Percolation of a strongly connected component in simple directed random graphs with a given degree distribution

Femke van Ieperen[1] and Ivan Kryven[1,*]

[1]Mathematical Institute, Utrecht University, PO Box 80010, 3508 TA Utrecht, the Netherlands

December 14, 2020

Abstract

We study site and bond percolation on directed simple random graphs with a given degree distribution and derive the expressions for the critical value of percolation probability above which the giant strongly connected component emerges and the fraction of vertices in this component.

Keywords: random graphs, directed graphs, percolation, connected components.

1 Introduction

Percolation on infinite graphs is typically studied in a setting where edges (or vertices) are removed uniformly at random and some connectivity-related property is being traced as a function of the percolation probability $\pi \in (0,1)$ that a randomly chosen edge (or vertex) is present. The traced property is often chosen to describe connected components or clusters. A connected component is a maximal set of vertices, such that any pair of them is connected with a path. Many results about sizes of connected components, are known for percolation in infinite *regular lattices*, see, for example [1]. More recently, *random graphs* [2, 3] became an object of study for percolation too. Fountoulakis [4] and Jason [5] studied percolation in random graphs independently using different techniques, both of which rely on Molloy and Reed's theorem [6, 7], which indicates whether a simple undirected random graph with a given degree distribution contains a giant component and how large it is. Remarkably, Fountoulakis showed that a percolated random graph can be again viewed as a random graph generated by the configuration model, albeit with a modified degree distribution. The percolation thresholds can also be determined by applying Molloy and Reed's existence theorem to the modified degree distribution. Janson [5] derived the percolation threshold based on a technique called *exploding*. Due to this approach, auxiliary vertices of degree 1 are added to the percolated graph.

In *digraphs*, there exist several non-equivalent definitions for a connected component, all of which give rise to an interesting percolation problem. Let $G = (V, E)$ be a simple digraph and $n = |V|$. We say that $\mathcal{C} \subset V$ is a strongly connected component (SCC) if for all $v_1, v_2 \in \mathcal{C}$ there are directed paths that connect v_1 with v_2 and v_2 with v_1, and no other vertex from V can be added to $\mathcal{C}$ without loosing this property. Suppose G_n is uniformly sampled from the set of all digraphs with a fixed graphic degree sequence

$$\mathbf{d}^n := \big((d_1^-, d_1^+), (d_2^-, d_2^+), \dots, (d_n^-, d_n^+)\big), \tag{1}$$

where d_v^- and d_v^+ indicate correspondingly the in- and out-degree of $v \in V$. Let

$$\mu = \frac{1}{n}\sum_{v\in V} d_v^- = \frac{1}{n}\sum_{v\in V} d_v^+ < \infty$$

be expected number of edges per vertex and

$$\mu_{11} = \frac{1}{n}\sum_{v\in V} d_v^- d_v^+ < \infty.$$

Several authors formulated the existence criteria for the giant component in the context of directed graphs, see for example Penrose [8], Coulson [9] and Cooper and Frieze [10]. Cooper and Frieze considered sequences $(\mathbf{d}^n)_{n\in\mathbb{N}}$ and showed that certain regularity conditions guarantee convergence of $(G_n)_{n\in\mathbb{N}}$. Furthermore, these authors showed that if the limiting degree sequence has a positive fraction of dead-ends (vertices with one in- and no out-edges or vice versa) and, in average, large enough degrees, $\mu_{11} - \mu > 0$, then the size of the largest strongly connected component $\mathcal{C}_1(G_n)$ is of the order n,

$$\lim_{n\to\infty} \frac{|\mathcal{C}_1(G_n)\,|}{n} = c > 0. \tag{2}$$

Moreover, a SCC with this property is unique in the sense that the size of the second largest SCC is $o(n)$. If the latter limit holds, we say the random graph contains the giant strongly connected component (GSCC). Likewise, if the sign of the inequality is flipped, $\mu_{11} - \mu < 0$, then the size of all SCCs is $o(n)$, and random graph contains no GSCC. This result shows that there are two classes of limiting degree sequences, those that correspond to the size of the largest SCC being $O(n)$ and those for which the size is $o(n)$.

In this paper we show that, if the GSCC exists, removing a positive fraction of edges (or vertices) can modify the degree distribution just in the right way to flip the sign of the inequality, while keeping the percolated graph to be uniform in the a set of all graphs with a fixed (modified) degree distribution. By combining the latter observation with the theory of Cooper and Frieze, we show that the 'phase transition' from $O(n)$ to $o(n)$ takes place at a critical value $\pi_c = \frac{\mu}{\mu_{11}}$, such that only for $\pi > \pi_c$, G_n contains a GSCC with high probability (w.h.p.). Remarkably, the critical threshold π_c is the same for bond and site percolation, and the expressions for $c(\pi)$, as used in (2), are closely related, $c^{\text{site}}(\pi) = \pi c^{\text{bond}}(\pi)$. This work and the related proofs, are both largely inspired by the results for percolation on undirected graphs by Fountoulakis[4], which, in turn, rely upon Molloy and Reed's theorem [6] for the existence of the giant connected component.

2 Main result

This section introduces the theorem for the percolation threshold of GSCC. We consider two types of percolation processes on a simple digraph $G_n = (V_n, E_n)$, $n = |V_n|$ that result in a random subgraph G_n^π on the same vertex set:

- *Bond percolation,* fix a percolation probability $\pi \in (0, 1)$. Each edge of G_n is removed independently of the other edges with probability $1 - \pi$.

- *Site percolation,* fix a percolation probability $\pi \in (0,1)$. For each vertex of G_n all the edges incident to this vertex are removed with probability $1-\pi$, independently of the other vertices. Such a vertex is then referred to as a *deleted vertex.*

It should be clear from the context which type of percolation is discussed. Strictly speaking, existence of the GSCC is a limiting property of a sequence of graphs $(G_n)_{n\in\mathbb{N}}$, in which each element is defined by a finite graphic degree sequence $\mathbf{d}^n$. Thus we refer to an infinite sequence of degree sequences, $(\mathbf{d}^n)_{n\in\mathbb{N}}$, as the *degree progression*, where n is the index and the number of vertices in each element of this progression. Although our ultimate goal is to make statements about random graphs satisfying a specific degree distribution in the limit $n\to\infty$, the bulk of the paper is spent on determining whether a growing degree progression $(\mathbf{d}^n)_{n\in\mathbb{N}}$ maintains or acquires some property of interest on the way to such a limit. This provides a major difference between this work and percolation on infinite trees, which features the same percolation thresholds for the directed case [11, 12, 13].

To make valid statements about GSCC, we need to impose several requirements on the degree progression. We are interested in simple graphs, which indicates that each degree sequence in the progression has to be *graphic* as required by the equivalent of Erdős-Gallai theorem for directed graphs, Theorem 3.2. In Sections 3 and 4, we progressively add several more technical constraints on $(\mathbf{d}^n)_{n\in\mathbb{N}}$, namely Definitions 3.1, 3.3, and 4.8, which we jointly refer to as the requirements for a *proper degree progression.* The latter condition guarantees sufficient regularity of the degree progression to allow us to reason about the limiting behaviour of the corresponding random graphs.

Definition 2.1. The *percolation threshold* of a proper degree progression $(\mathbf{d}^n)_{n\in\mathbb{N}}$ is given by

$$\pi_c = \sup\left\{\pi\in(0,1)\ \middle|\ \forall\varepsilon>0,\ \lim_{n\to\infty}\mathbb{P}\left[\frac{|\mathcal{C}_1(G_n^\pi)|}{n}\geq\varepsilon\right]=0\right\}, \tag{3}$$

where superscripts are used to further specify the type of percolation, hence π_c^{bond} or π_c^{site}.

For each n, the probability in this definition is taken with respect to probability spaces G_n^π – random graphs that remain after percolation on uniform simple random graphs obeying $(\mathbf{d}^n)_{n\in\mathbb{N}}$.

Theorem 2.2. *Let $(\mathbf{d}^n)_{n\in\mathbb{N}}$ be a proper degree progression and $\mu_{11}-\mu>0$ for each element of this progression. Then the percolation threshold for the existence of a giant strongly connected component is given by*

$$\pi_c = \pi_c^{bond} = \pi_c^{site} = \frac{\mu}{\mu_{11}} < 1. \tag{4}$$

Let $N_{j,k}(n)$ be the number of vertices with degree (j,k) in G_n,

$$N_{j,k}(n) := \left|\{i\in V_n | d_i^- = j, d_i^+ = k\}\right|, \tag{5}$$

and

$$p_{j,k} := \lim_{n\to\infty}\frac{N_{j,k}(n)}{n} \tag{6}$$

be the corresponding limiting degree distribution. Let additionally,

$$U_\pi(x,y) := \sum_{j,k\geq 0} p_{j,k}(1-\pi+\pi x)^j(1-\pi+\pi y)^k,$$

$$U_\pi^-(y) := (\pi\mu)^{-1}\frac{\partial}{\partial x}U_\pi(x,y)|_{x=1},$$

and

$$U_\pi^+(x) := (\pi\mu)^{-1}\frac{\partial}{\partial y}U_\pi(x,y)|_{y=1}$$

be formal power series in x and y, having $\pi \in (0,1)$ as a parameter.

Theorem 2.3. *Let $(\mathbf{d}^n)_{n\in\mathbb{N}}$ be a proper degree progression and $\pi \in (\pi_c, 1)$, then there are unique values $c^{bond}(\pi)$ and $c^{site}(\pi)$, such that for all $\varepsilon_1, \varepsilon_2 > 0$,*

$$\lim_{n\to\infty}\mathbb{P}\left[|\frac{|\mathcal{C}_1(G_n^\pi)|}{n} - c^{bond}(\pi)| \geq \varepsilon_1\right] = 0,$$

for the bond percolation process, and

$$\lim_{n\to\infty}\mathbb{P}\left[|\frac{|\mathcal{C}_1(G_n^\pi)|}{n} - c^{site}(\pi)| \geq \varepsilon_2\right] = 0,$$

for the site percolation process, where

$$c^{bond}(\pi) = 1 - U_\pi(x^*, 1) - U_\pi(1, y^*) + U_\pi(x^*, y^*),$$

$$c^{site}(\pi) = \pi c^{bond}(\pi)$$

and $x^*, y^* \in (0,1)$ are the unique solutions of

$$x^* = U_\pi^+(x^*), \tag{7}$$

$$y^* = U_\pi^-(y^*). \tag{8}$$

Our Theorems 2.2, and 2.3 can be regarded a generalisation of Theorem 1.1 in Ref. [4], which determines the percolation threshold for the existence of a giant connected component in undirected simple graphs for the bond and site percolation processes.

The reminder of the paper is structured as follows: we start with a technical premise and introduce in Section 3 the working framework for the directed random graph, namely the *directed configuration model*, which is based on random matchings of half-edge configurations. We then introduce definitions of the giant components and recap the necessary existence theorems for these objects in Section 4. The latter section also contains Definition 4.8, introducing proper degree progressions. Section 5 proves our main result for percolation, Theorems 2.2 and 2.3. The section starts with several concentration inequalities and formulates Lemmas 5.5 and 5.7, stating that conditional on the degree sequence after bond/site percolation, a random configuration remains uniform. The section concludes with two separate proofs for bond and site percolation, provided in Subsections 5.2 and 5.3 respectively.

3 Random digraphs with a given degree distribution

A degree sequence of a digraph, as introduced in equation (1), can be uniquely defined by adopting the lexicographic order. Let us denote the set of *all* multigraphs obeying degree sequence $\mathbf{d}^n$ by $\mathcal{G}_{\mathbf{d}^n}$. Given a degree sequence $\mathbf{d}^n$, one may define a 'model' according to which a graph is chosen uniformly at random from $\mathcal{G}_{\mathbf{d}^n}$. Such a model can be viewed as a graph-valued alternative to the uniform random variable. When choosing the degree sequence $\mathbf{d}^n$ at random, there is no guarantee that there is a graph that corresponds to it, as it could happen that $\mathcal{G}_{\mathbf{d}^n} = \emptyset$. Additionally, to enable us to study its limiting behaviour, a degree progression $(\mathbf{d}^n)_{n\in\mathbb{N}}$ must satisfy some regularity constraints. Therefore, we introduce a notion of a *graphical* degree sequence and a *feasible* degree progression.

Definition 3.1. A degree sequence $\mathbf{d}^n$ is *valid* if for all $i \in \{1, 2, \dots, n\}$, $d_i^-, d_i^+ \in \mathbb{N}_0$ and

$$m := \sum_{i=1}^{n} d_i^- = \sum_{i=1}^{n} d_i^+, \tag{9}$$

where $m = |E|$ is the number of edges.

Since the sum of the in-degrees equals the sum of the out-degrees, it will always be possible to draw edges such that the graph obeys the desired degree sequence if self-loops and multi-edges are allowed. Thus, so long as multigraphs are of concern, the set $\mathcal{G}_{\mathbf{d}^n}$ is non-empty for any valid degree sequence.

A random *simple* graph that obeys a given degree sequence will be denoted by $G_{\mathbf{d}^n}$, and random *multigraph* obeying the same degree sequence, $\tilde{G}_{\mathbf{d}^n}$. If there is a simple graph obeying a degree sequence, the sequence is called *graphical.* The following generalisation of the Erdős-Gallai theorem gives necessary and sufficient conditions for a degree sequence to be graphical.

Theorem 3.2 (Fulkerson)**.** *[14, Theorem 4] Let $\mathbf{d}^n$ be a valid degree sequence and $\overline{\mathbf{d}^n}$ a positive lexicographical ordering of $\mathbf{d}^n$, that is $d_i^+ \geq d_{i+1}^+$ and $d_i^- \geq d_{i+1}^-$ if $d_i^+ = d_{i+1}^+$ for all $i \in \{1, 2, \dots, n-1\}$. Furthermore let $\underline{\mathbf{d}^n}$ be negative lexicographical ordering of $\mathbf{d}^n$, that is $d_i^- \geq d_{i+1}^-$ and $d_i^+ \geq d_{i+1}^+$ if $d_i^- = d_{i+1}^-$. Then the degree sequence is graphical, i.e. it can be represented by a simple graph, if and only if for all $k \in \{1, 2, \dots, n\}$:*

$$\text{(i)}\ \sum_{i=1}^{k} \min[\overline{d_i^-}, k-1] + \sum_{i=k+1}^{n} \min[\overline{d_i^-}, k] \geq \sum_{i=1}^{k} \overline{d_i^+}, \tag{10}$$

$$\text{(ii)}\ \sum_{i=1}^{k} \min[\underline{d_i^+}, k-1] + \sum_{i=k+1}^{n} \min[\underline{d_i^+}, k] \geq \sum_{i=1}^{k} \underline{d_i^-}. \tag{11}$$

Since our primary goal is to study connected components, we rely on additional assumptions on the degree distribution, as formalised by the following definition:

Definition 3.3. A degree progression $(\mathbf{d}^n)_{n\in\mathbb{N}}$ is called *feasible* if for any $n \in \mathbb{N}$, $\mathbf{d}^n$ is graphical and there exists a bivariate probability distribution $p_{j,k}$ that is independent of n and its first moment are equal, *i.e.*

$$\sum_{j,k=0}^{\infty} j p_{j,k} = \sum_{j,k=0}^{\infty} k p_{j,k}. \tag{12}$$

Furthermore, for this probability distribution the following must hold:

(i) for every $j, k \geq 0$, $\lim\limits_{n\to\infty} \frac{N_{j,k}(n)}{n} = p_{j,k} < \infty$;

(ii) $\lim\limits_{n\to\infty} \sum_{j,k=0}^{\infty} \frac{j N_{j,k}(n)}{n} = \sum_{j,k=0}^{\infty} j p_{j,k} \in (0, \infty)$;

(iii) $\lim_{n\to\infty} \sum_{j,k=0}^{\infty} \frac{jkN_{j,k}(n)}{n} = \sum_{j,k=0}^{\infty} jkp_{j,k} \in (0,\infty)$;

(iv) $\lim_{n\to\infty} \sum_{j,k=0}^{\infty} \frac{j^2N_{j,k}(n)}{n} = \sum_{j,k=0}^{\infty} j^2p_{j,k} \in (0,\infty)$;

(v) $\lim_{n\to\infty} \sum_{j,k=0}^{\infty} \frac{k^2N_{j,k}(n)}{n} = \sum_{j,k=0}^{\infty} k^2p_{j,k} \in (0,\infty)$.

We refer to the probability distribution $p_{j,k}$ as the degree distribution of a feasible degree progression and denote its moments by

$$\mu_{il} := \sum_{j,k=0}^{\infty} j^i k^l p_{j,k},\ i,l \in \{0,1,2\}, \tag{13}$$

and the probability generating function by

$$U(x,y) := \sum_{j,k=0}^{\infty} p_{j,k} x^j y^k. \tag{14}$$

Then, condition (12) can be now written out as $\mu := \mu_{10} = \mu_{01}$.
We consider the following construction process for a graph progression: We draw a degree sequence $\mathbf{d}^n$ from the bivariate degree distribution $p_{j,k}$, which is equivalent to saying that the limit

$$p_{j,k} = \lim_{n\to\infty} \frac{N_{j,k}(n)}{n}$$

converges for all $j,k \geq 0$. Suppose further, we uniformly choose a multigraph from $\mathcal{G}_{\mathbf{d}^n}$. When collected together for $n = 1,2\dots$, such degree sequences result in degree $(\mathbf{d}^n)_{n\in\mathbb{N}}$ and *multigraph* $\left(\tilde{G}_{\mathbf{d}^n}\right)_{n\in\mathbb{N}}$ progressions. Generally speaking, we are interested in the statements of the form

$$\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}_{\mathbf{d}^n} \in \mathcal{A}(\mathbf{d}^n)\right] = 1,$$

where $\mathcal{A}(\mathbf{d}^n)$ is the set of all graphs, which obey the degree sequence $\mathbf{d}^n$ and additionally satisfy some desired property. If the limit exists for a given property $\mathcal{A}$, then we say that the random graph has this property with high probability (w.h.p) or asymptotically almost surely (a.a.s.).

3.1 Directed configuration model

We will now show that the directed configuration model can be used to study the behaviour of *simple* directed random graphs, despite the fact that this model generates multigraphs.

Definition 3.4. Let $\mathbf{d}^n$ be a valid degree sequence. For all $i \in \{1,2,\dots,n\}$ define a set of *in-stubs* W_i^- consisting of d_i^- unique elements and a set *out-stub* W_i^+ containing d_i^+ elements. Let $W^- = \cup_{i\in\{1,2,\dots,n\}} W_i^-$ and $W^+ = \cup_{i\in\{1,2,\dots,n\}} W_i^+$. Then a *configuration* is a random perfect bipartite matching of W^- and W^+, that is a set of tuples (a,b) such that each tuple contains one element from W^- and one from W^+ and each element of W^- and W^+ appears in exactly one tuple.

A configuration $\mathcal{M}$ prescribes a matching for all stubs, and therefore, defines a multigraph $\tilde{G}_{\mathbf{d}^n}$ with vertices $V = \{1, 2, \ldots, n\}$ and edges

$$E = \{(i, j) \mid (a, b) \in \mathcal{M}, a \in W_i^+, b \in W_j^-\}. \tag{15}$$

The resulting multigraph will satisfy the degree sequence $\mathbf{d}^n$, as each vertex has $d_i^- = |W_i^-|$ incoming edges and $d_i^+ = |W_i^+|$ outgoing edges.

We will now study the probability that the configuration model generates a specific multigraph $\tilde{G}_{\mathbf{d}^n}$. Note that multiple configurations may result in the same graph. As an example, let $a, a' \in W_i^+$, $b \in W_j^-$ and $c \in W_k^-$. Consider a matching $\mathcal{M}$ with $(a, b), (a', c) \in \mathcal{M}$ and define $\mathcal{M}' = \mathcal{M} \setminus ((a, b), (a', c)) \cup ((a', b), (a, c))$. Then the multigraph induced by $\mathcal{M}$ is the same as the one induced by $\mathcal{M}'$. Since the configuration is chosen uniformly at random, the probability that the configuration model generates $\tilde{G}_{\mathbf{d}^n}$ depends on the number of configurations that induce this multigraph. Let $\mathrm{CM}_n(\mathbf{d}^n)$ be the random variable producing the outcome of the configuration model.

Proposition 3.5. *Let $\tilde{G}_{\mathbf{d}^n}$ be a multigraph with degree sequence $\mathbf{d}^n$. For all pairs $i, j \in V$ let x_{ij} denote the number of copies of the edge (i, j) in the graph. Then there holds*

$$\mathbb{P}\left[CM_n(\mathbf{d}^n) = \tilde{G}_{\mathbf{d}^n}\right] = \frac{1}{m!} \frac{\prod_{i=1}^n d_i^-! \prod_{i=1}^n d_i^+!}{\prod_{1 \le i,j \le n} x_{ij}!}. \tag{16}$$

Proof. This proposition and its proof are adapted for directed graphs from [15, Proposition 7.4], which is presented for the undirected case. First, we determine the number of different configurations, which is equal to the number of perfect bipartite matchings between W^- and W^+. Suppose we sequentially choose a match in W^+ for each element of W^-. Each element of W^- receives a match amongst the unmatched elements of W^+. The first element of W^- has m choices for its match. Then the second element can choose its match from the remaining $m-1$ unmatched elements of W^+. Continuing in this fashion, we find $m!$ different perfect matchings. Thus there are $m!$ different configurations. As the configuration is chosen uniformly at random, this implies that

$$\mathbb{P}\left[\mathrm{CM}_n(\mathbf{d}^n) = \tilde{G}_{\mathbf{d}^n}\right] = \frac{1}{m!} N\left(\tilde{G}_{\mathbf{d}^n}\right),$$

with $N\left(\tilde{G}_{\mathbf{d}^n}\right)$ being the number of different configurations $\mathcal{M}$ inducing the graph $\tilde{G}_{\mathbf{d}^n}$. From equation (15) it follows that the exact matching of the stubs does not matter. As long as an element of W_i^+ is matched to an element of W_j^-, the graph gets an edge (i, j). In other words permuting the stub labels leads to a different configuration that induces the same multigraph. There are $\prod_{i=1}^n d_i^-! \prod_{i=1}^n d_i^+!$ of such permutations. However some permutations lead to the same configuration $\mathcal{M}$. For $a, a' \in W_i^+$ and $b, b' \in W_j^-$ with $(a, b), (a', b') \in \mathcal{M}$, any permutation swapping a with a' and b with b' leads to the same configuration as the permutation acting on a, a', b and b' as the identity. We have to compensate for this by a factor $x_{ij}!$ for all edges. This leads to

$$N\left(\tilde{G}_{\mathbf{d}^n}\right) = \frac{\prod_{i=1}^n d_i^-! \prod_{i=1}^n d_i^+!}{\prod_{1 \le i,j \le n} x_{ij}!},$$

completing the proof. ∎

So, despite the fact that the configuration model generates uniformly random configurations, it does not generate uniformly random multigraphs. That being said, it does generate all simple graphs with equal probability. To see this remark that x_{ij} is $0, 1$ for every pair i, j in a simple graph. Therefore, conditional on the event that configuration model generates a simple graph, an element of $\mathcal{G}_{\mathbf{d}^n}$ is chosen uniformly.

The configuration model allows us to sample a random element of $\mathcal{G}_{\mathbf{d}^n}$. Surprisingly, it can be shown that results on uniformly random configurations and their induced multigraphs can be transferred to uniformly generated simple graphs. Let $\mathcal{A}(\mathbf{d}^n)$ be the set of all multigraphs, which obey the degree sequence $\mathbf{d}^n$ while additionally satisfying some property. The goal of the remainder of this section is to show that

$$\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}_{\mathbf{d}^n} \in \mathcal{A}(\mathbf{d}^n)\right] = 1,$$

implies that

$$\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}_{\mathbf{d}^n} \in \mathcal{A}(\mathbf{d}^n) \mid \tilde{G}_{\mathbf{d}^n} \text{ is simple}\right] = 1.$$

The first step is showing that the probability that the configuration model generates a simple graph is bounded away from zero. To show this, we need to pose restrictions on the maximum degree, which we define as:

$$d_{\max} = \max\left\{\max\{d_1^-, d_2^-, \ldots, d_n^-\}, \max\{d_1^+, d_2^+, \ldots, d_n^+\}\right\}.$$

Theorem 3.6. *[16, Theorem 4.3] Let $(\mathbf{d}^n)_{n\in\mathbb{N}}$ be a feasible degree progression with $d_{\max} = \mathcal{O}(\sqrt{n})$. Then the probability that the configuration model generates a simple graph is asymptotically*

$$e^{-\frac{\mu_{11}}{\mu} - \frac{(\mu_{20}-\mu)(\mu_{02}-\mu)}{\mu}} > 0.$$

Proof. The proof of the Theorem follows from the proof [16, Theorem 4.3] which is based on [16, Proposition 4.2]. The main difference is that in [16] the in-degree of a vertex is independent of its out-degree, *i.e.* the bivariate degree distribution is the product of two univariate distributions. It suffices to replace Condition 4.1 and Lemma 5.2 from Ref. [16] with the requirement of a feasible degree progression obeying $d_{\max} = \mathcal{O}(\sqrt{n})$ to generalise the proof to the case of an arbitrary bivariate degree distribution. □

Lemma 3.7. *Let $(\mathbf{d}^n)_{n\in\mathbb{N}}$ be a feasible degree progression with $d_{\max} = \mathcal{O}(\sqrt{n})$ and let $\mathcal{A}(\mathbf{d}^n)$ be a set of multigraphs, all obeying the degree sequence $\mathbf{d}^n$. If for a random multigraph $\tilde{G}_{\mathbf{d}^n}$ generated by the configuration model there holds*

$$\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}_{\mathbf{d}^n} \in \mathcal{A}(\mathbf{d}^n)\right] = 0,$$

then it is also true that

$$\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}_{\mathbf{d}^n} \in \mathcal{A}(\mathbf{d}^n) \mid \tilde{G}_{\mathbf{d}^n} \textit{ is simple}\right] = 0.$$

Proof. By Bayes'rule

$$\mathbb{P}\left[\tilde{G}_{\mathbf{d}^n} \in \mathcal{A}\left(\mathbf{d}^n\right) \mid \tilde{G}_{\mathbf{d}^n} \text{ is simple}\right] \leq \frac{\mathbb{P}\left[\tilde{G}_{\mathbf{d}^n} \in \mathcal{A}\left(\mathbf{d}^n\right)\right]}{\mathbb{P}\left[\tilde{G}_{\mathbf{d}^n} \text{ is simple}\right]}. \tag{17}$$

Because we have a feasible degree progression with $d_{\max} = \mathcal{O}\left(\sqrt{n}\right)$, Theorem 3.6 assures that

$$\mathbb{P}\left[\tilde{G}_{\mathbf{d}^n} \text{ is simple}\right] = (1 + o(1))\, e^{-\frac{\mu_{11}}{\mu} - \frac{(\mu_{20}-\mu)(\mu_{02}-\mu)}{\mu}}.$$

Hence

$$\lim_{n\to\infty} \inf \mathbb{P}\left[\tilde{G}_{\mathbf{d}^n} \text{ is simple}\right] > 0.$$

This completes the proof as the numerator in equation (17) converges to zero. ☐

Corollary 3.8. *Let $(\mathbf{d}^n)_{n\in\mathbb{N}}$ be a feasible degree progression with $d_{\max} = \mathcal{O}\left(\sqrt{n}\right)$. Take $\mathcal{A}\left(\mathbf{d}^n\right)$ to be a set of multigraphs, all obeying the degree sequence $\mathbf{d}^n$. Let $\tilde{G}_{\mathbf{d}^n}$ be a random multigraph generated by the configuration model. If there holds*

$$\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}_{\mathbf{d}^n} \in \mathcal{A}\left(\mathbf{d}^n\right)\right] = 1,$$

then it is also true that

$$\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}_{\mathbf{d}^n} \in \mathcal{A}\left(\mathbf{d}^n\right) \mid \tilde{G}_{\mathbf{d}^n} \textit{ is simple}\right] = 1.$$

Proof. Let $\mathcal{G}_{\mathbf{d}^n}$ denote the set of all multigraphs on n vertices obeying the degree sequence $\mathbf{d}^n$ and define

$$\overline{\mathcal{A}}\left(\mathbf{d}^n\right) := \mathcal{G}_{\mathbf{d}^n} \setminus \mathcal{A}\left(\mathbf{d}^n\right).$$

As there holds that $\tilde{G}_{\mathbf{d}^n} \in \mathcal{G}_{\mathbf{d}^n}$ by definition and $\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}_{\mathbf{d}^n} \in \mathcal{A}\left(\mathbf{d}^n\right)\right] = 1$ by assumption, the law of total probability implies $\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}_{\mathbf{d}^n} \in \overline{\mathcal{A}}\left(\mathbf{d}^n\right)\right] = 0$. Then Lemma 3.7 implies that

$$\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}_{\mathbf{d}^n} \in \overline{\mathcal{A}}\left(\mathbf{d}^n\right) \mid \tilde{G}_{\mathbf{d}^n} \text{ is simple}\right] = 0.$$

Again using that $\mathcal{A}\left(\mathbf{d}^n\right) \cup \overline{\mathcal{A}}\left(\mathbf{d}^n\right)$ is the set of all multigraphs obeying the given degree sequence, the claim follows. ☐

4 Giant strongly connected component in a directed graph

In simple graphs, a connected component is a maximal subset $C \subset V$ such that there is a path between any pair of vertices $u, v \in C$. Here maximal means that no vertex can be added to C without destroying the property that all vertices are connected by a path. In directed graphs, a path may also be directed.

Definition 4.1. Let $G_n = (V, E)$ be a digraph. A pair of vertices $v_1, v_k \in V$ is connected by a *directed path* if there exist distinct vertices $v_2, v_3, \ldots, v_{k-1} \in V$ such that for all $i \in \{2, 3, \ldots, k\}$ $(v_{i-1}, v_i) \in E$. We refer to such a sequence as a directed $v_1 - v_k$ path.

We therefore adapt the definition of a connected component to be based on a directed path.

Definition 4.2. (Strongly connected component) Consider a directed graph G_n. The *strongly connected components* of G_n are the maximal subsets of V such that between any pair of vertices u, v directed $u - v$ and $v - u$ paths exist simultaneously.

Definition 4.3. Consider a directed graph $G_n = (V, E)$ and take $v \in V$. Then

- the *in-component* of v, denoted by $\mathrm{In}\,(v)$, consists of v itself and all vertices $u \in V$ such that a directed $u - v$ path exists;
- the *out-component* of v, denoted by $\mathrm{Out}\,(v)$, consists of v itself and all vertices $w \in V$ for which a directed $v - w$ path exists;
- the *strong-component* of v, denoted by $\mathrm{SCC}\,(v)$, consists of v itself and all vertices $w \in V$ for which both $v - w$ and $v - w$ paths exist.

Note that $u \in \mathrm{In}\,(v)$ does not imply $\mathrm{In}\,(u) = \mathrm{In}\,(v)$. The same observation holds for the out-component as well. This means that these types of components do not partition the vertices of the graph. Furthermore, we note that $\mathrm{In}\,(v)\,, \mathrm{Out}\,(v) \supset \mathrm{SCC}\,(v)$. These components allow to characterize the strongly connected component in the following way.

Lemma 4.4. *Consider a directed graph $G_n = (V, E)$. For any $v \in V$ there holds:*

$$In\,(v) \cap Out\,(v) = SCC\,(v)\,.$$

Proof. To prove that $\mathrm{In}\,(v) \cap \mathrm{Out}\,(v) = \mathrm{SCC}\,(v)$, it suffices to show $\mathrm{SCC}\,(v) \subset \mathrm{In}\,(v) \cap \mathrm{Out}\,(v)$ and $\mathrm{SCC}\,(v) \supset \mathrm{In}\,(v) \cap \mathrm{Out}\,(v)$. Take $u \in \mathrm{SCC}\,(v)$. By definition of the strongly connected component this implies that directed $u - v$ and $v - u$ paths exist. Hence $u \in \mathrm{In}\,(v)$ and $u \in \mathrm{Out}\,(v)$, which shows that $\mathrm{SCC}\,(v) \subset \mathrm{In}\,(v) \cap \mathrm{Out}\,(v)$. Next take $u \in \mathrm{In}\,(v) \cap \mathrm{Out}\,(v)$. This implies that directed $u - v$ and $v - u$ paths exist. For u to be an element of $\mathrm{SCC}\,(v)$ it must also hold that for any $w \in \mathrm{SCC}\,(v)$ directed $w - u$ and $u - w$ paths exist. Because $w \in \mathrm{SCC}\,(v)$, directed $v - w$ and $w - v$ paths are present in G_n. This implies existence of a $u - w$ path by concatenation of the $u - v$ and $v - w$ paths. Similarly the $w - u$ path can be formed by concatenating the $w - v$ path with the $v - u$ path. Thus $u \in \mathrm{SCC}\,(v)$, which shows that $\mathrm{SCC}\,(v) \supset \mathrm{In}\,(v) \cap \mathrm{Out}\,(v)$. □

By denoting the largest strongly connected component of G_n by $\mathcal{C}_1(G_n)$, the notion of a giant strongly connected component can be defined as follows.

Definition 4.5. The graph progression $\left(G_{\mathbf{d}^n}\right)_{n\in\mathbb{N}}$ contains a *giant strongly connected component* (GSCC) if

$$\lim_{n\to\infty} \frac{|\mathcal{C}_1(G_{\mathbf{d}^n})|}{n} = c_{scc} > 0. \tag{18}$$

Here c_{scc} represents the fraction of the vertices engaged in the GSCC, which we simply refer to as the size of the GSCC.

Definition 4.6. The graph progression $\left(G_{\mathbf{d}^n}\right)_{n\in\mathbb{N}}$ contains a *giant in-component* (GIN) if for a uniformly random vertex v

$$\lim_{n\to\infty} \frac{|\mathrm{In}\,(v)|}{n} = c_{in} > 0. \tag{19}$$

Definition 4.7. The graph progression $\left(G_{\mathbf{d}^n}\right)_{n\in\mathbb{N}}$ contains a *giant out-component* (GOUT) if for a uniformly random vertex v

$$\lim_{n\to\infty} \frac{|\mathrm{Out}\,(v)|}{n} = c_{out} > 0. \tag{20}$$

4.1 Existence of GIN, GOUT & GSCC

In this section we introduce a theorem that determines whether a random graph progression $\left(G_{\mathbf{d}^n}\right)_{n\in\mathbb{N}}$ w.h.p. contains a GSCC. Cooper and Frieze showed that the out-component of each vertex either contains $\mathcal{O}\left(d_{\max}^2 \ln(n)\right)$ vertices or w.h.p. contains $a_0^+ n$ vertices, for some constant $a_0^+ > 0$. Similarly they show that the in-component of each vertex either contains $\mathcal{O}\left(d_{\max}^2 \ln(n)\right)$ vertices or w.h.p. contains $a_0^- n$ vertices, for some $a_0^- > 0$. Denote the set of all vertices with out-component of size $a_0^+ n$ by L^+ and the set of all vertices with in-component of size $a_0^- n$ by L^-. Lemma 4.4 implies that a strongly connected component is the intersection of the in-component and the out-component of some vertex. Thus a vertex v must be in L^+ and L^- to be in the GSCC, if the GSCC exists.

Definition 4.8. A degree progression $(\mathbf{d}^n)_{n\in\mathbb{N}}$ is *proper* if it is feasible and additionally satisfies

1. $d_{\max} \leq \frac{n^{1/12}}{\ln(n)}$;
2. $\rho := \max\left(\sum_{j,k=0}^{\infty} \frac{j^2 k N_{j,k}(n)}{\mu n}, \sum_{j,k=0}^{\infty} \frac{j k^2 N_{j,k}(n)}{\mu n}\right) = o(d_{\max})$.

In what follows, we first present an intuitive explanation for the probability that a random vertex is in L^+ (respectively L^-) by investigating the out-components (in-components) of a random graph and then formalise the explanation with a rigours theorem. To investigate the out-component of a random graph, we take a proper degree progression $(\mathbf{d}^n)_{n\in\mathbb{N}}$ with underlying degree distribution $(p_{j,k})_{j,k\in\mathbb{N}}$. In Section 3.1 we showed that results on multigraphs induced by uniformly random configurations can be transferred to uniformly random simple graphs. Thus we will consider multigraphs generated by the configuration model. As discussed in Section 3.1, the configuration model generates a uniformly random configuration of W^- and W^+. This gives us the freedom to construct the configuration in such a way that the out-component of a vertex v is constructed first: Pick a vertex v, then choose any out-stub of v and match it to a uniformly random unmatched in-stub. Suppose this in-stub is an element of W_w^-. This implies that we add w to $\mathrm{Out}\,(v)$. Now also the out-stubs of w can extend the out-component of v. Hence at the next step, we take any unmatched out-stub from the set $W_v^+ \cup W_w^+$ and match it to a uniformly random unmatched in-stub. This process continues until all out-stubs of $\cup_{w\in\mathrm{Out}(v)} W_w^+$ are matched. At that point the out-component of v is completed. Randomly match the remaining unmatched in-stubs with the unmatched out-stubs to complete the configuration.

The construction of the out-component of v, as described above, is similar to a Galton-Watson branching process. A Galton-Watson branching process is a discrete time process. It starts with one individual. At each time-step a random living individual generates an integer amount of off-spring and dies. The amount of off-spring generated is governed by the off-spring distribution $Z(k)$. Such a process either continues on indefinitely, *i.e.* at each time-step at least one individual remains alive, or it becomes extinct at some point, *i.e.* at some time-step all individuals are dead. Let ξ be the probability that the branching process becomes extinct at some point.

Lemma 4.9. *Consider a Galton-Watson branching process with offspring distribution $Z(k)$. The extinction probability ξ of this process is*

- 1 *if* $\mathbb{E}[Z(k)] < 1$*;*
- *the unique solution in* $[0,1)$ *of the equation* $x = \sum_{k=0}^{\infty} \mathbb{P}[Z(k) = k]\, x^k$ *if* $\mathbb{E}[Z(k)] > 1$.

The parallel between the branching process and the construction of the out-component becomes apparent when we identify an individual with an unmatched out-stub. An unmatched out-stub 'dies' when it is matched to a random in-stub b. The amount of off-spring it generates, equals d_w^+, where w is the vertex such that $b \in W_w^-$. There is one exception: in case w was already an element of $\mathrm{Out}(v)$, the amount of off-spring is zero. This means that the off-spring distribution is approximately $\left(p_k^+\right)_{k\in\mathbb{N}}$, with p_k^+ being the probability that by following a uniformly random chosen in-stub, we find a vertex with out-degree k. This latter probability is given by

$$p_k^+ = \sum_{j=0}^{\infty} \frac{j}{\mu} p_{j,k}. \tag{21}$$

Here the division by μ ensures that the probability distribution is normalized. We expect that $v \in L^+$ if the corresponding Galton-Watson branching with off-spring distribution $\left(p_k^+\right)_{k\in\mathbb{N}}$ continues on indefinitely. Applying Lemma 4.9 to the distribution, we find that this can happen only if

$$\sum_{k=0}^{\infty} k p_k^+ = \frac{\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} jk p_{j,k}}{\mu} = \frac{\mu_{11}}{\mu} > 1.$$

Thus we expect that L^+ is empty unless $\frac{\mu_{11}}{\mu} > 1$. When $\frac{\mu_{11}}{\mu} > 1$, the probability that the branching process does not become extinct is η^+, where $1-\eta^+$ is the unique solution in $[0,1)$ to

$$1 - \eta^+ = \sum_{k=0}^{\infty} p_k^+ \left(1-\eta^+\right)^k. \tag{22}$$

Recall that the Galton-Watson process always starts with one individual. However the generation of the out-component of v starts with d_v^+ out-stubs. Assuming that each out-stub of v generates a disjoint subset of $\mathrm{Out}(v)$, this can be regarded as d_v^+ independent copies of the branching process. This collection of processes terminates if and only if all of the individual branching processes become extinct. Therefore the probability that the process corresponding to the generation of the out-component of v becomes extinct is $(1-\eta^+)^{d_v^+}$. Thus for a random vertex v the probability that $v \in L^+$ is approximately ζ^+, where $1-\zeta^+$ is given by

$$1 - \zeta^+ = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} p_{j,k} \left(1-\eta^+\right)^k. \tag{23}$$

The same idea can be applied to the in-component, but with a different off-spring distribution. In this case we need to probability that following a uniformly random chosen out-stub, we find a vertex with out-degree j, i.e.

$$p_j^- = \sum_{k=0}^{\infty} \frac{k}{\mu} p_{j,k}. \tag{24}$$

A Galton-Watson branching process with off-spring distribution $\left(p_j^-\right)_{j\in\mathbb{N}}$ has an extinction probability smaller than 1 if

$$\sum_{j=0}^{\infty} jp_j^- = \frac{\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} jkp_{j,k}}{\mu} = \frac{\mu_{11}}{\mu} > 1.$$

Note that this is the same condition as for the off-spring distribution of the out-component. In other words L^+ and L^- are expected to become non-empty under the same conditions. The probability that this branching process does not become extinct is η^-, where $1-\eta^-$ is the unique solution in $[0,1)$ to

$$1-\eta^- = \sum_{j=0}^{\infty} p_j^- \left(1-\eta^-\right)^j. \tag{25}$$

Thus we expect a random vertex v to be in L^- with probability ζ^-, with $1-\zeta^-$ defined by

$$1-\zeta^- = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} p_{j,k} \left(1-\eta^-\right)^j. \tag{26}$$

Using these probabilities for a random vertex to be in L^- or L^+ and Lemma 4.4, we can determine whether a giant strongly connected component is present in the graph. For a vertex $v \in L^-$ and a vertex $u \in L^+$ it is very likely that there exists an edge from a vertex in $\text{Out}(u)$ to one in $\text{In}(v)$, i.e. $v \in \text{Out}(u)$ and $u \in \text{In}(v)$. Thus if $u, v \in L^-$ and $u, v \in L^+$, it is likely that u and v are in each others in-component and out-component. By Lemma 4.4 this implies that u and v lie in the same strongly connected component. Hence we expect a GSCC consisting of the vertices $L^+ \cap L^-$. This set is expected to be non-empty if $\frac{\mu_{11}}{\mu} > 1$. To determine the probability for a random vertex to be in $L^+ \cap L^-$, i.e. to approximate the size of the GSCC, define

$$\psi = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} p_{j,k} \left(1-\eta^-\right)^j \left(1-\eta^+\right)^k. \tag{27}$$

This approximates the probability that the branching process corresponding to the generation of the in-component and the out-component both become extinct. Hence the probability that both process go on indefinitely is given by

$$c = \zeta^+ + \zeta^- + \psi - 1. \tag{28}$$

The above intuitive exposition is rigorously proven by Cooper and Frieze for proper degree progressions.

Theorem 4.10 (Existence of a GIN, GOUT and GSCC)**.** *[10, Theorem 1 and 2] Consider a proper degree progression* $(\mathbf{d}^n)_{n\in\mathbb{N}}$*. Take a uniformly random sequence of simple graphs* $\left(G_{\mathbf{d}^n}\right)_{n\in\mathbb{N}}$ *obeying this degree progression. Then the following statements hold.*

1. *If* $\frac{\mu_{11}}{\mu} < 1$*, with high probability the size of the in-component and the out-component of each vertex is* $\mathcal{O}\left(d_{\max}^2 \ln(n)\right)$*.*

2. *If* $\frac{\mu_{11}}{\mu} > 1$ *and* $p_0^+, p_0^- > 0$*, with high probability*

- *There are $\zeta^+ n$ vertices with an out-component containing $\zeta^- n$ vertices;*
- *There are $\zeta^- n$ vertices with an in-component containing $\zeta^+ n$ vertices;*
- *There is a unique giant strongly connected component with vertex set $L^+ \cap L^-$ of size $(\zeta^+ + \zeta^- + \psi - 1)\, n$.*

Note that in case $\frac{\mu_{11}}{\mu} < 1$, the theorem only regards the size of the in-components and out-components. Applying Lemma 4.4 it follows that the size of the strongly connected component to which a vertex belongs is upper bounded by the minimum of the size of its in-component and out-component. Hence the fact that w.h.p. the in-component and the out-component do not scale linear in n for any vertex, implies that w.h.p. no GSCC is present. Furthermore we remark in case $\frac{\mu_{11}}{\mu} > 1$ and $p_0^+, p_0^- > 0$ the theorem assures that a GIN and GOUT exist w.h.p.

5 Proofs of Theorems 2.2 and 2.3

5.1 Concentration inequalities

We consider several concentration results that will be used in the proof of Theorem 2.2. Concentration inequalities are important as we deal with two sources of randomness: the initial multigraph is random and percolation randomly removes edges.

Theorem 5.1. *(Hoeffding's inequality)[17] Let $X_1, X_2, \dots, X_n$ be independent random variables. Suppose that $a_i \le X_i \le b_i$ for all $i \in \{1, 2, \dots, n\}$ and define $c_i = b_i - a_i$. Furthermore define $S_n = \sum_{i=1}^n X_i$. Then there holds*

$$\mathbb{P}\left[|S_n - \mathbb{E}\left[S_n\right]| > t\right] \le 2 \exp\left(-\frac{2t^2}{\sum_{i=1}^n c_i^2}\right). \tag{29}$$

The following concentration inequality is a corollary of a theorem by McDiarmid.

Theorem 5.2. *[18, Theorem 7.4] Let (V, d) be a finite metric space. Suppose there exists a sequence $\mathcal{P}_0, \mathcal{P}_1, \dots, \mathcal{P}_s$ of increasingly refined partitions with $\mathcal{P}_0$ the trivial partition consisting of V and $\mathcal{P}_s$ the partition where each element of V is a partition element on its own. Take a sequence of positive integers $c_0, c_1, \dots, c_s$ such that for all $k \in \{1, 2, \dots, s\}$ and any $A, B \in \mathcal{P}_k$ with C satisfying $A, B \subset C \in \mathcal{P}_{k-1}$ there exists a bijection $\phi : A \to B$ with $d\left(x, \phi(x)\right) \le c_k$ for all $x \in A$. Let the function $f : V \to \mathbb{R}$ satisfy $|f(x) - f(y)| \le d(x, y)$ for all $x, y, \in V$. Then for X uniformly distributed over V and any $t > 0$ there holds*

$$\mathbb{P}\left[|f(X) - \mathbb{E}\left[f(X)\right]| > t\right] \le 2 \exp\left(-\frac{2t^2}{\sum_{k=0}^s c_k^2}\right).$$

Corollary 5.3. *Consider two finite sets A_0 and A_1 with $|A_0| = a_0$ and $|A_1| = a_1$. Let $S := \cup_{i \in \{0,1\}} \{(x, i) \mid x \in A_i\}$. A subset of S containing b_0 elements with $i = 0$ and b_1 elements with $i = 1$ is called a (b_0, b_1)-subset of S. Let V be the space of all (b_0, b_1)-subsets of S. Let $f : V \to \mathbb{R}$ be a function such that for any $B, B' \in V$ there holds $|f(B) - f(B')| \le |B \triangle B'|$. Here $B \triangle B'$ denotes the symmetric difference, i.e. $B \triangle B' = (B \cup B') \setminus (B \cap B')$. Then for X distributed uniformly over V and any $t > 0$ there holds*

$$\mathbb{P}\left[|f(X) - \mathbb{E}\left[f(X)\right]| > t\right] \le 2 \exp\left(-\frac{t^2}{8(b_0 + b_1)}\right). \tag{30}$$

Proof. Consider a (b_0, b_1)-subset of S. Assign each element a unique number from the index set $\{1, 2, \ldots, b_0 + b_1\}$, such that for all elements with $i = 0$ this number is smaller than $b_0 + 1$. Note that this implies that for each element with $i = 1$ its index is larger than b_0. A (b_0, b_1)-subset of S with such a numbering is called a (b_0, b_1)-ordering of S. Define W to be the set of all (b_0, b_1)-orderings of S. The function $f : V \to \mathbb{R}$ can be extended to a function $f : W \to \mathbb{R}$ by regarding each (b_0, b_1)-ordering as (b_0, b_1)-subset. This extension respects the relation $|f(x) - f(y)| \leq x \triangle y$, *i.e.* it holds for $x, y \in W$ as well. This is true since for any two orderings x, y their symmetric difference as (b_0, b_1)-orderings is bounded from bellow by their symmetric difference as (b_0, b_1)-subsets. The next step in proving equation (30) is applying Theorem 5.2 to the metric space $(W, \triangle)$.

We will now define a sequence of refined partitions on W using the notion of an i-prefix. An i-prefix determines the first i elements of an ordering. This allows for all $k \in \{0, 1, \ldots, b_0 + b_1\}$ to construct the partition $\mathcal{P}_k$ by defining its elements to be the sets of orderings with the same k-prefix. The partition $\mathcal{P}_0$ is the trivial partition consisting of W. As each (b_0, b_1)-ordering has $b_0 + b_1$ elements, $\mathcal{P}_{b_0+b_1}$ will be the partition where each element is a single ordering. Next the values c_k need to be determined. Take $B, D \in \mathcal{P}_k$ with C satisfying $B, D \subset C \in \mathcal{P}_{k-1}$. This implies that any ordering in B has the same $k - 1$-prefix as an ordering in D. Furthermore these orderings must differ at the k^{th} element. The remaining $b_0 + b_1 - k$ elements can be any element that is not present in the k-prefix that lead to a valid ordering. Denote the k^{th} element of any ordering in B by $a_{B,k}$. Similarly let $a_{D,k}$ denote the k^{th} element of any ordering in D. Define the bijection $\phi : B \to D$ by taking $x \in B$ and mapping its k^{th} element to $a_{D,k}$. If x contains $a_{D,k}$ at some position $l > k$, map the l^{th} element of x to $a_{B,k}$. All the other elements are unchanged by the bijection. By definition this is an element of D.

According to definition of ϕ for any $x \in B$ we have $|x \triangle \phi(x)| \leq 4$. Thus we may take $c_k = 4$ for all $k \in \{1, 2, \ldots, b_0 + b_1\}$. Applying Theorem 5.2 we find that for distributing X uniformly random over W and any $t > 0$ there holds

$$\mathbb{P}\left[|f(X) - \mathbb{E}\left[f(X)\right]| > t\right] \leq 2 \exp\left(-\frac{t^2}{8(b_0 + b_1)}\right).$$

Remark that each element in V gives rise to $b_0! + b_1!$ different orderings. All these orderings have the same value under f. Thus the probability that $f(X) = c$ does not change when we take X to be a uniformly random element of V instead of W. Together with the above equation, this proves the claim. □

5.2 Bond percolation

Theorem 2.2 is proven separately for bond and site percolation, by using similar techniques as in the proof of [4, Theorem 1.1], which determines the percolation threshold in undirected graphs. Although Theorem 2.2 gives the percolation threshold for simple graphs, we work with uniformly random configurations and their induced multigraphs that obey a proper degree progression $(\mathbf{d}^n)_{n \in \mathbb{N}}$. The results on the multigraphs are then transferred to the simple graphs using Corollary 3.8, as for any proper degree progression there holds $d_{\max} \leq \frac{n^{1/12}}{\ln(n)} = \mathcal{O}(\sqrt{n})$. This means that Theorem 4.10 is applied to configurations and their induced multigraphs rather than simple graphs.

Let $\mathbf{D}^n$ be the random variable for the degree sequence after percolation and $\mathbf{d}^n_\pi$ a possible value of $\mathbf{D}^n$. Percolation removes edges in a graph, thus in a configuration it removes in-stubs together with their matched out-stubs. Let $W^{-,\pi}$ and $W^{+,\pi}$ denote the in-stubs and out-stubs surviving percolation. Conditional on $\mathbf{D}^n = \mathbf{d}^n_\pi$ these stubs are in one-to-one correspondence with $W^-_{\mathbf{d}^n_\pi}$ and $W^+_{\mathbf{d}^n_\pi}$, the sets of stubs inducing the degree sequence $\mathbf{d}^n_\pi$. The one-to-one correspondence follows from the fact that all vertices have the same amount of in-stubs (respectively out-stubs) in $W^{-,\pi}(W^{+,\pi})$ as in $W^-_{\mathbf{d}^n_\pi}(W^+_{\mathbf{d}^n_\pi})$. Thus any mapping sending an in-stub of vertex i of $W^{-,\pi}$ to an in-stub of the same vertex in $W^-_{\mathbf{d}^n_\pi}$, as well as an out-stub of i of $W^{+,\pi}$ to an out-stub of i in $W^+_{\mathbf{d}^n_\pi}$, induces a bijection. Let us fix such a bijection. This induces a one-to-one correspondence between the configuration on $(W^{-,\pi}, W^{+,\pi})$ and the configuration on $\left(W^-_{\mathbf{d}^n_\pi}, W^+_{\mathbf{d}^n_\pi}\right)$. Let D_n be the probability space containing all degree sequences $\mathbf{d}^n_\pi$ that can be obtained by applying percolation to a random configuration on (W^-, W^+). The probability assigned to each $\mathbf{d}^n_\pi$ is the probability that it is induced by $(W^{-,\pi}, W^{+,\pi})$. The probability space for the degree progression $\left(\mathbf{d}^n_\pi\right)_{n\in\mathbb{N}}$ is the product space $D = \prod_{n=1}^\infty D_n$ with the product measure ν.

In the reminder of this section, we show in Section 5.2.1 that conditional on the degree sequence after percolation, each configuration on $W^-_{\mathbf{d}^n_\pi}$ and $W^+_{\mathbf{d}^n_\pi}$ is equally likely. In Section 5.2.2, we determine the limit of the expected number of vertices with degree (j,k) after percolation and show that there is always a positive fraction of dead ends, $p_0^+, p_0^- > 0$. Combining these results in Section 5.2.3, the proof of Theorem 2.2 is completed by showing that an element of D is ν-a.s. proper, which allows us to apply Theorem 4.10.

5.2.1 A percolated configuration is a uniformly random configuration

Consider a uniformly random configuration on (W^-, W^+). In this section we show that that conditional on the degree sequence after percolation, the configuration on $(W^{-,\pi}, W^{+,\pi})$ is also a uniformly random one. The proof is split into two lemma's.

Lemma 5.4. *Suppose that l out of m edges survive bond percolation applied to a uniformly random configuration $\mathcal{M}$ on (W^-, W^+). Then the surviving stubs $W^{-,\pi} \subset W^-$ and $W^{+,\pi} \subset W^+$ are uniformly distributed amongst all pairs of subsets of W^- and W^+ of size l.*

Proof. As l edges survive percolation, there holds $|W^{-,\pi}| = |W^{+,\pi}| = l$. The probability that $W^{-,\pi} \subset W^-$ and $W^{+,\pi} \subset W^+$ are the stubs surviving percolation equals the probability that all points in $W^{-,\pi}$ have their match in $W^{+,\pi}$ and that exactly these l matches survive percolation. Since the graph contains m matches of which l survive percolation, the probability that exactly those l matches remain is $\frac{1}{\binom{m}{l}}$.

It is left to investigate the probability that all stubs in $W^{-,\pi}$ have their match in $W^{+,\pi}$. This implies that $\mathcal{M}$ must decompose into a perfect bipartite matching of $W^{-,\pi}$ with $W^{+,\pi}$ and a perfect bipartite matching of $W^- \setminus W^{-,\pi}$ with $W^+ \setminus W^{+,\pi}$. Between two sets of size l there are $l!$ perfect bipartite matchings, hence the probability that $\mathcal{M}$ decomposes as desired, is $l!(m-l)!/m!$. Thus the probability that $(W^{-,\pi}, W^{+,\pi})$ are the stubs surviving percolation is

$$\frac{l!\,(m-l)!}{m!}\frac{1}{\binom{m}{l}} = \frac{1}{\binom{m}{l}^2}.$$

This is the probability that $W^{-,\pi}$ is a uniformly random subset of size l of W^- and $W^{+,\pi}$ is a uniformly random subset of W^+ of size l. $\square$

Lemma 5.5. *Apply bond percolation to a uniformly random configuration on* (W^-, W^+) *obeying the degree sequence* $\mathbf{d}^n$*. Conditional on having degree sequence* $\mathbf{d}^n_\pi$ *after bond percolation, i.e.* $\mathbf{D}^n = \mathbf{d}^n_\pi$*, all configurations of* $W^-_{\mathbf{d}^n_\pi}$ *with* $W^+_{\mathbf{d}^n_\pi}$ *are equally likely.*

Proof. The goal is to show that all configurations on $W^-_{\mathbf{d}^n_\pi}$ and $W^+_{\mathbf{d}^n_\pi}$ have equal probability, given that $\mathbf{D}^n = \mathbf{d}^n_\pi$. This implies that for any perfect bipartite matching $\mathcal{M}^\pi$ of $W^-_{\mathbf{d}^n_\pi}$ with $W^+_{\mathbf{d}^n_\pi}$ there must hold

$$\mathbb{P}\left[\mathcal{M}^\pi | \mathbf{D}^n = \mathbf{d}^n_\pi\right] = \frac{1}{l!}.$$

Here l denotes the sum of the in-degrees of $\mathbf{d}^n_\pi$. First, rewrite this probability using $\mathbb{P}\left[|W^{-,\pi}| = l \middle| \mathbf{D}^n = \mathbf{d}^n_\pi\right] = 1$ and due to the law of total probability obtain

$$\mathbb{P}\left[\mathcal{M}^\pi | \mathbf{D}^n = \mathbf{d}^n_\pi\right] = \mathbb{P}\left[\mathcal{M}^\pi \middle| |W^{-,\pi}| = l, \mathbf{D}^n = \mathbf{d}^n_\pi\right].$$

Applying Bayes' formula to the right hand side of the previous equation gives

$$\mathbb{P}\left[\mathcal{M}^\pi | \mathbf{D}^n = \mathbf{d}^n_\pi\right] = \frac{\mathbb{P}\left[\mathcal{M}^\pi \cap \mathbf{D}^n = \mathbf{d}^n_\pi \middle| |W^{-,\pi}| = l\right]}{\mathbb{P}\left[\mathbf{D}^n = \mathbf{d}^n_\pi \middle| W^{-,\pi}| = l\right]}. \tag{31}$$

It will now be shown that this expression equals $\frac{1}{l!}$. First determine the value of $\mathbb{P}\left[\mathbf{D}^n = \mathbf{d}^n_\pi \middle| W^{-,\pi}| = l\right]$. Let $S(\mathbf{d}^n_\pi)$ be the collection of pairs of subsets of (W^-, W^+) that induce the degree sequence $\mathbf{d}^n_\pi$. Recalling that $|W^{-,\pi}| = |W^{+,\pi}|$, we see that for any pair of subsets in $S(\mathbf{d}^n_\pi)$, both sets must contain l elements. In combination with Lemma 5.4 this implies

$$\mathbb{P}\left[\mathbf{D}^n = \mathbf{d}^n_\pi \middle| |W^{-,\pi}| = l\right] = \frac{|S(\mathbf{d}^n_\pi)|}{\binom{m}{l}^2}.$$

Next we investigate $\mathbb{P}\left[\mathcal{M}^\pi \cap \mathbf{D}^n = \mathbf{d}^n_\pi \middle| |W^{-,\pi}| = l\right]$. By definition of $S(\mathbf{d}^n_\pi)$, $\mathbf{D}^n = \mathbf{d}^n_\pi$ implies that $(W^{-,\pi}, W^{+,\pi}) \in S(\mathbf{d}^n_\pi)$. So let $(W^{-,\pi}, W^{+,\pi}) \in S(\mathbf{d}^n_\pi)$, then we aim to find the probability that the configuration on these stubs induces the configuration $\mathcal{M}^\pi$ on $\left(W^-_{\mathbf{d}^n_\pi}, W^+_{\mathbf{d}^n_\pi}\right)$. As we fixed a bijection between $(W^{-,\pi}, W^{+,\pi})$ and $\left(W^-_{\mathbf{d}^n_\pi}, W^+_{\mathbf{d}^n_\pi}\right)$, exactly one configuration of $(W^{-,\pi}, W^{+,\pi})$ induces the configuration $\mathcal{M}^\pi$ on $\left(W^-_{\mathbf{d}^n_\pi}, W^+_{\mathbf{d}^n_\pi}\right)$. However we are free to choose the configuration on $(W^- \setminus W^{-,\pi}, W^+ \setminus W^{+,\pi})$. Thus assuming that $(W^{-,\pi}, W^{+,\pi}) \in S(\mathbf{d}^n_\pi)$, the probability that it induces the configuration $\mathcal{M}^\pi$ on $(W^{-,\pi}, W^{+,\pi})$ is

$$\frac{(m-l)!}{m!}.$$

Thus for any collection of remaining stubs that induce the right degree sequence, it has probability $\frac{(m-l)!}{m!}$ to induce the right matching. However, as we condition only on the size of $W^{-,\pi}$, we also must take into account the probability that exactly the desired l edges survive percolation. Hence, we have:

$$\mathbb{P}\left[\mathcal{M}^\pi \cap \mathbf{D}^n = \mathbf{d}^n_\pi \middle| |W^{-,\pi}| = l\right] = \frac{(m-l)!}{m!} \frac{|S(\mathbf{d}^n_\pi)|}{\binom{m}{l}}.$$

Plugging our findings back in equation (31), we obtain:

$$\mathbb{P}\left[\mathcal{M}^\pi | \mathbf{D}^n = \mathbf{d}^n_\pi\right] = \frac{1}{l!}.$$

□

5.2.2 The expected number of vertices with degree (j,k) after bond percolation

Let $N^{\pi}_{j,k}(n)$ be the number of vertices in the percolated graph (or configuration) with in-degree j and out-degree k. In this section we show that the limit

$$\lim_{n\to\infty} \frac{\mathbb{E}\left[N^{\pi}_{j,k}(n)\right]}{n} := p^{\text{bond}}_{j,k} \tag{32}$$

exists and determine its value for $j,k\in\mathbb{N}_0$.

For all values of (j,k) with $j>d_{\max}$ or $k>d_{\max}$, or both, the limit of equation (32) is easily evaluated to be $p^{\text{bond}}_{j,k}=0$. Let $j,k\le d_{\max}$, remark that

$$\mathbb{E}\left[N^{\pi}_{j,k}(n)\right] = \sum_{l=0}^{m} \mathbb{E}\left[N^{\pi}_{j,k}(n)\,\middle|\,|W^{-,\pi}|=l\right]\mathbb{P}\left[|W^{-,\pi}|=l\right]. \tag{33}$$

This conditional expectation of $N^{\pi}_{j,k}(n)$ can be written as

$$\mathbb{E}\left[N^{\pi}_{j,k}(n)\,\middle|\,|W^{-,\pi}|=l\right] = \sum_{d^-=j}^{d_{\max}}\sum_{d^+=k}^{d_{\max}} N_{d^-,d^+}(n)\,\mathbb{P}\left[(d^-,d^+)\to(j,k)\,\middle|\,|W^{-,\pi}|=l\right].$$

Here (d^-,d^+) is the degree of a vertex before percolation and (j,k) is its degree after percolation. As Lemma 5.4 implies that the surviving stubs are chosen uniformly at random, conditional on the size of $W^{-,\pi}$, there holds

$$\mathbb{P}\left[(d^-,d^+)\to(j,k)\,\middle|\,|W^{-,\pi}|=l\right] = \binom{d^-}{j}\frac{\binom{m-d^-}{l-j}}{\binom{m}{l}}\binom{d^+}{k}\frac{\binom{m-d^+}{l-k}}{\binom{m}{l}}.$$

This value can be further approximated for $l\in I := \left[m\pi-\ln(n)\sqrt{n}, m\pi+\ln(n)\sqrt{(n)}\right]$. As the edges are removed independently of each other, the size of $W^{-,\pi}$ is the sum of m independent Bernoulli variables, each having expectation π. Applying Theorem 5.1 yields

$$\mathbb{P}\left[\big|\,|W^{-,\pi}|-m\pi\big|>\ln(n)\sqrt{n}\right]\le \exp\left[-\Omega(\ln^2(n))\right]. \tag{34}$$

This implies that $\mathbb{P}[l\notin I]=o\left(\frac{1}{n^3}\right)$. Fountoulakis [4] showed that for $d_{\max}\le n^{1/9}$ and $l\in I$ there holds

$$\frac{\binom{2m-d}{2l-j}}{\binom{2m}{2k}} = \pi^j(1-\pi)^{d-j}\left(1+\mathcal{O}\left(\frac{\ln(n)}{n^{7/18}}\right)\right).$$

As we consider proper degree progressions, there holds $d_{\max}\le\frac{n^{1/12}}{\ln(n)}$, and since $\frac{n^{1/12}}{\ln(n)}<n^{1/9}$ for all $n\ge 3$, an analogous argument to that of Fountoulakis gives:

$$\mathbb{P}\left[\,(d^-,d^+)\to(j,k)\,|\,|W^{-,\pi}|=l\right] = \binom{d^-}{j}\binom{d^+}{k}\pi^{j+k}(1-\pi)^{d^-+d^+-j-k}\left(1+\mathcal{O}\left(\frac{\ln(n)}{n^{7/18}}\right)\right),$$

for all $d^-,d^+\le d_{\max}$ and $l\in I$. This allows to determine $\mathbb{E}\left[N^{\pi}_{j,k}(n)\,|\,|W^{-,\pi}|=l\right]$ for all $l\in I$. In combination with equation (33), $\mathbb{P}[l\notin I]=o\left(\frac{1}{n^3}\right)$, and the fact that $N^{\pi}_{j,k}(n)\le n$, we find

$$\mathbb{E}\left[N^{\pi}_{j,k}(n)\right] = \left(1+\mathcal{O}\left(\frac{\ln(n)}{n^{7/18}}\right)\right)\sum_{d^-=j}^{d_{\max}}\sum_{d^+=k}^{d_{\max}} N_{d^-,d^+}(n)\binom{d^-}{j}\binom{d^+}{k}\pi^{j+k}(1-\pi)^{d^-+d^+-j-k}+o\left(\frac{1}{n^3}\right), \tag{35}$$

using same argument as for $\mathbb{E}\left[D_i'(n)\right]$ in [4, p. 344]. With this approximation of $\mathbb{E}\left[N_{j,k}^{\pi}(n)\right]$, we can show that the limit of equation (32) exists. This requires for all $\epsilon > 0$ the existence of $\kappa(\epsilon)$ and $N(\epsilon)$ such that for all $n > N$,

$$\frac{1}{n}\sum_{\substack{(d^-,d^+)=(0,0)\\ d^-\geq\kappa+1 \text{ or } d^+\geq\kappa+1}}^{(d_{\max},d_{\max})} N_{d^-,d^+}(n)\binom{d^-}{j}\binom{d^+}{k}\pi^{j+k}(1-\pi)^{d^-+d^+-j-k} \leq \frac{1}{n}\sum_{\substack{(d^-,d^+)=(0,0)\\ d^-\geq\kappa+1 \text{ or } d^+\geq\kappa+1}}^{(d_{\max},d_{\max})} N_{d^-,d^+}(n) < \epsilon. \quad (36)$$

The left inequality follows from the binomial theorem, which implies that $\sum_{j=0}^{d^-}\binom{d^-}{j}\pi^j(1-\pi)^{d^--j} = \sum_{k=0}^{d^+}\binom{d^+}{k}\pi^k(1-\pi)^{d^+-k} = 1$, yielding that $\binom{d^-}{j}\binom{d^+}{k}p^{j+k}(1-p)^{d^-+d^+-j-k} \leq 1$ for all $j \leq d^-, k \leq d^+$. The right inequality holds since $\lim_{n\to\infty}\frac{N_{j,k}(n)}{n} = p_{j,k}$ for $j,k \geq 0$, which follows from the degree progression being proper. Equations (35) and (36) combined together are sufficient for the existence of the limit in equation (32). Moreover, the value of the limit is:

$$\lim_{n\to\infty}\frac{\mathbb{E}\left[N_{j,k}^{\pi}(n)\right]}{n} = p_{j,k}^{\text{bond}} = \sum_{d^-=j}^{\infty}\sum_{d^+=k}^{\infty} p_{d^-,d^+}\binom{d^-}{j}\binom{d^+}{k}\pi^{j+k}(1-\pi)^{d^--j+d^+-k}, \quad (37)$$

and the generating function for $p_{j,k}^{\text{bond}}$ is given by:

$$U_{\pi}(x,y) = U(1-\pi+\pi x, 1-\pi+\pi y). \quad (38)$$

We will now show that $p_{j,k}^{\text{bond}}$ is normalized and obeys equation (12). The normalization follows from the binomial theorem and that fact that $p_{j,k}$ is normalized:

$$\begin{aligned}
&\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{d^-=j}^{\infty}\sum_{d^+=k}^{\infty} p_{d^-,d^+}\binom{d^-}{j}\binom{d^+}{k}\pi^{j+k}(1-\pi)^{d^--j+d^+-k}\\
&\quad= \sum_{d^-=0}^{\infty}\sum_{d^+=0}^{\infty} p_{d^-,d^+}\sum_{j=0}^{d^-}\binom{d^-}{j}\pi^j(1-\pi)^{d^--j}\sum_{k=0}^{d^+}\binom{d^+}{k}\pi^k(1-\pi)^{d^+-k}\\
&\quad= \sum_{d^-=0}^{\infty}\sum_{d^+=0}^{\infty} p_{d^-,d^+} = 1.
\end{aligned}$$

By using the equality $\sum_{k=0}^{n}\binom{n}{k}kx^ky^{n-k} = xn(x+y)^{n-1}$, which can be obtained by applying $x\frac{d}{dx}$ to the binomial theorem, we find that

$$\begin{aligned}
\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} j p_{j,k}^{\text{bond}} &= \sum_{j=0}^{\infty} j \sum_{k=0}^{\infty}\sum_{d^-=j}^{\infty}\binom{d^-}{j}\pi^j(1-\pi)^{d^--j}\sum_{d^+=k}^{\infty}\binom{d^+}{k}\pi^k(1-\pi)^{d^+-k}p_{d^-,d^+}\\
&= \sum_{d^-=0}^{\infty}\sum_{d^+=0}^{\infty} p_{d^-,d^+}\sum_{j=0}^{d^-} j\binom{d^-}{j}\pi^j(1-\pi)^{d^--j}\sum_{k=0}^{d^+}\binom{d^+}{k}\pi^k(1-\pi)^{d^+-k}\\
&= \pi\sum_{d^-=0}^{\infty}\sum_{d^+=0}^{\infty} d^- p_{d^{\cdot}d^+},
\end{aligned}$$

and by symmetry,

$$\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} k p_{j,k}^{\text{bond}} = \pi \sum_{d^-=0}^{\infty}\sum_{d^+=0}^{\infty} d^+ p_{d^, d^+}.$$

This proves that

$$\mu^{\pi,\text{bond}} := \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} j p_{j,k}^{\text{bond}} = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} k p_{j,k}^{\text{bond}} = \pi\mu. \tag{39}$$

Thus $p_{j,k}^{\text{bond}}$ indeed is normalized and satisfies equation (12). In analogy to (39), we also find that

$$\mu_{11}^{\pi,\text{bond}} := \sum_{j,k=0}^{\infty} jk p_{j,k}^{\text{bond}} = \pi^2 \sum_{d^-=0}^{\infty}\sum_{d^+=0}^{\infty} d^- d^+ p_{d^-,d^+} = \pi^2 \mu_{11}. \tag{40}$$

We now see that after percolation with $0 < \pi < 1$ the fraction of dead-end vertices having no out edges is positive:

$$p_0^{+,bond} = \sum_{j=0}^{\infty} \frac{j p_{j,0}^{\text{bond}}}{\mu^{\pi,\text{bond}}} = \frac{1}{\pi\mu} \sum_{d^+,d^-\geq 0} d^- \pi (1-\pi)^{d^+} p_{d^-,d^+} > 0, \tag{41}$$

and a similar bound holds for vertices with no in edges, $p_0^{-,bond} > 0$. Therefore, assuming that $p_{j,k}^{\text{bond}}$ is the degree distribution of the percolated graph, Equations (41), (39) and (40) together justify applying Theorem 4.10 to $p_{j,k}^{\text{bond}}$ and hence formally defining the percolation threshold as such a value of $\pi = \hat{\pi}^{\text{bond}}$ that

$$\sum_{j,k=0}^{\infty} jk p_{j,k}^{\text{bond}} = \sum_{j,k=0}^{\infty} j p_{j,k}^{\text{bond}}. \tag{42}$$

By plugging equations 40 and (39) into the latter equation gives:

$$\hat{\pi}^{\text{bond}} = \frac{\sum_{d^-=0}^{\infty}\sum_{d^+=0}^{\infty} d^- p_{d^-,d^+}}{\sum_{d^-=0}^{\infty}\sum_{d^+=0}^{\infty} d^- d^+ p_{d^-,d^+}} = \frac{\mu}{\mu_{11}}.$$

In Section 5.2.3 we show that $p_{j,k}^{\text{bond}}$ is indeed ν-a.s the degree distribution of the percolated graph, and therefore, $\pi_c^{\text{bond}} = \hat{\pi}^{\text{bond}}$.

5.2.3 Determining π_c^{bond} and c^{bond}

To prove the equality $\pi_c^{\text{bond}} = \hat{\pi}^{\text{bond}}$ it is sufficient to apply Theorem 4.10 to the percolated multigraph progression $\left(\tilde{G}_{\mathbf{d}^n}^{\pi}\right)_{n\in\mathbb{N}}$. Theorem 4.10 requires $\left(\tilde{G}_{\mathbf{d}^n}^{\pi}\right)_{n\in\mathbb{N}}$ to obey a *proper* degree progression. We make use of Lemma 5.5, stating that instead of actually removing edges, one may view the percolated multigraph $\tilde{G}_{\mathbf{d}^n}^{\pi}$ as a uniformly random *configuration* obeying the percolated degree sequence $\mathbf{d}_{\pi}^n$, which requires $\left(\mathbf{d}_{\pi}^n\right)_{n\in\mathbb{N}}$ to be proper. We show that $\left(\mathbf{d}_{\pi}^n\right)_{n\in\mathbb{N}}$ is indeed ν-a.s. proper as an element of D, and hence, Theorem 4.10 is applicable to $\left(\tilde{G}_{\mathbf{d}^n}^{\pi}\right)_{n\in\mathbb{N}}$.

This means that Theorem 4.10 may be then applied to almost all degree progressions $\left(\mathbf{d}_\pi^n\right)_{n\in\mathbb{N}}$ to determine π_c^{bond} and c^{bond} for random multigraphs. Finally we apply a variant of Lemma 3.7 and Corollary 3.8 and show that similar assertions hold for a graph progression $(G_n^\pi)_{n\in\mathbb{N}}$ for percolated *simple* graphs obeying $(\mathbf{d}^n)_{n\in\mathbb{N}}$, hence proving Theorem 2.3 for the case of bond percolation. In the remainder of this section, we make the above-stated argument formal.

Definition 4.8 implies that each proper degree progression must be feasible. For any feasible $\left(\mathbf{d}_\pi^n\right)_{n\in\mathbb{N}} \in D$, it can be shown that it is also proper, using the fact that $(\mathbf{d}^n)_{n\in\mathbb{N}}$ is proper. So we will first prove that any $\left(\mathbf{d}_\pi^n\right)_{n\in\mathbb{N}} \in D$ is ν-a.s. feasible. By Definition 3.3 a degree progression $\left(\mathbf{d}_\pi^n\right)_{n\in\mathbb{N}}$ is feasible if $\frac{N_{j,k}^\pi(n)}{n}$ and its first, first mixed, and second moments converge to those of a bivariate distribution obeying equation (12). For now, we replace the constraint of each degree sequence being graphical with each degree sequence being valid, as we consider multigraphs at this point. The assertion that each $\mathbf{d}_\pi^n$ is valid, is a direct consequence of the fact that $\mathbf{d}^n$ is valid, which holds when $(\mathbf{d}^n)_{n\in\mathbb{N}}$ is proper. This implies that Theorem 4.10 can be readily applied to multigraphs and configurations. When conditioning on the graph before percolation being simple later on in the proof, we implicitly replace valid by graphical. Conditioning on the graph before percolation being simple will ensure that the percolated graph is simple as well.

Now we will show that ν-a.s. $\frac{N_{j,k}^\pi(n)}{n}$ and its first, first mixed and second moments converge to $p_{j,k}^{\text{bond}}$ and its corresponding moments. The first step is showing that

$$\lim_{n\to\infty} \frac{N_{j,k}^\pi(n)}{n} = p_{j,k}^{\text{bond}},\ \nu\text{-a.s. for } j,k\in\mathbb{N}_0, \tag{43}$$

for which, it suffices to show (see e.g. [19, Lemma 6.8]) that for all $\epsilon > 0$,

$$\sum_{n=1}^{\infty} \mathbb{P}\left[\left|\frac{1}{n} N_{j,k}^\pi(n) - p_{j,k}^{\text{bond}}\right| > \epsilon\right] < \infty. \tag{44}$$

By definition of $p_{j,k}^{\text{bond}}$, for any fixed $\epsilon > 0$ there is K such that for all $n > K$

$$\left|\frac{1}{n}\mathbb{E}\left[N_{j,k}^\pi(n)\right] - p_{j,k}^{\text{bond}}\right| \le \frac{\epsilon}{2}.$$

This implies that

$$\mathbb{P}\left[\left|\frac{1}{n} N_{j,k}^\pi(n) - p_{j,k}^{\text{bond}}\right| > \epsilon\right] \le \mathbb{P}\left[\frac{1}{n}\left|N_{j,k}^\pi(n) - \mathbb{E}\left[N_{j,k}^\pi(n)\right]\right| > \frac{\epsilon}{2}\right].$$

Lemma 5.4 states that conditional on $|W^{-,\pi}| = l$, the stubs surviving percolation $(W^{-,\pi}, W^{+,\pi})$ are uniformly distributed amongst all pairs of subsets of (W^-, W^+) of size l. The value $N_{j,k}^\pi(n)$ is a function of $W^{-,\pi} \cup W^{+,\pi}$. Furthermore for two sets $W^{-,\pi} \cup W^{+,\pi}$ and $W^{-,\pi'} \cup W^{+,\pi'}$ their values of $N_{j,k}^\pi(n)$ differ by at most the number of elements in the symmetric difference of $W^{-,\pi} \cup W^{+,\pi}$ and $W^{-,\pi'} \cup W^{+,\pi'}$. This implies that for $A_0 = W^-, A_1 = W^+, b_0 = b_1 = l$ and $N_{j,k}^\pi(n)$ as function f the requirements of Corollary 5.3 are fulfilled. Applying this corollary gives

$$\mathbb{P}\left[\left|N_{j,k}^\pi(n) - \mathbb{E}\left[N_{j,k}^\pi(n)\right]\right| > \frac{n\epsilon}{2} \mid |W^{-,\pi}| = l\right] \le 2\exp\left(\frac{\epsilon^2 n^2}{64 l^2}\right).$$

If $l \in I$, this probability is $o\left(\frac{1}{n^3}\right)$. By equation (34) the probability that $l \notin I$ is $o\left(\frac{1}{n^3}\right)$. Combining these observations we find that for all $\epsilon > 0$ the terms in (44) are vanishing:

$$\mathbb{P}\left[\left|N^{\pi}_{j,k}(n) - \mathbb{E}\left[N^{\pi}_{j,k}(n)\right]\right| > n\epsilon\right] = o\left(\frac{1}{n^3}\right). \tag{45}$$

which in turn proves that the limit in equation (43) holds ν-a.s.

To prove that $\left(\mathbf{d}^n_{\pi}\right)_{n\in\mathbb{N}}$ is ν-a.s. feasible, it remains to show that the first, first mixed and second moments of $\frac{N^{\pi}_{j,k}(n)}{n}$ converge ν-a.s. to those of $p^{\text{bond}}_{j,k}$. In Section 5.2.2 we showed that $\sum_{j,k=0}^{\infty} jp^{\text{bond}}_{j,k} = \sum_{j,k=0}^{\infty} kp^{\text{bond}}_{j,k}$, which implies that $\sum_{j,k=0}^{\infty} jN^{\pi}_{j,k}(n) = \sum_{j,k=0}^{\infty} kN^{\pi}_{j,k}(n)$. We can thus restrict ourself to showing that

$$Q := \lim_{n\to\infty} \frac{\sum_{j,k=0}^{\infty} jN^{\pi}_{j,k}(n)}{n} = \pi\mu,\ \nu\text{-a.s} \tag{46}$$

to show that both first moments converge. To determine the limit in equation (46), let us define

$$Q'_n := \frac{1}{n}\sum_{j,k=0}^{\infty} jN^{\pi}_{j,k}(n),$$
$$X_{\kappa,n} := \frac{1}{n}\sum_{j=0}^{\kappa}\sum_{k=0}^{\kappa} jN^{\pi}_{j,k}(n),$$

and remark that $X_{\kappa,n} \le Q'_n$. Since $(\mathbf{d}^n)_{n\in\mathbb{N}}$ is a proper degree progression, for all $\epsilon > 0$ there exists $\widetilde{\kappa}(\epsilon), m(\epsilon)$ such that for all $\kappa > \widetilde{\kappa}$ and $n > m$

$$\frac{1}{n}\sum_{\substack{(d^-,d^+)=(0,0)\\ d^->\kappa \text{ or } d^+>\kappa}}^{(d_{\max},d_{\max})} jN_{j,k}(n) < \epsilon. \tag{47}$$

Thus for $\kappa > \widetilde{\kappa}$ there holds $X_{\kappa,n} \le Q'_n \le X_{\kappa,n} + \epsilon$. This implies that if ν-a.s.

$$\lim_{n\to\infty} X_{\kappa,n} = \sum_{j=0}^{\kappa}\sum_{k=0}^{\kappa} jp^{\text{bond}}_{j,k} := \tilde{X}_{\kappa}, \tag{48}$$

then $\lim_{n\to\infty} Q'_n = Q$ ν-a.s. as well [4, (3.5)]. Thus the goal is to prove equation (48). Applying Lemma 6.8 [19] we can show that this limit ν-a.s. holds, if for any $\epsilon > 0$ there holds

$$\sum_{n=1}^{\infty}\mathbb{P}\left[\left|X_{\kappa,n} - \tilde{X}_{\kappa}\right| > \epsilon\right] < \infty. \tag{49}$$

We show this analogous to the proof of equation (44). By the definition of $X_{\kappa,n}, \tilde{X}_{\kappa}$ and $p^{\text{bond}}_{j,k}$, for all $\epsilon > 0$ there exists $\widetilde{N}$ such that for all $n > \widetilde{N}$

$$\left|\mathbb{E}\left[X_{\kappa,n}\right] - \tilde{X}_{\kappa}\right| < \frac{\epsilon}{2}.$$

Combing this with the reverse triangle inequality, we find for any $\epsilon > 0$

$$\mathbb{P}\left[\left|X_{\kappa,n} - \tilde{X}_{\kappa}\right| > \epsilon\right] \le \mathbb{P}\left[\left|X_{\kappa,n} - \mathbb{E}\left[X_{\kappa,n}\right]\right| > \frac{\epsilon}{2}\right].$$

Remark that

$$|X_{\kappa,n} - \mathbb{E}\left[X_{\kappa,n}\right]| = \frac{1}{n}\sum_{j=0}^{\kappa}\sum_{k=0}^{\kappa} j\left(N_{j,k}^{\pi}\left(n\right) - \mathbb{E}\left[N_{j,k}^{\pi}\left(n\right)\right]\right). \tag{50}$$

This implies that for $\epsilon' = \frac{\epsilon}{2\sum_{j\le\kappa} j}$ there holds

$$\mathbb{P}\left[|X_{\kappa,n} - \mathbb{E}\left[X_{\kappa,n}\right]| > \frac{\epsilon}{2}\right] \le \sum_{j\le\kappa,k\le\kappa} \mathbb{P}\left[\frac{1}{n}|N_{j,k}^{\pi}\left(n\right) - \mathbb{E}\left[N_{j,k}^{\pi}\left(n\right)\right]| > \epsilon'\right].$$

Using equation (45) we find

$$\mathbb{P}\left[|X_{\kappa,n} - \mathbb{E}\left[X_{\kappa,n}\right]| > \frac{\epsilon}{2}\right] \le \sum_{j\le\kappa,k\le\kappa} o\left(\frac{1}{n^3}\right) \le o\left(\frac{1}{n^{2\frac{7}{9}}}\right).$$

Here we used the fact that $d_{\max} = O\left(n^{1/9}\right)$ and that for $j > d_{\max}$ or $k > d_{\max}$ or both, there holds $N_{j,k}^{\pi}\left(n\right) = \mathbb{E}\left[N_{j,k}^{\pi}\left(n\right)\right] = 0$. This shows equation (49) and hence proves that Q_n' converges ν-a.s. to Q.

By redefining $Q_n', Q, X_{\kappa,n}, \tilde{X}_\kappa$ and setting

$$\epsilon' := \min\left\{\frac{\epsilon}{2}, \frac{\epsilon}{2\sum_{j\le\kappa} j}, \frac{\epsilon}{2\sum_{j\le\kappa,k\le\kappa} jk}, \frac{\epsilon}{2\sum_{j\le\kappa} j^2}, \frac{\epsilon}{2\sum_{k\le\kappa} k^2}\right\},$$

similar derivations hold for the first mixed moment and the second moments, that is one may show that all the moments of interest and the distribution itself converge simultaneously ν-a.s. for an element of D. Thus we have shown that $\left(\mathbf{d}_\pi^n\right)_{n\in\mathbb{N}}$ is ν-a.s. feasible.

To prove that $\left(\mathbf{d}_\pi^n\right)_{n\in\mathbb{N}}$ is ν-a.s. proper according to Definition (4.8), we will now show that for each feasible $\left(\mathbf{d}_\pi^n\right)_{n\in\mathbb{N}} \in D$, there holds:

$$d_{\max}^{\pi} \le \frac{n^{1/12}}{\ln(n)}$$

and

$$\rho^{\pi} = \max\left\{\frac{\sum_{j,k=0}^{\infty} j^2 k N_{j,k}^{\pi}\left(n\right)}{\mu^{\pi} n}, \frac{\sum_{j,k=0}^{\infty} jk^2 N_{j,k}^{\pi}\left(n\right)}{\mu^{\pi} n}\right\} = o\left(\frac{n^{1/12}}{\ln(n)}\right),$$

where $d_{\max}^{\pi}$ is the maximum degree of the percolated degree sequence. As the degree progression before percolation $(\mathbf{d}^n)_{n\in\mathbb{N}}$ is proper, there holds $d_{\max} \le \frac{n^{1/12}}{\ln(n)}$. Percolation can only decrease the in-degree and the out-degree of a vertex, implying $d_{\max}^{\pi} \le d_{\max}$. Together these observations show that $d_{\max}^{\pi} \le \frac{n^{1/12}}{\ln(n)}$ for any $\left(\mathbf{d}_\pi^n\right)_{n\in\mathbb{N}} \in D$. It remains to show $\rho^{\pi} = o\left(\frac{n^{1/12}}{\ln(n)}\right)$. Consider the sum $\sum_{j,k=0}^{\infty} j^2 k N_{j,k}^{\pi}\left(n\right)$ (respectively $\sum_{j,k=0}^{\infty} jk^2 N_{j,k}^{\pi}\left(n\right)$). Each vertex of degree (j,k) contributes j^2k (or jk^2) to the total sum. As the in-degree and the out-degree can only decrease due to percolation, this implies that

$$\sum_{j,k=0}^{\infty} j^2 k N_{j,k}^{\pi}\left(n\right) \le \sum_{j,k=0}^{\infty} j^2 k N_{j,k}\left(n\right) \quad \text{and} \quad \sum_{j,k=0}^{\infty} jk^2 N_{j,k}^{\pi}\left(n\right) \le \sum_{j,k=0}^{\infty} jk^2 N_{j,k}\left(n\right).$$

Recall that $\mu^{\pi,\text{bond}} = \pi\mu$ and π is a constant, i.e. it is assigned a fixed value before percolation. These observations together with the fact that $(\mathbf{d}^n)_{n\in\mathbb{N}}$ is proper imply

$$\rho^\pi \le \max\left\{\frac{\sum_{j,k=0}^{\infty} j^2 k N_{j,k}(n)}{\pi\mu n}, \frac{\sum_{j,k=0}^{\infty} j k^2 N_{j,k}(n)}{\pi\mu n}\right\} = \frac{\rho}{\pi} = o\left(\frac{n^{1/12}}{\ln(n)}\right).$$

Thus any $\left(\mathbf{d}_\pi^n\right)_{n\in\mathbb{N}} \in D$ is ν-a.s. proper.

Let $E \subset D$ be the event over which the degree progression is proper. As any element of D is ν-a.s. proper, there holds $\nu(E) = 1$. For any $\left(\mathbf{d}_\pi^n\right)_{n\in\mathbb{N}} \in E$ we may apply Theorem 4.10 to a sequence of random multigraphs $\left(\tilde{G}^\pi_{\mathbf{d}^n}\right)_{n\in\mathbb{N}}$ arising from uniformly random configurations. Recall that Lemma 5.5 implies that this is the case for all n if we condition on $\mathbf{D}^n = \mathbf{d}_\pi^n$. We will now fix $\left(\mathbf{d}_\pi^n\right)_{n\in\mathbb{N}} \in E$ and apply Theorem 4.10 to $\left(\tilde{G}^\pi_{\mathbf{d}^n}\right)_{n\in\mathbb{N}}$, distinguishing two cases π: $\pi < \hat{\pi}^{\text{bond}}$ and $\pi > \hat{\pi}^{\text{bond}}$, with $\hat{\pi}^{\text{bond}} = \frac{\mu}{\mu_{11}}$.

Let $\pi < \hat{\pi}^{\text{bond}}$. Define $\mathcal{A}_\epsilon\left(\mathbf{d}_\pi^n\right)$ to be the set of all multigraphs obeying $\mathbf{d}_\pi^n$ for which the largest strongly connected component contains no more than ϵn vertices, for $\epsilon \in (0,1)$. As $\frac{\mu_{11}^\pi}{\mu^\pi} = \pi\frac{\mu_{11}}{\mu} < 1$, Theorem 4.10 implies that for all ϵ:

$$\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}^\pi_{\mathbf{d}^n} \in \mathcal{A}_\epsilon\left(\mathbf{d}_\pi^n\right) \mid \mathbf{D}^n = \mathbf{d}_\pi^n\right] = 1. \tag{51}$$

Next consider $\pi > \hat{\pi}^{\text{bond}}$. Define $\mathcal{B}_\epsilon\left(\mathbf{d}_\pi^n\right)$ to be the set of all graphs whose largest strongly connected component contains ϵn vertices, for $\epsilon \in (0,1)$. As $\frac{\mu_{11}^\pi}{\mu^\pi} = \pi\frac{\mu_{11}}{\mu} > 1$, Theorem 4.10 implies that there exists a unique ϵ, such that

$$\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}^\pi_{\mathbf{d}^n} \in \mathcal{B}_\epsilon\left(\mathbf{d}_\pi^n\right) \mid \mathbf{D}^n = \mathbf{d}_\pi^n\right] = 1. \tag{52}$$

Not only does the theorem imply existence of such of ϵ, it also determines its value. This value plays the same role for the degree distribution $p_{j,k}^{\text{bond}}$ as $(\zeta^+ + \zeta^- + \psi - 1)$ for $p_{j,k}$. We will express this value in terms of $p_{j,k}$. This requires us to determine the probability that a uniformly at random chosen in-stub is attached to a vertex of out-degree k in the percolated graph. In analogy to equation (21) this equals

$$\begin{aligned} p_k^{+,\text{bond}} &= \frac{1}{\mu^{\pi,\text{bond}}}\sum_{j=0}^{\infty} j p_{j,k}^{\text{bond}} = \frac{1}{\pi\mu}\sum_{d^+=k}^{\infty}\sum_{d^-=0}^{\infty} p_{d^-,d^+}\sum_{j=0}^{d^-} j\binom{d^-}{j}\pi^j(1-\pi)^{d^--j}\binom{d^+}{k}\pi^k(1-\pi)^{d^+-k} \\ &= \frac{\pi}{\pi\mu}\sum_{d^+=k}^{\infty}\sum_{d^-=0}^{\infty} d^- p_{d^-,d^+}\binom{d^+}{k}\pi^k(1-\pi)^{d^+-k} = \sum_{d^+=k}^{\infty} p_{d^+}^+\binom{d^+}{k}\pi^k(1-\pi)^{d^+-k}, \end{aligned} \tag{53}$$

having generating function

$$U_\pi^+(x) := (\pi\mu)^{-1}\frac{\partial}{\partial y}U_\pi(x,y)|_{y=1}.$$

Similarly the probability that a uniformly random out-stub is attached to a vertex of in-degree j in the percolated graph becomes

$$p_j^{-,\text{bond}} = \sum_{d^-=j}^{\infty} p_{d^-}^-\binom{d^-}{j}\pi^j(1-\pi)^{d^--j}, \tag{54}$$

which corresponds to generating function

$$U_\pi^+(x) := (\pi\mu)^{-1}\,\frac{\partial}{\partial y}U_\pi(x,y)|_{y=1}.$$

The distributions $\left(p_j^{-,\mathrm{bond}}\right)_{j\in\mathbb{N}_0}$ and $\left(p_j^{+,\mathrm{bond}}\right)_{j\in\mathbb{N}_0}$ both have expected value $\pi\frac{\mu_{11}}{\mu} > 1$, and, $p_0^{-,\mathrm{bond}}, p_0^{+,\mathrm{bond}} > 0$, we have unique fixed points $x^*, y^* \in (0,1)$, see [10, Lemma 1]:

$$x^* = U_\pi^+(x^*), \tag{55}$$

$$y^* = U_\pi^-(y^*). \tag{56}$$

Using x^* and y^* we can determine the analogues of ζ^+, ζ^- and ψ, as used in Theorem 4.10, for the degree distribution $p_{j,k}^{\mathrm{bond}}$. Using equations (26), (23) and (27) these are defined as

$$\zeta^{-,\mathrm{bond}} := 1 - \sum_{j,k=0}^{\infty} p_{j,k}^{\mathrm{bond}}\,(x^*)^j\,, \quad \zeta^{+,\mathrm{bond}} := 1 - \sum_{j,k=0}^{\infty} p_{j,k}^{\mathrm{bond}}\,(y^*)^k \tag{57}$$

and

$$\psi^{\mathrm{bond}} := \sum_{j,k=0}^{\infty} p_{j,k}^{\mathrm{bond}}\,(x^*)^j\,(y^*)^k\,. \tag{58}$$

We can now define

$$c^{\mathrm{bond}} := \zeta^{-,\mathrm{bond}} + \zeta^{+,\mathrm{bond}} + \psi^{\mathrm{bond}} - 1 = 1 - U_\pi(x^*,1) - U_\pi(1,y^*) + U_\pi(x^*,y^*). \tag{59}$$

Hence, $\epsilon = c^{\mathrm{bond}}$ is the unique value required by equation (52).
To finalise the proofs for Theorem 2.2 and 2.3, we need to supplement Equations (52) and (51) with two minor observations. First, the theorem is stated for a percolated multigraph progression $\left(\tilde{G}_{\mathbf{d}^n}^{\pi}\right)_{n\in\mathbb{N}}$ without conditioning on the degree progression of percolated graphs. As $\nu(E) = 1$, the argument of Fountoulakis [4, p. 348] can be applied to show that:

- $\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}_{\mathbf{d}^n}^{\pi} \in \mathcal{A}_\epsilon\left(\mathbf{d}_\pi^n\right)\right] = 1$ for all $\epsilon \in (0,1)$ if $\pi < \hat{\pi}^{\mathrm{bond}}$;
- $\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}_{\mathbf{d}^n}^{\pi} \in \mathcal{B}_{c^{\mathrm{bond}}}\left(\mathbf{d}_\pi^n\right)\right] = 1$ and $\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}_{\mathbf{d}^n}^{\pi} \in \mathcal{B}_\epsilon\left(\mathbf{d}_\pi^n\right)\right] = 0$ for all $\epsilon \in (0,1), \epsilon \neq c^{\mathrm{bond}}$ if $\pi > \hat{\pi}^{\mathrm{bond}}$.

Second, Theorems 2.2 and 2.3 make assertions about uniformly random *simple* graphs instead of random multigraphs. Replace the the graph $\tilde{G}_{\mathbf{d}^n}$ in Lemma 3.7 and Corollary 3.8 by the graph $\tilde{G}_{\mathbf{d}^n}^{\pi}$ and condition on the graph to which percolation is applied ($G_{\mathbf{d}^n}$) being simple. This yields slightly different variants of the lemma and corollary that do not require additional changes to the proof. Now applying this variant of Lemma 3.7 and Corollary 3.8 to the above limits, we deduce that

- $\lim_{n\to\infty} \mathbb{P}\left[G_n^{\pi} \in \mathcal{A}_\epsilon\left(\mathbf{d}_\pi^n\right)\right] = 1$ for all $\epsilon \in (0,1)$ if $\pi < \hat{\pi}^{\mathrm{bond}}$;
- $\lim_{n\to\infty} \mathbb{P}\left[G_n^{\pi} \in \mathcal{B}_{c^{\mathrm{bond}}}\left(\mathbf{d}_\pi^n\right)\right] = 1$ and $\lim_{n\to\infty} \mathbb{P}\left[G_n^{\pi} \in \mathcal{B}_\epsilon\left(\mathbf{d}_\pi^n\right)\right] = 0$ for all $\epsilon \in (0,1), \epsilon \neq c^{\mathrm{bond}}$ if $\pi > \hat{\pi}^{\mathrm{bond}}$,

completing the proofs of Theorems 2.2 and 2.3 for the case of bond percolation.

5.3 Site percolation

The proofs of Theorems 2.2 and 2.3 for site percolation have similar structures as those for bond percolation. Hence, we will refer back to Section 5.2 where applicable. As in the case of bond percolation, the proof is split into three steps. First, in Section 5.3.1 we show that applying site percolation to a uniformly random configuration results in another uniformly random configuration, if we condition on the degree sequence after percolation. Second, we determine the limit of the expected number of vertices with degree (j, k) after site percolation, see Section 5.2.2. The proof is completed in Section 5.3.3 by combining the first two steps with results of Section 5.2.

Recall from Section 2 that deleting a vertex means that we remove all edges adjacent to this vertex. In the setting of the configuration model this implies that all stubs attached to a deleted vertex are removed. Let us denote these stubs by $(W^{-,r}, W^{+,r})$. As site percolation removes any edges adjacent to a vertex, the match of any stub in $(W^{-,r}, W^{+,r})$ will be removed too. A stub in $(W^{-,r}, W^{+,r})$ may or may not have its match in the same set, as it might happen that both endpoints of one edge are deleted. Let $(W^{-,m}, W^{+,m})$ contain all the matches of stubs in $(W^{-,r}, W^{+,r})$ that are not connected to a deleted vertex. Thus $W^{-,r} \cup W^{-,m}$ (respectively $W^{+,r} \cup W^{+,m}$) are all in-stubs (out-stubs) removed by site percolation. The stubs that survive percolation are still denoted by $(W^{-,\pi}, W^{+,\pi})$. Remark that this implies

$$W^{-} = W^{-,\pi} \cup W^{-,r} \cup W^{-,m} \quad \text{and} \quad W^{+} = W^{+,\pi} \cup W^{+,r} \cup W^{+,m}. \tag{60}$$

These definitions of $(W^{-,r}, W^{+,r})$ and $(W^{-,m}, W^{+,m})$ will be important throughout the proof.

5.3.1 A percolated configuration is a uniformly random configuration

As in the case of bond percolation, conditional on $\mathbf{d}_\pi^n$ being the degree sequence after percolation, applying site percolation to a uniformly random configuration on (W^-, W^+) results in a uniformly random configuration obeying $\mathbf{d}_\pi^n$. We shown in Lemma 5.6 that conditional on the stubs that are removed by site percolation, the matching on the surviving stubs is uniformly random.

Lemma 5.6. *Apply site percolation to a uniformly random configuration $\mathcal{M}$ on (W^-, W^+). Conditional on the elements of $(W^{-,r}, W^{+,r})$ and $(W^{-,m}, W^{+,m})$, each configuration on $(W^{-,\pi}, W^{+,\pi})$ is equally likely.*

Proof. According to equation (60), fixing the elements of $(W^{-,r}, W^{+,r})$ and $(W^{-,m}, W^{+,m})$, uniquely determines the elements of $(W^{-,\pi}, W^{+,\pi})$. Choosing the elements of $(W^{-,r}, W^{+,r})$ and $(W^{-,m}, W^{+,m})$ furthermore implies that the configuration $\mathcal{M}$ is the union of a configuration on $(W^{-,r} \cup W^{-,m}, W^{+,r} \cup W^{-,m})$ with the one on $(W^{-,\pi}, W^{+,\pi})$. As $\mathcal{M}$ is a uniformly random configuration obeying this split and the elements of $(W^{-,\pi}, W^{+,\pi})$ are fixed, the configuration on $(W^{-,\pi}, W^{+,\pi})$ will be a uniformly random one.

□

This lemma allows us to prove that conditional on the degree sequence after percolation, there remains a uniformly random configuration.

Lemma 5.7. *Apply site percolation to a uniformly random configuration on (W^-, W^+). Conditional on $\mathbf{D}^n = \mathbf{d}_\pi^n$, any configuration on $\left(W^-_{\mathbf{d}_\pi^n}, W^+_{\mathbf{d}_\pi^n}\right)$ is equally likely.*

Proof. Define $l = |W^{-,\pi}|$ and let $S(\mathbf{d}_\pi^n)$ contains all sets of surviving stubs $(W^{-,\pi}, W^{+,\pi})$ that induce the degrees sequence $\mathbf{d}_\pi^n$. Fix a matching $\mathcal{M}^\pi$ of $\left(W^-_{\mathbf{d}_\pi^n}, W^+_{\mathbf{d}_\pi^n}\right)$. Then it holds that

$$\mathbb{P}\left[\mathcal{M}^\pi \middle| \mathbf{D}^n = \mathbf{d}_\pi^n\right] = \sum_{(A,B)\in S(\mathbf{d}_\pi^n)} \mathbb{P}\left[\mathcal{M}^\pi \middle| \mathbf{D}^n = \mathbf{d}_\pi^n, \left(W^{-,\pi}, W^{+,\pi}\right) = (A,B)\right] \times \\ \mathbb{P}\left[\left(W^{-,\pi}, W^{+,\pi}\right) = (A,B) \middle| \mathbf{D}^n = \mathbf{d}_\pi^n\right].$$

Remark that $\mathbb{P}\left[\mathcal{M}^\pi \middle| \mathbf{D}^n = \mathbf{d}_\pi^n, (W^{-,\pi}, W^{+,\pi}) = (A,B)\right] = \mathbb{P}\left[\mathcal{M}^\pi | (W^{-,\pi}, W^{+,\pi}) = (A,B)\right]$ as $(A,B) \in S(\mathbf{d}_\pi^n)$ implies that $(W^{-,\pi}, W^{+,\pi})$ must induce the degree sequence $\mathbf{d}_\pi^n$. Using Lemma 5.6 and the bijection between $(W^{-,\pi}, W^{+,\pi})$ and $\left(W^-_{\mathbf{d}_\pi^n}, W^+_{\mathbf{d}_\pi^n}\right)$ we find:

$$\mathbb{P}\left[\mathcal{M}^\pi \middle| \left(W^{-,\pi}, W^{+,\pi}\right) = (A,B)\right] = \frac{1}{l!}.$$

Furthermore, combining these observations with

$$\sum_{(A,B)\in S(\mathbf{d}_\pi^n)} \mathbb{P}\left[\left(W^{-,\pi}, W^{+,\pi}\right) = (A,B) \middle| \mathbf{D}^n = \mathbf{d}_\pi^n\right] = 1$$

following from the definition of $S(\mathbf{d}_\pi^n)$, we obtain

$$\mathbb{P}\left[\mathcal{M}^\pi \middle| \mathbf{D}^n = \mathbf{d}_\pi^n\right] = \frac{1}{l!} \sum_{(A,B)\in S(\mathbf{d}_\pi^n)} \mathbb{P}\left[\left(W^{-,\pi}, W^{+,\pi}\right) = (A,B) \middle| \mathbf{D}^n = \mathbf{d}_\pi^n\right] = \frac{1}{l!},$$

completing the proof. ☐

5.3.2 The expected number of vertices with degree (j,k) after site percolation

The next step in the proof of Theorem 2.2 for site percolation is proving existence of the limit

$$p_{j,k}^{\text{site}} := \lim_{n\to\infty} \frac{\mathbb{E}\left[N_{j,k}^\pi(n)\right]}{n}, \tag{61}$$

and determining its value for $j,k \in \mathbb{N}_0$. Then, given $p_{j,k}^{\text{site}}$, $\hat{\pi}^{\text{site}}$ is determined analogously to $\hat{\pi}^{\text{bond}}$, and in Section 5.3.3, $\hat{\pi}^{\text{site}}$ is shown to be the desired threshold for site percolation.

If the in-degree j or out-degree k, or both, are larger than $d_{\max}$, then

$$p_{j,k}^{\text{site}} = \lim_{n\to\infty} \frac{\mathbb{E}\left[N_{j,k}^\pi(n)\right]}{n} = 0.$$

Let $0 \le j,k \le d_{\max}$. We will now bound the value of $\mathbb{E}\left[N_{j,k}^\pi(n)\right]$. Let $N_{d^-,d^+}^{\pi,r}(n)$ denote the number of vertices of degree (d^-, d^+) before percolation that are not deleted. Thus $N_{d^-,d^+}(n) - N_{d^-,d^+}^{\pi,r}(n)$ equals the number of vertices of degree (d^-, d^+) that are deleted. Each vertex is deleted with probability $1-\pi$ independently of other vertices, hence:

$$\mathbb{E}\left[N_{d^-,d^+}^{\pi,r}(n)\right] = \pi N_{d^-,d^+}(n), \tag{62}$$

$$\mathbb{E}\left[N_{d^-,d^+}(n) - N_{d^-,d^+}^{\pi,r}(n)\right] = (1-\pi) N_{d^-,d^+}(n). \tag{63}$$

A deleted vertex will have degree $(0,0)$ after percolation with probability 1. Let $P_{j,k}(d^-,d^+)$ be the probability that a non-deleted vertex of degree (d^-,d^+) has degree (j,k) after percolation. For $(j,k)=(0,0)$ we have

$$\mathbb{E}\left[N^{\pi}_{0,0}(n)\right] = \sum_{d^-=0}^{d_{\max}}\sum_{d^+=0}^{d_{\max}} \left((1-\pi)\,N_{d^-,d^+}(n) + \pi P_{0,0}\left(d^-,d^+\right) N_{d^-,d^+}(n)\right), \tag{64}$$

and otherwise,

$$\mathbb{E}\left[N^{\pi}_{j,k}(n)\right] = \sum_{d^-=j}^{d_{\max}}\sum_{d^+=k}^{d_{\max}} \pi P_{j,k}\left(d^-,d^+\right) N_{d^-,d^+}(n). \tag{65}$$

We will now derive the expression for $P_{j,k}(d^-,d^+)$. Let $s^- = |W^{-,\pi}\cup W^{-,m}|$, $s^+ = |W^{+,\pi}\cup W^{+,m}|$, $r^- = |W^{-,m}|$ and $r^+ = |W^{+,m}|$. Note that there must hold $s^- - r^- = s^+ - r^+$ as $s^- - r^- = |W^{-,\pi}|$, $s^+ - r^+ = |W^{+,\pi}|$ and the remaining configuration on $(W^{-,\pi}, W^{+,\pi})$ forms a directed graph. Let $P_{j,k}(d^-,d^+,s^-,s^+,r^-,r^+)$ denote the probability $P_{j,k}(d^-,d^+)$ conditional on the values s^-,s^+,r^-,r^+. We will now determine this conditional probability. Site percolation combines the independent random processes of deleting vertices and creating a uniformly random configuration on (W^-,W^+). As these processes are independent, we may first determine the elements of $(W^{-,r},W^{+,r})$ and then randomly create a configuration on (W^-,W^+). Thus conditional on the value r^- (respectively r^+), each subset of $W^-\setminus W^{-,r}(W^+\setminus W^{+,r})$ of this size is equally likely to be $W^{-,m}$ (or $W^{+,m}$). This implies that

$$P_{j,k}\left(d^-,d^+,r^-,r^+,s^-,s^+\right) = \binom{d^-}{d^- - j}\binom{d^+}{d^+ - k}\frac{\binom{s^- - d^-}{r^- - d^- + j}}{\binom{s^-}{r^-}}\frac{\binom{s^+ - d^+}{r^+ - d^+ + k}}{\binom{s^+}{r^+}}. \tag{66}$$

To approximate this probability we will show that with high probability s^-,s^+ will in some bounded interval I' and r^-,r^+ in I both. This enables us to determine $P_{j,k}(d^-,d^+,r^-,r^+,s^-,s^+)$ for s^-,s^+,r^-,r^+ in these intervals. First consider s^- and s^+. By using equation (62), we obtain:

$$\mathbb{E}\left[s^-\right] = \sum_{d^-=0}^{d_{\max}}\sum_{d^+=0}^{d_{\max}} \pi d^- N^{\pi,r}_{d^-,d^+}(n) = m\pi \quad\text{and}\quad \mathbb{E}\left[s^+\right] = \sum_{d^-=0}^{d_{\max}}\sum_{d^+=0}^{d_{\max}} \pi d^+ N^{\pi,r}_{d^-,d^+}(n) = m\pi.$$

Using $d_{\max}\le n^{1/9}$ and Hoeffding's inequality we also find that

$$\mathbb{P}\left[\left|s^- - \mathbb{E}\left[s^-\right]\right| > n^{2/3}\ln(n)\right] \le e^{-\Omega\left(\ln^2(n)\right)} \quad\text{and}\quad \mathbb{P}\left[\left|s^+ - \mathbb{E}\left[s^+\right]\right| > n^{2/3}\ln(n)\right] \le e^{-\Omega\left(\ln^2(n)\right)}. \tag{67}$$

This implies that

$$s^-,s^+ \in I' := \left[m\pi - n^{2/3}\ln(n), m\pi + n^{2/3}\ln(n)\right]$$

with probability $1-e^{-\Omega\left(\ln^2(n)\right)}$. The following Lemma specifies such an interval for r^- and r^+.

Lemma 5.8. *Conditional on $s^-,s^+\in I'$, there holds*

$$r^+,r^- \in I := \left[m\pi(1-\pi) - n^{2/3}\ln^2(n), m\pi(1-\pi) + n^{2/3}\ln(n)^2\right]$$

with probability $1-e^{-\Omega\left(\ln^2(n)\right)}$ for both values separately.

Proof. We present the proof for r^-. The proof for r^+ is identical to the one for r^- up to switching the roles of in-stubs and out-stubs. Since we consider a uniformly random configuration on (W^-, W^+), the probability that any in-stub is matched to an out-stub in $W^{+,r}$ is $\frac{m-s^+}{m} = (1-\pi)\left(1+\mathcal{O}\left(n^{-1/3}\ln(n)\right)\right)$ as $s^-, s^+ \in I'$. Since r^- equals the number of in-stubs in $W^-\backslash W^{-,r}$ with a match in $W^{+,r}$, this implies

$$\mathbb{E}\left[r^-\right] = s^- \frac{m-s^+}{m} = m\pi\left(1-\pi\right)\left(1+\mathcal{O}\left(n^{2/3}\ln(n)\right)\right).$$

To complete the proof, we will now show that

$$\mathbb{P}\left[\left|r^- - \mathbb{E}\left[r^-\right]\right| > n^{2/3}\ln^2(n)\right] \le e^{-\Omega\left(\ln^2(n)\right)}.$$

This is realized by applying Theorem 5.2 to the space of configurations on (W^-, W^+) with the symmetric difference as the metric. The value of r^- plays the role of the function f. To partition this space, we order the in-stubs of W^-. Define an i-prefix to be the first i in-stubs together with their match. An element of the partition $\mathcal{P}_k$ consists of all configurations with the same k-prefix for all $k \in \{0,1,\ldots,m\}$. For any $A, B \in \mathcal{P}_k$ such that $A, B \subset C \in \mathcal{P}_{k-1}$ a bijection $\phi : A \to B$ can be defined. Denote the k^{th} pair of a configuration in A by (x, y_A) and the k^{th} pair of a configuration in B by (x, y_B). Then ϕ maps $\mathcal{M} \in A$ to the configuration in B with (x, y_A) replaced by (x, y_B) and with y_A the match of the in-stub in $\mathcal{M}$ matched to y_B. By definition of ϕ it follows that $c_k := |\mathcal{M} - \phi(\mathcal{M})| = 4$ for all $k \in \{1,2,\ldots,m\}$. As the value of r^- also changes by at most the symmetric difference of the two matchings, Theorem 5.2 implies

$$\mathbb{P}\left[\left|r^- - \mathbb{E}\left[r^-\right]\right| > n^{2/3}\ln^2(n)\right] \le 2\exp\left(\frac{n^{4/3}\ln^2(n)}{2m}\right) = e^{-\Omega\left(\ln^2(n)\right)},$$

as $m \le nd_{\max} \le n^{10/9}$. □

Fountoulakis [4, Section 4] shows that for $d_{\max} \le n^{1/9}$ there holds uniformly for $r \in I$ and $s \in I'$:

$$\binom{d}{d-i}\frac{\binom{s-d}{r-d+i}}{\binom{s}{r}} = \binom{d}{d-i}(1-\pi)^{d-i}\pi^i\left(1+\mathcal{O}\left(\frac{\ln^2(n)}{n^{1/3}}\right)\right).$$

Applying this to equation 66 implies that uniformly for all $s^-, s^+ \in I'$ and $r^-, r^+ \in I$ there holds

$$P_{j,k}\left(d^-, d^+, r^-, r^+, s^-, s^+\right) = \binom{d^-}{d^- - j}\binom{d^+}{d^+ - k}\pi^{j+k}\left(1-\pi\right)^{d^- + d^+ - j - k}\left(1+\mathcal{O}\left(\frac{\ln^2(n)}{n^{1/3}}\right)\right).$$

However we cannot yet determine this probability if at least one of the following conditions is violated: $s^-, s^+ \in I'$, $r^-, r^+ \in I$. Instead of determining the probability in these cases, we show that such violations are unlikely, that is, instead of bounding the probability

$$\mathbb{P}\left[s^- \notin I' \vee s^+ \notin I' \vee r^- \notin I \vee r^+ \notin I\right],$$

we add a condition on $N^{\pi,r}_{d^-,d^+}(n)$, allowing us to bound the value of $\mathbb{E}\left[N^{\pi}_{j,k}(n)\right]$. Theorem 5.1 implies that

$$\mathbb{P}\left[N^{\pi,r}_{d^-,d^+}(n) - \mathbb{E}\left[N^{\pi,r}_{d^+,d^-}(n)\right]| > \sqrt{n}\ln(n)\right] < e^{-\Omega\left(\ln^2(n)\right)}. \tag{68}$$

In combination with equation (62) this implies that

$$N^{\pi,r}_{d^-,d^+}(n) \in I''(d^-,d^+) = \left[\max\left\{\pi N_{d^-,d^+}(n) - \sqrt{n}\ln(n), 0\right\}, \pi N_{d^-,d^+}(n) + \sqrt{n}\ln(n)\right],$$

with probability $1 - e^{-\Omega\left(\ln^2(n)\right)}$. Together with equation (67) and Lemma 5.8 there follows:

$$\begin{aligned}
&\mathbb{P}\left[s^- \notin I' \text{ or } s^+ \notin I' \text{ or } r^- \notin I \text{ or } r^+ \notin I \text{ or } N^{\pi,r}_{d^-,d^+}(n) \notin I''\left(d^-,d^+\right)\right] \\
\le &\mathbb{P}\left[s^- \notin I'\right] + \mathbb{P}\left[s^+ \notin I'\right] + \mathbb{P}\left[r^- \notin I\right] + \mathbb{P}\left[r^+ \notin I\right] + \mathbb{P}\left[N^{\pi,r}_{d^-,d^+}(n) \notin I''\left(d^-,d^+\right)\right] \\
= &o\left(\frac{1}{n^3}\right) + \mathbb{P}\left[r^- \notin I\right] + \mathbb{P}\left[r^+ \notin I\right].
\end{aligned}$$

By the law of total probability

$$\begin{aligned}
\mathbb{P}\left[r^- \notin I\right] = &\mathbb{P}\left[r^- \notin I | s^- \in I', s^+ \in I'\right]\mathbb{P}\left[s^- \in I', s^+ \in I'\right] + \\
&\mathbb{P}\left[r^- \notin I | s^- \notin I', s^+ \in I'\right]\mathbb{P}\left[s^- \notin I', s^+ \in I'\right] + \\
&\mathbb{P}\left[r^- \notin I | s^- \in I', s^+ \notin I'\right]\mathbb{P}\left[s^- \in I', s^+ \notin I'\right] + \\
&\mathbb{P}\left[r^- \notin I | s^- \notin I', s^+ \notin I'\right]\mathbb{P}\left[s^- \notin I', s^+ \notin I'\right] = o\left(\frac{1}{n^3}\right).
\end{aligned}$$

In a similar fashion, it is shown that $\mathbb{P}\left[r^+ \notin I\right] = o\left(\frac{1}{n^3}\right)$. Thus there holds

$$\mathbb{P}\left[s^- \notin I' \text{ or } s^+ \notin I' \text{ or } r^- \notin I \text{ or } r^+ \notin I \text{ or } N^{\pi,r}_{d^-,d^+}(n) \notin I''\left(d^-,d^+\right)\right] = o\left(\frac{1}{n^3}\right). \tag{69}$$

This allows to determine a lower and upper bound for the value $\mathbb{E}\left[N^{\pi}_{j,k}(n)\right]$. As $N^{\pi,r}_{d^-,d^+}(n) \le N_{d^-,d^+}(n)$ and $(\mathbf{d}^n)_{n\in\mathbb{N}}$ is proper, for all $\epsilon > 0$ there exist $\kappa(\epsilon)$ and $N(\epsilon)$ such that for all $n > N$:

$$\sum_{\substack{(d^-,d^+)=(0,0)\\ d^-\ge\kappa+1 \text{ or } d^+\ge\kappa+1}}^{(d_{\max},d_{\max})} P_{j,k}\left(d^-,d^+\right) N^{\pi,r}_{d^-,d^+}(n) \le \sum_{\substack{(d^-,d^+)=(0,0)\\ d^-\ge\kappa+1 \text{ or } d^+\ge\kappa+1}}^{(d_{\max},d_{\max})} N_{d^-,d^+}(n) \le \epsilon n. \tag{70}$$

In combination with equation (65) this implies for $(j,k) \neq (0,0)$

$$\sum_{d^-=j}^{\kappa}\sum_{d^+=k}^{\kappa} P_{j,k}\left(d^-,d^+\right) N^{\pi,r}_{d^-,d^+}(n) \le \mathbb{E}\left[N^{\pi}_{j,k}(n)\right] \le \sum_{d^-=j}^{\kappa}\sum_{d^+=k}^{\kappa} P_{j,k}\left(d^-,d^+\right) N^{\pi,r}_{d^-,d^+}(n) + \epsilon n. \tag{71}$$

Using equation (69) on the left-hand side of the above equation we find

$$\begin{aligned}
\mathbb{E}\left[N^{\pi}_{j,k}(n)\right] \ge \sum_{d^-=j}^{\kappa}\sum_{d^+=k}^{\kappa}\sum_{\tilde{r}^-\in I'}\sum_{\tilde{r}^+\in I'}\sum_{\tilde{s}^-\in I}\sum_{\tilde{s}^+\in I}\sum_{\tilde{d}_{d^-,d^+}\in I''(d^-,d^+)} &\tilde{d}_{d^-,d^+} P_{j,k}\left(d^-,d^+,\tilde{r}^-,\tilde{r}^+,\tilde{s}^-,\tilde{s}^+\right) \times \\
&\mathbb{P}\left[r^- = \tilde{r}^-, r^+ = \tilde{r}^+, s^- = \tilde{s}^-, s^+ = \tilde{s}^+, N^{\pi,r}_{d^-,d^+}(n) = \tilde{d}_{d^-,d^+}\right] + o\left(\frac{1}{n^2}\right).
\end{aligned}$$

As equation (68) implies that

$$\sum_{\tilde{d}_{d^-,d^+}\in I''(d^-,d^+)} \tilde{d}_{d^-,d^+}\mathbb{P}\left[N^{\pi,r}_{d^-,d^+}(n)=\tilde{d}_{d^-,d^+}\right]=\mathbb{E}\left[N^{\pi,r}_{d^-,d^+}(n)\right]+o\left(\frac{1}{n^2}\right),$$

following [4] we obtain the lower bound:

$$\mathbb{E}\left[N^{\pi}_{j,k}(n)\right]\geq o\left(\frac{1}{n^2}\right)+\pi\sum_{d^-=j}^{\kappa}\sum_{d^+=k}^{\kappa}N_{d^-,d^+}(n)\binom{d^-}{d^- - j}\binom{d^+}{d^+ - k}\times$$
$$\pi^{j+k}(1-\pi)^{d^-+d^+-j-k}\left(1+\mathcal{O}\left(\frac{\ln^2(n)}{n^{1/3}}\right)\right).$$

In a similar fashion we can show, using the right-hand side of equation (71), that the upper bound is

$$\mathbb{E}\left[N^{\pi}_{j,k}(n)\right]\leq \epsilon n+o\left(\frac{1}{n^2}\right)+\pi\sum_{d^-=j}^{\kappa}\sum_{d^+=k}^{\kappa}N_{d^-,d^+}(n)\binom{d^-}{d^- - j}\binom{d^+}{d^+ - k}\times$$
$$\pi^{j+k}(1-\pi)^{d^-+d^+-j-k}\left(1+\mathcal{O}\left(\frac{\ln^2(n)}{n^{1/3}}\right)\right).$$

Combining the upper and lower bounds together proves convergence of the limit for $j,k>0$:

$$\lim_{n\to\infty}\frac{\mathbb{E}\left[N^{\pi}_{j,k}(n)\right]}{n}=\pi\sum_{d^-=j}^{\infty}\sum_{d^+=k}^{\infty}p_{d^-,d^+}\binom{d^-}{j}\binom{d^+}{k}\pi^{j+k}(1-\pi)^{d^- -j+d^+-k}=p^{\text{site}}_{j,k}. \tag{72}$$

For $(j,k)=(0,0)$, we need to use equation (64) instead of equation (65). Since $N^{\pi,r}_{d^-,d^+}(n)\leq N_{d^-,d^+}(n)$ and $(\mathbf{d}^n)_{n\in\mathbb{N}}$ is proper implies that for all $\epsilon>0$ there exist $\kappa(\epsilon)$ and $N(\epsilon)$ such that for all $n>N$

$$\sum_{\substack{(d^-,d^+)=(0,0)\\ d^-\geq\kappa+1\,\text{or}\,d^+\geq\kappa+1}}^{(d_{\max},d_{\max})}\left(N_{d^-,d^+}(n)-N^{\pi,r}_{d^-,d^+}(n)\right)\leq\sum_{\substack{(d^-,d^+)=(0,0)\\ d^-\geq\kappa+1\,\text{or}\,d^+\geq\kappa+1}}^{(d_{\max},d_{\max})}N_{d^-,d^+}(n)\leq\epsilon n.$$

Thus the equivalent of (71) for $(j,k)=(0,0)$ becomes

$$\sum_{d^-=0}^{\kappa}\sum_{d^+=0}^{\kappa}\left[\left(N_{d^-,d^+}(n)-N^{\pi,r}_{d^-,d^+}(n)\right)+P_{0,0}\left(d^-,d^+\right)N^{\pi,r}_{d^-,d^+}(n)\right]\leq\mathbb{E}\left[N^{\pi}_{0,0}(n)\right]\leq$$
$$\sum_{d^-=0}^{\kappa}\sum_{d^+=0}^{\kappa}\left[\left(N_{d^-,d^+}(n)-N^{\pi,r}_{d^-,d^+}(n)\right)+P_{0,0}\left(d^-,d^+\right)N^{\pi,r}_{d^-,d^+}(n)\right]+2\epsilon n,$$

and the analogous argument as for $(j,k)\neq(0,0)$ is applied to obtain:

$$\lim_{n\to\infty}\frac{\mathbb{E}\left[N^{\pi}_{0,0}(n)\right]}{n}=(1-\pi)+\pi\sum_{d^-=j}^{\infty}\sum_{d^+=k}^{\infty}p_{d^-,d^+}\binom{d^-}{j}\binom{d^+}{k}\pi^{j+k}(1-\pi)^{d^- -j+d^+-k}=p^{\text{site}}_{0,0}. \tag{73}$$

Comparing equations (72) and (73) with equation (37) we find:

$$p_{j,k}^{\text{site}} = \begin{cases} \pi p_{j,k}^{\text{bond}}, & (j,k) \neq (0,0), \\ \pi p_{0,0}^{\text{bond}} + 1 - \pi, & (j,k) = (0,0). \end{cases} \tag{74}$$

To guarantee that $p_{j,k}^{\text{site}}$ must be normalized, we exploit the connection between $p_{j,k}^{\text{site}}$ and $p_{j,k}^{\text{bond}}$:

$$\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} p_{j,k}^{\text{site}} = 1 - \pi + \pi \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} p_{j,k}^{\text{bond}} = 1.$$

Using equation (74), we find that

$$\mu^{\pi,\text{site}} = \pi \mu^{\pi,\text{bond}} = \pi^2 \mu \quad \text{and} \quad \mu_{11}^{\pi,\text{site}} = \pi \mu_{11}^{\pi,\text{bond}} = \pi^3 \mu_{11}. \tag{75}$$

This link between the two distribution also implies that $p_{j,k}^{\text{site}}$ satisfies equation (12).

It is left to determine $\hat{\pi}^{\text{site}}$. As we expect the percolated graph to obey the degree distribution $p_{j,k}^{\text{site}}$, from Theorem 4.10 we expect that the percolation threshold is the value of π such that

$$\sum_{j,k=0}^{\infty} jk p_{j,k}^{\text{site}} = \sum_{j,k=0}^{\infty} j p_{j,k}^{\text{site}}. \tag{76}$$

Denote this value by $\hat{\pi}^{\text{site}}$. Combing equations (76) and (75) we find $\hat{\pi}^{\text{site}^2} \mu = \hat{\pi}^{\text{site}^3} \mu_{11}$, which implies that

$$\hat{\pi}^{\text{site}} = \frac{\mu}{\mu_{11}} = \hat{\pi}^{\text{bond}}.$$

Hence, we expect that the percolation thresholds for site and bond percolation are equal. This can be explained by remarking that the expected degree distribution after site percolation is a rescaled version the degree distribution after bond percolation, expect for $(0,0)$. Hence one expects a GSCC to appear under the same conditions. However the GSCC after site percolation is expected to contain fewer vertices, because the probability to find an isolated vertex is larger. Note that, as in the case of bond percolation, see equation (41), we have a positive fraction of dead ends for $0 < \pi < 1$, which fulfils one of the prerequisites for Theorem 4.10.

5.3.3 Determining π_c^{site} and c^{site}

To finalise the proofs of Theorems 2.2 and 2.3 for site percolation, it remains to show that $\pi_c^{\text{site}} = \hat{\pi}^{\text{site}}$ and to determine c^{site}. This is done analogously to the proof of Theorem 2.2 for bond percolation in Section 5.2.3. Because of the similarity between these proofs, we only explain the changes that are made in Section 5.2.3 to convert it into the proof for site percolation.

First of all, we need to replace $p_{j,k}^{\text{bond}}$ with $p_{j,k}^{\text{site}}$. Lemma 5.7 proves exactly the same statement for site percolation as Lemma 5.5 for bond percolation, therefore substituting this lemma in Section 5.2.3 will suffice. However, equation (49) requires a different proof. Conditional on a certain realisation of $(W^{-,r}, W^{+,r})$ and the values $s^-, s^+ \in I'$, $r^-, r^+ \in I$, the value of $N_{j,k}^{\pi}(n)$ is determined by the random choice of $(W^{-,m}, W^{+,m})$. By changing one element of $(W^{-,m}, W^{+,m})$ the value of $N_{j,k}^{\pi}(n)$ changes by at most 2. Thus Corollary 5.3 can be applied to obtain

$$\mathbb{P}\left[|N_{j,k}^{\pi}(n) - \mathbb{E}\left[N_{j,k}^{\pi}(n)\right]| > \sqrt{n}\ln^2(n) \mid s^-, s^+, r^-, r^+, \left(W^{-,r}, W^{+,r}\right)\right]$$
$$\leq 2\exp\left(\frac{n\ln^2(n)}{\left(m(1-\pi)\pi + n^{2/3}\ln^2(n)\right)} \right) = e^{-\Omega\left(\ln^2(n)\right)}.$$

Using Lemma 5.8 and equation (67) there follows

$$\mathbb{P}\left[\left|N^{\pi}_{j,k}(n) - \mathbb{E}\left[N^{\pi}_{j,k}(n)\right]\right| > \sqrt{n}\ln^2(n)\right] = o\left(\frac{1}{n^3}\right),$$

and, as κ is bounded, this completes the proof of equation (49).

The last change we need, is related to the fact that Theorem 4.10 is now applied to a proper degree progression with $p^{\text{site}}_{j,k}$ as degree distribution instead of $p^{\text{bond}}_{j,k}$. Hence $\hat{\pi}^{\text{bond}}$ and c^{bond} must be replaced by $\hat{\pi}^{\text{site}}$ and c^{site}. In Section 5.3.2 we already found that $\hat{\pi}^{\text{site}} = \frac{\mu}{\mu_{11}}$. Thus it remains to determine c^{site}. This value is derived analogously to the derivation of c^{bond}, expect for replacing $p^{\text{bond}}_{j,k}$ by $p^{\text{site}}_{j,k}$. This implies that we first need to determine the probability that a uniformly random out-stub (respectively in-stub) is attached to a vertex with in-degree j (out-degree k) in the configuration after applying site percolation. In analogy to equations (54) and (53) these probabilities are given by

$$p^{-,\text{site}}_j = \sum_{k=0}^{\infty} \frac{k}{\mu^{\pi,\text{site}}} p^{\text{site}}_{j,k} = \sum_{d^-=j}^{\infty} p^-_{d^-}\binom{d^-}{j}\pi^j(1-\pi)^{d^- - j}$$

and

$$p^{+,\text{site}}_k = \sum_{d^+=k}^{\infty} p^+_{d^+}\binom{d^+}{k}\pi^k(1-\pi)^{d^+ - k}.$$

Note that $p^{-,\text{site}}_j = p^{-,\text{bond}}_j$ and $p^{+,\text{site}}_k = p^{+,\text{bond}}_k$. While this might seem surprising, there is a logical explanation. A vertex of degree $(0,0)$ does not play any role in this distribution, as it will never be encountered by following a uniformly random in-stub or out-stub. Equation (74) implies that for all other degrees there holds $p^{\text{site}}_{j,k} = \pi p^{\text{bond}}_{j,k}$. Hence, after normalization the value of $p^{-,\text{site}}_j$ (respectively $p^{+,\text{site}}_k$) equals $p^{-,\text{bond}}_j$ (or $p^{+,\text{bond}}_k$) for all j (or k). Since these distributions are equal, they have the same fixed points, x^* and y^*, as in equation (55). Therefore, the difference between the two types of percolation is only in the definitions of $\zeta^{-,\text{site}}, \zeta^{+,\text{site}}$ and ψ^{site}:

$$\zeta^{-,\text{site}} := 1 - \sum_{j,k=0}^{\infty} p^{\text{site}}_{j,k}(x^*)^j = \pi\left(1 - \zeta^{-,\text{bond}}\right) + 1 - \pi, \tag{77}$$

$$\zeta^{+,\text{site}} := 1 - \sum_{j,k=0}^{\infty} p^{\text{site}}_{j,k}(y^*)^k = \pi\left(1 - \zeta^{+,\text{bond}}\right) + 1 - \pi, \tag{78}$$

and

$$\psi^{\text{site}} := \sum_{j,k=0}^{\infty} p^{\text{site}}_{j,k}(x^*)^j(y^*)^k = \pi\psi^{\text{bond}} + 1 - \pi. \tag{79}$$

Applying Theorem 4.10 and elementary transformations to the above equations, we obtain

$$c^{\text{site}} = \zeta^{-,\text{site}} + \zeta^{+,\text{site}} + \psi^{\text{site}} - 1 = \pi\left(\zeta^{-,\text{bond}} + \zeta^{+,\text{bond}} + \psi^{\text{bond}} - 1\right) = \pi c^{\text{bond}}. \tag{80}$$

This simple relation between c^{site} and c^{bond} also can be intuitively explained using equation (74). The main difference in the distributions is in vertices with degree $(0,0)$. A vertex of degree $(0,0)$ forms its own strongly connected component. Hence these vertices are not in the GSCC. As $p_{j,k}^{\text{site}} = \pi p_{j,k}^{\text{bond}}$ and $p_{0,0}^{\text{site}} = \pi p_{j,k}^{\text{bond}} + 1 - \pi$, hence we could already have predicted that $c^{\text{site}} = \pi c^{\text{bond}}$. This is the last change that needs to be made to Section 5.2.3 to complete the proofs of Theorems 2.2 and 2.3 for site percolation. This completes the proofs of Theorems 2.2 and 2.3.

References

[1] Béla Bollobás, Bela Bollobás, Oliver Riordan, and O Riordan. *Percolation.* Cambridge University Press, 2006.

[2] Joel Spencer. The giant component: The golden anniversary. *Notices of the AMS*, 57(6):720–724, 2010.

[3] Béla Bollobás and Oliver Riordan. An old approach to the giant component problem. *Journal of Combinatorial Theory, Series B*, 113:236–260, 2015.

[4] Nikolaos Fountoulakis. Percolation on sparse random graphs with given degree sequence. *Internet Mathematics*, 4(4):329–356, 2007.

[5] Svante Janson et al. On percolation in random graphs with given vertex degrees. *Electronic Journal of Probability*, 14:86–118, 2009.

[6] Michael Molloy and Bruce Reed. A critical point for random graphs with a given degree sequence. *Random structures & algorithms*, 6(2-3):161–180, 1995.

[7] Michael Molloy and Bruce Reed. The size of the giant component of a random graph with a given degree sequence. *Combinatorics, probability and computing*, 7(3):295–305, 1998.

[8] Mathew D Penrose et al. The strong giant in a random digraph. *Journal of Applied Probability*, 53(1):57–70, 2016.

[9] Matthew Coulson. On the largest component of the critical random digraph. *Acta Mathematica Universitatis Comenianae*, 88(3):567–572, 2019.

[10] Colin Cooper and Alan Frieze. The size of the largest strongly connected component of a random digraph with a given degree sequence. *Combinatorics, Probability and Computing*, 13(3):319–337, 2004.

[11] Mark EJ Newman, Steven H Strogatz, and Duncan J Watts. Random graphs with arbitrary degree distributions and their applications. *Physical review E*, 64(2):026118, 2001.

[12] Ivan Kryven. Emergence of the giant weak component in directed random graphs with arbitrary degree distributions. *Physical Review E*, 94(1):012315, 2016.

[13] Ivan Kryven. Finite connected components in infinite directed and multiplex networks with arbitrary degree distributions. *Physical Review E*, 96(5):052304, 2017.

[14] M Drew LaMar. Directed 3-cycle anchored digraphs and their application in the uniform sampling of realizations from a fixed degree sequence. In *Proceedings of the 2011 Winter Simulation Conference (WSC)*, pages 3348–3359. IEEE, 2011.

[15] Remco Van Der Hofstad. *Random graphs and complex networks*, volume 1. Cambridge university press, 2016.

[16] Ningyuan Chen and Mariana Olvera-Cravioto. Directed random graphs with given degree distributions. *Stochastic Systems*, 3(1):147–186, 2013.

[17] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press, 2013.

[18] Colin McDiarmid. On the method of bounded differences. *Surveys in combinatorics*, 141(1):148–188, 1989.

[19] Valentin V Petrov. Limit theorems of probability theory: sequences of independent random variables. Technical report, Oxford, New York, 1995.