm{BP}}(p) = \frac{1 - 2p(1-p) - |2p - 1|}{2p^2}, \qquad (1.41)$$

so that $\zeta_{\mathrm{BP}}(p) = 1$ for $p \in [0, 1/2]$, while, for $p \in [1/2, 1]$,

$$\zeta_{\mathrm{BP}}(p) = \frac{1 - 2p(1-p) + (1 - 2p)}{2p^2} = \frac{2 - 4p + 2p^2}{2p^2} = \Big(\frac{1-p}{p}\Big)^2. \qquad (1.42)$$

As a result, we have the explicit form $\theta_{\mathrm{BP}}(p) = 0$ for $p \in [0, 1/2]$ and

$$\theta_{\mathrm{BP}}(p) = 1 - \Big(\frac{1-p}{p}\Big)^2 = \frac{2p - 1}{p^2}, \qquad (1.43)$$

for $p \in [1/2, 1]$, so that $p_c = 1/2$. In particular, $p \mapsto \theta_{\mathrm{BP}}(p)$ is continuous, and, for $p \downarrow p_c$,

$$\theta_{\mathrm{BP}}(p) = 8(p - p_c)(1 + o(1)). \qquad (1.44)$$

It is not hard to see that (1.44) together with (1.31) implies that

$$\theta(p) = 12(p - p_c)(1 + o(1)). \qquad (1.45)$$

Thus, for $r = 3$, the percolation function is continuous, and $\beta = 1$ in the asymptotic sense. It is not hard to extend the asymptotic analysis in (1.44)–(1.45) to $r \geq 4$, for which $p_c(\mathbb{T}_r) = p_{|sssT}(\mathbb{T}_r) = 1/(r - 1)$, but we omit the details.

In order to study $\chi^{\mathrm{f}}_{\mathrm{BP}}(p) = \mathbb{E}_p[|C_{\mathrm{BP}}(x)|\mathbb{1}_{\{|C_{\mathrm{BP}}(x)|<\infty\}}]$ for $p > p_c = 1/(r-1)$, we make use of the fact that

$$\chi^{\mathrm{f}}_{\mathrm{BP}}(p) = (1 - \theta_{\mathrm{BP}}(p))\mathbb{E}_p[|C_{\mathrm{BP}}(x)|\,\big|\,|C_{\mathrm{BP}}(x)| < \infty], \qquad (1.46)$$

and the conditional law of percolation on the tree given that $|C_{\mathrm{BP}}(x)| < \infty$ is percolation on a tree with p replaced by the *dual* percolation probability p_d given by

$$p_d = p(1 - \theta_{\mathrm{BP}}(p)). \qquad (1.47)$$

The crucial fact is that $p_d < p_c(\mathbb{T}_r)$, which follows from the equality $1 - \theta_{\mathrm{BP}}(p) = \zeta_{\mathrm{BP}}(p)$, (1.38) and the fact that

$$
\begin{aligned}
(r-1)p\zeta_{\mathrm{BP}}(p) &= (r-1)p\big(1 - p + p\zeta_{\mathrm{BP}}(p)\big)^{r-1} < (r-1)p\big(1 - p + p\zeta_{\mathrm{BP}}(p)\big)^{r-2} \\
&= \frac{\mathrm{d}}{\mathrm{d}s}(1 - p + ps)^{r-1}\Big|_{s=\zeta_{\mathrm{BP}}(p)},
\end{aligned} \qquad (1.48)
$$

which, since $\zeta_{\mathrm{BP}}(p)$ is the smallest solution of $(1 - p + ps)^{r-1} = s$, implies that the derivative of $(1 - p + ps)^{r-1}$ at $s = \zeta_{\mathrm{BP}}(p)$ is strictly bounded above by 1 for $p > p_c(\mathbb{T}_r)$. Thus, by conditioning a supercritical cluster in percolation on a tree to die out, we obtain a subcritical cluster at an appropriate subcritical p_d which is related to the original percolation parameter. This fact is sometimes called the *discrete duality principle*.

We conclude that

$$\chi^{\mathrm{f}}_{\mathrm{BP}}(p) = (1 - \theta_{\mathrm{BP}}(p))\mathbb{E}_p[|C|\,\big|\,|C| < \infty] = (1 - \theta_{\mathrm{BP}}(p))\frac{1}{1 - (r-1)p(1 - \theta_{\mathrm{BP}}(p))}. \qquad (1.49)$$

Using that $\beta = 1$ in the asymptotic sense then gives that

$$\chi^{\mathrm{f}}_{\mathrm{BP}}(p) = \frac{C_{\gamma'} + o(1)}{p - p_c}. \qquad (1.50)$$

By (1.31), this can easily be transferred to $\chi^{\mathrm{f}}(p)$, so that $\gamma' = 1$ in the asymptotic sense.

To compute ν, we note that, by (1.30) and (1.35), we have that, for $p < 1/(r-1)$,

$$\xi(p)^2 = \frac{1}{\chi^{\mathrm{f}}_{\mathrm{BP}}(p)}\sum_{n=1}^{\infty} nr(r-1)^{n-1}p^n = \frac{1}{\chi^{\mathrm{f}}_{\mathrm{BP}}(p)}\frac{r}{[1 - (r-1)p]^2}, \qquad (1.51)$$

so that $\nu = 1/2$ in the asymptotic sense by (1.50). We refrain from proving that $\nu' = \nu$, and only remark that this follows again by using the duality used in (1.47).

To compute η, we note that

$$\mathbb{E}_{p_c}\big[|C_{\mathrm{BP}}(x) \cap B(n)|\big] = \sum_{m=1}^{n^2} (r-1)^m p_c^m = n^2, \qquad (1.52)$$

so that $\eta = 0$. We can compute δ by using the *random walk hitting time theorem*, see (Grimmett, 1999, Prop. 10.22) and (van der Hofstad and Keane, 2007), where a very simple proof is given applying to general branching processes. This result yields that

$$\mathbb{P}_p(|C_{\mathrm{BP}}(x)| = k) = \mathbb{P}(X_1 + \ldots + X_k = k - 1), \qquad (1.53)$$

where $\{X_i\}_{i=1}^{\infty}$ is an i.i.d. sequence of binomial random variables with parameter $r-1$ and success probability p. Thus,

$$\mathbb{P}_p(|C_{\mathrm{BP}}(x)| = k) = \binom{k(r-1)}{k-1} p^{k-1}(1-p)^{k(r-1)-(k-1)}. \qquad (1.54)$$

To prove that $\delta = 2$, we note that for $p = p_c = 1/(r-1)$, by a local limit theorem, we obtain

$$\mathbb{P}_{p_c}(|C_{\mathrm{BP}}(x)| = k) = (C_\delta + o(1))\frac{1}{\sqrt{k^3}}, \qquad (1.55)$$

so that, by summing over $k \geq n$, we obtain $\delta = 2$ in an asymptotic sense. We can compute ρ by noting that

$$\theta_n = \mathbb{P}_{p_c}(\exists y \in C_{\mathrm{BP}}(x) \text{ such that } h(y) = n + h(x)) \qquad (1.56)$$

satisfies the recursion relation

$$1 - \theta_n = (1 - p_c\theta_{n-1})^{r-1}. \qquad (1.57)$$

It is not hard to see that (1.57) together with $p_c(\mathbb{T}_r) = 1/(r-1)$ implies that $\theta_n = (C_\rho + o(1))/n$. Thus $\rho = 1/2$, since $|x| = \sqrt{h(x)}$ and $\mathbb{P}_{p_c}(o \longleftrightarrow \partial B(n)) = \theta_{\sqrt{n}} \sim (C_\rho + o(1))/\sqrt{n}$.

$$\square$$

The computation of the key objects for percolation on a tree is feasible due to the close relationship to branching processes, a topic which has attracted substantial interest in the probability community. See (Athreya and Ney, 1972), (Harris, 1963) and (Jagers, 1975) for detailed discussions about branching processes. As it turns out, the computations on a tree also have direct consequences for percolation on general graphs, with and without loops. We shall now discuss some of these consequences, the first being that $p_c > 0$ on any graph with bounded degree:

Theorem 1.2. (Percolation threshold is strictly positive) *Let $\mathbb{G} = (\mathbb{V}, \mathbb{E})$ be a graph for which the degree of every vertex is bounded by r. Then, for every $x \in \mathbb{V}$, and $p < 1/(r-1)$,*

$$\mathbb{E}_p[|C(x)|] < \infty. \qquad (1.58)$$

In particular, for transitive graphs $\mathbb{G}$ with degree equal to r, $p_c(\mathbb{G}) \geq p_T(\mathbb{G}) \geq 1/(r-1)$.

Proof Let $\omega = (\omega_1, \ldots, \omega_n)$ be a nearest-neighbor path in $\mathbb{G}$. We call ω *self-avoiding* when $\omega_i \neq \omega_j$ for all $0 \leq i < j \leq n$. We let $c_n(x, y)$ denote the number of n-step nearest-neighbor self-avoiding walks starting at x with endpoint y. Then, we note that if $x \longleftrightarrow y$, then there must be a self-avoiding walk path consisting of occupied bonds. As a result,

$$\tau_p(x, y) \leq \sum_{n=0}^{\infty} p^n c_n(x, y). \tag{1.59}$$

Therefore,

$$\mathbb{E}_p[|C(x)|] = \sum_{y \in \mathbb{V}} \tau_p(x, y) \leq \sum_{y \in \mathbb{V}} \sum_{n=0}^{\infty} p^n c_n(x, y) = \sum_{n=0}^{\infty} p^n c_n(x), \tag{1.60}$$

where $c_n(x)$ denotes the number of n-step self-avoiding walk paths starting at x. Since the degree of $\mathbb{G}$ is bounded by r, we have that, uniformly in $x \in \mathbb{V}$,

$$c_n(x) \leq r(r-1)^{n-1}. \tag{1.61}$$

Thus, we arrive at

$$\mathbb{E}_p[|C(x)|] \leq \sum_{n=0}^{\infty} p^n r(r-1)^{n-1} = \frac{r}{r-1} \frac{1}{1 - p(r-1)}, \tag{1.62}$$

so that $\mathbb{E}_p[|C(x)|] < \infty$ for $p < 1/(r-1)$. This completes the proof. $\square$

For percolation on a tree, it is not hard to see that the number of infinite clusters N equals infinity a.s. To see this, fix a root o of the tree $\mathbb{T}_r$. For any vertex v unequal to o, let u be the unique vertex in $\mathbb{T}_r$ that is closer to the root. Then, the probability that the bond (u, v) is vacant is strictly positive. If $p > p_c = 1/(r-1)$, then with probability $\theta_{\mathrm{BP}}(p) > 0$, the vertex v will lie in an infinite component. Thus, with strictly positive probability, there will be at least two bonds (o, e_1) and (o, e_2) that are vacant, and of which e_1 and e_2 lie in an infinite cluster. Thus, $\mathbb{P}_p(N \geq 2) > 0$, so that, since $N \in \{0, 1, \infty\}$ a.s., we must have that $N = \infty$ a.s.

We conclude that percolation on a tree is very well understood, that all of its critical exponents exist in an asymptotic sense and can be explicitly identified. Moreover, for $p \leq p_c(\mathbb{T}_r) = 1/(r-1)$, there is no infinite connected component, while, for $p \in (p_c, 1)$, there are infinitely many of them. For $p = 1$, there is a unique infinite cluster. We now proceed to study percolation on graphs with loops, starting with the paradigm example of $\mathbb{G} = \mathbb{Z}^d$.

1.2.3 *Percolation on $\mathbb{Z}^d$.*

In this section, we study percolation on $\mathbb{Z}^d$. We start by proving that the phase transition is non-trivial:

Theorem 1.3. (Phase transition on $\mathbb{Z}^d$ is non-trivial) *For nearest-neighbor percolation on $\mathbb{Z}^d$ with $d \geq 2$, $p_c(\mathbb{Z}^d) \in [\frac{1}{2d-1}, 1)$.*

Proof The lower bound on $p_c(\mathbb{Z}^d)$ is immediate from Theorem 1.2, so that we are left to prove that $p_c(\mathbb{Z}^d) < 1$. For this, we must prove that, on $\mathbb{Z}^d$, $\theta(p) > 0$ for p sufficiently close to 1. We first show that it suffices to prove this for $d = 2$. For this, we fix $p \in (0, 1)$ and note that $|C| = |C(o)|$ for percolation on $\mathbb{Z}^d$ with parameter p is stochastically larger than $|C| = |C(o)|$ for percolation on $\mathbb{Z}^2$ with parameter p, since by only using the nearest-neighbor edges in $\mathbb{Z}^2 \times \{\vec{0}_{d-2}\}$, where $\vec{0}_{d-2} \in \mathbb{Z}^{d-2}$ denotes the origin in $\mathbb{Z}^{d-2}$, we obtain a cluster which is not larger than the one using all nearest-neighbor edges in $\mathbb{Z}^d$. Thus, it suffices to prove that $\theta(p) > 0$ for p sufficiently close to 1 for $d = 2$.

We shall make use of *duality*, a notion which is crucial for $d = 2$ (see also Section 1.2.4 below). For a set of vertices A, let the *boundary edges* of A be $\partial_e A = \{\{x, y\} : x \in A, y \notin A\}$. Clearly, for $C(o)$ the cluster of the origin, we have that all edges in $\partial_e C(o)$ are vacant. Define the *dual lattice* $\mathbb{L}^*$ consist of the vertices $\mathbb{V}_d = \{x + (1/2, 1/2) : x \in \mathbb{Z}^2\}$ and an edge exists between $x + (1/2, 1/2)$ and $y + (1/2, 1/2)$ if and only if x and y are nearest-neighbors in $\mathbb{Z}^2$. We note that each edge in $\mathbb{L}^*$ intersects *precisely* one nearest-neighbor edge in $\mathbb{Z}^d$. We perform percolation on $\mathbb{L}^*$ by identifying the status of the edge $e \in \mathbb{L}^*$ by the status of the unique edge it intersects in $\mathbb{Z}^d$. Then, the set of vacant edges $\partial_e C$ is identified with a set of vacant edges in $\mathbb{L}$. We next study the structure of the set of edges.

We call a path $\omega = (\omega_0, \ldots, \omega_n)$ a *self-avoiding polygon of length n* on $\mathbb{L}$ when $\omega_i \in \mathbb{V}_d$, $\omega_0 = \omega_n$ and when $\omega_i \neq \omega_j$ for every $0 \leq i < j \leq n$ for which $i, j \neq 0, n$. A self-avoiding polygon separates $\mathbb{L}^*$ into the outside and the inside of the polygon, and we say that a self-avoiding polygon surrounds a point $x \in \mathbb{Z}^2$ when the point lies on the inside of the polygon. For $\mathbb{Z}^2$, we observe that if $|C(o)| < \infty$, then there must be a self-avoiding polygon $\omega = (\omega_1, \ldots, \omega_n)$ of length $n \geq 4$ on $\mathbb{L}^*$ which surrounds the origin $o = (0, 0)$ of which each edge is vacant in $\mathbb{L}^*$. This is worked out in detail in (Kesten, 1982, Page 386). Thus,

$$1 - \theta(p) \leq \sum_{n=4}^{\infty} m_n (1 - p)^n, \tag{1.63}$$

where m_n is the number of self-avoiding loops of length n on $\mathbb{L}$ surrounding the origin $o = (0, 0)$. Clearly, we have that $m_n \leq n3^{n-1}$, since there are at most n possible positions where the self-avoiding loop crosses the half-line $\{(x, 0) : x \geq 0\}$, and the number of self-avoiding loops starting from any vertex is bounded by 3^{n-1}, so that

$$1 - \theta(p) \leq \sum_{n=4}^{\infty} n3^{n-1}(1 - p)^n < 1, \tag{1.64}$$

when $p < 1$ is sufficiently close to 1. Thus, for such p, we have that $\theta(p) > 0$, proving that $p > p_c(\mathbb{Z}^2)$. □

Having established that the phase transition is non-trivial, the natural question is what the critical value is. Below, we shall give some results on critical values,

particularly in 2 dimensions (see Section 1.2.4). An excellent reference to both numerical values as well as rigorous bounds is (Hughes, 1996, Chapter 3). For example, see (Hughes, 1996, Table 3.3) for some exact values of critical values, and (Hughes, 1996, Table 3.6) for numerical values of $p_c(\mathbb{Z}^d)$ on the nearest-neighbor lattice, showing that the inequality $p_c(\mathbb{Z}^d) \geq 1/(2d-1)$ is only a few percent off in dimensions $d \geq 5$. We now move on to the problem of the *uniqueness of the phase transition*, i.e., whether $p_c(\mathbb{Z}^d) = p_T(\mathbb{Z}^d)$:

Theorem 1.4. (Phase transition on $\mathbb{Z}^d$ is unique) *For nearest-neighbor percolation on $\mathbb{Z}^d$ with $d \geq 2$, $p_c(\mathbb{Z}^d) = p_T(\mathbb{Z}^d)$. In particular, $\chi(p) < \infty$ for all $p < p_c(\mathbb{Z}^d)$.*

The proof of Theorem 1.4 was given independently in (Menshikov, 1986) and (Aizenman and Barsky, 1987). Menshikov investigates $\mathbb{P}_p(A_n)$, where A_n is the event that there exists an open path from o to $\partial S(n)$ and $S(n) = \{x \in \mathbb{Z}^d : |x_1| + \ldots + |x_d| \leq n\}$. Since $p < p_c$, we have that $\mathbb{P}_p(A_n) \downarrow 0$ as $n \to \infty$. In order to prove that $\chi(p) < \infty$, it suffices to prove that $\mathbb{P}_p(A_n) \downarrow 0$ converges to zero sufficiently fast, since

$$\chi(p) = \sum_{n=0}^{\infty} \mathbb{P}_p(|C(o)| > n) \leq \sum_{n=0}^{\infty} \mathbb{P}_p(A_{\lceil n^{1/d} \rceil}). \qquad (1.65)$$

Menshikov shows that $\mathbb{P}_p(A_n)$ is close to exponentially small in n for rather general graphs, (Grimmett, 1999, Theorem 5.4) improves this to

$$\mathbb{P}_p(A_n) \leq e^{-n\psi(p)} \qquad (1.66)$$

for some $\psi(p)$ which is strictly positive for $p < p_c$. The proof of (1.66) makes clever use of Russo's formula, together with an appropriate renewal type argument to bound the expected number of pivotal bonds for A_n.

The proof in (Aizenman and Barsky, 1987) makes use of *differential inequalities*, a powerful technique which can also be used to prove bounds on critical exponents. In more detail, (Aizenman and Barsky, 1987) prove that for any p for which $\chi^f(p) = \infty$, we have that either $\theta(p) > 0$ or $\theta(p) = 0$ and $\theta(p') \geq \frac{1}{p'}(p'-p)$ for all $p' > p$. This shows that $\chi(p) < \infty$ for $p < p_c$. Indeed, suppose that $\chi(p) = \infty$. Then, since $p < p_c$, we have that

$$\chi^f(p) = \mathbb{E}_p[|C(o)|\mathbb{1}_{\{|C(o)|<\infty\}}] = \mathbb{E}_p[|C(o)|] = \infty, \qquad (1.67)$$

where we made use of the fact that since $p < p_c$, we have $\theta(p) = 0$. Then, by the above result and again since $\theta(p) = 0$, $\theta(p') \geq \frac{1}{p'}(p'-p) > 0$ for all $p' > p$, so that $p \geq p_c = p_c(\mathbb{Z}^d)$. This is in contradiction to $p < p_c$. We shall not present further details of the proof of Theorem 1.4, as this is beyond the scope of this paper. We do mention that both the proofs in (Aizenman and Barsky, 1987) and in (Menshikov, 1986) are valid more generally. The proof in (Aizenman and Barsky, 1987) holds for general partially-oriented models, including also finite-range models as well as oriented percolation models. The proof in (Menshikov,

1986) does not rely on a lattice structure, but rather on the fact that the boundary of a ball of size n grows not faster than e^{cn^a} for some $a < 1$. These extensions will be relevant below.

We have seen that, in general, the number of infinite clusters N is a.s. constant, and that $N \in \{0, 1, \infty\}$ a.s. When $p < p_c$, then clearly $N = 0$ a.s. This leaves the question what the value of N is in the supercritical regime. On the tree, we have seen that $N = \infty$ a.s., and the main question is whether this can also occur on finite-dimensional lattices. This question was beautifully answered by (Burton and Keane, 1989) using a geometric argument. A nice survey of the implications of the Burton-Keane proof can be found in (Meester, 1994).

Theorem 1.5. (Uniqueness of infinite cluster on $\mathbb{Z}^d$) *For nearest-neighbor percolation on $\mathbb{Z}^d$ with $d \geq 2$, there is at most one infinite cluster. More precisely, for p such that $\theta(p) > 0$,*

$$\mathbb{P}_p(\text{there is exactly one infinite cluster}) = 1. \tag{1.68}$$

A first proof of this result appeared in (Aizenman, Kesten and Newman, 1987).

Proof As discussed on page 3, we only need to rule out the possibility that N, the number of infinite clusters, is equal to $+\infty$ a.s. Indeed, when $\theta(p) = 0$, then $N = 0$ a.s., while if $\theta(p) > 0$, then the proportion of vertices in infinite clusters will be strictly positive, so that $N \geq 1$ a.s. Since N is a.s. constant and $N \in \{0, 1, \infty\}$ a.s., we conclude that $N = 1$ a.s. when $N < \infty$.

We will argue by contradiction. We call $x \in \mathbb{V}$ a *trifurcation* when it satisfies

(1) $|C(x)| = \infty$;
(2) there are *precisely* three occupied edges incident to x;
(3) the removal of the three edges incident to x splits the infinite cluster $C(x)$ into precisely three disjoint infinite clusters, and no finite clusters.

We write I_x for the indicator that x is a trifurcation, and $\mathbb{P}_p(I_x) = \mathbb{P}_p(I_0)$ for every $x \in \mathbb{V}$, and any reference point 0. We shall first argue that when $N = \infty$ a.s., then $\mathbb{P}_p(I_0) > 0$. Let M_B denote the number of infinite clusters touching B, and $M_B(0)$ the number of infinite clusters touching B when all edges with both vertices in B are declared vacant. Since $M_B(0) \geq M_B$, and $\mathbb{P}_p(M_{B(n)} \geq 3) \to \mathbb{P}_p(N \geq 3) = 1$, by our assumption, we can take n so large that $\mathbb{P}_p(M_{B(n)}(0) \geq 3) \geq 1/2$. Let ω be such that $\{M_{B(n)}(0) \geq 3\}$ occurs, and fix $x, y, z \in \partial B(n)$ such that x, y, z lie in three *distinct* infinite clusters in the configuration $\tilde{\omega}$ for which all edges with both vertices in B are declared vacant. We make the choice of x, y, z unique in an arbitrary way. Denote by $K_{x,y,z}$ the event that $M_{B(n)}(0) \geq 3$ *and* the three points are x, y, z.

Then, there exist 3 self-avoiding paths that only cross in the origin, and have only x, y, z, respectively, as elements in $\partial B(n)$. Let $J_{x,y,z}$ be the event that all the edges along the 3 paths are occupied, while all other edges are vacant. Since

$B(n)$ is finite, and letting $R(n)$ denote the number of edges in $B(n)$, we obtain that

$$\mathbb{P}_p(J_{x,y,z}) \geq (\min\{p, 1-p\})^{R(n)} = \varepsilon > 0. \qquad (1.69)$$

Finally, since, for all fixed x, y, z, the events $J_{x,y,z}$ and $K_{x,y,z}$ are independent, we obtain that

$$\mathbb{P}_p(I_0) \geq \sum_{x,y,z \in \partial B(n)} \mathbb{P}_p(J_{x,y,z} \cap K_{x,y,z}) = \sum_{x,y,z \in \partial B(n)} \mathbb{P}_p(J_{x,y,z})\mathbb{P}_p(K_{x,y,z})$$

$$\geq \varepsilon \sum_{x,y,z \in \partial B(n)} \mathbb{P}_p(K_{x,y,z}) = \varepsilon \mathbb{P}_p(M_{B(n)}(0) \geq 3) \geq \varepsilon/2 > 0, \qquad (1.70)$$

i.e., the probability that x is a trifurcation is strictly positive. By translation invariance, we therefore obtain that

$$\mathbb{E}_p\Big[\sum_{x \in B(n)} I_x \Big] \geq |B(n)|\varepsilon/2. \qquad (1.71)$$

We now derive a contradiction by noting that the number of trifurcations in $B(n)$ can be no more than $\partial B(n+1)$. Let τ denote the number of trifurcations and $u_1, u_2, \ldots, u_\tau$ denote the trifurcations in $B(n)$. Next, we informally argue that we can identify to each trifurcation u_k a point $y_k \in \partial B(n+1)$ such that $y_i \neq y_j$ for each $1 \leq i < j \leq \tau$. This can be argued by noting that each trifurcation $u_i \in B(n)$ in a cluster K splits the intersection $K \cap \partial B(n+1)$ into three disjoint parts. This can be formalized into a family of so-called *compatible families of distinct 3-partitions* of $K \cap \partial B(n+1)$, and their number is restricted by $|K \cap \partial B(n+1)| - 2$, see (Grimmett, 1999, Lemma 8.5). As a result, the number of trifurcations in $B(n)$ of a cluster K is bounded by $|K \cap \partial B(n+1)| - 2 \leq |K \cap \partial B(n+1)|$. Summing out over the distinct clusters yields

$$\sum_{x \in B(n)} I_x \leq |\partial B(n+1)|. \qquad (1.72)$$

By (1.71), we obtain

$$|B(n)|\varepsilon/2 \leq |\partial B(n+1)|, \qquad (1.73)$$

for some $\varepsilon > 0$, which causes a contradiction when $n \to \infty$ whenever $|\partial B(n+1)|/|B(n)| \to 0$. As we see, this argument applies more generally than merely on $\mathbb{Z}^d$. $\qquad \square$

Theorems 1.3, 1.4 and 1.5 answer many of the important questions for percolation, namely, the phase transition is non-trivial, it is unique in the sense that $p_c(\mathbb{Z}^d) = p_T(\mathbb{Z}^d)$ and the infinite cluster, when it exists, is unique. However, the critical behavior is less well understood. One of the fundamental questions in percolation is whether there is an infinite component at criticality, i.e., is $\theta(p_c) = 0$, so that $p \mapsto \theta(p)$ is continuous? Remarkably, this question is in general wide open, even though in $d = 2$ and in high dimensions, it is affirmatively

answered. We shall discuss these special cases in more detail below. In general, the behavior of percolation systems at criticality is still ill understood, and there are no general arguments for the existence of critical exponents. Proofs so far concentrate on proving the existence by computing their values.

Interestingly, several relations can be drawn between the critical exponents on the tree and the ones on $\mathbb{Z}^d$. This is the content of the following theorem:

Theorem 1.6. (Inequalities for critical exponents on $\mathbb{Z}^d$) *For nearest-neighbor bond percolation on $\mathbb{Z}^d$, $\beta \leq 1, \gamma \geq 1$ and $\beta \geq 2/\delta$, when, for each of the inequalities, we assume that the critical exponent in question exists.*

We note that on $\mathbb{T}_r$, the bounds in Theorem 1.6 hold with equality, since $\beta = \gamma = 1, \delta = 2$. In general, the equalities are thus changed to inequalities. We shall prove that $\gamma \geq 1$ below, making use of Russo's formula as well as the BK-inequality. See (Hughes, 1996, Sections 4.3 and 4.4) for a more complete account of inequalities for critical exponents.

Proof that $\gamma \geq 1$. We use Russo's formula (1.8) to obtain

$$\frac{d}{dp}\tau_p(x) = \sum_{(u,v)\in\mathbb{E}} \mathbb{P}_p((u,v) \text{ is pivotal for } o \longleftrightarrow x). \qquad (1.74)$$

Summing over x yields

$$\frac{d}{dp}\chi(p) = \sum_{x\in\mathbb{V}} \sum_{(u,v)\in\mathbb{E}} \mathbb{P}_p((u,v) \text{ is pivotal for } o \longleftrightarrow x). \qquad (1.75)$$

We next note that if (u,v) is pivotal for $o \longleftrightarrow x$, then there exist two disjoint paths of occupied bonds connecting 0 and u, and v and x, respectively. Thus, $\{o \longleftrightarrow u\} \circ \{v \longleftrightarrow x\}$ occurs. By the BK-inequality (1.7), we obtain

$$\frac{d}{dp}\chi(p) \leq \sum_{x\in\mathbb{V}} \sum_{(u,v)\in\mathbb{E}} \tau_p(u)\tau_p(x-v) = r\chi(p)^2, \qquad (1.76)$$

where r is the degree of the graph $\mathbb{G}$. To make (1.76) precise, one has to overcome a few technicalities, since Russo's formula in (1.8) can only be applied to events E depending on a *finite* number of bonds, which is not the case for $\{o \longleftrightarrow x\}$. For details, see e.g., (Heydenreich, van der Hofstad and Sakai, 2008).

To complete the argument, note that we can rewrite (1.76) as

$$\frac{d}{dp}\chi(p)^{-1} \geq -r. \qquad (1.77)$$

Integrating (1.77) over $[p, p_c]$, and noting that $\chi(p_c)^{-1} = 0$ by Theorem 1.4, we obtain that
$$\chi(p_c)^{-1} - \chi(p)^{-1} = -\chi(p)^{-1} \geq -r(p_c - p), \qquad (1.78)$$

so that

$$\chi(p) \geq \frac{1}{r(p_c - p)}. \tag{1.79}$$

This proves that $\gamma \geq 1$ if γ exists. $\qquad\square$

As remarked before, no general arguments exist that critical exponents exist. As we shall explain in more detail below, there are results that guarantee their existence in special cases, and that the critical exponents satisfy the scaling relations.

For percolation on $\mathbb{Z}^d$, apart from the scaling relations in (1.28), theoretical physics arguments predict that for low enough dimensions, two further scaling relations going under the name *hyperscaling relations*, which involve the dimension d, hold:

$$d\rho = \delta + 1, \qquad d\nu = 2 - \alpha. \tag{1.80}$$

More precisely, it is predicted that the scaling relations in (1.28) hold for *all d*, while the hyperscaling relations in (1.80) only hold for d below the so-called *upper critical dimension*. Above the upper critical dimension, the critical exponents are equal to the ones on the tree as appearing in Theorem 1.1. There has been some work in the direction of establishing the scaling and hyperscaling relations. In (Borgs, Chayes, Kesten and Spencer, 1999), it is shown that *if* the critical exponent ρ exists in the logarithmic sense, and if the probability at criticality $p = p_c$ that there is an easy-way crossing in the cube $[0, n] \times [0, 3n]^{d-1}$ is uniformly bounded by $1 - \varepsilon$ for some $\varepsilon > 0$, then also δ and η exist in the logarithmic sense, and satisfy $d\rho = \delta + 1$ and $2 - \eta = d(\delta - 1)/(\delta + 1)$. The latter equation is, when we assume the scaling relations in (1.28), equivalent to $d\nu = 2 - \alpha$.

In (Tasaki, 1987a), (Tasaki, 1987b) and (Chayes and Chayes, 1987) it is shown that when the critical exponents take on the values on a tree, then $d \geq 6$, suggesting that the upper critical dimension is at least 6.

After this discussion on percolation on $\mathbb{Z}^d$ in general dimension, we now move to two special cases where much more is rigorously known, two dimensions and high dimensions.

1.2.4 *Percolation in two dimensions*

In this section, we study percolation in the plane. We start with the fact that $p_c(\mathbb{Z}^2) = 1/2$ for bond percolation on the square lattice:

Theorem 1.7. (Harris-Kesten Theorem: $p_c(\mathbb{Z}^2) = 1/2$) *For nearest-neighbor bond percolation on $\mathbb{Z}^2$, $\theta(1/2) = 0$ and $\theta(p) > 0$ for $p > 1/2$.*

The fact that $\theta(1/2) = 0$, which implies that $p_c(\mathbb{Z}^2) \geq 1/2$, is sometimes called *Harris' Theorem* and was proved in (Harris, 1960). The proof that $\theta(p) > 0$ for $p > 1/2$ is sometimes called *Kesten's Theorem* (Kesten, 1980), and is considerably more involved. We shall prove Harris' Theorem below, making use of an important tool in 2-dimensional percolation going under the name of *Russo-Seymour-Welsh* or RSW-theory (Russo, 1981), (Seymour and Welsh, 1978). We

state the version of the RSW Theorem from (Russo, 1981). We first introduce some notation. We let $R_{n,k}(p)$ be the probability that there is an occupied path from left to right crossing the rectangle $[0, kn] \times [0, n]$.

Theorem 1.8. (RSW Theorem) *For any $p \in (0, 1)$ and $n, k \geq 1$ with n even, and for percolation on the square lattice $\mathbb{Z}^2$, the following bounds hold:*

$$R_{n,2}(p) \geq R_{n,1}(p)\Big[1 - \sqrt{1 - R_{n,1}(p)}\Big]^6, \quad R_{n,3}(p) \geq R_{n,1}(p)^3 \Big[1 - \sqrt{1 - R_{n,1}(p)}\Big]^{12}. \tag{1.81}$$

We start by proving the RSW Theorem:

Proof of the RSW Theorem 1.8. The main tool will be the Harris-inequality in (1.5). We start by deducing the second inequality from the first. We note that if we take the two rectangles $[0, 2n] \times [0, n]$ and $[n, 3n] \times [0, n]$, then their intersection is $[n, 2n] \times [0, n]$. Further, if there are left-to-right crossings in $[0, 2n] \times [0, n]$ and in $[n, 3n] \times [0, n]$, and a top-to-bottom crossing in $[n, 2n] \times [0, n]$, then there is also a left-to-right crossing in $[0, 3n] \times [0, n]$. Let $\mathrm{LR}_{n,1}, \mathrm{LR}_{n,2}, \mathrm{TB}_{n,3}$ be the events that the three respective crossings exist. Then,

$$R_{n,3}(p) \geq \mathbb{P}_p(\mathrm{LR}_{n,1} \cap \mathrm{LR}_{n,2} \cap \mathrm{TB}_{n,3}). \tag{1.82}$$

Now, $\mathrm{LR}_{n,1}, \mathrm{LR}_{n,2}$ and $\mathrm{TB}_{n,3}$ are all increasing events with

$$\mathbb{P}_p(\mathrm{LR}_{n,1}) = \mathbb{P}_p(\mathrm{LR}_{n,2}) = R_{n,2}(p), \quad \text{and} \quad \mathbb{P}_p(\mathrm{TB}_{n,3}) = R_{n,1}(p). \tag{1.83}$$

Thus, we arrive at

$$R_{n,3}(p) \geq R_{n,1}(p) R_{n,2}(p)^2, \tag{1.84}$$

showing that the first inequality in (1.81) implies the second.

To see the first inequality in (1.81), in exactly the same way as above, we can show that

$$R_{n,2}(p) \geq R_{n,1}(p) R_{n,3/2}(p)^2. \tag{1.85}$$

We shall frequently make use of a clever consequence of Harris' inequality, sometimes called the *square root trick*. This trick states that, for any two increasing events A_1 and A_2 with equal probability $\mathbb{P}_p(A_1) = \mathbb{P}_p(A_2)$, we have

$$\mathbb{P}_p(A_1) \geq 1 - \sqrt{1 - \mathbb{P}_p(A_1 \cup A_2)}. \tag{1.86}$$

To see (1.86), we note that

$$[1 - \mathbb{P}_p(A_1)]^2 = \mathbb{P}_p(A_1^c)^2 = \mathbb{P}_p(A_1^c)\mathbb{P}_p(A_2^c) \leq \mathbb{P}_p(A_1^c \cap A_2^c) = 1 - \mathbb{P}_p(A_1 \cup A_2), \tag{1.87}$$

so that (1.86) follows by taking the square root and rearranging terms. We shall now apply (1.86). For this, let H_u be the event that there exists a left-to-right crossing in $[0, n] \times [0, n]$ starting in the line $a_u = \{0\} \times [n/2, n]$ and H_l the event that there exists a left-to-right crossing in $[0, n] \times [0, n]$ starting in the line

$a_l = \{0\} \times [0, n/2]$. Then, by (1.86), and the fact that there is a left-to-right crossing in $[0, n] \times [0, n]$ precisely when $H_u \cup H_l$ holds, we have

$$\mathbb{P}_p(H_u) \geq 1 - \sqrt{1 - R_{n,1}(p)}. \tag{1.88}$$

Now we introduce some more notation. Let $s = (s_1, \ldots, s_m)$ be a path connecting the left and right sides of $[0, n] \times [0, n]$, starting on the left side and taking values in $[0, n] \times [0, n]$. We let S_l denote the set of such paths for which $s^a \in a_l$. Let E_s be the event that s is the *lowest* occupied path connecting the left and right sides of $[0, n] \times [0, n]$. For such an s, and with $a = a_u \cup a_l = \{0\} \times [0, n]$, we let s^a be the last intersection with a, and we write $s^r = (s^a, \ldots, s_m)$ for the part of the path s after its last visit to a, and $s^{r'}$ for the reflection of s^r in the line $\{n\} \times [0, n]$. Let F_s be the event that there exists a path in $[n/2, 3n/2] \times [0, n]$ connecting the top of $[n/2, 3n/2] \times [0, n]$ to the path s^r, and which always remains above $s^r \cup s^{r'}$. It is not hard to see that, for every s, we have that $E_s \cap F_s \cap H_u \subset \mathrm{LR}_{n,3/2}$, where we write $\mathrm{LR}_{n,3/2}$ for the event that there is a left-to-right crossing in $[0, 3n/2] \times [0, n]$. Thus, with

$$G = \bigcup_{s \in S_l} E_s \cap F_s, \tag{1.89}$$

we obtain that $G \cap H_u \subset \mathrm{LR}_{n,3/2}$, which implies that

$$R_{n,3/2}(p) \geq \mathbb{P}_p(G \cap H_u). \tag{1.90}$$

Applying the Harris' inequality once more and noting that both G and H_u are increasing, we obtain that

$$R_{n,3/2}(p) \geq \mathbb{P}_p(G)\mathbb{P}_p(H_u). \tag{1.91}$$

We have already derived a lower bound on $\mathbb{P}_p(H_u)$ in (1.88), so we are left to lower bound $\mathbb{P}_p(G)$. Since G is a disjoint union, we have that

$$\mathbb{P}_p(G) = \sum_{s \in S_l} \mathbb{P}_p(E_s \cap F_s). \tag{1.92}$$

Now, for a fixed s, F_s only depends on edges in $[0, 3n/2] \times [0, n]$ that are above $s^r \cup s^{r'}$, while E_s only depends on edges in $[0, n] \times [0, n]$ that are below $s^r \cup s^{r'}$, so that E_s and F_s are independent. Thus,

$$\mathbb{P}_p(G) \geq \sum_{s \in S_l} \mathbb{P}_p(F_s)\mathbb{P}_p(E_s). \tag{1.93}$$

For fixed s, denote by F'_s the event that there exists a path in $[n/2, 3n/2] \times [0, n]$ connecting the top of $[n/2, 3n/2] \times [0, n]$ to the path $s^{r'}$, and which always remains

above $s^r \cup s^{r'}$. Then, clearly, $\mathbb{P}_p(F_s) = \mathbb{P}_p(F'_s)$ and both F_s and F'_s are increasing, so that by the square root trick (1.86), we have

$$\mathbb{P}_p(F_s) \geq 1 - \sqrt{1 - \mathbb{P}_p(F_s \cup F'_s)} \geq 1 - \sqrt{1 - R_{n,1}(p)}, \qquad (1.94)$$

since $\mathbb{P}_p(F_s \cup F'_s) \leq R_{n,1}(p)$. Thus, using that the union over $s \in S_l$ of E_s equals H_u,

$$\mathbb{P}_p(G) \geq \Big[1 - \sqrt{1 - R_{n,1}(p)}\Big] \sum_{s \in S_l} \mathbb{P}_p(E_s) = \Big[1 - \sqrt{1 - R_{n,1}(p)}\Big] \mathbb{P}_p(H_u). \ (1.95)$$

Combining (1.91) with (1.95) yields,

$$R_{n,3/2}(p) \geq \Big[1 - \sqrt{1 - R_{n,1}(p)}\Big] \mathbb{P}_p(H_u)^2 \geq \Big[1 - \sqrt{1 - R_{n,1}(p)}\Big]^3, \qquad (1.96)$$

which, combined with (1.85), yields the first claim in (1.81). □

We continue by discussing *duality*, a notion which has been extremely important in two-dimensional percolation. We shall assume that we are working on a two-dimensional *planar lattice* $\mathbb{L}$, i.e., a graph that can be embedded into $\mathbb{R}^2$ in such a way that different edges can only meet at the vertices of the lattice. We shall assume that the graph is translation invariant, and that the embedding of the lattice divides $\mathbb{R}^2$ into an infinite tiling of identical and bounded faces. Then, the vertices of the dual lattice $\mathbb{L}^*$ are the *faces* of the embedding of $\mathbb{L}$, and we connect two vertices in $\mathbb{L}^*$ when their corresponding faces share a boundary, which is a bond in $\mathbb{L}$. Thus, to each bond in $\mathbb{L}^*$ we can identify a *unique* bond in $\mathbb{L}$. Thus, to a bond percolation configuration on $\mathbb{L}$, we have identified a bond percolation configuration on $\mathbb{L}^*$ by identifying the occupation status of a bond on $\mathbb{L}^*$ to the one of the bond on $\mathbb{L}$ to which it is identified. It is sometimes convenient to identify the vertices of the dual lattice with the *centers* of the faces, and in this representation, we have that the nearest-neighbor square lattice on $\mathbb{Z}^2$ is dual to the nearest-neighbor square lattice on $\mathbb{Z}^2 + (1/2, 1/2)$, as discussed in Section 1.2.3. As a result, percolation on the nearest-neighbor square lattice $\mathbb{Z}^2$ is self-dual.

We now investigate the event LR_n that there is a left-right crossing of occupied bonds in the rectangle $[0, n+1] \times [0, n]$. Also, denote by TB_n^* the event that there is a top to bottom crossing of *vacant* bonds in the dual lattice on $[1/2, n+1/2] \times [-1/2, n+1/2]$. Clearly, one of the two must happen, so that

$$\mathbb{P}_p(\mathrm{LR}_n) + \mathbb{P}_p(\mathrm{TB}_n^*) = 1. \qquad (1.97)$$

By construction, we have that $\mathbb{P}_p(\mathrm{TB}_n^*) = \mathbb{P}_{1-p}(\mathrm{LR}_n)$, so that we obtain, for all $p \in [0, 1]$ and $n \geq 0$,

$$\mathbb{P}_p(\mathrm{LR}_n) + \mathbb{P}_{1-p}(\mathrm{LR}_n) = 1. \qquad (1.98)$$

When we pick $p = 1/2$, we thus obtain that $\mathbb{P}_{1/2}(\mathrm{LR}_n) = 1/2$ for every $n \geq 1$. Thus, $R_{n,1}(1/2) \geq \mathbb{P}_{1/2}(\mathrm{LR}_n) = 1/2$ for every $n \geq 1$, so that $R_{n,2}(1/2) \geq$

$[1 - \sqrt{1/2}]^6/2 > 0$ uniformly in $n \geq 1$. We shall see that this is sufficient to show that $\theta(1/2) = 0$. Followed by the proof of $\theta(1/2) = 0$, we shall discuss some further consequences of duality.

Proof that $\theta(1/2) = 0$ for bond percolation on the two-dimensional square lattice. Fix $n \geq 1$, and let G_n be the event that there exists an occupied path from $\partial[-n, n] \times [-n, n]$ to $\partial[-3n, 3n] \times [-3n, 3n]$, where we write $\partial[a, b] \times [c, d]$ for the boundary of the rectangle $[a, b] \times [c, d]$. Then,

$$\mathbb{P}_{1/2}(o \longleftrightarrow \partial B(3^k)) \leq \mathbb{P}_{1/2}(\cap_{l=1}^k G_{3^l}) = \prod_{l=1}^k \mathbb{P}_{1/2}(G_{3^l}) = \prod_{l=1}^k [1 - \mathbb{P}_{1/2}(G_{3^l}^c)],$$
(1.99)

since the events $\{G_{3^l}\}_{l=1}^\infty$ depend on the occupation status of disjoint sets of bonds, and are therefore independent. By duality, the event G_n^c occurs *precisely* when there is a dual path of vacant edges in $[-3n, 3n] \times [-3n, 3n] \setminus [-n, n] \times [-n, n]$ surrounding the square $[-n, n] \times [-n, n]$. Denote this event by $O(n)$. Then, $\mathbb{P}_{1/2}(G_n^c) = \mathbb{P}_{1/2}(O(n))$. The event $O(n)$ is a subset of the event that there exists left-to-right crossings in the rectangles $[-3n, 3n] \times [n, 3n]$ and $[-3n, -3n] \times [-n, -3n]$, and top-to-bottom crossings in the rectangles $[n, 3n] \times [-3n, 3n]$ and $[-n, -3n] \times [-3n, 3n]$. The probability of each of these events is equal to $R_{n,3}(1/2)$. By Harris' inequality, the crossings are positively correlated:

$$\mathbb{P}_{1/2}(O(n)) \geq R_{n,3}(1/2)^4.$$
(1.100)

By Theorem 1.8 and since $R_{n,1}(1/2) \geq 1/2$, we obtain that $R_{n,3}(1/2) \geq a$ for some explicit $a > 0$, and uniformly in $n \geq 1$, so that

$$\mathbb{P}_{1/2}(o \longleftrightarrow \partial B(3^k)) \leq \prod_{l=1}^k (1 - a) = (1 - a)^k.$$
(1.101)

As a result, we even obtain that $\mathbb{P}_{1/2}(o \longleftrightarrow \partial B(n)) \leq (1 - a)^{\lfloor \log n / \log 3 \rfloor} = O(n^{-a/\log 3})$, so that, if $\rho > 0$ exists, we obtain that $\rho \leq \log 3/a$. In particular, we have that $\theta(1/2) \leq \mathbb{P}_{1/2}(o \longleftrightarrow \partial B(n))$ for each $n \geq 1$, which tends to 0 when $n \to \infty$. $\square$

Duality can be used in several more convenient ways. For example, we can define an alternative *correlation length* $\tilde{\xi}(p)$ by the limit

$$\tilde{\xi}(p)^{-1} = -\lim_{n \to \infty} \frac{\log \tau_p^{\mathrm{f}}(0, ne_1)}{n}.$$
(1.102)

It can be expected that $\tilde{\xi}(p)$ is of the same order of magnitude as $\xi(p)$ in (1.15) when p is close to critical, as they both describe the maximal distance between vertices for which there is dependence between their clusters. Then, (Chayes,

Chayes, Grimmett, Kesten and Schonmann, 1989) use duality to show that $\tilde{\xi}(p) = \tilde{\xi}(1-p)/2$ for $p > 1/2$. Thus, in particular, if ν exists, then so does ν' and it takes the same value.

Using arguments such as duality and RSW-theory, the critical values of several other lattices in two dimensions have been established. An example is $p_c^{\mathrm{site}}(T) = 1/2$, where T is the triangular lattice, which will play an important role later on. Also site percolation on the triangular lattice turns out to be self-dual, which explains why its critical value is $1/2$.

Recently, Bollobás and Riordan adapted the RSW-ideas in such a way that they are more generally applicable. For example, in (Bollobás and Riordan, 2006c), they used this new methodology to give a simpler proof of the Harris-Kesten Theorem. The adapted RSW-methods have also been crucial in order to prove that certain critical values for site percolation on certain tesselations equal $1/2$. For example, take a Poisson point process. For each vertex in $\mathbb{R}^2$, attach it to the closest point(s) in the Poisson point process. This divides $\mathbb{R}^2$ in cells, and we draw an edge between two cells when they share a line of their respective boundaries. Then, when we color the cells independently green with probability p and yellow with probability $1-p$, (Bollobás and Riordan, 2006a) shows that the critical value of the occurrence of an infinite green connected component again is $1/2$. In (Bollobás and Riordan, 2008), this result is extended to other two-dimensional tesselations, such as the Johnson-Mehl tesselation and two-dimensional slices of three-dimensional Voronoi tesselations.

In the remainder of this section, we shall work with site percolation on the triangular lattice, for which in the past decade tremendous and remarkable progress has been made. We start by giving the exact values of the critical exponents, the values of which have been predicted early on in the physics community (see (Nienhuis, 1984)) for general two-dimensional lattices:

Theorem 1.9. (Critical exponents on the triangular lattice) *For site percolation on the two-dimensional triangular lattice, the critical exponents $\beta, \gamma, \nu, \eta, \delta$ and ρ exist in the logarithmic sense, and take on the values*

$$\beta = \frac{5}{36}, \quad \gamma = \gamma' = \frac{43}{18}, \quad \nu = \nu' = \frac{4}{3}, \quad \eta = \frac{5}{24}, \quad \delta = \frac{91}{5} \quad \rho = \frac{48}{5}. \quad (1.103)$$

Theorem 1.9 is one of the major breakthroughs in modern probability theory, particularly since its proof has shed light not only on the existence and scaling limit of two-dimensional percolation, but rather of the critical behavior of a wide class of two-dimensional statistical physical models. So far, this technology has not only been used for percolation, but also for loop-erased random walk and uniform spanning trees (Schramm, 2000), and a proof for the Ising model has been announced by Smirnov. The proof of Theorem 1.9 is a consequence of the connection between critical percolation and so-called *stochastic Loewner evolution* (SLE), a topic which we will discuss in some detail below. The history is that Schramm (Schramm, 2000) first identified a class of continuous models, so-called SLE, which are conformally invariant models in the plane of which the

properties depend on its parameter $\kappa > 0$. Schramm continued by noting that *if the scaling limit of two-dimensional percolation would be conformally invariant, then it must be equal to SLE with parameter $\kappa = 6$.* Smirnov (Smirnov, 2001) proved that indeed the scaling limit of critical percolation on the triangular lattice is conformally invariant. This is the celebrated result by Smirnov (Smirnov, 2001), which we shall discuss in more detail below. Schramm already noted that SLE with parameter $\kappa = 6$ has similar critical exponents as in (1.103), when defined in an appropriate way. Smirnov and Werner (Smirnov and Werner, 2001) realized that the values listed in (1.103) follow from the two statements

$$\mathbb{P}_{1/2}(A_R^1) = R^{-5/48+o(1)} \tag{1.104}$$

and

$$\mathbb{P}_{1/2}(A_R^2) = R^{-5/4+o(1)}, \tag{1.105}$$

where A_R^1 is the event that the origin is connected to the boundary of a ball of radius R, while A_R^2 is the probability that there exist two neighbors of the origin of which one has a green connection and another has a yellow connection to the boundary of the ball of width R. The first of these identities simply states that ρ exists and takes the value $48/5$. The fact that these two statements imply the existence and values of the critical exponents listed in (1.103) is non-trivial and due to (Kesten, 1987). The asymptotics in (1.104) was shown in (Lawler, Schramm and Werner, 2002), the one in (1.105) in (Smirnov and Werner, 2001). The equalities $\gamma = \gamma'$ and $\nu = \nu'$ follow from the self-duality of site percolation on the triangular lattice. Theorem 1.9 identifies almost all critical exponents, an exception being α, which has always remained to be somewhat mysterious. While Δ does not appear in Theorem 1.9, we believe that its derivation should be easier than that of α.

We now discuss the recent work on the scaling limit of critical percolation on the triangular lattice in more detail. In order to do so, we must start with the notion of *conformal invariance*, a notion which is crucial in the study of two-dimensional critical systems. We work on $\mathbb{C}$, and we let $D \subset \mathbb{C}$ be a simply connected domain. We say that a map $f \colon \mathbb{C} \to \mathbb{C}$ is *conformally invariant* when it preserves angles. In order to explain when a map f is angle-preserving we introduce some notation. Let $t \mapsto \gamma_1(t)$ and $t \mapsto \gamma_2(t)$ be two crossing curves with $\gamma_1(t), \gamma_2(t) \in \mathbb{C}$ for all t. Suppose that γ_1 and γ_2 are sufficiently smooth, then for t close to s, the curves $\gamma_1(t)$ and $\gamma_2(t)$ cross each other at a certain angle. f is called a *conformal map* or *preserves angles* when the curves $f(\gamma_1(t))$ and $f(\gamma_2(t))$ cross at the same angle as γ_1 and γ_2 for every pair of crossing curves γ_1 and γ_2 in D. Important examples of conformal maps are *Möbius transformations* given by

$$f(z) = \frac{az+b}{cz+d}, \tag{1.106}$$

where $a, b, c, d \in \mathbb{C}$ with $ad - bc \neq 0$. Let $D \subset \mathbb{C}$ be a (sufficiently smooth) domain, with four points P_1, P_2, P_3 and P_4 on the boundary which are such

that P_i is in between P_{i-1} and P_{i+1} (where, by convention, $P_5 \equiv P_1$). We call $D_4 = (D; P_1, P_2, P_3, P_4)$ a 4-marked domain.

We shall investigate general lattices $\mathbb{L}$ in two dimensions, and we shall rescale the lattice with a small factor δ that shall later tend to zero (note that this δ has nothing to do with the critical exponent δ in (1.25)). For example, for the triangular lattice, we can think of δ as being the width of the edges in the lattice. Then, (Langlands, Pouliot and Saint-Aubin, 1994) studied crossing probabilities of the form that the boundary of D between P_1 and P_2 has an occupied path to the part of the boundary between P_3 and P_4. Denote this event by $C_\delta(D_4)$. When we work on a rectangle with $P_1 = (0,0), P_2 = (0,n), P_3 = (n,m), P_4 = (m,0)$, then this is nothing but the statement that the rectangle has a left-to-right crossing. In (Langlands, Pouliot and Saint-Aubin, 1994), the hypothesis was made that

$$P(D_4) = \lim_{\delta \downarrow 0} \mathbb{P}_{p_c}(C_\delta(D_4)) \tag{1.107}$$

exists and lies in $(0,1)$ when the points P_1, P_2, P_3, P_4 are different. These assumptions are already highly non-trivial, but the main assumption in (Langlands, Pouliot and Saint-Aubin, 1994) is that the limit $P(D_4)$ is *conformally invariant*. This is what is often meant with the assumption that the scaling limit of percolation is conformally invariant. Let us now explain what this assumption means in more detail. The limit $P(D_4)$ is conformally invariant when, for D_4' being the image under a conformal map of D_4, we have that that $P(D_4) = P(D_4')$. This means that if we would consider the intersection of D_4' with the discretized lattice of width δ, and we compute the limit

$$P(D_4') = \lim_{\delta \downarrow 0} \mathbb{P}_{p_c}(C_\delta(D_4')), \tag{1.108}$$

then in fact this limit exists and equals $P(D_4') = P(D_4)$.

The above 'hypothesis' now goes under the name of *conformal invariance* of percolation, and is in fact what the celebrated paper (Smirnov, 2001) has proved on the triangular lattice. Since there are many conformal maps, the conformal invariance hypothesis is actually quite strong. In fact, Cardy (Cardy, 1992) used it to make a prediction of the exact limit of crossing probabilities in various domains, using mathematically non-rigorous arguments from conformal field theory. In order to explain Cardy's conjecture, we note that, since the limit of crossing probabilities is invariant under conformal maps, and for any 4-marked domain D_4 there exists a conformal map that maps D to the circle $B(1)$ and P_i to z_i on the boundary of the circle, the limit of crossing probabilities is determined by the crossing probabilities on the circle. For such special 4-marked domains, we define the *cross-ratio* η (which again has nothing to do with the critical exponent η defined in (1.27)) by

$$\eta = \frac{(z_4 - z_3)(z_2 - z_1)}{(z_4 - z_2)(z_3 - z_1)}. \tag{1.109}$$

It turns out that $\eta \in (0,1)$ and that two 4-marked domains on the circle can be mapped to one another by a conformal map *if and only if* they have the same cross-ratio η. Thus, 4-marked domains on the circle can be characterized by their cross-ratios. In turn, there is a unique conformal map mapping any 4-marked domain to a 4-marked domain on the circle, so that we can define the cross-ratio of a general 4-marked domain to be the cross-ratio of the image under the unique conformal map to the circle. Thus, we see that two 4-marked domains are conformally equivalent precisely when their cross-ratios are equal, and we can reformulate the hypothesis of (Langlands, Pouliot and Saint-Aubin, 1994) to say that the limiting crossing probabilities are a function of their cross-ratio, i.e., there exists a function $f\colon (0,1) \mapsto (0,1)$ such that $P(D_4) = f(\eta(D_4))$ when $\eta(D_4)$ is the cross-ratio of the 4-marked domain D_4. Based on this assumption, (Cardy, 1992) shows that in fact

$$P(D_4) = \frac{3\Gamma(2/3)}{\Gamma(1/3)^2}\eta^{1/3}\ {}_2F_1(1/3, 2/3; 4/3; \eta), \qquad (1.110)$$

where ${}_2F_1$ is a hypergeometric function. Carleson noted that Cardy's conjecture takes a particularly appealing form on an equilateral triangle, i.e., we take $P_1 = (1,0), P_2 = (1/2, \sqrt{3}/2), P_3 = (0,0)$ and $P_4 = (x,0)$ where $x \in (0,1)$, and D is the interior of the equilateral triangle spanned by P_1, P_2, P_3. Then, (1.110) is equivalent to stating that, for all $x \in (0,1)$,

$$P(D_4) = x. \qquad (1.111)$$

Note that any 4-marked domain can be conformally mapped to the equilateral triangle, the only degree of freedom being the value of x in $P_x = (x, 0)$. In his seminal paper, Smirnov (Smirnov, 2001) showed (1.111). We shall not go into the proof in (Smirnov, 2001), as this would take up many more pages than were allotted to us. Instead, we give some more references to the literature. The first idea of a possible scaling limit of critical percolation is in the seminal paper of Schramm, which introduce the limiting stochastic process, which goes under the name of *Stochastic Loewner Evolution*, or sometimes *Schramm Loewner Evolution* (SLE). Now, SLE has developed into the main tool for studying two-dimensional critical systems, and it is likely that much more progress shall be made in this direction in the coming years. We refer to (Bollobás and Riordan, 2006*b*, Chapter 7) for an expository discussion of Smirnov's proof, as well as its consequences. Reviews on SLE and its consequences can be found in (Lawler, 2004; Lawler, 2005; Kager and Nienhuis, 2004; Werner, 2004; Werner, 2005), and we refer there for more detailed discussions. We close this discussion by noting that, while the physics community has always predicted that the *same* critical behavior should be valid for a wide range of two-dimensional critical percolation models, the proof on the triangular lattice is, to date, still basically the only proof of conformal invariance for two-dimensional percolation models. In particular, the corresponding result for two-dimensional bond percolation on

the square lattice is unknown. One reason for this is that the proof in (Smirnov, 2001) makes essential use of the three-fold rotational symmetry of the triangular lattice, and it is, up to date, unclear how these symmetries can be replaced by the different sets of symmetries on the square lattice.

SLE has also proved useful to understand the so-called *near-critical phase* of percolation, where $p = 1/2 + \theta \delta^{4/3}$, a nice survey of these results, as well as of the proof of convergence of the percolation exploration process which explored the boundary of clusters can be found in (Camia, 2008).

The two-dimensional percolation problem is, after percolation on the tree, the best understood percolation problem, and the results described above give a rather complete overview of the depth and wealth of two-dimensional percolation theory. Yet, several results are not yet known and are worth considering: (a) proof of existence of *scaling functions* (see e.g., (Hughes, 1996, (4.295))); (b) a closer investigation of near-critical percolation (see (Camia, 2008) for an overview); (c) improvement of our understanding of *universality* in two-dimensional percolation, for example, by proving conformal invariance of two-dimensional bond percolation on the square lattice.

1.2.5 *Percolation in high dimensions*

In this section, we study percolation in high-dimensions. We consider $\mathbb{G} = (\mathbb{V}, \mathbb{E})$ with $\mathbb{V} = \mathbb{Z}^d$ and with edge set $\mathbb{E}$ either the nearest-neighbor bonds in sufficiently high dimension, or the *spread-out* bonds

$$\mathbb{E} = \{\{x, y\} : \|x - y\|_\infty \leq L\} \tag{1.112}$$

for some L sufficiently large. The main result in high dimensions is the following:

Theorem 1.10. (Mean-field critical exponents for high-d percolation)
For percolation on $\mathbb{Z}^d$, for either d sufficiently large in the nearest-neighbor model, or $d > 6$ and L sufficiently large in the spread-out model, $\beta = \gamma = 1$, $\nu = 1/2$ and $\delta = \Delta = 2$ in the bounded-ratio sense, while $\eta = 0$ in the asymptotic sense.

It is believed, by invoking the paradigm of *universality*, that the critical exponents for any finite-range system which has sufficient symmetry are equal. Thus, Theorem 1.10 suggests that also $\beta = \gamma = 1$, $\nu = 1/2$ and $\delta = \Delta = 2$ for the nearest-neighbor model with $d > 6$. However, since the paradigm of universality is not mathematically rigorous, we cannot conclude this. Note that Theorem 1.10 *does* support the prediction of universality, since, in particular, the values of the critical exponents do not depend on the precise values of L, when L is sufficiently large. Also, for d sufficiently large, the critical values agree for *all* values of L.

The reason for the fact that L or d needs to be big in Theorem 1.10 is that the proof of Theorem 1.10 makes use of a perturbation expansion called the *lace expansion*. We shall now first discuss the history of the proof, before discussing the details.

In (Aizenman and Newman, 1984), it was proved that $\gamma = 1$ when the so-called *triangle condition*, a condition on the percolation model, holds. The triangle condition states that

$$\nabla(p_c) = \sum_{x,y \in \mathbb{V}} \tau_{p_c}(0, x)\tau_{p_c}(x, y)\tau_{p_c}(y, 0) < \infty, \qquad (1.113)$$

where we recall that $\tau_p(x, y) = \mathbb{P}_p(x \longleftrightarrow y)$ is the two-point function. In (Barsky and Aizenman, 1991), it was shown that, under the same condition, $\beta = 1$ and $\delta = 2$. Needless to say, without the actual *verification* of the triangle condition, this would not prove anything. The triangle condition was proved to hold in the setting in Theorem 1.10 in (Hara and Slade, 1990) by the use of the lace expansion, a method which has since proved to be extremely powerful in order to characterise mean-field behavior of various models in high dimensions. Later, the results for ν, Δ and η were proved in (Hara, 1990), (Nguyen, 1987) and (Hara, 2005; Hara, van der Hofstad and Slade, 2003), again using lace expansion arguments. Several related results on high-dimensional percolation, in particular suggesting that the scaling limit of large critical clusters is a process called *Integrated Brownian excursion*, can be found in (Hara and Slade, 2000a; Hara and Slade, 2000b), where also the fact that $\delta = 2$ in the asymptotic sense was proved.

We now discuss the methodology in high dimensions. We start with the proof that $\gamma = 1$ if the triangle condition holds. Recall the argument that shows that $\gamma \geq 1$ below Theorem 1.6, and (1.75) in particular. The BK-inequality gives an upper bound on (1.75), and, in order to prove that $\gamma = 1$, a matching lower bound needs to be obtained. For this, we can use the independence of the occupation status of the bonds to explicitly write

$$\mathbb{P}_p((u, v) \text{ is pivotal for } o \longleftrightarrow x) = \mathbb{E}_p[\mathbb{1}_{\{o \longleftrightarrow u\}} \tau^{\tilde{C}^{(u,v)}(o)}(v, x)], \qquad (1.114)$$

where, for a set of sites A, the restricted two-point function $\tau^A(v, x)$ is the probability that v is connected to x using only bonds with both endpoints outside A, and $\tilde{C}^{(u,v)}(o)$ consists of those sites which are connected to 0 without using the bond (u, v). Clearly, $\tau^{\tilde{C}^{(u,v)}(o)}(v, x) \leq \tau(v, x)$, and this reproves the upper bound previously obtained using the BK-inequality. We note that

$$\tau(v, x) - \tau^A(v, x) = \mathbb{P}_p(v \xleftrightarrow{A} x), \qquad (1.115)$$

where we write that $v \xleftrightarrow{A} x$ when every path of occupied bonds from v to x contains a bond containing a vertex in A. Thus,

$$\frac{d}{dp}\chi(p) = \sum_{x\in\mathbb{V}}\sum_{(u,v)\in\mathbb{E}}\mathbb{E}_p[\mathbb{1}_{\{o\longleftrightarrow u\}}]\tau_p(v,x)$$

$$-\sum_{x\in\mathbb{V}}\sum_{(u,v)\in\mathbb{E}}\mathbb{E}_p[\mathbb{1}_{\{o\longleftrightarrow u\}}\mathbb{P}_p(v\overset{\tilde{C}^{(u,v)}(o)}{\longleftrightarrow}x)]$$

$$= r\chi(p)^2 - \sum_{x\in\mathbb{V}}\sum_{(u,v)\in\mathbb{E}}\mathbb{E}_p[\mathbb{1}_{\{o\longleftrightarrow u\}}\mathbb{P}_p(v\overset{\tilde{C}^{(u,v)}(o)}{\longleftrightarrow}x)]. \qquad (1.116)$$

Now, for any $A\subseteq\mathbb{Z}^d$,

$$\mathbb{P}_p(v\overset{A}{\longleftrightarrow}x) \le \sum_{a\in A}\mathbb{P}_p(\{v\longleftrightarrow a\}\circ\{a\longleftrightarrow x\}), \qquad (1.117)$$

which leads to

$$\frac{d}{dp}\chi(p) \ge r\chi(p)^2 - \chi(p)\sum_{(u,v)\in\mathbb{E}}\mathbb{E}_p[\mathbb{1}_{\{o\longleftrightarrow u,a\}}]\mathbb{P}_p(v\longleftrightarrow a). \qquad (1.118)$$

Applying the BK-inequality yields that

$$\mathbb{P}_p(o\longleftrightarrow u,a) \le \sum_z\mathbb{P}_p(\{o\longleftrightarrow z\}\circ\{z\longleftrightarrow u\}\circ\{z\longleftrightarrow a\}) \qquad (1.119)$$

$$\le \sum_z\tau_p(z)\tau_p(u-z)\tau_p(a-z),$$

so that

$$\sum_{(u,v)\in\mathbb{E}}\mathbb{P}_p(o\longleftrightarrow u,a)\mathbb{P}_p(v\longleftrightarrow a) \le \chi(p)[\nabla(p)-1], \qquad (1.120)$$

implying that

$$\frac{d}{dp}\chi(p) \ge \chi(p)^2[2-\nabla(p_c)]. \qquad (1.121)$$

If we know that $\nabla(p_c) < 2$, then we can integrate the above equation in a similar fashion as around (1.77) to obtain that $\gamma = 1$. When we only have the *finiteness* of the triangle, then some more work is necessary to make the same conclusion (see (Aizenman and Newman, 1984) for details).

In the lace expansion, the above argument is adapted to deal with $\tau_p(x)$ directly, using rewrites as in (1.114)–(1.115) repeatedly, instead of using the upper bound in (1.117). We refer to (Hara and Slade, 1990) or the monograph (Slade, 2006) for detailed derivations of the lace expansion.

The lace expansion can also be used to prove asymptotics of the critical value in high dimensions, either when $L\to\infty$ for $d > 6$ fixed or for the nearest-neighbor model and $d\to\infty$. In (van der Hofstad and Sakai, 2005), the asymptotics of the critical point for percolation, as well as for self-avoiding walk, the

contact process and oriented percolation, were investigated for $d > 6$ and $L \to \infty$. It was shown that

$$p_c(L, \mathbb{Z}^d) = \frac{1 + c_d L^d + o(L^{-d})}{(2L + 1)^d - 1},\tag{1.122}$$

for some explicit constant $c_d > 0$, and where $p_c(L, \mathbb{Z}^d)$ is the critical value of spread-out percolation with edge set $\mathbb{E}$ in (1.112). The best asymptotics of $p_c(\mathbb{Z}^d)$ when $d \to \infty$ are in (Hara and Slade, 1993; Hara and Slade, 1995; van der Hofstad and Slade, 2005; van der Hofstad and Slade, 2006), where it is shown that, when $d \to \infty$, $p_c(\mathbb{Z}^d)$ has an asymptotic expansion in terms of inverse powers of $(2d)$ with rational coefficients, i.e., for each n, we can write

$$p_c(\mathbb{Z}^d) = \sum_{i=1}^{n} a_i (2d)^{-i} + O((2d)^{n+1}),\tag{1.123}$$

where the a_i are rational coefficients with $a_1 = a_2 = 1, a_3 = 7/2$. We refer to the references in (van der Hofstad and Sakai, 2005) and (van der Hofstad and Slade, 2006) for the literature on asymptotics of percolation critical points.

We close this section by discussing finite-size scaling in high-dimensional percolation. In (Aizenman, 1997), it was assumed that a version of $\eta = 0$ holds (more precisely, that $\tau_{p_c}(x)$ is bounded above and below by positive and finite constants times $|x|^{-(d-2)}$) in order to show that, at criticality, the largest intersection of a cluster with a cube of width $2r + 1$ grows like r^4 times logarithmic corrections. This corresponds to the *bulk* boundary condition. The condition used was verified in (Hara, 2005; Hara, van der Hofstad and Slade, 2003) in the setting of Theorem 1.10.

Aizenman proceeds to conjecture that at criticality, with *periodic* boundary conditions, the largest critical cluster grows like $r^{2d/3}$, i.e., like $V^{2/3}$, where $V = (2r + 1)^d$ is the volume of the cube. This was proved in (Heydenreich and van der Hofstad, 2007) making crucial use of the combined results in (Borgs, Chayes, van der Hofstad, Slade and Spencer, 2005a; Borgs, Chayes, van der Hofstad, Slade and Spencer, 2005b). Such behavior is dubbed *random graph asymptotics*, as $V^{2/3}$ growth at criticality is best known for the Erdős-Rényi random graph discussed in Section 1.3 below. Related results in this direction can be found in (van der Hofstad and Luczak, 2006) and (Nachmias, 2007), using alternative methods. An interesting question is what proper definitions of the critical value or window are in the general context of percolation on finite graphs, and what the proper conditions on the graph are such that percolation on it has random graph asymptotics close to criticality. While the combined examples in (Borgs, Chayes, van der Hofstad, Slade and Spencer, 2005a; Borgs, Chayes, van der Hofstad, Slade and Spencer, 2005b; Heydenreich and van der Hofstad, 2007; Nachmias, 2007) provide some initial ideas, the general picture is not yet clear.

Despite the fact that detailed results are available in high dimensions, several results are not yet known and are worth considering: (a) proof of existence and mean-field values of ν', γ', ρ (particularly the supercritical regime in high-dimensions is still ill understood); (b) proof of existence of *scaling functions* (see

e.g., (Hughes, 1996, (4.295)); (c) improvement of our understanding of *univer-sality* in high-dimensional percolation, by, for example, showing that bond per-colation on the nearest-neighbor lattice has the critical exponents in Theorem 1.10 for any $d > 6$.

1.2.6   *Oriented percolation*

In this section, we study so-called *oriented* or *directed* percolation. In this case, $\mathbb{G} = (\mathbb{V}, \mathbb{E})$ is given by $\mathbb{V} = \mathbb{Z}^d \times \mathbb{Z}_+$, and $\mathbb{E} = \{((x,n),(y,n+1) : |x - y| = 1\}$, and $\mathbb{G}$ is considered as a *directed* graph, i.e., we remove all bonds independently with probability p and $(x,m) \longrightarrow (y,n)$ is only possible when $m \geq n$. Thus, we can only traverse edges in the direction of increasing last coordinate, and this last coordinate has the convenient interpretation of *time*. We define the *forward cluster* $C(x,n)$ of $(x,n) \in \mathbb{Z}^d \times \mathbb{Z}_+$ to be

$$C(x,n) = \{(z,l) : (x,n) \longrightarrow (z,l)\}, \tag{1.124}$$

so that, in particular, $C(x,n) \subset \mathbb{Z}^d \times \{n, n+1, \ldots\}$.

In some cases, we shall, similarly to the setting in Section 1.2.5, also deal with the *spread-out model*, in which $\mathbb{E} = \{((x,n),(y,n+1) : \|x - y\|_\infty \leq L\}$ for some $L \geq 1$. While one might expect that percolation on oriented lattices is quite similar to percolation on unoriented lattices, this turns out not to be the case:

Theorem 1.11. (Continuity of oriented percolation) *For oriented perco-lation on $\mathbb{Z}^d \times \mathbb{Z}_+$, for $d \geq 1$, there is no infinite cluster at $p = p_c(\mathbb{Z}^d \times \mathbb{Z}_+)$, i.e., $\theta(p_c(\mathbb{Z}^d \times \mathbb{Z}_+)) = 0$.*

Theorem 1.11 was first proved in (Bezuidenhout and Grimmett, 1990) for di-rected percolation, which is a slight variation of the model defined here. The results were extended to the oriented percolation setting described above in (Grimmett and Hiemer, 2002).

The proof of Theorem 1.11 makes use of a block renormalization which was also used in (Barsky, Grimmett and Newman, 1991*a*; Barsky, Grimmett and Newman, 1991*b*) to prove that percolation does not occur in half-spaces. The proof in (Bezuidenhout and Grimmett, 1990) also applies to the *contact process*, a continuous-time adaptation of oriented percolation. The deep relation between the contact process and oriented percolation has proved to be quite useful, and results in one model can typically also be proved for the other.

In (Durrett, 1980), the one-dimensional contact process and oriented perco-lation models were studied, focussing on the growth of the vertices in the cluster of the origin $(0,0) \in \mathbb{Z}^d \times \mathbb{Z}_+$ at time n. These results basically show that when the cluster of the origin is infinite, then the part of it at time n grows linearly in n with a specific growth constant. In (Sakai, 2002), the hyperscaling inequalities for oriented percolation and the contact process were proved, indicating that mean-field critical exponents can only occur for $d > 4$, thus suggesting that the upper critical dimension of oriented percolation equals $d_c = 4$. Indeed, as proved

thereafter, the orientation of the percolation problem implies that mean-field behavior already occurs for $d > 4$:

Theorem 1.12. (Mean-field critical exponents for oriented percolation)
For oriented percolation on $\mathbb{Z}^d \times \mathbb{Z}_+$, for either d sufficiently large in the nearest-neighbor model, or $d > 4$ and L sufficiently large in the spread-out model, $\beta = \gamma = 1$ and $\delta = \Delta = 2$ in the bounded-ratio sense, while $\eta = 0$ in the Fourier-asymptotic sense.

The proof of Theorem 1.12 is given in (Nguyen and Yang, 1993; Nguyen and Yang, 1995; van der Hofstad and Slade, 2003) and follows a similar strategy as the proof of Theorem 1.10 by employing the results in (Aizenman and Barsky, 1987; Aizenman and Newman, 1984) assuming the triangle condition, and using the lace expansion as in (Hara and Slade, 1990). In (van der Hofstad, den Hollander and Slade, 2007; van der Hofstad, den Hollander and Slade, 2007) it is proved that in the spread-out setting, for $d > 4$, the probability that there is an occupied path at criticality connecting $(0,0)$ to $\{(x,n) : x \in \mathbb{Z}^d\}$ is asymptotic to $1/(Bn)(1+o(1))$. This can be seen as a version of the statement that the critical exponent ρ exists and takes the mean-field value $\rho = 1/2$. In (Sakai, 2002), hyperscaling inequalities are shown that imply that critical exponents cannot take their mean-field values when $d < 4$. The main results in (van der Hofstad and Slade, 2003) make a connection between clusters at criticality for the spread-out oriented percolation model above 4 dimensions, and a measure-valued process called *super-Brownian motion*, a model which can be seen as the scaling limit of critical branching random walk. See (Dawson, 1993; Dynkin, 1994; Etheridge, 2000; Le Gall, 1999; Perkins, 2002) for expositions on super-processes.

It would be of interest to prove that *scaling functions* exist for high-dimensional percolation, and to prove further results concerning the critical exponents. For example, we do not know ν, or that γ' exists and $\gamma' = \gamma = 1$ for oriented percolation above four dimensions.

1.2.7 *Percolation on non-amenable graphs.*

We start by defining what an *amenable* graph is. For a finite set of vertices V, we denote its *edge boundary* by

$$\partial_E V = \{(u,v) : u \in V, v \notin V\}. \tag{1.125}$$

The notion of amenability is all about whether the size of $\partial_E V$ is of equal order as that of V, or is much smaller. To formalize this, we denote the *Cheeger constant* of a graph $\mathbb{G}$ by

$$h(\mathbb{G}) = \inf_{V \subset \mathbb{V}: |V| < \infty} \frac{|\partial_E V|}{|V|}. \tag{1.126}$$

A graph is called amenable when $h(\mathbb{G}) = 0$, and is it called non-amenable otherwise. Key examples of amenable graphs are finite-range translation invariant graphs $\mathbb{G}$ with vertex set $\mathbb{Z}^d$, the simplest example of a non-amenable graph is

the regular tree $\mathbb{T}_r$ with $r \geq 3$. For the regular tree $\mathbb{T}_r$ with $r \geq 3$, it is not hard to see that $h(\mathbb{T}_r) = r - 2$. (Benjamini and Schramm, 1996) contains certain preliminary results of percolation on non-amenable graphs, and many open questions, some of which have been settled in the mean time. For example, (Benjamini and Schramm, 1996, Theorem 1) shows that $p_c(\mathbb{G}) \leq 1/(h(\mathbb{G}) + 1)$, so that $p_c(\mathbb{G}) < 1$ for every non-amenable graph.

A related definition of non-amenability can be given in terms of the *spectral radius* of a graph. Let $p_n(u, v)$ be the probability that simple random walk on $\mathbb{G}$ starting at $u \in \mathbb{V}$ is at time n at $v \in \mathbb{V}$. The spectral radius of $\mathbb{G}$ is defined as

$$\rho(\mathbb{G}) = \lim_{n \to \infty} p_{2n}(u, u)^{1/(2n)}. \tag{1.127}$$

By Kesten's Theorem (Kesten, 1959a; Kesten, 1959b), see also (Dodziuk, 1984), when $\mathbb{G}$ has bounded degree, $\rho(\mathbb{G}) < 1$ precisely when $h(\mathbb{G}) > 0$. This exemplifies the fact that there is a close relationship between graph theoretic properties, and the behavior of stochastic processes on the graph. A similar relation between the existence of invariant site percolation and amenability of Cayley graphs is proved in (Benjamini, Lyons, Peres and Schramm, 1999b, Theorem 1.1).

As we have seen for percolation on $\mathbb{Z}^d$ in Theorem 1.5, in the super-critical regime, the infinite cluster is unique. It turns out (see e.g., the discussion following (Kesten, 2002, Theorem 4)) that the uniqueness of the infinite cluster is valid for all amenable graphs. As the proof of Theorem 1.5 shows, there is a close relation between the ratio of the size of the boundary and its volume and the uniqueness of the infinite cluster, which helps to explain the uniqueness for all amenable graphs. On the other hand, for trees, the number of infinite components equals $N = \infty$ a.s. in the supercritical phase, which can be attributed to the fact that it we remove one edge, then a tree falls apart into two infinite graphs which will each have at least one infinite component a.s., so that, in total there will be infinitely many infinite clusters. Thus, this phenomenon is more related to there not being any cycles rather than the boundary being large.

In order to investigate the number of infinite clusters, we define the uniqueness critical value by

$$p_u = p_u(\mathbb{G}) = \inf\{p : \mathbb{P}_p - \text{a.s. there is a unique infinite cluster}\}. \tag{1.128}$$

For the regular tree with $r \geq 3$, $p_u = 1$, while for $\mathbb{Z}^d$, $p_u = p_c$. Below, we shall give examples where $p_c < p_u < 1$.

While the existence of an infinite cluster is clearly an increasing event, the uniqueness of the infinite cluster is not. Therefore, it is a priori not obvious that for all $p > p_u$, the infinite cluster will be unique. This is the main content of the following theorem. In its statement, we will write $N(p)$ for the number of infinite clusters in the coupling of the percolation models for all $p \in [0, 1]$ described above (1.4).

Theorem 1.13. (Uniqueness transition) *For percolation on a connected, quasi-transitive, infinite graph of bounded degree, a.s.,*

$$N(p) = \begin{cases} 0 & for\ p \in [0, p_c); \\ \infty & for\ p \in (p_c, p_u); \\ 1 & for\ p \in (p_u, 1]. \end{cases} \qquad (1.129)$$

The proof of this theorem can be found in (Schonmann, 1999), and related results appeared in (Häggström and Peres, 1999; Häggström, Peres and Schonmann, 1999). Note that, in general, not much is known for the critical cases $p = p_c$ and $p = p_u$.

(Benjamini and Schramm, 1996, Theorem 4) gives a criterion in terms of the spectral radius which implies that $p_c < p_u$. Indeed, it shows that if $\rho(\mathbb{G})pr < 1$, where r is the maximal degree of $\mathbb{G}$, then there are a.s. infinitely many infinite clusters. Thus, if $\rho(\mathbb{G})p_c r < 1$, then $p_c < p_u$. Interesting examples arise by looking at Cartesian products of graphs. Let $\mathbb{G}_1 = (\mathbb{V}_1, \mathbb{E}_1)$ and $\mathbb{G}_2 = (\mathbb{V}_1, \mathbb{E}_1)$ be two graphs, and let $\mathbb{G} = \mathbb{G}_1 \times \mathbb{G}_2$ have vertex set $\mathbb{V} = \mathbb{V}_1 \times \mathbb{V}_2$ and edge set

$$\mathbb{E} = \big\{ \{(u_1, u_2), (v_1, v_2)\} : (u_1, v_1) \in \mathbb{E}_2 \text{ or } (u_2, v_2) \in \mathbb{E}_2 \big\}. \qquad (1.130)$$

Then, clearly, when the maximal degrees of $\mathbb{G}_1$ and $\mathbb{G}_2$, respectively, are denoted by r_1 and r_2 respectively, the maximal degree of $\mathbb{G} = \mathbb{G}_1 \times \mathbb{G}_2$ is $r_1 + r_2$. Also, $p_c(\mathbb{G}_1 \times \mathbb{G}_2) \leq p_c(\mathbb{G}_2)$. (Benjamini and Schramm, 1996, Corollary 1) gives many examples of graphs with $p_c < p_u$ by looking at $\mathbb{G} = \mathbb{G}_1 \times \mathbb{T}_r$, where $\mathbb{G}_1$ is quasi transitive and $\mathbb{T}_r$ is the r-regular tree of degree $r \geq 3$. Indeed, by (Benjamini and Schramm, 1996, Theorem 4), and $p_c(\mathbb{G}_1 \times \mathbb{T}_r) \leq p_c(\mathbb{T}_r) = (r-1)^{-1}$, we obtain that $\rho(\mathbb{G})p_c r < 1$ when r is sufficiently large since $\rho(\mathbb{G}) \to 0$ as $r \to \infty$. A simple example of a graph where $p_c < p_u < 1$ is $\mathbb{Z} \times \mathbb{T}_r$ as proved in (Grimmett and Newman, 1990).

We continue by studying the nature of the phase transition on non-amenable graphs. The first result concerns the continuity of the phase transition for *Cayley graphs of non-amenable groups*. We start by defining what a Cayley graph is. Let Γ be a group, and let $S = \{g_1, \ldots, g_n\} \cup \{g_1^{-1}, \ldots, g_n^{-1}\}$ be a finite set of generators. The Cayley graph $\mathbb{G} = \mathbb{G}(\Gamma)$ has vertex set $\mathbb{V} = \Gamma$, and edge set $\mathbb{E} = \{\{g, h\} : g^{-1}h \in S\}$.

Theorem 1.14. (Continuity on non-amenable Cayley graphs) *For percolation on a Cayley graph of a finitely generated non-amenable group, there is no infinite cluster at $p = p_c(\mathbb{G})$, i.e., $\theta(p_c(\mathbb{G})) = 0$.*

This result was proved in (Benjamini, Lyons, Peres and Schramm, 1999*a*; Benjamini, Lyons, Peres and Schramm, 1999*b*), and generalized earlier work by (Wu, 1993) on $\mathbb{Z} \times \mathbb{T}_r$ with $r \geq 7$. The proof makes use of the *mass-transport technique* to a clever choice of the mass-transport function.

We complete this section by describing some results on critical exponents. For this, we need to introduce the notions of *planar* and *unimodular* graphs. A graph

is called *planar* when it can be embedded into $\mathbb{R}^2$ with vertices being represented by points in $\mathbb{R}^2$ and edges by lines between the respective vertices such that the edges only intersect at their end-points. For $x \in \mathbb{V}$, let the *stabilizer of x $S(x)$* be the set of automorphisms of $\mathbb{G}$ that keep x fixed, i.e., $S(x) = \{\gamma : \gamma(x) = x\}$. The graph $\mathbb{G}$ is called *unimodular* if $|\{\gamma(y) : \gamma \in S(x)\}| = |\{\gamma(x) : \gamma \in S(y)\}|$ for every $x, y \in \mathbb{V}$. Unimodularity turns out to be an extremely useful notion, particularly since it turns out to be imply the so-called *mass-transport principle* (MTP). Indeed, let $\mathbb{G}$ be a transitive unimodular graph. We say that $f \colon \mathbb{V} \times \mathbb{V} \to [0, \infty)$ is diagonally invariant under the automorphisms of $\mathbb{G}$ when $f(x, y) = f(\gamma(x), \gamma(y))$ for every automorphism $\gamma \colon \mathbb{V} \to \mathbb{V}$. Then, the MTP states that

$$\sum_{y \in \mathbb{V}} f(x, y) = \sum_{y \in \mathbb{V}} f(y, x). \tag{1.131}$$

For a proof, see e.g. (Benjamini, Lyons, Peres and Schramm, 1999a). In most cases, the MTP in (1.131) is used in the following way. We take $\phi \colon \mathbb{V} \times \mathbb{V} \times 2^{\mathbb{G}} \to [0, \infty)$ such that $\phi(x, y, \omega) = \phi(\gamma(x), \gamma(y), \gamma \circ \omega)$, for any configuration ω and where $\gamma \circ \omega$ is the configuration $(\gamma \circ \omega)(x) = \omega(\gamma(x))$. We interpreted $\phi(x, y, \omega)$ as the mass which x sends to y in the configuration κ. Then, we take $f(x, y) = \mathbb{E}_p[\phi(x, y, \omega)]$, and the MTP implies that the mass sent out by x is equal to the amount of mass x receives, which explains the name mass transport principle. For non-unimodular graphs, an adaptation of (1.131) holds, where the left hand side is multiplied by $w(x)$ and the right hand side by $w(y)$, where $w(x) = |S_x o|/|S_o x|$, but this relation is not as powerful. All amenable graphs are unimodular, an example of a non-unimodular graph is the so-called *grandmother graph*, which is obtained by adding a connection between any vertex of the tree to its grandmother (i.e., the unique vertex which is two steps close to the root than the vertex itself). See (Timár, 2006) for a wealth of related examples of non-unimodular graphs.

Finally, the *number of ends* of a graph $\mathbb{G}$ is

$$\mathcal{E}(\mathbb{G}) = \sup_{S \subset \mathbb{V} : |S| < \infty} \{\text{number of infinite connected components of } \mathbb{G} \setminus S\}. \tag{1.132}$$

Then, (Schonmann, 2001; Schonmann, 2002) prove that percolation has mean-field critical exponents in the following cases:

Theorem 1.15. (Mean-field critical exponent on non-amenable graphs)
For percolation on a locally finite, connected, transitive, non-amenable graph $\mathbb{G}$, $\beta = \gamma = 1, \delta = \Delta = 2$ in the bounded-ratio sense, in the following cases:

1. *Graphs $\mathbb{G}$ for which $h(\mathbb{G}) > (\sqrt{2r^2 - 1} - 1)/2$, where r is the degree of the graph;*
2. *Graphs $\mathbb{G}$ which are planar and have one end;*
3. *Graphs $\mathbb{G}$ which are unimodular and have finitely many ends.*

Since percolation in high-dimensions is known to have mean-field critical exponents, which are the critical exponents on the tree, one would expect that,

in great generality, percolation on non-amenable graphs do so too. Theorem 1.15 is a step into the direction of proving this belief, but a general result to this extent is still missing. It would be of interest to investigate this form of universality in more detail, as well as the existence of scaling functions for general non-amenable graphs.

An application of percolation on non-amenable graphs to image analysis can be found in (Kendall and Wilson, 2003).

1.2.8 *Continuum percolation.*

Continuum percolation is a close brother of percolation, where instead of working on a lattice, we work in the continuum. While there are many possible models, we shall restrict to the simplest version, the so-called *Boolean model*. For details on the model, see the monograph (Meester and Roy, 1996), or (Penrose, 2003, Section 9.6), and the references therein. In the Boolean model, we start with a Poisson point process PPP of a given intensity $\lambda > 0$, and each point $x \in \text{PPP}$ is assigned a radius. The radii of the different vertices are independent random variables, an important special case is when all radii are fixed. For $x \in \text{PPP}$, let R_x be its corresponding radius. We create an *occupied region* by looking at all vertices contained in the union over x of the balls of radius R_x, i.e., the occupied region is given by

$$O = \bigcup_{x \in \text{PPP}} B(x, R_x), \tag{1.133}$$

where $B(x, r)$ is the ball of radius r centered at $x \in \mathbb{R}^d$. We denote by $C(x)$ the connected part of O that contains x, we let $\theta(\lambda)$ be the probability that $C(x)$ is unbounded, and we define

$$\chi(\lambda) = \mathbb{E}_\lambda[|C(o)|], \tag{1.134}$$

where, for a region $C \subset \mathbb{R}^d$, we write $|C|$ for the Lebesgue measure of C. The function $\theta(\lambda)$ plays a similar role in continuum percolation as the percolation function in (discrete) percolation, while $\chi(\lambda)$ plays a similar role as the expected cluster size. With the above definitions at hand, continuum percolation as a model is quite similar to discrete percolation described above, and most of the results discussed above for percolation on $\mathbb{Z}^d$ carry over to continuum percolation on $\mathbb{R}^d$. In fact, many proofs make crucial use of the discrete result, by an appropriate discretization procedure. However, by varying the random radii, certain phenomena arise that are not present in bond percolation, such as the possibility that cluster have, a.s., finitely many Poisson points, yet the expected number of Poisson points is infinite.

Continuum percolation is an important model from the point of applications, as it can be seen as a simple model for a communication network where transmitters have a finite transmitting distance. When the points in the Poisson point process PPP mark the locations of sensors in an ad hoc network, and the radii correspond to their transmission distance, then the fact that $|C(x)| = \infty$ corresponds to the fact that the sensors can jointly transmit over an unbounded

domain. In this light, continuum percolation is becoming an important tool in the investigation of various telecommunication networks. For examples of the application of continuum percolation ideas to communications, see e.g. (Dousse, Franceschetti, Macris, Meester and Thiran, 2006), (Baccelli and Błaszczyszyn, 2001) or the recent book (Franceschetti and Meester, 2008). In (Grossglauser and Thiran, 2006), you can find a discussion on the relation between percolation problems and the engineering of wireless telecommunication models, touching upon navigability of networks, and the relation between connectivity and capacity.

Random geometric graphs are obtained by taking only a bounded domain and performing a similar strategy as described above on it. More precisely, random geometric graphs are characterized by two parameters and can be constructed as follows. We consider the square $[0,1]^2$ and we put n points in it, uniformly at random. We can think of these n points as being a population of adhoc or peer-to-peer network users, who wish to communicate data to each other. Then we connect pairs of points within distance r for some appropriate r (possibly depending on n). Questions of interest are how large the radius should be as a function of the intensity in order for all points to be connected, and what the minimal degree of the vertices in this graph is. The connectivity of the graph is essential for each of the network users to be able to transmit data to each other. Clearly, this model is only a mere caricature of reality, and a better understanding of the real-world properties of networks is necessary. However, this caricature model does already shed light on some of the basic problems in geometric wireless networks. We refer to (Penrose, 2003) for detailed results on this model, as well as on the literature.

1.3   Random graphs

1.3.1   *Motivation*

Real-world networks.   In the past decade, many examples have been found of real-world networks that are *small worlds*, *scale-free* and *highly clustered*. We shall start by discussing these notions one by one.

The small-world phenomenon states that distances in a network are relatively small. This is related to the well-known 'Six degrees of separation' paradigm in social networks, which conjectures that any pair of individuals in the world can be connected by a chain of persons knowing each other on a first name basis, the chain consisting of at most 6 intermediary individuals. This paradigm has attracted considerable attention, see e.g. (Newman, Watts and Barabási, 2006) for a historical account, including the original papers. The notion of real networks being small worlds is inherently a bit imprecise, as we do not define what 'relatively' small is. For the time being one can have in mind that distances in many real networks are at most 10-20, later, we shall give a more precise definition what it means for a process of random graphs to be a small world process.

The scale-free phenomenon states that the degree sequences in a network satisfies a power law. In more detail, the degree sequence of a network of size n, which we denote by $\{p_k^{(n)}\}_{k=0}^\infty$, is given by

$$p_k^{(n)} = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{d_i = k\}}, \tag{1.135}$$

where d_i is the degree of vertex i in the network, i.e., the number of neighbors of vertex i. Then, a real network is scale free when $p_k^{(n)}$ is approximately proportional to a power law with a certain exponent $\tau \in [1, \infty)$, i.e., for k sufficiently large, we have that there is a constant $C > 0$ such that

$$p_k^{(n)} \approx C k^{-\tau}. \tag{1.136}$$

Naturally, also the notion of a real network being scale free is inherently somewhat vague. For example, how precise must the approximation in (1.136) be? Equation (1.136) can only be valid up to a certain point, since, for a simple graph of size n (i.e., a graph without self-loops and multiple edges), the maximal degree is equal to $n - 1$, so that the left hand side of (1.136) equals 0 for $k \geq n$, while the right hand side remains positive for *all* k. In practice, (1.136) is verified by taking the logarithm on both sides and noticing that $\log p_k^{(n)}$ is close to linear in $\log k$:

$$\log p_k^{(n)} \approx \log C - \tau \log k, \tag{1.137}$$

i.e., a loglog plot of the degree sequence should be close to linear, and the slope is given by $-\tau$, where τ denotes the power-law exponent.

A network is highly clustered when two typical neighbors of an arbitrary vertex are more likely to be connected to each other as well than an arbitrary pair of vertices, i.e., many wedges are in fact closed to become triangles. When we draw two different vertices uniformly at random from a graph of size n, the probability that there is an edge between the drawn vertices is equal to

$$\frac{2E}{n(n-1)}, \tag{1.138}$$

where E is the number of edges. Note that $2E/n = \bar{d}$, where

$$\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i \tag{1.139}$$

is the average degree of all vertices in the network. Thus, we see that $\frac{2E}{n(n-1)} = \bar{d}/(n-1)$. In most real-world networks, the average degree is much smaller than the size of the network, so that the probability that two uniformly drawn vertices

share an edge is rather small. The *clustering coefficient* $C_{\mathbb{G}}$ of a graph $\mathbb{G} = (\mathbb{V}, \mathbb{E})$ is defined by

$$C_{\mathbb{G}} = \frac{\sum_{i,j,k \in \mathbb{V}} \mathbb{1}_{\{ij,ik,jk \in \mathbb{E}\}}}{\sum_{i,j,k \in \mathbb{V}} \mathbb{1}_{\{ij,ik \in \mathbb{E}\}}}, \tag{1.140}$$

i.e., the proportion of wedges that forms a triangle. A network is highly clustered when $C_{\mathbb{G}}$ is much larger than $\frac{2E}{n(n-1)}$. Again, this notion is inherently imprecise, as we do not define *how much* larger $C_{\mathbb{G}}$ needs to be. The reason why many networks are highly clustered is that there often is a certain group structure. For example, in a collaboration network, if a mathematician has published with another mathematician, then they are likely to be in the same community. Thus, when a mathematician has published with two other mathematicians, then they are likely to all be in the same community, which increases the likelihood that the two mathematicians have also published together. Similar effects play a role in the World-Wide Web and in social networks.

As explained above, all these notions are merely empirical, and we shall give a proper mathematical definition before we introduce the mathematical random graph models that we shall consider in this section, and which are aimed at describing real networks. The aim of this section is not to define these notions precisely for empirical networks, but instead to define these notions precisely for random graph models for them. See (Albert and Barabási, 2002; Dorogovtsev and Mendes, 2002; Newman, 2003) for reviews on complex networks, and (Barabási, 2002) for a more expository account. We do notice that these real-world complex networks are not at all like classical random graphs (see (Alon and Spencer, 2000; Bollobás, 2001; Janson, Łuczak and Rucinski, 2000) and the references therein), particularly since the classical models do not have power-law degrees. As a result, the empirical findings of real-world networks have ignited enormous research on adaptations of the classical random graph that do have power-law degree sequences. We shall survey some of these results, and we shall start by defining the notions of small-world, scale-free and highly-clustered random graphs in a precise mathematical way.

Small-world, scale-free and highly-clustered random graph processes. As described in the above motivation, many real-world complex networks are large. Many of them, such as the World-Wide Web and collaboration networks, grow in size as time proceeds. Therefore, it is reasonable to consider graphs of growing size, and to define the notions of scale-free, small-world and highly-clustered random graphs as a limiting statement when the size of the random graphs tends to infinity. This naturally leads us to study graph sequences. We shall denote a sequence of random graphs by $\{\mathbb{G}_n\}_{n=1}^{\infty}$, where n denotes the size of the graph $\mathbb{G}_n$, i.e., the number of vertices in $\mathbb{G}_n$.

Denote the proportion of vertices with degree k by $P_k^{(n)}$, i.e.,

$$P_k^{(n)} = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{D_i^{(n)} = k\}}, \tag{1.141}$$

where $D_i^{(n)}$ denotes the degree of vertex $i \in \{1, \ldots, n\}$ in the graph $\mathbb{G}_n$, and recall that the degree sequence of $\mathbb{G}_n$ is given by $\{P_k^{(n)}\}_{k=0}^{\infty}$. We use capital letters in our notation to indicate that we are dealing with *random variables*, due to the fact that $\mathbb{G}_n$ is a random graph. This explains why there are capitals in (1.141), but not in (1.135). Now we are ready to define what it means for a random graph process $\{\mathbb{G}_n\}_{n=1}^{\infty}$ to be scale free.

We call a random graph process $\{\mathbb{G}_n\}_{n=1}^{\infty}$ *sparse* when

$$\lim_{n \to \infty} P_k^{(n)} = p_k, \tag{1.142}$$

for some *deterministic* limiting probability distribution $\{p_k\}_{k=0}^{\infty}$. Since the limit p_k in (1.142) is deterministic, the convergence in (1.142) can equivalently be taken as convergence in probability or in distribution. Also, since $\{p_k\}_{k=0}^{\infty}$ sums up to one, for large n, most of the vertices have a bounded degree, which explains the phrase sparse random graphs.

We further call a random graph process $\{\mathbb{G}_n\}_{n=1}^{\infty}$ *scale free with exponent τ* when it is sparse and when

$$\lim_{k \to \infty} \frac{\log p_k}{\log (1/k)} = \tau \tag{1.143}$$

exists. Thus, for a scale-free random graph process its degree sequence converges to a limiting probability distribution as in (1.142), and the limiting distribution has asymptotic power-law tails described in (1.143). This gives a precise mathematical meaning to a random graph process being scale free. In some cases, the definition in (1.143) is a bit too restrictive, particularly when the probability mass function $k \mapsto p_k$ is not very smooth. Instead, we can also replace it by

$$\lim_{k \to \infty} \frac{\log [1 - F(k)]}{\log (1/k)} = \tau - 1, \tag{1.144}$$

where $F(x) = \sum_{y \leq x} p_y$ denotes the distribution function corresponding to the probability mass function $\{p_k\}_{k=0}^{\infty}$. In particular models, we shall use the version that is most appropriate in the setting under consideration.

We say that a graph process $\{\mathbb{G}_n\}_{n=1}^{\infty}$ is *highly clustered* when

$$\lim_{n \to \infty} C_{\mathbb{G}_n} = C_{\mathbb{G}_\infty} > 0. \tag{1.145}$$

We finally define what it means for a graph process $\{\mathbb{G}_n\}_{n=1}^{\infty}$ to be a small world. Intuitively, a small world should have distances that are much smaller than those in a lattice or torus. When we consider the nearest-neighbor torus $\mathbb{T}_{r,d}$, then, and we draw two vertices uniformly at random, their distance will be of the order r. Denote the size of the torus by $n = (2r + 1)^d$, then the typical distance between two uniformly chosen vertices is of the order $n^{1/d}$, so that it grows as a positive power of n. We shall be dealing with random graph processes $\{\mathbb{G}_n\}_{n=1}^{\infty}$ for which $\mathbb{G}_n$ is not necessarily connected. Let H_n denote the distance between

two uniformly chosen *connected* vertices, i.e., we pick a pair of vertices uniformly
at random from all pairs of connected vertices, and we let H_n denote the graph
distance between these two vertices. We shall call H_n the *typical distance of* $\mathbb{G}_n$.
Then, we say that a random graph process $\{\mathbb{G}_n\}_{n=1}^{\infty}$ is a *small world* when there
exists a constant K such that

$$\lim_{n\to\infty} \mathbb{P}(H_n \leq K \log n) = 1. \qquad (1.146)$$

Note that, for a graph with a *bounded* degree $d_{\max}$, the typical distance is *at least*
$(1-\varepsilon)\log n/\log d_{\max}$, with high probability, so that a random graph process with
bounded degree is a small world precisely when the order of the typical distance
is optimal.

For a graph $\mathbb{G}_n$, let $\mathrm{diam}(\mathbb{G}_n)$ denote the *diameter* of $\mathbb{G}_n$, i.e., the maximal
graph distance between any pair of connected vertices. Then, we could also have
chosen to replace H_n in (1.146) by $\mathrm{diam}(\mathbb{G}_n)$. However, the diameter of a graph
is a rather sensitive object which can easily be changed by making small changes
to a graph in such a way that the scale-free nature and the typical distance H_n
do not change. For example, by adding a sequence of m vertices in a line, which
are not connected to any other vertex, the diameter of the graph becomes at
least m, whereas, if m is much smaller than n, H_n is not changed very much.
This explain why we have a preference to work with the typical distance H_n
rather than with the diameter $\mathrm{diam}(\mathbb{G}_n)$.

In some models, we shall see that typical distances can be even much smaller
than $\log n$, and this is sometimes called an *ultra-small world*. More precisely, we
say that a random graph process $\{\mathbb{G}_n\}_{n=1}^{\infty}$ is an *ultra-small world* when there
exists a constant K such that

$$\lim_{n\to\infty} \mathbb{P}(H_n \leq K \log\log n) = 1. \qquad (1.147)$$

There are many models for which (1.147) is satisfied, but for which at the same
time $\mathrm{diam}(\mathbb{G}_n)/\log n$ converges in probability to a positive limit. This once more
explains our preference to work with the typical graph distance H_n.

We have given precise mathematical definitions for the notions of random
graphs being scale free, highly clustered and small worlds. This has not been
done in the literature so far, and our definitions are based upon a summary of
the relevant results proved for random graph models. We believe it to be a sound
step forward to make the connection between the theory of random graphs and
the empirical findings on real-life networks.

The remainder of this section is organised as follows. In Section 1.3.2, we
study three models for random graphs without geometry that can have rather
general degree sequences, namely, inhomogeneous random graphs, the configu-
ration model and preferential attachment models. We discuss results concerning
the phase transitions and distances in such models. In Section 1.3.5, we shall
discuss random graphs with geometry. The results in Section 1.3.5 are closely

related to percolation questions discussed in Section 1.2. The main distinction between the random networks discussed in this section and the percolation networks discussed in Section 1.2 is that the random graphs discussed here shall be *finite*, while the networks in Section 1.2 are all *infinite*. This raises interesting new questions, such as how the phase transition can be defined (a cluster can never be infinite), to what extent the phase transition is *unique*, and what the distance between two uniformly chosen vertices is.

1.3.2 *Models without geometry*

Extensive discussions of scale-free random graphs are given in (Chung and Lu, 2006a; Durrett, 2007), monographs on classical random graphs are (Bollobás, 2001; Janson, Łuczak and Rucinski, 2000). We now discuss three particular examples of random graphs with power-law degree sequences, namely, the inhomogeneous random graph, the configuration model, and preferential attachment models.

(a) Inhomogeneous random graphs. The simplest imaginable random graph is the so-called *Erdős-Rényi random graph*, which consists of n vertices and each of the $n(n-1)/2$ edges is present or occupied with probability p, independently of the occupation status of the other edges. Denote the resulting graph by $G(n,p)$. This model was introduced by (Gilbert, 1959), while (Erdős and Rényi, 1959) introduced a model where a fixed number of edges is chosen uniformly at random and without replacement. The two models are quite comparable, and most asymptotic results in one of the two models can easily be transferred to asymptotic results in the other. A model with a fixed number of edges being chosen *with* replacement, so that possibly multiple edges between vertices arise, can be found in (Austin, Fagen, Penney and Riordan, 1959). The name Erdős-Rényi random graph is given to this class of models due to the fact that the first rigorous results were derived in the seminal paper (Erdős and Rényi, 1960), which can be seen as having founded the field of random graphs, and which has inspired research questions for decades to follow (see also the books (Bollobás, 2001; Janson, Łuczak and Rucinski, 2000)). The above model with independent edges can be viewed as *percolation on the complete graph*, the main difference to the theory in Section 1.2 being that the graph is *finite*.

One of the charming features of the Erdős-Rényi random graph is the fact that its vertices are completely exchangeable. For example, every vertex $v \in [n]$, where we write $[n] = \{1, \ldots, n\}$ for the vertex set, has a degree that is distributed as a binomial random variable with parameters $n - 1$ and p. Thus, when $np \to \infty$, the average degree tends to infinity, while if $np = \lambda$, for some λ, the average degree remains uniformly bounded. When $p = \lambda/n$, the average degree of each vertex is roughly equal to λ, and the degree is a vertex converges in distribution to a Poisson random variable with parameter λ. It can be seen that also the proportion of vertices with degree k, as defined in (1.141), converges in probability to the Poisson probability mass function $p_k = e^{-\lambda}\lambda^k/k!$. Thus, the Erdős-Rényi random graph process is *sparse*.

Note that the tails of a Poisson distribution are quite thin, even subexponentially thin. As a result, the Erdős-Rényi random graph process is *not* scale free. This problem can be overcome by stepping away from the assumption that the edge probabilities are *equal*, instead taking them unequal. This is the celebrated *inhomogeneous random graph* (IRG), about which the seminal paper (Bollobás, Janson and Riordan, 2007) proves substantial results in full generality. See also the references in (Bollobás, Janson and Riordan, 2007) for several examples which have been studied in the literature, and which they generalize. We shall not go into the precise definition of the model in (Bollobás, Janson and Riordan, 2007), but rather look at some simpler examples which already allow for general degree sequences.

To give the general setting, we let $G(n, \mathbf{p})$ denote a general inhomogeneous random graph, where $\mathbf{p} = \{p_{ij}\}_{1 \leq i < j \leq n}$ is such that p_{ij} is the probability that the edge $ij = (i, j)$ is occupied, and where different edges are independent. The Erdős-Rényi random graph is retrieved when taking $p_{ij} = p$ for all ij. We now generalize the definition in such a way that power-law degree sequences can be obtained. We assign *weights* $\{w_i\}_{i=1}^n$ to the vertices, w_i being the weight of vertex i. Let $l_n = \sum_{i=1}^n w_i$ be the total weight. Then, we can take

$$p_{ij} = 1 - e^{-w_i w_j / l_n}. \tag{1.148}$$

In this way, we retrieve the *Poisson random graph* (Norros and Reittu, 2006). Alternatively, for $p_{ij} = \max\{w_i w_j / l_n, 1\}$ we retrieve the *random graph with given expected degree* as studied in detail by Chung and Lu (Chung and Lu, 2002a; Chung and Lu, 2002b; Chung and Lu, 2003; Chung and Lu, 2006b; Chung and Lu, 2006a; Chung, Lu and Vu, 2004). Note that, in this model, if we assume that $\max_{i=1}^n w_i^2 < n$, and if we allow for a single self-loop at vertex i with probability w_i^2 / l_n, the expected degree of vertex i is precisely equal to w_i, which explains the name of this model. A final example occurs when we take $p_{ij} = w_i w_j / (l_n + w_i w_j)$, which is called the *generalized random graph* (Britton, Deijfen and Martin-Löf, 2006). In (Janson, 2008), conditions are given as to when two inhomogeneous random graph processes $\{G(n, \mathbf{p})\}_{n=1}^\infty$ and $\{G(n, \mathbf{p}')\}_{n=1}^\infty$ are equivalent, i.e., when events have asymptotically the same probability for $G(n, \mathbf{p})$ and $G(n, \mathbf{p}')$ as $n \to \infty$.

Since the expected degree of vertex i is close to w_i, in order to obtain a specified degree sequence in the graph, we need to pick the weights $\{w_i\}_{i=1}^n$ appropriately. In the sequel, we shall take

$$w_i = w_i^{(n)} = [1 - F]^{-1}(i/n), \qquad i \in [n], \tag{1.149}$$

where $[1 - F]^{-1}$ is the generalized inverse function of the survival function $x \mapsto 1 - F(x)$ given by

$$(1 - F)^{-1}(u) = \inf\{s : [1 - F](s) \leq u\}, \qquad u \in (0, 1). \tag{1.150}$$

We call $G(n, \mathbf{p})$, with $\mathbf{p}$ as in (1.148) and $\{w_i\}_{i=1}^n$ as in (1.149) the *rank-1 inhomogeneous random graph with deterministic weights according to F*. In this

case, the degrees of $G(n, \mathbf{p})$ can be seen to have a mixed Poisson distribution with mixing distribution F (Britton, Deijfen and Martin-Löf, 2006), i.e., the asymptotic probability that a uniform vertex has degree k is

$$p_k = \mathbb{E}[\frac{W^k}{k!}e^{-W}], \qquad (1.151)$$

where W has distribution function F (see, for example, (Britton, Deijfen and Martin-Löf, 2006, Theorem 3.1) or (van der Hofstad, 2008, Chapter 6)). Thus, the rank-1 inhomogeneous random graph is sparse. It is not hard to see, using concentration techniques on a Poisson random variable, that the mixed-Poisson distribution in (1.151) satisfies (1.143) when W is a continuous random variable with density f_W satisfying

$$\lim_{w \to \infty} \frac{\log f_W(w)}{\log (1/w)} = \tau. \qquad (1.152)$$

The condition in (1.152) is sufficient, but not necessary. However, it does illustrate that there is a close relation between the tails of a mixed Poisson random variable and that of its mixing distribution.

Instead of taking the weights deterministic as in (1.149), we can also take $w_i = W_i$, where $\{W_i\}_{i=1}^n$ is an i.i.d. sequence of random variables with distribution function F. We call $G(n, \mathbf{p})$ with $\mathbf{p}$ as in (1.148) and $\{W_i\}_{i=1}^n$ an i.i.d. sequence of random variables with distribution F, the *rank-1 inhomogeneous random graph with random weights according to F*. The advantage of random degrees is that the vertices in the resulting graph are exchangeable, the disadvantage is that the edges are no longer independent, their dependence being moderated by their random weights.

IRGs give a flexible class of models for random graphs with flexible degree sequences of mixed Poisson form.

(b) The configuration model. The class of IRGs is a flexible class of random graphs, allowing for rather general degree sequences. One slight disadvantage of the model is the fact that one cannot specify the degree sequence more precisely. For example, one cannot have fixed degrees (so-called random regular graphs), or force the degrees to be at least 1 (which might be appropriate when using the model for certain real networks). An alternative model, allowing for more general degree sequences, is the configuration model (CM). We shall first introduce the model.

Fix an integer n. Consider a sequence of degrees $d_1, d_2, \ldots, d_n$. We will construct an undirected graph with n vertices where vertex j has degree d_j. Obviously, we will assume that $l_n = \sum_{j=1}^n d_j$ is even, since in any graph, the total degree is twice the number of edges.

To construct the graph, we have n separate vertices and incident to vertex j, we take d_j stubs. All stubs need to be connected to another stub to build the graph. The stubs are numbered in an arbitrary order from 1 to l_n. We start by

connecting at random the first stub with one of the $l_n - 1$ remaining stubs. Once paired, two stubs form a single edge of the graph. We continue the procedure of randomly choosing and pairing the stubs without replacement, until all stubs are connected. Unfortunately, vertices having self-loops may occur. However, self-loops are scarce when $n \to \infty$. See (Janson, 2006) for the best results on the simplicity of the CM. In general, it turns out that the vector of the number of cycles of all lengths (i.e., the vector consisting of the number of self-loops, the number of multiple edges, the number of triangles, the number of squares, etc) converges in distribution to an independent sequence of Poisson random variables, when $\nu_n = \frac{1}{l_n} \sum_{j=1}^{n} d_j(d_j - 1)$ converges to a certain $\nu < \infty$, and $\max_i d_i = o(\sqrt{n})$. See also (Bollobás, 2001, Section 2.4). The CM with fixed degrees has a long history. It was introduced in (Bender and Canfield, 1978) to study uniform random graphs with a given degree sequence (see also (Bollobás, 2001, Section 2.4)). One specific example is to take the degrees all equal, in which case we speak of a *random regular graph.*

Similarly to the rank-1 case of IRG described above, there are two cases which have been studied for the CM, depending on whether the degrees are deterministic or random. In the case of deterministic degrees, a possible choice to obtain a degree sequence with distribution function F is to take F the distribution function of an integer random variable, and to take the number of vertices with degree k to be equal to $\lceil nF(k) \rceil - \lceil nF(k-1) \rceil$. We shall call this the *configuration model with deterministic degrees according to F*. For random degrees, we can take the degrees $\{D_i\}_{i=1}^{n}$ to be i.i.d. random variables with distribution function F, which we shall call the *configuration model with random degrees according to F*. In the latter case, unless D_1 is *even* with probability 1, the sum $L_n = \sum_{i=1}^{n} D_i$ will be *odd* with probability exponentially close to $1/2$. To avoid this problem, we can increase the degree of vertex n by 1, so that the degree of vertex n equals $D_n + \mathbb{1}_{\{D_n \text{ is odd}\}}$. This hardly makes any difference to the properties of the graph under consideration, and we will ignore this effect in the sequel.

In (Britton, Deijfen and Martin-Löf, 2006, Theorem 2.1), it is shown that for the CM with random degrees according to F, where F is a distribution function of a random variable with finite $(1 + \varepsilon)^{\text{th}}$ moment, if we erase all self-loops and contract all multiple edges to a single edge, the resulting graph is simple and sparse (recall (1.142)), where $p_k = F(k) - F(k-1)$. Thus, multiple edges and self-loops only form a small proportion of the total number of edges. Thus, in order to obtain a scale-free random graph process, we need to take $p_k = F(k) - F(k-1)$ to satisfy (1.143). In (Britton, Deijfen and Martin-Löf, 2006), also other constructions to create a CM with a prescribed degree sequence are analyzed.

(c) Preferential attachment models.   While inhomogeneous random graphs and the configuration model can have power-law degree sequences when the edge probabilities or degrees are chosen appropriately, they do not explain *why* many

real-world networks are scale free. A possible explanation was given by Barabási and Albert (Barabási and Albert, 1999) by a phenomenon called *preferential attachment*. Preferential attachment models the growth of the network in such a way that new vertices are more likely to add their edges to already present vertices with a high degree. For example, in a social network, a newcomer is more likely to get acquainted to a person who is socially active, and, therefore, is likely to already possess a high degree. Thus, large degrees are likely to become even larger, which explains why this model is sometimes called the *Rich-get-Richer model*.

Interestingly, in certain cases, preferential attachment models (PAMs) have power-law degree sequences, and, therefore, preferential attachment offers a convincing explanation why many real-world networks have power-law degree sequences. As a result, many papers appeared that study such models. See e.g. (Aiello, Chung and Lu, 2002; Bollobás, Borgs, Chayes and Riordan, 2003; Bollobás and Riordan, 2003*a*; Bollobás and Riordan, 2003*b*; Bollobás and Riordan, 2004*a*; Bollobás and Riordan, 2004*b*; Bollobás, Riordan, Spencer and Tusnády, 2001; Cooper and Frieze, 2003) and the references therein. The literature primarily focusses on three main questions. The first is to prove that such random graphs are indeed scale free (Aiello, Chung and Lu, 2002; Bollobás, Borgs, Chayes and Riordan, 2003; Bollobás and Riordan, 2003*a*; Bollobás and Riordan, 2003*b*; Bollobás, Riordan, Spencer and Tusnády, 2001; Cooper and Frieze, 2003). The second is to show that the resulting models are small worlds by investigating the distances in them. See for example (Bollobás and Riordan, 2004*b*) for a result on the diameter. In non-rigorous work, it is often suggested that many of the scale-free models, such as the CM or IRGs, the models in (Bollobás, Janson and Riordan, 2007) and the PAMs, have similar properties for their distances. Distances in the CM have been shown to depend on the number of finite moments of the degree distribution. The natural question is therefore whether the same applies to PAMs. A third key question for PAMs is their vulnerability, for example to deliberate attack (Bollobás and Riordan, 2003*b*) or to the spread of a disease (Berger, Bollobás, Borgs, Chayes and Riordan, 2003).

We now formulate the preferential attachment model, which is an extension of the Barabási-Albert model, formulated rigorously in (Bollobás, Riordan, Spencer and Tusnády, 2001), by constructing a graph process $\{G_m(n)\}_{n=1}^{\infty}$. We start by formulating the model for $m = 1$, for which we start with $G_1(1)$ consisting of a single vertex with a single self-loop. We denote the vertices of the graph $G_m(n)$ by $v_1^{(m)}, v_2^{(m)}, \ldots$. We denote the degree of vertex $v_i^{(m)}$ by $D_i^{(n)}$, where a self-loop increases the degree by 2.

Then, for $m = 1$, and conditionally on $G_1(n)$, the growth rule to obtain $G_1(n+1)$ is as follows. We add a single vertex $n + 1$ having a single edge. The other end of this edge is equal to $i = n + 1$ with probability proportional to $1 + \delta$, and to $i \in [n]$ with probability proportional to $D_i^{(n)} + \delta$, where $\delta \geq -1$ is a parameter of the model. Thus,

$$\mathbb{P}\big(v_{n+1}^{(1)} \to v_i^{(1)}\big|G_1(n)\big) = \begin{cases} \frac{1+\delta}{n(2+\delta)+(1+\delta)}, & \text{for } i = n+1, \\ \frac{D_i^{(n)}+\delta}{n(2+\delta)+(1+\delta)}, & \text{for } i = 1, 2, \dots, n. \end{cases} \qquad (1.153)$$

The model with integer $m > 1$, is defined in terms of the model for $m = 1$ as follows. We take $\delta \geq -m$, and then start with $G_1(mn)$, with $\delta' = \delta/m \geq -1$ by identifying the vertices $v_1^{(1)}, v_2^{(1)}, \dots, v_m^{(1)}$ in $G_1(mn)$ to be vertex $v_1^{(m)}$ in $G_m(n)$, and for $1 < j \leq n$, the vertices $v_{(j-1)m+1}^{(1)}, \dots, v_{jm}^{(1)}$ in $G_1(mn)$ to be vertex $v_j^{(m)}$ in $G_m(n)$; in particular the degree $D_j(n)$ of vertex $v_j^{(m)}$ in $G_m(n)$ is equal to the sum of the degrees of the vertices $v_{(j-1)m+1}^{(1)}, \dots, v_{jm}^{(1)}$ in $G_1(mn)$. This defines the model for integer $m \geq 1$. Observe that the range of δ is $[-m, \infty)$.

The resulting graph $G_m(n)$ has precisely mn edges and n vertices at time n, but is not necessarily connected. For $\delta = 0$, we obtain the original model studied in (Bollobás, Riordan, Spencer and Tusnády, 2001), and further studied in (Bollobás and Riordan, 2003b; Bollobás and Riordan, 2004a; Bollobás and Riordan, 2004b). The extension to $\delta \neq 0$ is crucial in our setting, as we shall explain in more detail below.

There are several related ways in which we can define the model. For example, we disallow self-loops when $m = 1$ by setting the probability that $v_{n+1}^{(1)}$ connects to $v_{n+1}^{(1)}$ to be 0 in (1.153), and construct the model for $m > 1$ as in the paragraph above. Alternatively, we can let the m edges incident to vertex n to be attached *independently* of each other (in particular, in this case, vertex $v_{n+1}^{(m)}$ cannot to connect to itself, so that the graph is connected). For many of the results, this precise choice is irrelevant, and we shall stick to the model in (1.153). The last two versions have the nice feature that they lead to *connected* random graphs.

We continue to discuss the degree sequence of the above preferential attachment model. Recall (1.141) for the definition of the degree sequence. Much of the available literature on PAMs centers around the proof that the asymptotic degree sequence obeys a power law, where the exponent τ depends in a sensitive way on the parameters of the model. Thus, the PAM is scale free. For the PAM considered here, the power-law exponent equals $\tau = 3 + \delta/m$, so that it can take any value $\tau \in (2, \infty)$ by adjusting the parameter $\delta > -m$. It is here that we rely on the choice of the model in (1.153). A form of bias in growing networks towards vertices with higher degree is, from a practical point of view, quite likely to be present in various real networks, but it is unclear why the PA scheme should be *affine* as in (1.153). However, only affine PA schemes give rise to power-law degree sequences. See (Oliveira and Spencer, 2005; Rudas, Tóth and Valkó, 2007) for examples of PAMs with (possibly) non-linear PA-mechanisms and their degree sequences. We now explain how the affine PA-mechanism in (1.153) gives rise to power-law degree sequences and highlight the proof. We start by introducing some notation.

For $m \geq 1$ and $\delta > -m$, we define $\{p_k\}_{k=0}^{\infty}$ to be the probability distribution given by $p_k = 0$ for $k = 0, \dots, m-1$ and, for $k \geq m$,

$$p_k = (2 + \delta/m) \frac{\Gamma(k+\delta)\Gamma(m+2+\delta+\delta/m)}{\Gamma(m+\delta)\Gamma(k+3+\delta+\delta/m)} \qquad (1.154)$$

Then the main result on the scale-free nature of preferential attachment models is the following:

Theorem 1.16. (Degree sequence in the PAM) *Fix $\delta > -m$ and $m \geq 1$. Then, there exists a constant $C > 0$ such that, as $n \to \infty$,*

$$\mathbb{P}\left(\max_k |P_k^{(n)} - p_k| \geq C\sqrt{\frac{\log n}{n}} \right) = o(1). \qquad (1.155)$$

Furthermore, there exists a constant $C = C(m,\delta) > 0$ such that, as $k \to \infty$,

$$p_k = Ck^{-\tau}(1 + O(\frac{1}{k})), \qquad (1.156)$$

where

$$\tau = 3 + \frac{\delta}{m} > 2. \qquad (1.157)$$

In particular, Theorem 1.16 implies that the PA-random graph process is scale free. Theorem 1.16 appears in many forms in various settings. The statement which is closest to Theorem 1.16 is (Deijfen, van den Esker, van der Hofstad and Hooghiemstra, 2007, Theorem 1.3), where also the setting where each vertex enters the graph process with a *random* number of edges is considered. The first proof of a result as in Theorem 1.16 appeared in (Bollobás, Riordan, Spencer and Tusnády, 2001), they show a slightly weaker version of Theorem 1.16 when $\delta = 0$.

Virtually all proofs of asymptotic power laws in preferential attachment models consist of two steps: one step where it is proved that the degree sequence is concentrated around its mean, and one where the mean degree sequence is identified. We shall now give an intuitive explanation of Theorem 1.16.

Let $N_k^{(n)} = nP_k^{(n)}$ be the number of vertices with degree k in $G_m(n)$ (recall (1.141)). We are interested in the limiting distribution of $P_k^{(n)}$ as $n \to \infty$. This distribution arises as the solution of a certain recurrence relation, of which we will now give a short heuristic derivation. First note that, obviously,

$$\mathbb{E}[N_k^{(n+1)}|G_m(n)] = N_k^{(n)} + \mathbb{E}[N_k^{(n+1)} - N_k^{(n)}|G_m(n)]. \qquad (1.158)$$

Asymptotically, for n large, it is quite unlikely that a vertex will be hit by more than one of the m edges added upon the addition of vertex n. Let us hence ignore this possibility for the moment. The difference $N_k^{(n+1)} - N_k^{(n)}$ between the number of vertices with degree k at time $n + 1$ and time n respectively, is then obtained as follows:

(a) Vertices with degree k in $G_m(n)$ that are hit by one of the m edges emanating from vertex n are subtracted from $N_k^{(n)}$. The conditional probability that a fixed edge is attached to a vertex with degree k is $(k +$

$\delta) N_k^{(n)}/(n(2m+\delta))$, so that (ignoring multiple attachments to a single vertex) the mean number of vertices to which this happens is approximately $m(k+\delta)N_k^{(n)}/(n(2m+\delta))$. We note that we have replaced the numerator, which is $n(2m+\delta)+(e-1)(2+\delta/m)+1+\delta$ in the attachment of the eth edge emanating from vertex n, by its approximate value $n(2m+\delta)$ for large n.

  (b) Vertices with degree $k-1$ in $G_m(n)$ that are hit by one of the m edges emanating from vertex n are added to $N_k^{(n)}$. By reasoning as above, it follows that the mean number of such vertices is approximately $m(k-1+\delta)N_k^{(n)}/(n(2m+\delta))$.

  (c) The new vertex n should be added if it has degree k. When we ignore the case that vertex n attaches edges to itself, this happens precisely when $k=m$.

Combining this gives

$$\mathbb{E}\left[N_k^{(n+1)}-N_k^{(n)}|G_m(n)\right] \approx \frac{m(k-1+\delta)}{n(2m+\delta)}N_{k-1}^{(n)} - \frac{m(k+\delta)}{n(2m+\delta)}N_k^{(n)} + \mathbb{1}_{\{k=m\}}. \tag{1.159}$$

Substituting (1.159) into (1.158) and taking expectations, we arrive at

$$\mathbb{E}\left[N_k^{(n+1)}\right] \approx \mathbb{E}[N_k^{(n)}] + \frac{m(k-1+\delta)}{n(2m+\delta)}\mathbb{E}[N_{k-1}^{(n)}] - \frac{m(k+\delta)}{n(2m+\delta)}\mathbb{E}[N_k^{(n)}] + \mathbb{1}_{\{k=m\}}. \tag{1.160}$$

Now assume that $P_k^{(n)}$ converges to some limit p_k as $n \to \infty$, so that hence $N_k^{(n)} \sim np_k$. Then, for $n \to \infty$, and observing that $\mathbb{E}[N_k^{(n+1)}] - \mathbb{E}[N_k^{(n)}] \to p_k$ implies that $\frac{1}{n}\mathbb{E}[N_k^{(n)}] \to p_k$, for all k, yields the recursion

$$p_k = \frac{m(k-1+\delta)}{2m+\delta}p_{k-1} - \frac{m(k+\delta)}{2m+\delta}p_k + \mathbb{1}_{\{k=m\}}, \tag{1.161}$$

or, equivalently,

$$p_k = \frac{k-1+\delta}{k+2+\delta+\delta/m}p_{k-1} + \frac{2+\delta/m}{k+2+\delta+\delta/m}\mathbb{1}_{\{k=m\}}, \tag{1.162}$$

By iteration, it can be seen that this recursion is solved by $p_k = 0$ when $k < m$ and

$$p_k = \frac{2+\delta/m}{k+2+\delta+\delta/m}\prod_{j=m}^{k}\frac{k-j+\delta}{k-j+2+\delta+\delta/m}, \qquad k \geq m. \tag{1.163}$$

By rewriting the products in terms of Gamma-functions, we see that (1.163) is equal to (1.154). It is not hard to see that, when $k \to \infty$, $p_k \sim Ck^{-\tau}$ with $\tau = 3 + \delta/m$. This explains the occurrence of power-law degree sequences in

affine PAMs. The above argument can be made rigorous in order to show that $\max_k |\mathbb{E}[N_k^{(n)}] - np_k|$ remains uniformly bounded (see e.g., (Deijfen, van den Esker, van der Hofstad and Hooghiemstra, 2007)).

In order to prove concentration of $N_k^{(n)}$, all proofs in the literature make use of a clever martingale argument from (Bollobás, Riordan, Spencer and Tusnády, 2001). Define the Doob martingale M_t by

$$M_t = \mathbb{E}[N_k^{(n)}|G_m(t)]. \tag{1.164}$$

Then, $M_0 = \mathbb{E}[N_k^{(n)}]$ while $M_n = N_k^{(n)}$, so that $N_k^{(n)} - \mathbb{E}[N_k^{(n)}] = M_n - M_0$. The key ingredient is the observation that, for all $t \in [n]$, $|M_t - M_{t-1}| \leq 2m$ a.s., since the only vertices that are affected by the information of $G_m(t)$ instead of $G_m(t-1)$ are the vertices affected by the attachment of the edges incident to vertex t. Together, the concentration and the asymptotic mean give that $G_m(n)$ has an asymptotic degree sequence $\{p_k\}_{k \geq m}$, where p_k is close to a power-law for k large.

A prediction of universality. In non-rigorous work, it is often suggested the various scale-free random graph models, such as the CM or various models with conditional independence of the edges as in (Bollobás, Janson and Riordan, 2007), behave similarly. For scale-free random graph processes, this informal statement can be made precise by conjecturing that the phase transition or distances have the same behavior in graphs with the same power-law degree exponent. We shall discuss some of the results in this direction below.

1.3.3 *Phase transition in models without geometry*

In this section, we study the *phase transition* in random graphs. We first introduce some notation. For the CM with deterministic or random degrees according to F, we define

$$\nu = \frac{\mathbb{E}[D(D-1)]}{\mathbb{E}[D]}, \tag{1.165}$$

where D has distribution function F. For the rank-1 inhomogeneous random graph with deterministic or random weights according to F, we define

$$\nu = \frac{\mathbb{E}[W^2]}{\mathbb{E}[W]}, \tag{1.166}$$

where W has distribution function F. For the PAM, we let $\nu = m$. Below, we say that a sequence of events $\{E_n\}_{n=1}^{\infty}$ occurs *with high probability* (**whp**) when $\lim_{n \to \infty} \mathbb{P}(E_n) = 1$. It turns out that $\nu = 1$ plays the role of a critical value for all these random graphs:

Theorem 1.17. (Phase transition in random graphs) *(a-b) For the configuration model with deterministic or random degrees according to F, and for the rank-1 inhomogeneous random graph with deterministic or random weights according to F, the largest connected component has,* **whp***, size*

      $o(n)$ when $\nu \leq 1$, and size $\zeta n(1 + o(1))$ for some $\zeta > 0$ when $\nu > 1$, where n is the size of the graph.

  (c) For the PAM of size n, **whp***, the largest connected component has size $o(n)$ when $\nu = m = 1$, while the probability that the PAM is connected converges to 1 for $n \to \infty$ when $\nu = m > 1$.*

We write Theorem 1.17(a-b) to indicate that the result holds both for the IRG (which is model (a)) and for the CM (which is model (b)). This notation shall be used frequently below.

    The result for IRG is a special case of (Bollobás, Janson and Riordan, 2007, Theorem 3.1). Earlier versions for the random graph with given expected degrees appeared in (Chung and Lu, 2002b; Chung and Lu, 2006b) (see also the monograph (Chung and Lu, 2006a)). For the CM, the first result in the generality of Theorem 1.17 appeared in (Molloy and Reed, 1995; Molloy and Reed, 1998) under stronger conditions than mentioned here. For the sharpest result, see (Janson and Luczak, 2007). The connectivity of PAMs was investigated for $\delta = 0$ in (Bollobás and Riordan, 2004b), it was extended to all $\delta > -m$ in (van der Hofstad, 2008, Chapter 11).

    In (Deijfen, van den Esker, van der Hofstad and Hooghiemstra, 2007), also the setting where the number of edges with which a vertex enters the random graph is *random*. Indeed, denote the number of edges of vertex t by W_t, then in (Deijfen, van den Esker, van der Hofstad and Hooghiemstra, 2007) it is assumed that $\{W_t\}_{t=1}^{\infty}$ is an i.i.d. sequence. In general, such models are also scale-free with power-law exponent $\tau = \min\{\tau_{\mathrm{P}}, \tau_{\mathrm{W}}\}$, where $\tau_{\mathrm{P}} = 3 + \delta/\mu$ and $\mu = \mathbb{E}[W_t]$ is the "PA exponent", while τ_{W} is the power-law exponent of the weight distribution, i.e.,

$$\lim_{k \to \infty} \frac{\log \mathbb{P}(W_1 \geq k)}{\log (1/k)} = \tau_{\mathrm{W}} - 1. \tag{1.167}$$

Thus, one can summarize this by the fact that the effect with the least corresponding power-law exponent determines the power-law exponent of the graph. It would be of interest to study the phase transition for such more general models, and to verify r under what condition a giant component exists.

1.3.4 *Distances in models without geometry*

In this section, we summarize the results on distances in power-law random graphs. We combine the results in the three models discussed in Section 1.3.2 by the value of their respective power-law exponent. We define H_n to be the typical distance in the graph of size n, i.e., the number of edges in the shortest path between two uniformly chosen connected vertices. Note that even in a fixed graph, H_n is a *random variable*, as it depends on the uniform vertices which are chosen. We shall also discuss results on the *diameter* of the graph, which is the maximum of the shortest path distances between any pair of connected vertices. Both give information about distances in graphs, the typical distance being a more robust and informative feature of the graph than the diameter.

Distances in random graphs with finite variance degrees. The main results on distances in power-law random graphs with power-law exponent $\tau > 3$ are summarized in the following theorem:

Theorem 1.18. (Distances in graphs with finite variance degrees) *(a-b)*
For the configuration model and the rank-1 inhomogeneous random graph of size n, $H_n/\log n$ converges in probability to $1/\log\nu$, where ν is given by (1.165) for the CM, and by (1.166) for the rank-1 IRG, when F in the definition of the models satisfies that there exist $c > 0$ and $\tau > 3$ such that

$$1 - F(x) \leq cx^{-(\tau-1)}. \tag{1.168}$$

*(c) For the affine PAM of size n with $\delta > 0$, so that $\tau = 3 + \delta/m > 3$, **whp**, $H_n/\log n$ is bounded above and below by positive and finite constants.*

The result for the rank-1 IRG can be found in (van den Esker, van der Hofstad and Hooghiemstra, 2006*a*), where it is also shown that the fluctuations of H_n around $\log_\nu n$ remain bounded in probability, both in the case of i.i.d. degrees as well as for deterministic weights under a mild further condition on the distribution function. The first result in this direction was proved in (Chung and Lu, 2002*a*; Chung and Lu, 2003) for the expected degree random graph, in the case of rather general deterministic weights. A special case of the IRG with finite variance degrees is the Erdős-Rényi random graph with edge probability $p = \lambda/n$, for which $\nu = \lambda$.

The result for the CM can be found in (van der Hofstad, Hooghiemstra and Van Mieghem, 2005) in the case of i.i.d. degrees, where again also the fluctuations are determined. The results for deterministic degrees in the CM is conjectured in (van den Esker, van der Hofstad and Hooghiemstra, 2006*a*), but is not proved anywhere. We expect that the methodology in (van der Hofstad, Hooghiemstra and Van Mieghem, 2005) can be simply adapted to this case. The result for the affine PAM was proved in (van der Hofstad and Hooghiemstra, 2007*a*). Unfortunately the proof of convergence in probability is missing in this case. It would be of interest to identify the constant to which $H_n/\log n$ converges in this setting.

Distances in random graphs with finite mean and infinite variance degrees. When $\tau \in (2,3)$, the variance of the degrees becomes infinite, which is equivalent to the statement that, with $D_i^{(n)}$ denoting the degree of vertex i in the graph of size n, $\sum_{i=1}^{n}(D_i^{(n)})^2$ grows much faster than n. The following theorem shows that, in such cases, the distances are much smaller than $\log n$:

Theorem 1.19. (Distances in graphs with $\tau \in (2,3)$) *(a-b) For the configuration model and the rank-1 inhomogeneous random graph, $H_n/\log\log n$ converges in probability to $2/|\log(\tau-2)|$, when F in the definition of the models satisfies that there exist $c > 0$ and $\tau \in (2,3)$ such that*

$$1 - F(x) = cx^{-(\tau-1)}(1 + o(1)). \tag{1.169}$$

(c) *For the affine PAM with $\delta < 0$ and $m \geq 2$, so that $\tau = 3 + \delta/m \in (2,3)$,*
whp *$H_n/\log\log n$ is bounded above by a finite constant.*

Theorem 1.19 shows that all three models are ultra-small worlds when the power-law exponent τ satisfies $\tau \in (2,3)$.

The result for the rank-1 IRG is proved in (Chung and Lu, 2002a; Chung and Lu, 2003) for the expected degree random graph, in the case of certain deterministic weights. The result for the CM can be found in (van der Hofstad, Hooghiemstra and Znamenski, 2007) in the case of i.i.d. degrees, where again also the fluctuations are determined and are proved to be bounded. The restrictions on F are somewhat weaker than (1.169), as they also allow $x \mapsto x^{\tau-1}[1 - F(x)]$ to be slowly varying under certain conditions on the regularly varying function. The results in (Fernholz and Ramachandran, 2007) mentioned earlier apply in this case as well, and show that, when the proportion of vertices with degrees 1 and 2 is positive, the diameter divided by $\log n$ converges in probability to a positive constant. In (van der Hofstad and Hooghiemstra, 2007b), it is shown that the diameter in the CM is bounded above by a constant times $\log\log n$ when there are no vertices of degree 1 and 2.

The result for the affine PAM was proved in (van der Hofstad and Hooghiemstra, 2007a). Again, it would be of interest to identify the constant to which $H_n/\log\log n$ converges in this setting.

For the affine PA-model, we refer to (Bollobás and Riordan, 2004b) for a proof of the fact that the diameter of the PA-model $\mathrm{diam}(G_m(n))$, for $\delta = 0$, satisfies that $\mathrm{diam}(G_m(n)) \times \log\log n/\log n$ converges in probability to 1. This result is much sharper than the ones in Theorem 1.18(c) and Theorem 1.19(c), and it would be of interest to investigate whether the methodology used there can be adapted to the case where $\delta \neq 0$.

Distances in random graphs with infinite mean degrees. Only in the CM and the IRG, it is possible that the power-law exponent of the degrees or the weights of the vertices τ satisfies $\tau \in (1,2)$. In general, this is not very realistic, as it means that either there are extremely many multiple edges (in the CM) or the power-law exponent in the graph does not match the value of τ (in the IRG). Distances in infinite mean random graphs have been studied in (van den Esker, van der Hofstad, Hooghiemstra and Znamenski, 2006b; Norros and Reittu, 2006) and show that distances remain uniformly bounded by three. The intuition behind this is clear: all vertices are connected to vertices with extremely high degree, and these vertices form a complete graph, so that typical distances are at most 3. We refrain from a further discussion of random graphs with infinite mean degrees.

Conclusion on phase transition and distances. The main tool in order to study the phase transition and distances in the CM and IRGs is a comparison of the neighborhood of a vertex to a two-stage (multitype) branching process. In order to prove distance results, one then has to further investigate the growth of the number of vertices at a given distance using limit laws for branching processes.

When (1.168) holds, the number of vertices at a given distance k grows proportionally to ν^k, which suggests that distances are of the order $\log_\nu n$, as stated in Theorem 1.18.

Specifically, for the CM, with deterministic or i.i.d. degrees, the two-stage branching process, which we denote by $\{Z_k\}_{k\geq 0}$, starts from $Z_0 = 1$, has offspring distribution $\{f_n\}_{n=1}^\infty$, where $f_n = F(n) - F(n-1)$ are the jump sizes of the distribution F, in the first generation, and offspring distribution

$$g_n = \frac{(n+1)f_{n+1}}{\mathbb{E}[D]}, \qquad n \geq 0, \tag{1.170}$$

in the second and further generations. It is not hard to verify that the parameter ν in (1.165) is the expectation of the size-biased distribution $\{g_n\}_{n=0}^\infty$. For the rank-1 IRG with deterministic or random weights, the branching process $\{Z_k\}$ has a mixed-Poisson distribution with *random* parameter W in the first generation and a mixed Poisson distribution with random parameter W_e, which has the size-biased distribution of W, in the second and further generations. Thus, when W has a continuous density $w \mapsto f(w)$, the density of W_e is equal to $f_e(w) = wf(w)/\mathbb{E}[W]$. It can be seen that these two mixed Poisson distribution are again related through (1.170), and that again the parameter ν in (1.166) equals the expectation of the size-biased distribution used as offspring distribution in the second and further generations.

The condition $\nu > 1$ assures that the branching process $\{Z_k\}$ is supercritical, so that it can grow to a large size with positive probability (recall Theorem 1.17). Intuitively, all vertices for which the connected component are large (say larger than n^ε for some $\varepsilon > 0$) are connected and thus form a single giant component. The constant ζ in Theorem 1.17 is the survival probability of the two stage (multi-type) branching process $\{Z_k\}_{k\geq 0}$. Now, for a branching process $\{Z_k\}_{k\geq 0}$ for which the offspring distribution has finite $(1+\varepsilon)^{\text{th}}$ moment, we have that $Z_k\nu^{-k}$ converges a.s. to a limiting random variable W that is not identically 0 (by the Kesten-Stigum Theorem). The core of the proof is to use that H_n is the graph distance between two uniformly chosen connected vertices V_1 and V_2. Then, the neighborhood shells $Z_k^{(1,n)}$ and $Z_k^{(2,n)}$ consist of those vertices that are at graph distance *precisely* equal to k from V_1 and V_2, respectively. H_n is equal to $k_n + 1$, where k_n is the first time that any of the vertices in $Z_{\lfloor k_n/2\rfloor}^{(1,n)}$ connects to a vertex in $Z_{\lceil k_n/2\rceil}^{(2,n)}$. The proof is then completed by coupling these neighborhood shells to two independent two-stage branching processes as described above.

When (1.169) holds, then, by results of (Davies, 1978), the growth is *super-exponential*, i.e., $(\tau-2)^k \log(Z_k+1)$ converges a.s. to a limiting random variable Y, where $Y > 0$ precisely when the branching process survives. Thus, conditionally on $Y = y > 0$, the number of individuals in generation k grows like $e^{(\tau-2)^{-k}y(1+o(1))}$ suggesting that distances are of order $\log\log n/|\log(\tau-2)|$. The additional factor 2 in Theorem 1.19 is due to the fact that in order for two vertices to meet, *each* of their neighborhoods needs to have size at least n^ε for

some $\varepsilon > 0$, each of which can be expected to occur in a generation k with $k \approx \log \log n / |\log(\tau - 2)|$.

1.3.5 *Models with geometry*

In this section, we study some models of finite or bounded random graphs with geometry. We note that two of such models have been considered already in Section 1.2 as they are close relatives of the (infinite) percolation models. Indeed, in Section 1.2.5, the critical nature of percolation on high-dimensional tori was discussed, while in Section 1.2.8, random geometric graphs, which can be viewed as the restriction of continuum percolation to a bounded domain, have been described. In Section 2.5, models for random directed and on-line networks were discussed.

We now start by studying a spatial model for small worlds:

Small-world networks. The models described in Section 1.3.2 have flexible degree sequences and small distances, but they tend not to be very highly clustered. Also, these models do not incorporate geometry at all. An alternative approach of explaining the small-world phenomenon is to start with a finite torus, and to add random long range connections to them, independently for each pair of vertices. This gives rise to a graph which is a small perturbation of the original lattice, but has occasional long range connections that are crucial in order to shrink distances. From a practical point of view, we can think of the original graph as being the local description of acquaintances in a social network, while the shortcuts describe the occasional acquaintances in the population living far apart. The main idea is that, even though the shortcuts only form a tiny part of the connections in the graph, they are crucial in order to make it a small world.

There are various ways of adding long-range connections (for example by rewiring the existing edges), and we shall focus on the models in (Barbour and Reinert, 2001; Barbour and Reinert, 2004; Barbour and Reinert, 2006), for which the strongest mathematical results have been obtained. Small-world models were first introduced and analyzed in (Moore and Newman, 2000; Newman, Moore and Watts, 2000; Newman and Watts, 1999), and a non-rigorous mean-field analysis of distances in small-world models was performed in (Newman, Moore and Watts, 2000). See (Barbour and Reinert, 2001) for a discussion of the differences between the exact and mean-field analyses.

The simplest version of the model studied in (Barbour and Reinert, 2001) is obtained by taking the circle of circumference n, and adding a Poisson number of shortcuts with parameter $n\rho/2$, where the starting and endpoints of the shortcuts are chosen uniformly at random independently of each other. This model is called the *continuous circle model* in (Barbour and Reinert, 2001). Distance is measured as usual along the circle, and the shortcuts have, by convention, length zero. Thus, one can think of this model as the circle where the points along the random shortcut are identified, thus creating a puncture in the circle. Multiple shortcuts then lead to multiple puncturing of the circle, and the distance is then the usual distance along the punctured graph. Denote by D_n the distance

between two uniformly chosen points along the punctured circle. Then, (Barbour and Reinert, 2001, Theorem 3.9) states that as $n \to \infty$, $D_n(2\rho)/\log \rho n$ converges in probability to 1 when $n\rho \to \infty$, and that $\rho(D_n - \log \rho n/2)$ converges in distribution to a random variable T satisfying

$$\mathbb{P}(T > t) = \int_0^\infty \frac{e^{-y}dy}{1 + e^{2t}}. \qquad (1.171)$$

The random variable T can also be described by

$$\mathbb{P}(T > t) = \mathbb{E}[e^{-e^{2t}W^{(1)}W^{(2)}}], \qquad (1.172)$$

where $W^{(1)}$, $W^{(2)}$ are two independent exponential random variables with parameter 1. Alternatively, it can be see that $T = (G_1 + G_2 - G_3)/2$, where G_1, G_2, G_3 are three independent Gumbel distributions (see (Barbour and Reinert, 2006, Page 1242)).

Interestingly, the method of proof of (Barbour and Reinert, 2001, Theorem 3.9) is quite close to the method of proof for Theorem 1.18. Indeed, again the parts of the graph that can be reached in distance at most t are analyzed. Let P_1 and P_2 be two uniform points along the circle, so that D_n has the same distribution as the distance between P_1 and P_2. Denote by $R^{(1)}(t)$ and $R^{(2)}(t)$ the parts of the graph that can be reached within distance t. Then, $D_n = 2T_n$, where T_n is the first time that $R^{(1)}(t)$ and $R^{(2)}(t)$ have a non-zero intersection. The proof then consists of showing that, up to time T_n, the processes $R^{(1)}(t)$ and $R^{(2)}(t)$ are close to certain continuous-time branching processes, primarily due to the fact that the probability that there are two intervals that are overlapping in quite small. Then, $W^{(1)}$ and $W^{(2)}$ can be viewed as appropriate martingale limits of these branching processes. In (Barbour and Reinert, 2001, Theorem 4.2), also an extension to higher dimensions is given.

The proof was extended in (Barbour and Reinert, 2006) to deal with discrete tori where the shortcuts also contribute one to the graphs distance, so that distances are the usual distances on discrete graphs. For this, it was necessary that the average number of shortcuts per vertex $\rho \downarrow 0$, a restriction that does not appear in (Barbour and Reinert, 2001). It would be of interest to extend the results to the case of *fixed* ρ as well in the discrete setting.

A related model was considered in (Turova and Vallier, 2006). Indeed, (Turova and Vallier, 2006) studies a mixture between subcritical percolation on a finite cube and the Erdős-Rényi random graph. Using the methodology in (Bollobás, Janson and Riordan, 2007), it is shown that the phase transition is similar to the one described in Theorem 1.17. It would be of interest to verify whether the distance results in (Bollobás, Janson and Riordan, 2007) can also be used to prove that the distances grow like $\log_\nu n$, where n is the size of the graph, and $\nu > 1$ an appropriate constant.

A scale-free percolation network. In this section, we discuss the results in (Yukich, 2006) on an infinite scale-free percolation model. Note that, for a transitive

graph with fixed degree r and percolation with a fixed percolation parameter p, the degree of each vertex has a binomial distribution with parameters r and p. Since r is fixed, this does not allow for a power-law degree sequence. As a result, it is impossible to have a scale-free random graph when dealing with *independent* percolation, so that we shall abandon the assumption of independence of the different edges, while keeping the assumption of translation invariance.

The model considered in (Yukich, 2006) is on $\mathbb{Z}^d$, and, thus, the definition of a scale-free graph process does not apply so literally. We adapt the definition slightly by saying that an infinite random graph is scale-free when

$$p_k = \mathbb{P}(D_o = k), \tag{1.173}$$

where D_x is the degree of vertex $x \in \mathbb{Z}^d$ and $o \in \mathbb{Z}^d$ is the origin, satisfies (1.143). This is a reasonable definition, since if let $B_r = [-r, r]^d \cap \mathbb{Z}^d$ be a cube of width r around the origin, and denote $n = (2r + 1)^d$, then, for each $k \geq 0$,

$$P_k^{(n)} = \frac{1}{n} \sum_{x \in B_r} \mathbb{1}_{\{D_x = k\}}, \tag{1.174}$$

which, assuming translation invariance and ergodicity, converges to p_k.

We next describe the model in (Yukich, 2006). We start by taking an i.i.d. sequence $\{U_x\}_{x \in \mathbb{Z}^d}$ of uniform random variables on $[0, 1]$. Fix $\delta \in (0, 1]$ and $q \in (1/d, \infty)$. The edge $\{x, y\} \in \mathbb{Z}^d \times \mathbb{Z}^d$ appears in the random graph precisely when

$$|x - y| \leq \delta \min\{U_x^{-q}, U_y^{-q}\}. \tag{1.175}$$

We can think of the ball of radius δU_x^{-q} as being the region of influence of x, and two vertices are connected precisely when each of them lies into the region of influence of the other. This motivates the choice in (1.175). The parameter δ can be interpreted as the probability that nearest-neighbors are connected, and in the sequel we shall restrict ourselves to $\delta = 1$, in which case the infinite connected component equals $\mathbb{Z}^d$. We denote the resulting (infinite) random graph by $\mathbb{G}_q$.

We next discuss the properties of this model, starting with its scale-free nature. In (Yukich, 2006, Theorem 1.1), it is shown that, with $\tau = qd/(qd-1) \in (1, \infty)$, the limit

$$\lim_{k \to \infty} k^{\tau-1} \mathbb{P}(D_o \geq k) \tag{1.176}$$

exists, so that the model is scale free with degree power-law exponent τ (recall (1.144)). The intuitive explanation of (1.176) is as follows. Suppose we condition on the value of $U_o = u$. Then, the conditional distribution of D_o given that $U_o = u$ is equal to

$$D_o = \sum_{x \in \mathbb{Z}^d} \mathbb{1}_{\{|x| \leq \min\{U_o^{-q}, U_x^{-q}\}\}} = \sum_{x : |x| \leq u^{-q}} \mathbb{1}_{\{|x| \leq U_x^{-q}\}}. \tag{1.177}$$

Note that the random variables $\{\mathbb{1}_{\{|x| \leq U_x^{-q}\}}\}_{x \in \mathbb{Z}^d}$ are independent Bernoulli random variables with probability of success equal to

$$\mathbb{P}(\mathbb{1}_{\{|x|\leq U_x^{-q}\}} = 1) = \mathbb{P}(U \leq |x|^{-1/q}) = |x|^{-1/q}. \tag{1.178}$$

In order for $D_o \geq k$ to occur, for k large, we must have that $U_o = u$ is quite small, and, in this case, a central limit theorem should hold for D_o, with mean equal to

$$\mathbb{E}[D_o|U_o = u] = \sum_{x:|x|\leq u^{-q}} |x|^{-1/q} = cu^{-(qd-1)}(1 + o(1)), \tag{1.179}$$

for some explicit constant $c = c(q, d)$. Furthermore, the conditional variance of D_o given that $U_o = u$ is bounded above by its conditional expectation, so that the conditional distribution of D_o given that $U_o = u$ is highly concentrated. We omit the detail, and merely note that this can be made precise by using standard large deviations result. Assuming sufficient concentration, we obtain that the probability that $D_o \geq k$ is asymptotically equal to the probability that $U \leq u_k$, where u_k is determined by the equation that

$$\mathbb{E}[D_o|U_o = u_k] = cu_k^{-(qd-1)}(1 + o(1)) = k, \tag{1.180}$$

so that $u_k = (k/c)^{-1/(qd-1)}$. This suggests that

$$\mathbb{P}(D_o \geq k) = \mathbb{P}(U \leq u_k)(1 + o(1)) = (ck)^{-1/(qd-1)}(1 + o(1)), \tag{1.181}$$

which explains (1.176).

We next turn to distances in this scale-free percolation model. For $x, y \in \mathbb{Z}^d$, we denote by $d_{\mathbb{G}_q}(x, y)$ the graph distance (or chemical distance) between the vertices x and y, i.e., the minimal number of edges in $\mathbb{G}_q$ connecting x and y. The main result in (Yukich, 2006) is the following theorem:

Theorem 1.20. (Ultra-small distances for scale-free percolation) *For all $d \geq 1$ and all $q \in (1/d, \infty)$,* **whp** *as $|x| \to \infty$,*

$$d_{\mathbb{G}_q}(o, x) \leq 8 + 4\log\log|x|. \tag{1.182}$$

The result in Theorem 1.20 shows that distances in the scale-free percolation model are *much* smaller than those in normal percolation models. It would be of interest to investigate whether the limit $d_{\mathbb{G}_q}(o, x)/\log\log|x|$ exists, and, if so, what this limit is.

While Theorem 1.20 resembles the results in Theorem 1.19, there are a few essential differences. First of all, $\mathbb{G}_q$ is an *infinite* graph, whereas the models considered in Theorem 1.19 are all *finite*. It would be of interest to extend Theorem 1.20 to the setting on finite tori, where the Euclidean norm $|x - y|$ in (1.175) is replaced by the Euclidean norm on the torus, and the typical distance H_n is considered. This result is not immediate from the proof of Theorem 1.20. Secondly, in Theorems 1.18 and 1.19, it is apparent that the behavior for $\tau > 3$ is rather different compared to the behavior for $\tau \in (2, 3)$. This feature is missing

in Theorem 1.20. It would be of interest to find a geometric random graph model where the difference in behavior between $\tau > 3$ and $\tau \in (2, 3)$ also appears.

The result in Theorem 1.20 can be compared to similar results for *long-range percolation*, where edges are present independently, and the probability that the edge $\{x, y\}$ is present equals $|x - y|^{-s+o(1)}$ for some $s > 0$. In this case, detailed results exist for the limiting behavior of $d(o, x)$ depending on the value of s. For example, in (Benjamini, Kesten, Peres and Schramm, 2004), it is shown that the diameter of this infinite percolation model is equal to $\lceil d/(d - s) \rceil$ a.s. See also (Biskup, 2004) and the references therein.

Spatial preferential attachment models.   In the past years, several spatial preferential attachment models have been considered. We shall now discuss three of such models.

In (Flaxman, Frieze and Vera, 2006; Flaxman, Frieze and Vera, 2007), a class of geometric preferential attachment models that combines aspects of random geometric graphs and preferential attachment graphs is introduced and studied. Let $G_t = (V_t, E_t)$ denote the graph at time t. Let S be the sphere S in $\mathbb{R}^3$ with area equal to 1. Then, we let V_t be a subset of S of size t.

The process $\{G_t\}_{t=0}^{\infty}$ evolves as follows. At time $t = 0$, G_0 is the empty graph. At time $t+1$, given G_t, we obtain G_{t+1} as follows. Let x_{t+1} be chosen uniformly at random from S, and denote $V_{t+1} = V_t \cup \{x_{t+1}\}$. We assign m edges to the vertex x_{t+1}, which we shall connect independently of each other to vertices in $V_t(x_{t+1}) \equiv V_t \cap B_r(x_{t+1})$, where $B_r(u) = \{x \in S : \|x - u\| \leq r\}$ denotes the spherical cap of radius r around u. Let

$$D_t(x_{t+1}) = \sum_{v \in V_t(x_t)} D_v^{(t)}, \tag{1.183}$$

where $D_v^{(t)}$ denotes the degree of vertex $v \in V_t$ in G_t. The m edges are connected to vertices $(y_1, \ldots, y_m)$ conditionally independently given (G_t, x_{t+1}), so that, for all $v \in V_t(x_{t+1})$,

$$\mathbb{P}(y_i = v) = \frac{D_v^{(t)}}{\max(D_t(x_{t+1}), \alpha m A_r t)}, \tag{1.184}$$

while

$$\mathbb{P}(y_i = x_{t+1}) = 1 - \frac{D_t(x_{t+1})}{\max(D_t(x_{t+1}), \alpha m A_r t)}, \tag{1.185}$$

where A_r is the area of $B_r(u)$, $\alpha \geq 0$ is a parameter, and r is a radius which shall be chosen appropriately. Similarly to the situation of geometric random graphs, the parameter r shall depend on the size of the graph, i.e., we shall be interested in the properties of G_n when $r = r_n$ is chosen appropriately. The main result in (Flaxman, Frieze and Vera, 2006) is the study of the degree sequence of the arising model. Take $r_n = n^{\beta-1/2} \log n$, where $\beta \in (0, 1/2)$ is a constant. Finally, let $\alpha > 2$. Then, there exists a probability distribution $\{p_k\}_{k=m}^{\infty}$ such that, **whp**,

$$P_k^{(n)} = p_k(1 + o(1)), \tag{1.186}$$

where $\{p_k\}_{k=m}^{\infty}$ satisfies (1.143) with $\tau = 1 + \alpha \in (3, \infty)$. The precise result is in (Flaxman, Frieze and Vera, 2006, Theorem 1(a)) and is quite a bit sharper, as detailed concentration results are proved as well. Further results involve the proof of connectivity of G_n and an upper bound on the diameter when $r \geq n^{-1/2} \log n$, $m \geq K \log n$ for some large enough K and $\alpha \geq 0$ of order $O(\log(n/r))$. In (Flaxman, Frieze and Vera, 2007), these results were generalized to the setting where, instead of a unit ball, a smoother version is used.

In (Aiello, Bonato, Cooper, Janssen and Pralat, 2007), a spatial preferential attachment model with local influence regions is studied, as a model for the Web graph. The model is *directed*, but it can be easily adapted to an undirected setting. The idea behind the model in (Aiello, Bonato, Cooper, Janssen and Pralat, 2007) is that for normal preferential attachment models, new vertices should be aware of the degrees of the already present vertices. In reality, it is quite hard to observe the degrees of vertices, and, therefore, in (Aiello, Bonato, Cooper, Janssen and Pralat, 2007), vertices instead have a region of influence in some metric space, for example the torus $[0, 1]^m$ for some dimension m, for which the metric equals

$$d(x, y) = \min\{\|x - y + u\|_{\infty} : u \in \{0, 1, -1\}^m\}. \tag{1.187}$$

When the new vertex arrives, it is uniformly located somewhere in the unit cube, and it connects to each of the older vertices in which region of influence they land independently and with fixed probability p. These regions of influence evolve as time proceeds, in such a way that the volume of the influence region of the vertex i at time t is equal to

$$R(i, t) = \frac{A_1 D_i^{(t)} + A_2}{t + A_3}, \tag{1.188}$$

where now $D_i^{(t)}$ is the in-degree of vertex i at time t, and A_1, A_2, A_3 are parameters which are chosen such that $pA_1 \leq 1$. One of the main results of the paper is that this model is a scale-free graph process. Indeed, denote

$$p_k = \frac{p^k}{1 + kpA_1 + pA_2} \prod_{j=0}^{k-1} \frac{jA_1 + A_2}{1 + A_2 + pA_2}, \tag{1.189}$$

then (Aiello, Bonato, Cooper, Janssen and Pralat, 2007, Theorem 1.1) shows that **whp**, for $k \leq (n^{1/8}/\log n)^{4pA_1/(2pA_1+1)}$, the degree sequence of the graph of size n satisfies (recall (1.141))

$$P_k^{(n)} = p_k(1 + o(1)), \tag{1.190}$$

and $\{p_k\}_{k=0}^{\infty}$ satisfies (1.143) with $\tau = 1 + 1/(pA_1) \in [2, \infty)$. Further results involve the study of maximal in-degrees and the total number of edges.

For a relation between preferential attachment graphs with so-called fertility and aging, and a geometric competition-induced growth model for networks, we refer to (Berger, Borgs, Chayes, D'Souza and Kleinberg, 2004; Berger, Borgs, Chayes, D'Souza and Kleinberg, 2005) and the references therein.

1.4 Acknowledgements

REFERENCES

Aiello, W., Bonato, A, Cooper, C., Janssen, J., and Pralat, P. (Preprint (2007)). A spatial web graph model with local influence regions. To appear in *Internet Mathematics*.

Aiello, W., Chung, F., and Lu, L. (2002). Random evolution in massive graphs. In *Handbook of massive data sets*, Volume **4** of *Massive Comput.*, pp. 97–122. Kluwer Acad. Publ., Dordrecht.

Aizenman, M. (1997). On the number of incipient spanning clusters. *Nucl. Phys. B*, **485**, 551–582.

Aizenman, M. and Barsky, D.J. (1987). Sharpness of the phase transition in percolation models. *Commun. Math. Phys.*, **108**, 489–526.

Aizenman, M., Kesten, H., and Newman, C.M. (1987). Uniqueness of the infinite cluster and continuity of connectivity functions for short and long range percolation. *Commun. Math. Phys.*, **111**, 505–531.

Aizenman, M. and Newman, C.M. (1984). Tree graph inequalities and critical behavior in percolation models. *J. Stat. Phys.*, **36**, 107–143.

Albert, R. and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Rev. Modern Phys.*, **74**(1), 47–97.

Alon, N. and Spencer, J. (2000). *The probabilistic method* (Second edn). Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, New York.

Athreya, K. and Ney, P. (1972). *Branching processes*. Springer-Verlag, New York. Die Grundlehren der mathematischen Wissenschaften, Band 196.

Austin, T. L., Fagen, R. E., Penney, W. F., and Riordan, J. (1959). The number of components in random linear graphs. *Ann. Math. Statist*, **30**, 747–754.

Baccelli, F. and Błaszczyszyn, B. (2001). On a coverage process ranging from the Boolean model to the Poisson-Voronoi tessellation with applications to wireless communications. *Adv. in Appl. Probab.*, **33**(2), 293–323.

Barabási, A.-L. (2002). *Linked: The New Science of Networks*. Perseus Publishing, Cambridge, Massachusetts.

Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, **286**(5439), 509–512.

Barbour, A. D. and Reinert, G. (2001). Small worlds. *Random Structures Algorithms*, **19**(1), 54–74.

Barbour, A. D. and Reinert, Gesine (2004). Correction: "Small worlds" [Random Structures Algorithms **19** (2001), no. 1, 54–74; mr1848027]. *Random Structures Algorithms*, **25**(1), 115.

Barbour, A. D. and Reinert, G. (2006). Discrete small world networks. *Electron. J. Probab.*, **11**, no. 47, 1234–1283 (electronic).

Barsky, D.J. and Aizenman, M. (1991). Percolation critical exponents under

the triangle condition. *Ann. Probab.*, **19**, 1520–1536.

Barsky, D.J., Grimmett, G.R., and Newman, C.M. (1991*a*). Dynamic renormalization and continuity of the percolation transition in orthants. In *Spatial stochastic processes*, Volume **19** of *Progr. Probab.*, pp. 37–55. Birkhäuser Boston, Boston, MA.

Barsky, D.J., Grimmett, G.R., and Newman, C.M. (1991*b*). Percolation in half-spaces: equality of critical densities and continuity of the percolation probability. *Probab. Theory Related Fields*, **90**(1), 111–148.

Bender, E.A. and Canfield, E.R. (1978). The asymptotic number of labelled graphs with a given degree sequences. *Journal of Combinatorial Theory (A)*, **24**, 296–307.

Benjamini, I., Kesten, H., Peres, Y., and Schramm, O. (2004). Geometry of the uniform spanning forest: transitions in dimensions $4, 8, 12, \ldots$. *Ann. of Math. (2)*, **160**(2), 465–491.

Benjamini, I., Lyons, R., Peres, Y., and Schramm, O. (1999*a*). Critical percolation on any nonamenable group has no infinite clusters. *Ann. Probab.*, **27**(3), 1347–1356.

Benjamini, I., Lyons, R., Peres, Y., and Schramm, O. (1999*b*). Group-invariant percolation on graphs. *Geom. Funct. Anal.*, **9**(1), 29–66.

Benjamini, I. and Schramm, O. (1996). Percolation beyond $\mathbf{Z}^d$, many questions and a few answers. *Electron. Comm. Probab.*, **1**, no. 8, 71–82 (electronic).

van den Berg, J. and Keane, M. (1984). On the continuity of the percolation probability function. In *Conference in modern analysis and probability (New Haven, Conn., 1982)*, Volume **26** of *Contemp. Math.*, pp. 61–65. Amer. Math. Soc., Providence, RI.

Berg, J. van den and Kesten, H. (1985). Inequalities with applications to percolation and reliability. *J. Appl. Prob.*, **22**, 556–569.

Berger, N., Bollobás, B., Borgs, C., Chayes, J., and Riordan, O. (2003). Degree distribution of the FKP network model. In *Automata, languages and programming*, Volume **2719** of *Lecture Notes in Comput. Sci.*, pp. 725–738. Springer, Berlin.

Berger, N., Borgs, C., Chayes, J. T., D'Souza, R. M., and Kleinberg, R. D. ((2004)). Competition-induced preferential attachment. In *Automata, languages and programming*, Volume **3142** of *Lecture Notes in Comput. Sci.*, pp. 208–221. Springer, Berlin.

Berger, N., Borgs, C., Chayes, J. T., D'Souza, R. M., and Kleinberg, R. D. (2005). Degree distribution of competition-induced preferential attachment graphs. *Combin. Probab. Comput.*, **14**(5-6), 697–721.

Bezuidenhout, C. and Grimmett, G. (1990). The critical contact process dies out. *Ann. Probab.*, **18**, 1462–1482.

Biskup, M. (2004). On the scaling of the chemical distance in long-range percolation models. *Ann. Probab.*, **32**(4), 2938–2977.

Bollobás, B. (2001). *Random graphs* (Second edn), Volume **73** of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge.

Bollobás, B., Borgs, C., Chayes, J., and Riordan, O. (2003). Directed scale-free graphs. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms (Baltimore, MD, 2003)*, New York, pp. 132–139. ACM.

Bollobás, B., Janson, S., and Riordan, O. (2007). The phase transition in inhomogeneous random graphs. *Random Structures Algorithms*, **31**(1), 3–122.

Bollobás, B. and Riordan, O. (2003*a*). Mathematical results on scale-free random graphs. In *Handbook of graphs and networks*, pp. 1–34. Wiley-VCH, Weinheim.

Bollobás, B. and Riordan, O. (2003*b*). Robustness and vulnerability of scale-free random graphs. *Internet Math.*, **1**(1), 1–35.

Bollobás, B. and Riordan, O. (2004*a*). Coupling scale-free and classical random graphs. *Internet Math.*, **1**(2), 215–225.

Bollobás, B. and Riordan, O. (2004*b*). The diameter of a scale-free random graph. *Combinatorica*, **24**(1), 5–34.

Bollobás, B. and Riordan, O. (2006*a*). The critical probability for random Voronoi percolation in the plane is 1/2. *Probab. Theory Related Fields*, **136**(3), 417–468.

Bollobás, B. and Riordan, O. (2006*b*). *Percolation*. Cambridge University Press, New York.

Bollobás, B. and Riordan, O. (2006*c*). A short proof of the Harris-Kesten theorem. *Bull. London Math. Soc.*, **38**(3), 470–484.

Bollobás, B. and Riordan, O. (2008). Percolation on random Johnson-Mehl tessellations and related models. *Probab. Theory Related Fields*, **140**(3-4), 319–343.

Bollobás, B., Riordan, O., Spencer, J., and Tusnády, G. (2001). The degree sequence of a scale-free random graph process. *Random Structures Algorithms*, **18**(3), 279–290.

Borgs, C., Chayes, J., Hofstad, R. van der, Slade, G., and Spencer, J. (2005*a*). Random subgraphs of finite graphs. I. The scaling window under the triangle condition. *Random Structures Algorithms*, **27**(2), 137–184.

Borgs, C., Chayes, J., Hofstad, R. van der, Slade, G., and Spencer, J. (2005*b*). Random subgraphs of finite graphs. II. The lace expansion and the triangle condition. *Ann. Probab.*, **33**(5), 1886–1944.

Borgs, C., Chayes, J. T., Kesten, H., and Spencer, J. (1999). Uniform boundedness of critical crossing probabilities implies hyperscaling. *Random Structures Algorithms*, **15**(3-4), 368–413. Statistical physics methods in discrete probability, combinatorics, and theoretical computer science (Princeton, NJ, 1997).

Britton, T., Deijfen, M., and Martin-Löf, A. (2006). Generating simple random graphs with prescribed degree distribution. *J. Stat. Phys.*, **124**(6), 1377–1397.

Broadbent, S. R. and Hammersley, J. M. (1957). Percolation processes. I. Crystals and mazes. *Proc. Cambridge Philos. Soc.*, **53**, 629–641.

Burton, R. M. and Keane, M. (1989). Density and uniqueness in percolation. *Comm. Math. Phys.*, **121**(3), 501–505.

Camia, F. (Preprint (2008)). Scaling limits of two-dimensional percolation: an

overview.

Cardy, John L. (1992). Critical percolation in finite geometries. *J. Phys. A*, **25**(4), L201–L206.

Chayes, J. T. and Chayes, L. (1987). On the upper critical dimension of Bernoulli percolation. *Comm. Math. Phys.*, **113**(1), 27–48.

Chayes, J. T., Chayes, L., Grimmett, G. R., Kesten, H., and Schonmann, R. H. (1989). The correlation length for the high-density phase of Bernoulli percolation. *Ann. Probab.*, **17**(4), 1277–1302.

Chung, F. and Lu, L. (2002*a*). The average distances in random graphs with given expected degrees. *Proc. Natl. Acad. Sci. USA*, **99**(25), 15879–15882 (electronic).

Chung, F. and Lu, L. (2002*b*). Connected components in random graphs with given expected degree sequences. *Ann. Comb.*, **6**(2), 125–145.

Chung, F. and Lu, L. (2003). The average distance in a random graph with given expected degrees. *Internet Math.*, **1**(1), 91–113.

Chung, F. and Lu, L. (2006*a*). *Complex graphs and networks*, Volume **107** of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC.

Chung, F. and Lu, L. (2006*b*). The volume of the giant component of a random graph with given expected degrees. *SIAM J. Discrete Math.*, **20**, 395–411.

Chung, F., Lu, L., and Vu, V. (2004). The spectra of random graphs with given expected degrees. *Internet Math.*, **1**(3), 257–275.

Cooper, C. and Frieze, A. (2003). A general model of web graphs. *Random Structures Algorithms*, **22**(3), 311–335.

Davies, P. L. (1978). The simple branching process: a note on convergence when the mean is infinite. *J. Appl. Probab.*, **15**(3), 466–480.

Dawson, D.A. (1993). Measure-valued Markov processes. In *Ecole d'Eté de Probabilités de Saint-Flour 1991. Lecture Notes in Mathematics #1541*, Berlin. Springer.

Deijfen, M., Esker, H. van den, Hofstad, R. van der, and Hooghiemstra, G. (Preprint (2007)). A preferential attachment model with random initial degrees.

Dodziuk, J. (1984). Difference equations, isoperimetric inequality and transience of certain random walks. *Trans. Amer. Math. Soc.*, **284**(2), 787–794.

Dorogovtsev, S.N. and Mendes, J.F.F. (2002). Evolution of networks. *Advances in Physics*, **51**, 1079–1187.

Dousse, O., Franceschetti, M., Macris, N., Meester, R., and Thiran, P. (2006). Percolation in the signal to interference ratio graph. *J. Appl. Probab.*, **43**(2), 552–562.

Durrett, R. (1980). On the growth of one-dimensional contact processes. *Ann. Probab.*, **8**(5), 890–907.

Durrett, R. (2007). *Random graph dynamics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge.

Dynkin, E. B. (1994). *An introduction to branching measure-valued processes,*

Volume **6** of *CRM Monograph Series*. American Mathematical Society, Providence, RI.

Erdős, P. and Rényi, A. (1959). On random graphs. I. *Publ. Math. Debrecen*, **6**, 290–297.

Erdős, P. and Rényi, A. (1960). On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, **5**, 17–61.

Esker, H. van den, Hofstad, R. van der, and Hooghiemstra, G. (Preprint (2006*a*). Universality for the distance in finite variance random graphs.

Esker, H. van den, Hofstad, R. van der, Hooghiemstra, G., and Znamenski, D. (2006*b*). Distances in random graphs with infinite mean degrees. *Extremes*, **8**, 111–140.

Etheridge, A. M. (2000). *An Introduction to Superprocesses*, Volume **20** of *University Lecture Series*. American Mathematical Society, Providence, RI.

Fernholz, D. and Ramachandran, V. (2007). The diameter of sparse random graphs. *Random Structures Algorithms*, **31**(4), 482–516.

Flaxman, A., Frieze, A., and Vera, J. (2006). A geometric preferential attachment model of networks. *Internet Math.*, **3**(2), 187–205.

Flaxman, A., Frieze, A., and Vera, J. (2007). A geometric preferential attachment model of networks II. In *Proceedings of WAW 2007*.

Fortuin, C. M., Kasteleyn, P. W., and Ginibre, J. (1971). Correlation inequalities on some partially ordered sets. *Comm. Math. Phys.*, **22**, 89–103.

Franceschetti, M. and Meester, R. (2008). *Random networks for communication.* Cambridge University Press. from statistical physics to information systems.

Gilbert, E. N. (1959). Random graphs. *Ann. Math. Statist.*, **30**, 1141–1144.

Grimmett, G. (1999). *Percolation* (2nd edn). Springer, Berlin.

Grimmett, G. and Hiemer, P. (2002). Directed percolation and random walk. In *In and out of equilibrium (Mambucaba, 2000)*, Volume **51** of *Progr. Probab.*, pp. 273–297. Birkhäuser Boston, Boston, MA.

Grimmett, G. R. and Newman, C. M. (1990). Percolation in $\infty + 1$ dimensions. In *Disorder in physical systems*, Oxford Sci. Publ., pp. 167–190. Oxford Univ. Press, New York.

Grossglauser, M. and Thiran, P. (2006). Networks out of control: Models and methods for random networks.

Häggström, O. (2007) Problem solving is often a matter of cooking up an appropriate markov chain. *Scandinavian Journal of Statistics*, **34**, 768–780.

Häggström, O. and Peres, Y. (1999). Monotonicity of uniqueness for percolation on Cayley graphs: all infinite clusters are born simultaneously. *Probab. Theory Related Fields*, **113**(2), 273–285.

Häggström, O., Peres, Y., and Schonmann, R. H. (1999). Percolation on transitive graphs as a coalescent process: relentless merging followed by simultaneous uniqueness. In *Perplexing problems in probability*, Volume **44** of *Progr. Probab.*, pp. 69–90. Birkhäuser Boston, Boston, MA.

Hammersley, J. M. (1961). Comparison of atom and bond percolation processes. *J. Mathematical Phys.*, **2**, 728–733.

Hammersley, J. M. (1963). A Monte Carlo solution for percolation in a cubic lattice. In *Methods in Computational Physics, Vol. I*, pp. 281–298. Academic Press.

Hara, T. (1990). Mean field critical behaviour for correlation length for percolation in high dimensions. *Probab. Th. Rel. Fields*, **86**, 337–385.

Hara, T. (2005). Decay of correlations in nearest-neighbour self-avoiding walk, percolation, lattice trees and animals. Preprint. Available on `http://arxiv.org/abs/math-ph/0504021`.

Hara, T., Hofstad, R. van der, and Slade, G. (2003). Critical two-point functions and the lace expansion for spread-out high-dimensional percolation and related models. *Ann. Probab.*, **31**(1), 349–408.

Hara, T. and Slade, G. (1990). Mean-field critical behaviour for percolation in high dimensions. *Commun. Math. Phys.*, **128**, 333–391.

Hara, T. and Slade, G. (1993). Unpublished appendix to (Hara and Slade, 1995). Available on `http://www.ma.utexas.edu/mp_arc/index-93.html`.

Hara, T. and Slade, G. (1995). The self-avoiding-walk and percolation critical points in high dimensions. *Combinatorics, Probability and Computing*, **4**, 197–215.

Hara, T. and Slade, G. (2000*a*). The scaling limit of the incipient infinite cluster in high-dimensional percolation. I. Critical exponents. *J. Statist. Phys.*, **99**(5-6), 1075–1168.

Hara, T. and Slade, G. (2000*b*). The scaling limit of the incipient infinite cluster in high-dimensional percolation. II. Integrated super-Brownian excursion. *J. Math. Phys.*, **41**(3), 1244–1293.

Harris, T. (1963). *The theory of branching processes*. Die Grundlehren der Mathematischen Wissenschaften, Bd. 119. Springer-Verlag, Berlin.

Harris, T. E. (1960). A lower bound for the critical probability in a certain percolation process. *Proc. Cambridge Philos. Soc.*, **56**, 13–20.

Heydenreich, M. and Hofstad, R. van der (2007). Random graph asymptotics on high-dimensional tori. *Comm. Math. Phys.*, **270**(2), 335–358.

Heydenreich, M., Hofstad, R. van der, and Sakai, A. (Preprint (2008)). Mean-field behavior for long- and finite range Ising model, percolation and self-avoiding walk.

Hofstad, R. van der (2008). Random graphs and complex networks. In preparation, see `http://www.win.tue.nl/ rhofstad/NotesRGCN.pdf`.

Hofstad, R. van der, Hollander, F. den, and Slade, G. (2007). The survival probability for critical spread-out oriented percolation above 4+1 dimensions. I. Induction. *Probab. Theory Related Fields*, **138**(3-4), 363–389.

Hofstad, R. van der and Hooghiemstra, G. (Preprint (2007*a*)). Diameters in preferential attachment graphs.

Hofstad, R. van der and Hooghiemstra, G. (Preprint (2007*b*)). A phase transition for the diameter of the configuration model.

Hofstad, R. van der, Hooghiemstra, G., and Van Mieghem, P. (2005). Distances in random graphs with finite variance degrees. *Random Structures*

Algorithms, **27**(1), 76–123.

Hofstad, R. van der, Hooghiemstra, G., and Znamenski, D. (2007). Distances in random graphs with finite mean and infinite variance degrees. *Electron. J. Probab.*, **12**(25), 703–766 (electronic).

Hofstad, R. van der and Keane, M. (Preprint (2007)). An elementary proof of the hitting-time theorem.

Hofstad, R. van der and Luczak, M. (Preprint (2006)). Random subgraphs of the 2D Hamming graph: The supercritical phase.

Hofstad, R. van der and Sakai, A. (2005). Critical points for spread-out self-avoiding walk, percolation and the contact process above the upper critical dimensions. *Probab. Theory Related Fields*, **132**(3), 438–470.

Hofstad, R. van der and Slade, G. (2003). Convergence of critical oriented percolation to super-Brownian motion above $4 + 1$ dimensions. *Ann. Inst. H. Poincaré Probab. Statist.*, **39**(3), 413–485.

Hofstad, R. van der and Slade, G. (2005). Asymptotic expansions in n^{-1} for percolation critical values on the n-cube and $\mathbb{Z}^n$. *Random Structures Algorithms*, **27**(3), 331–357.

Hofstad, R. van der and Slade, G. (2006). Expansion in n^{-1} for percolation critical values on the n-cube and $\mathbb{Z}^n$: the first three terms. *Combin. Probab. Comput.*, **15**(5), 695–713.

Hughes, B. D. (1996). *Random walks and random environments. Vol. 2.* Oxford Science Publications. The Clarendon Press Oxford University Press, New York. Random environments.

Jagers, P. (1975). *Branching processes with biological applications.* Wiley-Interscience [John Wiley & Sons], London. Wiley Series in Probability and Mathematical Statistics—Applied Probability and Statistics.

Janson, S. (Preprint (2006)). The probability that a random multigraph is simple.

Janson, S. (Preprint (2008). Asymptotic equivalence and contiguity of some random graphs.

Janson, S. and Luczak, M. (Preprint (2007)). A new approach to the giant component problem.

Janson, S., Łuczak, T., and Rucinski, A. (2000). *Random graphs.* Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York.

Kager, Wouter and Nienhuis, Bernard (2004). A guide to stochastic Löwner evolution and its applications. *J. Statist. Phys.*, **115**(5-6), 1149–1229.

Kendall, W. S. and Wilson, R. G. (2003). Ising models and multiresolution quad-trees. *Adv. in Appl. Probab.*, **35**(1), 96–122. In honor of Joseph Mecke.

Kesten, H. (1959*a*). Full Banach mean values on countable groups. *Math. Scand.*, **7**, 146–156.

Kesten, H. (1959*b*). Symmetric random walks on groups. *Trans. Amer. Math. Soc.*, **92**, 336–354.

Kesten, H. (1980). The critical probability of bond percolation on the square

lattice equals $\frac{1}{2}$. *Comm. Math. Phys.*, **74**(1), 41–59.

Kesten, H. (1982). *Percolation theory for mathematicians*, Volume **2** of *Progress in Probability and Statistics*. Birkhäuser Boston, Mass.

Kesten, H. (1987). Scaling relations for 2D-percolation. *Comm. Math. Phys.*, **109**(1), 109–156.

Kesten, H. (2002). Some highlights of percolation. In *Proceedings of the International Congress of Mathematicians, Vol. I (Beijing, 2002)*, Beijing, pp. 345–362. Higher Ed. Press.

Langlands, R., Pouliot, P., and Saint-Aubin, Y. (1994). Conformal invariance in two-dimensional percolation. *Bull. Amer. Math. Soc. (N.S.)*, **30**(1), 1–61.

Lawler, G. F. (2004). An introduction to the stochastic Loewner evolution. In *Random walks and geometry*, pp. 261–293. Walter de Gruyter GmbH & Co. KG, Berlin.

Lawler, G. F. (2005). *Conformally invariant processes in the plane*, Volume **114** of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI.

Lawler, G. F., Schramm, O., and Werner, W. (2002). One-arm exponent for critical 2D percolation. Volume **7**, pp. no. 2, 13 pp. (electronic).

Le Gall, J.-F. (1999). *Spatial branching processes, random snakes, and partial differential equations*. Birkhäuser, Basel.

Meester, R. (1994). Uniqueness in percolation theory. *Statist. Neerlandica*, **48**(3), 237–252.

Meester, R. and Roy, R. (1996). *Continuum percolation*, Volume **119** of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge.

Menshikov, M.V. (1986). Coincidence of critical points in percolation problems. *Soviet Mathematics, Doklady*, **33**, 856–859.

Molloy, M. and Reed, B. (1995). A critical point for random graphs with a given degree sequence. In *Proceedings of the Sixth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science, "Random Graphs '93" (Poznań, 1993)*, Volume **6**, pp. 161–179.

Molloy, M. and Reed, B. (1998). The size of the giant component of a random graph with a given degree sequence. *Combin. Probab. Comput.*, **7**(3), 295–305.

Moore, C. and Newman, M.E.J. (2000). Epidemics and percolation in small-world networks. *Phys. Rev. E*, **61**, 5678–5682.

Nachmias, A. (Preprint (2007)). Mean-field conditions for percolation in finite graphs.

Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Rev.*, **45**(2), 167–256 (electronic).

Newman, M.E.J., Moore, C., and Watts, D.J. (2000). Mean-field solution of the small-world network model. *Phys. Rev. Lett.*, **84**, 3201–3204.

Newman, M.E.J. and Watts, D.J. (1999). Scaling and percolation in the small-world network model. *Phys. Rev. E*, **60**, 7332–7344.

Newman, M. E. J., Watts, D. J., and Barabási, A.-L. ((2006)). *The Structure and Dynamics of Networks*. Princeton Studies in Complexity. Princeton

University Press.

Nguyen, B. G. (1987). Gap exponents for percolation processes with triangle condition. *J. Statist. Phys.*, **49**(1-2), 235–243.

Nguyen, B.G. and Yang, W-S. (1993). Triangle condition for oriented percolation in high dimensions. *Ann. Probab.*, **21**, 1809–1844.

Nguyen, B.G. and Yang, W-S. (1995). Gaussian limit for critical oriented percolation in high dimensions. *J. Stat. Phys.*, **78**, 841–876.

Nienhuis, B. (1984). Critical behavior of two-dimensional spin models and charge asymmetry in the Coulomb gas. *J. Statist. Phys.*, **34**(5-6), 731–761.

Norros, I. and Reittu, H. (2006). On a conditionally Poissonian graph process. *Adv. in Appl. Probab.*, **38**(1), 59–75.

Oliveira, R. and Spencer, J. (2005). Connectivity transitions in networks with super-linear preferential attachment. *Internet Math.*, **2**(2), 121–163.

Penrose, M.D. (2003). *Random geometric graphs*, Volume **5** of *Oxford Studies in Probability*. Oxford University Press, Oxford.

Perkins, E. (2002). Dawson-Watanabe superprocesses and measure-valued diffusions. In *Lectures on probability theory and statistics (Saint-Flour, 1999)*, Volume **1781** of *Lecture Notes in Math.*, pp. 125–324. Springer, Berlin.

Reimer, D. (2000). Proof of the van den Berg-Kesten conjecture. *Combin. Probab. Comput.*, **9**(1), 27–32.

Rudas, A., Tóth, B., and Valkó, B. (2007). Random trees and general branching processes. *Random Structures Algorithms*, **31**(2), 186–202.

Russo, L. (1978). A note on percolation. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **43**(1), 39–48.

Russo, L. (1981). On the critical percolation probabilities. *Z. Wahrsch. Verw. Gebiete*, **56**(2), 229–237.

Sakai, A. (2002). Hyperscaling inequalities for the contact process and oriented percolation. *J. Statist. Phys.*, **106**(1-2), 201–211.

Schonmann, R. H. (2001). Multiplicity of phase transitions and mean-field criticality on highly non-amenable graphs. *Comm. Math. Phys.*, **219**, 271–322.

Schonmann, R. H. (1999). Stability of infinite clusters in supercritical percolation. *Probab. Theory Related Fields*, **113**(2), 287–300.

Schonmann, R. H. (2002). Mean-field criticality for percolation on planar non-amenable graphs. *Comm. Math. Phys.*, **225**(3), 453–463.

Schramm, O. (2000). Scaling limits of loop-erased random walks and uniform spanning trees. *Israel J. Math.*, **118**, 221–288.

Seymour, P. D. and Welsh, D. J. A. (1978). Percolation probabilities on the square lattice. *Ann. Discrete Math.*, **3**, 227–245. Advances in graph theory (Cambridge Combinatorial Conf., Trinity College, Cambridge, 1977).

Slade, G. (2006). *The lace expansion and its applications*, Volume **1879** of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin. Lectures from the 34th Summer School on Probability Theory held in Saint-Flour, July 6–24, 2004, Edited and with a foreword by Jean Picard.

Smirnov, S. (2001). Critical percolation in the plane: conformal invariance, Cardy's formula, scaling limits. *C. R. Acad. Sci. Paris Sér. I Math.*, **333**(3), 239–244.

Smirnov, S. and Werner, W. (2001). Critical exponents for two-dimensional percolation. *Math. Res. Lett.*, **8**(5-6), 729–744.

Tasaki, H. (1987*a*). Geometric critical exponent inequalities for general random cluster models. *J. Statist. Phys.*, **49**(3-4), 841–847.

Tasaki, H. (1987*b*). Hyperscaling inequalities for percolation. *Comm. Math. Phys.*, **113**(1), 49–65.

Timár, Á. (2006). Percolation on nonunimodular transitive graphs. *Ann. Probab.*, **34**(6), 2344–2364.

Turova, T.S. and Vallier, T. (Preprint (2006)). Merging percolation on $\mathbf{Z}^d$ and classical random graphs: Phase transition.

Werner, W. (2004). Random planar curves and Schramm-Loewner evolutions. In *Lectures on probability theory and statistics*, Volume **1840** of *Lecture Notes in Math.*, pp. 107–195. Springer, Berlin.

Werner, W. (2005). SLE, conformal restriction, loops. In *European Congress of Mathematics*, pp. 515–528. Eur. Math. Soc., Zürich.

Wu, C. C. (1993). Critical behavior or percolation and Markov fields on branching planes. *J. Appl. Probab.*, **30**(3), 538–547.

Yukich, J. E. (2006). Ultra-small scale-free geometric networks. *J. Appl. Probab.*, **43**(3), 665–677.