Probability in the Engineering and Informational Sciences (2024), **38:**2 268–289
doi:10.1017/S0269964823000128

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Percolation in simple directed random graphs with a given degree distribution

Femke van Ieperen and Ivan Kryven

Mathematical Institute, Utrecht University, PO Box 80010, 3508 TA Utrecht, The Netherlands
Corresponding author: Ivan Kryven; E-mail: i.kryven@uu.nl

Keywords: random graphs; directed graphs; percolation; connected components; bounded differences

Abstract
We study site and bond percolation in simple directed random graphs with a given degree distribution. We derive the percolation threshold for the giant strongly connected component and the fraction of vertices in this component as a function of the percolation probability. The results are obtained for degree sequences in which the maximum degree may depend on the total number of nodes n, being asymptotically bounded by $n^{\frac{1}{9}}$.

1. Introduction

Percolation in infinite graphs is typically studied in the 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. Conventionally, this property is chosen to describe *connected components*: maximal vertex sets in which any pair of vertices is connected with a path. Many results about sizes of connected components are known for percolation in infinite *lattices* [4] and *random graphs* [1, 3, 15, 16, 19]. Closely related to percolation are random graph models that depend on a real parameter, such as the well-studied Erdős Rényi random graph $G(n,p)$, and also various models for directed random graphs, often referred to as $D(n,p)$ or $\vec{G}(n,p)$ [17, 26, 29, 30].

Fountoulakis [16] and Janson [19] studied percolation in undirected random graphs with well-behaved degree sequences using techniques that rely on Molloy and Reed's existence theorem [22, 23], which indicates whether a simple undirected random graph with a given degree sequence contains a giant component and how large it is. These authors showed that if one starts with a simple random graph generated by the configuration model and then removes edges (vertices) uniformly at random, the resulting percolated graph can again be studied with the configuration model, albeit with a modified degree sequence. In this paper, we introduce a similar argument to directed graphs. Although there are several points of analogy, directed graphs generally require a distinct treatment from that of undirected graphs.

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 losing this property. Suppose G_n is uniformly sampled from the set of all digraphs with a fixed graphic degree sequence

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

where d_v^- and d_v^+ indicate correspondingly the in- and out-degree of vertex $v \in V$. Let additionally the following limits exist, $\mu := \lim_{n\to\infty} \mu(n)$ and $\mu_{11} := \lim_{n\to\infty} \mu_{11}(n)$, with $\mu(n) := n^{-1} \sum_{v\in V} d_v^- = n^{-1} \sum_{v\in V} d_v^+$ being the expected number of edges per vertex and $\mu_{11}(n) := n^{-1} \sum_{v\in V} d_v^- d_v^+$.

The existence criteria for the giant strongly connected component (GSCC) in directed graphs was formulated for slightly different models by Cooper and Frieze [10], Bloznelis, Götze, and Jaworski [2], and Penrose [24]. These results have been subsequently improved by weakening the requirement on the degree sequence by Graf [18] and Coulson [11], and more recently, further generalized by Cai and Perarnau [7] and Donderwinkel and Xie [14]. All in all, the abovementioned authors state that if $\mu_{11} - \mu > 0$, the size of the largest SCC $\mathcal{C}(G_n)$ is of the order n,

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

under some regularity conditions on the degree sequences $(\mathbf{d}^n)_{n\in\mathbb{N}}$. There are also results about the structure of the giant component [12], its diameter [8], and the properties of the dynamic processes on it [5, 6]. Sizes of *weakly* connected components were studied in [13].

If the limit in equation (2) holds, we say the random graph contains a GSCC. Likewise, if the sign of the inequality is flipped, $\mu_{11} - \mu < 0$, then the size of all SCCs is $\mathcal{O}(1)$, and the random graph is said to contain no GSCC. This result points out the existence of two classes of limiting degree sequences, those with the size of largest SCC being $\Theta(n)$ and those for which this size is $\mathcal{O}(1)$.

In this paper, we study a percolated graph G_n^π, in which each edge (vertex) in G_n is randomly removed with probability $1 - \pi$. We show that if the GSCC exists in the original graph G_n, 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 set of all graphs with a fixed (modified) degree distribution. By combining the latter observation with the theory of Cai and Perarnau [7], we show that the "phase transition" from $\mathcal{O}(1)$ to $\Theta(n)$ takes place at the critical probability $\pi_c = \mu/\mu_{11}$, such that only for $\pi > \pi_c$, G_n^π contains a GSCC with high probability (w.h.p.). The critical threshold π_c is the same for bond and site percolation, whereas the fractions of vertices in GSCC, as defined in Eq. (2), are closely related for bond, $c^{\text{bond}}(\pi)$, and site, $c^{\text{site}}(\pi)$, percolation models: $c^{\text{site}}(\pi) = \pi c^{\text{bond}}(\pi)$. This work and the related proofs are inspired by the results for percolation in undirected graphs by Fountoulakis [16] and are based on Chapter 4 of Thesis [28]. Similar conclusions were also obtained heuristically by Graf [18] by applying Janson's exploding method [19].

2. Main result

This section introduces our main theorems for the percolation threshold of the GSCC. We consider two types of percolation models 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 percolation probability $\pi \in (0, 1)$, then each edge of G_n is removed independently of the other edges with probability $1 - \pi$.
- *Site percolation,* fix percolation probability $\pi \in (0, 1)$, then for each vertex of G_n, all the edges incident to this vertex are removed together with probability $1-\pi$ independently of the other vertices. Such a vertex is referred to as a *deleted vertex*.

It should be clear from the context which type of percolation is discussed. Strictly speaking, the 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 the nth 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 degree progression $(\mathbf{d}^n)_{n\in\mathbb{N}}$ maintains or acquires some property of interest w.h.p. as $n \to \infty$.

To make valid statements about the 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 must be *graphic* as required by the equivalent of Erdős–Gallai theorem for directed graphs [20, Theorem 4]. In Section 3.1, we progressively add several more technical constraints on $(\mathbf{d}^n)_{n\in\mathbb{N}}$, namely Definitions 3.1, 3.2, and 3.3, which we jointly refer to as the requirements for a *proper* degree progression. Together, these conditions guarantee sufficient regularity of the degree progression, allowing us to reason about the limiting behavior of the corresponding random graphs and their connected components. In this condition, the most restricting requirement is that the maximum degree must be bounded, $d_{\max} \le n^{\frac{1}{9}}$.

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

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

with additional superscripts indicating the type of percolation, that is, π_c^{bond} or π_c^{site}.

For each n, the probability in this definition is taken with respect to G_n^π – the random graph that results from percolation on a uniform simple random graph that obeys $(\mathbf{d}^n)_{n\in\mathbb{N}}$.

The following theorems determine the percolation threshold for the existence of the GSCC, and, if this threshold exists, they additionally identify the fraction of the vertices in this component for bond and site percolation. The theorems can be regarded as the extension of [16, Thm. 1.1] to digraphs.

Theorem 2.2. *Consider progression $(\boldsymbol{d}^n)_{n\in\mathbb{N}}$ is proper and $\mu_{11} > \mu$, then the thresholds for the giant component in bond and site percolations are both equal to $\pi_c = \frac{\mu}{\mu_{11}} < 1$.*

Let $N_{j,k}(n)$ be the number of vertices with in-/out-degree (j,k) in G_n, and let $p_{j,k} := \lim_{n\to\infty} \frac{N_{j,k}(n)}{n}$ exist. Let additionally,

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

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

be formal power series in x and y, having probability $\pi \in (0,1)$ as a parameter. Here $U_\pi^{\text{bond}}(x,y)$ can be interpreted as the generating function of the degree of a randomly chosen vertex, and $U_\pi^+(y)$ (correspondingly $U_\pi^-(x)$) – the generating function for the degree of a vertex at the end of randomly chosen in- (or out-) edge.

Theorem 2.3. *If $(\boldsymbol{d}^n)_{n\in\mathbb{N}}$ is proper and $\pi \in (\pi_c, 1)$, then for all $\varepsilon > 0$,*

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

for bond percolation, and

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

for site percolation, with

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

and x^*, y^* *being the smallest positive solutions of* $x^* = U_\pi^-(x^*)$ *and* $y^* = U_\pi^+(y^*)$.

The remainder of the paper is structured as follows. Section 3 lays out the technical premise: Section 3.1 introduces different classes of degree progressions. Section 3.2 introduces the link between random digraphs and the directed configuration model by repurposing several theorems available for undirected graphs. Section 3.3 gives the definition of a GSCC and formulates Cai and Perarnau's existence theorem, Theorem 3.11. In Section 3.4, we derive a concentration inequality, Theorem 3.13, enabling the proof of the main result. Section 4 proves Theorems 2.2 and 2.3 separately for the cases of bond and site percolation in, respectively, Sections 4.1 and 4.2. Both subsections have a similar structure: first, we show that the configuration model introduced in Section 3.2 can be used to study percolated digraphs; second, we find the degree distribution after percolation; third, we show that the corresponding degree progression is almost surely *feasible* in large graphs, and therefore the existence theory from Section 3.3 is applicable.

3. Random digraphs

3.1. Degree sequence, degree progression, and degree distribution

The degree sequence of a given digraph can be uniquely defined by adopting the lexicographic order, as demonstrated in, for example, [20]. Let $\mathcal{G}_{\mathbf{d}^n}$ be the set of *all* directed multigraphs having degree sequence $\mathbf{d}^n$. Since we are interested in sampling from $\mathcal{G}_{\mathbf{d}^n}$, we want to be sure that for given $\mathbf{d}^n$, $\mathcal{G}_{\mathbf{d}^n} \neq \emptyset$. This is always the case for *valid* degree sequences.

Definition 3.1. *A degree sequence* $\boldsymbol{d}^n$ *is called valid if* $m := \sum_{i=1}^n d_i^- = \sum_{i=1}^n d_i^+$, *where* $m = |E|$ *is the number of edges in graph* $G_{\boldsymbol{d}^n}$.

If $\mathcal{G}_{\mathbf{d}^n}$ contains a simple graph, $\mathbf{d}^n$ is called *graphical*, and Theorem 4 in [20] gives necessary and sufficient criteria for this property. Suppose $\mathbf{d}^n$ is graphical and let $G_{\mathbf{d}^n} \in \mathcal{G}_{\mathbf{d}^n}$ be a uniformly chosen simple digraph. To be on the safe side, we need to ensure that the subset of simple graphs in $\mathcal{G}_{\mathbf{d}^n}$ is not vanishing for large n by imposing restrictions on the limiting behavior of $\mathbf{d}^n$. Let

$$d_{\max}(n) := \max\left\{\max\{d_1^-, d_2^-, \ldots, d_n^-\}, \max\{d_1^+, d_2^+, \ldots, d_n^+\}\right\}$$

be the largest degree in $\mathbf{d}^n$ for each index $n \in \mathbb{N}$. We will refer to this quantity as simply $d_{\max}$.

Definition 3.2. *A degree progression* $(\boldsymbol{d}^n)_{n\in\mathbb{N}}$ *is called feasible if (1) all* $\boldsymbol{d}^n$ *are graphical, (2) the following limits exist yielding bivariate degree distribution* $p_{j,k} := \lim_{n\to\infty} (N_{j,k}(n))/n$ *with finite partial moments up to order 2:*

$$\mu_{ab} := \lim_{n\to\infty} \sum_{j,k=0}^{\infty} \frac{j^a k^b N_{j,k}(n)}{n} = \sum_{j,k=0}^{\infty} j^a k^b p_{j,k} \in (0, \infty), \quad 0 \le a + b \le 2,$$

and (3) $d_{\max} = \mathcal{O}\left(\sqrt{n}\right)$. *Note that according to Definition 3.1, we have* $\mu := \mu_{10} = \mu_{01}$.

We refer to $p_{j,k}$ as the *degree distribution* of a feasible degree progression. The theorem of Cai and Perarnau [7] holds for feasible degree progressions. However, to show that the degree progression remains feasible after percolation, we need to narrow down the class of initial degree progressions further.

Definition 3.3. *A feasible degree progression* $(\boldsymbol{d}^n)_{n\in\mathbb{N}}$ *is called proper if* $d_{\max} \le n^{\frac{1}{9}}$.

Thus far, we have defined a chain of classes for $(\mathbf{d}^n)_{n\in\mathbb{N}}$: valid $\supset$ graphical $\supset$ feasible $\supset$ proper, where the definitions of valid and graphical are extended from $\mathbf{d}^n$ to $(\mathbf{d}^n)_{n\in\mathbb{N}}$ element-wisely.

3.2. Directed configuration model

The behavior of *simple* random digraphs can be studied with the *directed configuration model*, as defined below.

Definition 3.4. *Let* $\boldsymbol{d}^n$ *be a valid degree sequence. For all vertices enumerated with* $i \in [n]$*, let the set of in-stubs* W_i^- *consist of* d_i^- *unique elements and the set out-stubs* W_i^+ *contains* d_i^+ *elements. Let* $W^- = \cup_{i\in[n]} W_i^-$ *and* $W^+ = \cup_{i\in[n]} W_i^+$*. Then a configuration* $\mathcal{M}$ *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 between in-subs and out-stubs and hence defines a multigraph $\tilde{G}_{\mathbf{d}^n}$ with vertex set $V = [n]$ and edge multiset

$$E = [(i, j) \mid W_i^+ \ni a, W_j^- \ni b \text{ and } (a, b) \in \mathcal{M}]. \tag{4}$$

Note that multiple configurations may correspond to the same graph. We will now study the probability that the configuration model generates a specific multigraph $\tilde{G}_{\mathbf{d}^n}$.

Proposition 3.5. *Let* $\tilde{G}_{\boldsymbol{d}^n}$ *be a multigraph with degree sequence* $\boldsymbol{d}^n$ *and* $\Upsilon_{i,j}$ *be the multiplicity of edge* (i, j) *in* $\tilde{G}_{\boldsymbol{d}^n}$*. Then it holds that*

$$\mathbb{P}\left[CM_n(\boldsymbol{d}^n) = \tilde{G}_{\boldsymbol{d}^n}\right] = \frac{1}{m!}\frac{\prod_{i=1}^n d_i^-!\prod_{i=1}^n d_i^+!}{\prod_{1\le i,j\le n}\Upsilon_{i,j}!}. \tag{5}$$

Proof. This proof is based on [27, Prop. 7.4] formulated for undirected graphs. There are $m!$ different configurations. As the configuration is chosen uniformly at random,

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

with $N\left(\tilde{G}_{\mathbf{d}^n}\right)$ being the number of distinct configurations inducing $\tilde{G}_{\mathbf{d}^n}$. It follows from Eq. (4) that permuting the stub labels results in a different configuration that induces the same multigraph. There are $\prod_{i=1}^n d_i^-!\prod_{i=1}^n d_i^+!$ such permutations. 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' results in the same configuration. We compensate for this by a factor $\Upsilon_{i,j}!$ to obtain

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

□

Corollary 3.6. *Conditional on the event that configuration model generates a simple digraph, an element of* $\mathcal{G}_{\boldsymbol{d}^n}$ *is chosen uniformly.*

Generally speaking, we are interested in the statements of the form

$$\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}_{\mathbf{d}^n} \in \mathcal{A}\left(\mathbf{d}^n\right)\right] = 1, \tag{6}$$

where $\mathcal{A}\left(\mathbf{d}^n\right)$ is the set of all graphs that obey the degree sequence $\mathbf{d}^n$ and additionally satisfy some desired property. If this limit holds for a given property $\mathcal{A}$, then we say that the random graph has this property w.h.p or asymptotically almost surely (a. almost surely). The goal of the remainder of this section is to show that if Eq. (6) holds, then

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

First, we show that the probability that the configuration model generates a simple graph is bounded away from zero.

Proposition 3.7. Chen and Olvera-Cravioto [9, Thm. 4.3] *Let $(\boldsymbol{d}^n)_{n\in\mathbb{N}}$ be a feasible degree progression. The probability that the configuration model generates a simple graph is asymptotically* $\mathrm{e}^{-\frac{\mu_{11}}{\mu}-\frac{(\mu_{20}-\mu)(\mu_{02}-\mu)}{\mu}} > 0$.

Proof. The proof follows from the proof of [9, Thm. 4.3]. It suffices to replace Condition 4.1 and Lemma 5.2 from Ref. [9] with the requirement of a feasible degree progression. □

Lemma 3.8. *Let $(\boldsymbol{d}^n)_{n\in\mathbb{N}}$ be a feasible degree progression, and let $\mathcal{A}\left(\boldsymbol{d}^n\right)$ be a set of multigraphs all satisfying $\boldsymbol{d}^n$. Let $\tilde{G}_{\boldsymbol{d}^n}$ be a random multigraph generated by the configuration model.*

(1) If $\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}_{\boldsymbol{d}^n} \in \mathcal{A}\left(\boldsymbol{d}^n\right)\right] = 0$, *then* $\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}_{\boldsymbol{d}^n} \in \mathcal{A}\left(\boldsymbol{d}^n\right) \mid \tilde{G}_{\boldsymbol{d}^n} \text{ is simple}\right] = 0$.

(2) If $\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}_{\boldsymbol{d}^n} \in \mathcal{A}\left(\boldsymbol{d}^n\right)\right] = 1$, *then* $\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}_{\boldsymbol{d}^n} \in \mathcal{A}\left(\boldsymbol{d}^n\right) \middle| \tilde{G}_{\boldsymbol{d}^n} \text{ is simple}\right] = 1$.

3.3. Giant strongly connected component in a directed graph

A natural definition of a path in a directed graph requires that a path respects edge directions:

Definition 3.9. *Let $G = (V, E)$ be a digraph. A pair of vertices $v_1, v_k \in V$ is connected by a directed path if there are 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.*

This definition of connectivity can be extended to define connected components:

Definition 3.10. *SCC of a directed graph $G = (V, E)$ is a maximal subset of $C \subset V$ such that for any $u, v \in C$, there are directed paths $u - v$ and $v - u$.*

Let $\mathcal{C}(G_n)$ be the largest SCC. The notion of a *giant component* is introduced as a limiting property of the sequence of $\mathcal{C}(G_n)$. Let x^*, y^* be the smallest positive solutions of, respectively, $x^* = U^-(x^*)$ and $y^* = U^+(y^*)$, where

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

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

and $U(x, y) := \sum_{j,k=0}^{\infty} p_{j,k} x^j y^k$ is the generating function of $p_{j,k}$. Let further,

$$\zeta := 1 - U(x^*, 1) - U(1, y^*) + U(x^*, y^*).$$

Theorem 3.11. (Existence of GSCC [7, Theorem 1.2]). *Consider a feasible degree progression $(\boldsymbol{d}^n)_{n\in\mathbb{N}}$ and a uniformly random sequence of simple graphs $(G_{\boldsymbol{d}^n})_{n\in\mathbb{N}}$. Then,*

(1) If $\mu_{11}/\mu < 1$, the size of any SCC is $\mathcal{O}(1)$ with high probability.
(2) If $\mu_{11}/\mu > 1$, then with high probability, the largest SCC has vertex set of size $|\mathcal{C}(G_n)| \sim \zeta n$.

3.4. Concentration inequality

We introduce a concentration inequality based on McDiarmid's method of bounded differences, stated below. The inequality is used later in the proof of Theorem 2.2.

Theorem 3.12. [21, Theorem 7.4] *Let (V, d) be a finite metric space. Suppose there exists a sequence $\mathcal{P}_0, \mathcal{P}_1, \ldots, \mathcal{P}_s$ of increasingly refined partitions, with $\mathcal{P}_0$ being the trivial partition consisting of V and $\mathcal{P}_s$ consisting of singletons. Take a sequence of positive integers $c_1, \ldots, c_s$ such that for all $k \in \{1, 2, \ldots, s\}$, whenever $A, B \in \mathcal{P}_k$ and $A, B \subseteq C \in \mathcal{P}_{k-1}$ for some C, there exists a bijection $\phi : A \to B$ with $d(x, \phi(x)) \leq c_k$ for all $x \in A$. Let the function $f : V \to \mathbb{R}$ satisfy $|f(x) - f(y)| \leq d(x, y)$ for all $x, y, \in V$. Then, for X uniformly distributed over V and any $t > 0$,*

$$\mathbb{P}\left[|f(X) - \mathbb{E}[f(X)]| > t\right] \leq 2 \exp\left(-\frac{2t^2}{\sum_{k=1}^{s} c_k^2}\right).$$

Theorem 3.13. *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 $B, B' \in V$, it holds $|f(B) - f(B')| \leq |B \triangle B'|$. Here, $B \triangle B'$ denotes the symmetric difference, that is, $B \triangle B' = (B \cup B') \setminus (B \cap B')$. Then, for X distributed uniformly over V and any $t > 0$,

$$\mathbb{P}\left[|f(X) - \mathbb{E}[f(X)]| > t\right] \leq 2 \exp\left(-\frac{t^2}{8(b_0 + b_1)}\right). \tag{7}$$

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 at most b_0. We call a (b_0, b_1)-subset of S with such a numbering 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 $f : V \to \mathbb{R}$ naturally gives rise to a function $f : W \to \mathbb{R}$ by regarding each (b_0, b_1)-ordering as (b_0, b_1)-subset. Note that $|f(x) - f(y)| \leq x \triangle y$, that is, 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 below by their symmetric difference as (b_0, b_1)-subsets. The next step in proving Eq. (7) is applying Theorem 3.12 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. The partition $\mathcal{P}_{b_0+b_1}$ will be the partition where each element is a single ordering. Next, the values c_k need to be determined for $k \in \{1, 2, \ldots, b_0 + b_1\}$. 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 kth element. The remaining $b_0 + b_1 - k$

 https://doi.org/10.1017/S0269964823000128

elements can be any element that is not present in the k-prefix that leads to a valid ordering. Denote the kth element of any ordering in B by $a_{B,k}$. Similarly, let $a_{D,k}$ denote the kth element of any ordering in D. Define the bijection $\phi : B \rightarrow D$ by taking $x \in B$ and mapping its kth element to $a_{D,k}$. If x contains $a_{D,k}$ at some position $l > k$, map the lth 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 the 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 3.12, we find that Eq. (7) holds for any $t > 0$ and X distributed uniformly over W. 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 Eq. (7) being valid for a X uniformly random element of W, this proves the claim. □

4. Proofs of Theorems 2.2 and 2.3

Theorems 2.2 and 2.3 are proven separately for bond and site percolation using similar techniques as in the proof of [16, Theorem 1.1], which determines the percolation threshold in undirected graphs. Although these theorems are formulated for simple digraphs, we instead prove the statements on random directed multigraphs introduced in Section 3.2. The results on the multigraphs are then transferred to simple digraphs using Lemma 3.8. After percolation with probability π, $\mathbf{d}^n$ becomes a random variable $\mathbf{D}^n_\pi$, which will be considered separately for bond and site percolation.

4.1. Bond percolation

Bond percolation removes edges in a graph, thus in the 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. Given that the degree sequence after percolation is fixed, $\mathbf{D}^n_\pi = \mathbf{d}^n_\pi$, there is a bijection between the configurations on the surviving stubs $(W^{-,\pi}, W^{+,\pi})$ and the configurations on the stubs of $\mathbf{d}^n_\pi$, which we denote $\left(W^-_{\mathbf{d}^n_\pi}, W^+_{\mathbf{d}^n_\pi}\right)$. Let us fix such a bijection. Now, to show that the perfect matching on the remaining stubs (conditional on the degree sequence after percolation) is uniformly distributed on the set of perfect matchings on the stubs that survive percolation, it is sufficient to show that perfect matchings on $\left(W^-_{\mathbf{d}^n_\pi}, W^+_{\mathbf{d}^n_\pi}\right)$ appear equally likely. Let $\mathcal{D}_n$ 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 $\mathcal{D} = \prod_{n=1}^{\infty} \mathcal{D}_n$ with the product measure ν.

The remainder of this section is as follows: in Section 4.1.1, we show 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 4.1.2, we determine the limit of the expected number of vertices with degree (j, k) after percolation and show the concentration. Combining these results in Section 4.1.3, the proofs of Theorems 2.2 and 2.3 are completed by showing that an element of $\mathcal{D}$ is ν- almost surely feasible, which authorizes applying Theorem 3.11.

4.1.1. A percolated configuration is a uniformly random configuration

In this section, we show 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 lemmas.

Lemma 4.1. *Apply bond percolation to a uniformly random configuration $\mathcal{M}$ on (W^-, W^+). Conditional on the elements of $(W^{-,\pi}, W^{+,\pi})$, each configuration on these elements is equally likely.*

Proof. Choosing the elements of $(W^{-,\pi}, W^{+,\pi})$ implies that the configuration $\mathcal{M}$ is the union of a configuration on $(W^- \setminus W^{-,\pi}, W^+ \setminus W^{+,\pi})$ with 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. □

Lemma 4.2. *Conditional on having degree sequence $\boldsymbol{d}^n_\pi$ after bond percolation, that is, $\boldsymbol{D}^n_\pi = \boldsymbol{d}^n_\pi$, all configurations of $W^-_{\boldsymbol{d}^n_\pi}$ with $W^+_{\boldsymbol{d}^n_\pi}$ are equally likely.*

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

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

Remark that $\mathbb{P}\left[\mathcal{M}^\pi | \mathbf{D}^n_\pi = \mathbf{d}^n_\pi, (W^{-,\pi}, W^{+,\pi}) = (A,B)\right] = \mathbb{P}\left[\mathcal{M}^\pi | (W^{-,\pi}, W^{+,\pi}) = (A,B)\right]$ by definition of $S(\mathbf{d}^n_\pi)$. Using Lemma 4.1, we find

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

Furthermore, combining these observations with

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

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

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

completing the proof. □

4.1.2. *The expected number of vertices with degree* (j, k) *after bond percolation*

In this section, we derive an expression for the degree distribution after bond percolation. This uses the following auxiliary lemma.

Lemma 4.3. *Let 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 among all pairs of subsets of W^- and W^+ of size l.*

Proof. The graph contains m matches of which $l = |W^{-,\pi}| = |W^{+,\pi}|$ survive percolation. Thus, the probability that 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 $W^{+,\pi}$, that is, that $\mathcal{M}$ can be decomposed 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 $W^{-,\pi} \subset W^-$ and $W^{+,\pi} \subset W^+$ are uniformly random subsets both of size l. □

Lemma 4.4. *Let $N^{\pi}_{j,k}(n)$ be the number of vertices in the bond-percolated graph (or configuration) with in-degree j and out-degree k. The following limit exists*

$$p^{\text{bond}}_{j,k} := \lim_{n\to\infty} \frac{\mathbb{E}\left[N^{\pi}_{j,k}(n)\right]}{n} \quad \text{for } j,k \in \mathbb{N}_0 \tag{8}$$

and

$$p^{\text{bond}}_{j,k} = \sum_{d^-=j}^{\infty} \sum_{d^+=k}^{\infty} p_{d^-,d^+} \binom{d^-}{j}\binom{d^+}{k} \pi^{j+k} (1-\pi)^{d^- - j + d^+ - k}. \tag{9}$$

Proof. When $j,k > d_{\max}$, $p^{\text{bond}}_{j,k} = N^{\pi}_{j,k}(n) = N_{j,k}(n) = 0$. Let $j,k \le d_{\max}$ and remark that

$$\mathbb{E}\left[N^{\pi}_{j,k}(n)\right] = \sum_{l=0}^{m} \mathbb{E}\left[N^{\pi}_{j,k}(n) \,||W^{-,\pi}| = l\right] \mathbb{P}\left[|W^{-,\pi}| = l\right], \tag{10}$$

which, when conditioned on $|W^{-,\pi}| = l$, can be rewritten as

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

where event $(d^-,d^+) \to (j,k)$ is a shorthand for a vertex of degree (d^-,d^+) has degree (j,k) after percolation. Since Lemma 4.3 implies that the surviving stubs are chosen uniformly at random, conditional on $|W^{-,\pi}|$, we have

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

The edges are removed independently of each other; hence, $|W^{-,\pi}|$ is the sum of m independent Bernoulli variables, each having expectation π. Let

$$I_n := \left[m\pi - \ln n\sqrt{n}, m\pi + \ln n\sqrt{n}\right], \tag{11}$$

applying Hoefdding's inequality shows that $|W^{-,\pi}|$ concentrates

$$\mathbb{P}\left[|W^{-,\pi}| \notin I_n\right] \le \exp\left[-\Omega(\ln^2 n)\right]. \tag{12}$$

Note that as an immediate consequence of this inequality, we have $\mathbb{P}\left[l \notin I_n\right] = o\left(n^{\alpha}\right)$ for any $\alpha < 0$. Therefore, we have

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

uniformly for all $d^-, d^+ \le d_{\max} \le n^{\frac{1}{9}}$ and $l \in I_n$. In combination with Eq. (10) and the bound $N^{\pi}_{j,k}(n) \le n$, we find

$$\mathbb{E}\left[N^{\pi}_{j,k}(n)\right] = \left(1 + \mathcal{O}\left(\frac{\ln n}{n^{7/18}}\right)\right) \sum_{d^-=j}^{d_{\max}} \sum_{d^+=k}^{d_{\max}} N_{d^-,d^+}(n) \binom{d^-}{j}\binom{d^+}{k} \pi^{j+k} (1-\pi)^{d^- + d^+ - j - k} + o\left(\frac{1}{n^3}\right), \tag{13}$$

where we set $\alpha = 5$ in the estimate for $\mathbb{P}\left[l \notin I_n\right]$ when calculating the error term. We will now show that for all $\epsilon > 0$, there exist $\kappa(\epsilon)$ and $N(\epsilon)$ such that for all $n > N$,

$$\frac{1}{n}\sum_{\substack{(d^-,d^+)=(j,k)\\ d^-\geq\kappa+1 \text{ or } d^+\geq\kappa+1}}^{(d_{\max},d_{\max})} N_{d^-,d^+}(n)\binom{d^-}{j}\binom{d^+}{k}\pi^{j+k}(1-\pi)^{d^-+d^+-j-k} \leq \frac{1}{n}\sum_{\substack{(d^-,\,d^+)=(j,k)\\ d^-\geq\kappa+1 \text{ or } d^+\geq\kappa+1}}^{(d_{\max},d_{\max})} N_{d^-,\,d^+}(n) < \epsilon. \tag{14}$$

The left inequality follows from the binomial theorem, and the right – since $\lim\limits_{n\to\infty}\frac{N_{j,k}(n)}{n} = p_{j,k}$ for $j, k \geq 0$ and $(p_{j,k})$ is a probability distribution, which follows from the degree progression being proper. When combined together, Eqs. (13) and (14) prove the claim. □

Elementary calculations show that $p^{\text{bond}}_{j,k}$ is a probability distribution and has the following moments:

$$\begin{aligned}
\mu^{\pi,\text{bond}} &:= \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} jp^{\text{bond}}_{j,k} = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} kp^{\text{bond}}_{j,k} = \pi\mu,\\
\mu^{\pi,\text{bond}}_{11} &:= \sum_{j,k=0}^{\infty} jkp^{\text{bond}}_{j,k} = \pi^2\sum_{d^-=0}^{\infty}\sum_{d^+=0}^{\infty} d^-d^+p_{d^-,d^+} = \pi^2\mu_{11}.
\end{aligned} \tag{15}$$

The generating function U^{bond}_{π} of $p^{\text{bond}}_{j,k}$ and auxiliary functions U^+_{π}, U^-_{π} analogous to those used in Section 3.3 are given by

$$\begin{aligned}
U^{\text{bond}}_{\pi}(x,y) &:= U(1-\pi+\pi x, 1-\pi+\pi y),\\
U^+_{\pi}(x) &:= U^+(1-\pi+\pi x),\\
U^-_{\pi}(y) &:= U^-(1-\pi+\pi y).
\end{aligned} \tag{16}$$

The moments (15) and generating functions (16) provide sufficient information for applying Theorem 3.11 to $p^{\text{bond}}_{j,k}$ and hence formally defining the percolation threshold as such a value of $\pi = \hat{\pi}^{\text{bond}}$ that $\mu^{\pi,\text{bond}} = \mu^{\pi,\text{bond}}_{11}$, that is

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

In the following section, we will show that this quantity is indeed the desired threshold for bond percolation, $\pi^{\text{bond}}_c = \hat{\pi}^{\text{bond}}$. To this end, we show that $p^{\text{bond}}_{j,k}$ is indeed ν-a.s the limit of the degree distribution of the percolated graph, which itself is a configuration model with the percolated degree sequence.

4.1.3. Determining π^{bond}_c and c^{bond}

We make use of Lemma 4.2, stating that instead of actually removing edges, one may view the percolated multigraph $\tilde{G}^{\pi}_{\mathbf{d}^n}$ as a uniformly random *configuration* obeying the percolated degree sequence $\mathbf{d}^n_{\pi}$. We show that $\left(\mathbf{d}^n_{\pi}\right)_{n\in\mathbb{N}}$ is indeed ν- almost surely feasible (Lemma 4.5) and hence, Theorem 3.11 is applicable to $\left(\tilde{G}^{\pi}_{\mathbf{d}^n}\right)_{n\in\mathbb{N}}$. This means that Theorem 3.11 may be applied to almost all degree progressions $\left(\mathbf{d}^n_{\pi}\right)_{n\in\mathbb{N}}$ to determine π^{bond}_c and c^{bond} for random multigraphs. Conditioning on the graph before percolation being simple will ensure that the percolated graph is simple as well. Finally, we apply a variant of Lemma 3.8 and show that similar assertions hold for a graph progression $\left(G^{\pi}_n\right)_{n\in\mathbb{N}}$ 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.

Lemma 4.5. *The degree progression after bond percolation is feasible ν- almost surely*

Proof. By definition, sequence $\left(\mathbf{d}_\pi^n\right)_{n\in\mathbb{N}}$ is graphical. We need to show that it also satisfies the other requirements of Definition 3.2 ν- almost surely. To demonstrate that

$$\lim_{n\to\infty} \frac{N_{j,k}^{\pi}(n)}{n} = p_{j,k}^{\text{bond}}, \ \ \nu\text{–a.s.} \quad \text{for } j,k \in \mathbb{N}_0, \tag{17}$$

it suffices to show (see, e.g., [25, Lemma 6.8]) that for all $\epsilon > 0$,

$$\sum_{n=1}^{\infty} \mathbb{P}\left[\left|\frac{1}{n}N_{j,k}^{\pi}(n) - p_{j,k}^{\text{bond}}\right| > \epsilon\right] < \infty. \tag{18}$$

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 4.3 states that conditional on $|W^{-,\pi}| = l$, the stubs surviving percolation $(W^{-,\pi}, W^{+,\pi})$ are uniformly distributed among all pairs of subsets of (W^-, W^+) of size l. $N_{j,k}^{\pi}(n)$ 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 their symmetric difference. This implies that the requirements of Theorem 3.13 are fulfilled by setting $A_0 = W^-$, $A_1 = W^+$, $b_0 = b_1 = l$, and $N_{j,k}^{\pi}(n)$ as function f. Applying this theorem gives

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

For $l \in I_n$, defined in Eq. (11), this probability is $o\left(\frac{1}{n^3}\right)$. By Eq. (12), the probability that $l \notin I_n$ is $o\left(\frac{1}{n^3}\right)$. Combining these observations, we find that for all $\epsilon > 0$, the terms in Eq. (18) are vanishing:

$$\mathbb{P}\left[\left|N_{j,k}^{\pi}(n) - \mathbb{E}\left[N_{j,k}^{\pi}(n)\right]\right| > n\epsilon\right] = o\left(\frac{1}{n^3}\right), \tag{19}$$

which in turn proves that the limit in Eq. (17) holds ν- almost surely

It remains to show that the first, first mixed and second moments of $\frac{N_{j,k}^{\pi}(n)}{n}$ converge ν- almost surely to those of $p_{j,k}^{\text{bond}}$. In Section 4.1.2, we showed that $\sum_{j,k=0}^{\infty} j p_{j,k}^{\text{bond}} = \sum_{j,k=0}^{\infty} k p_{j,k}^{\text{bond}}$, and due to the graph context of the problem, $\sum_{j,k=0}^{\infty} j N_{j,k}^{\pi}(n) = \sum_{j,k=0}^{\infty} k N_{j,k}^{\pi}(n)$. Therefore, it is sufficient to show that only one of the first moments converges. Define $Q_n' := \frac{1}{n}\sum_{j,k=0}^{\infty} j N_{j,k}^{\pi}(n)$, we will then show that $Q := \lim_{n\to\infty} Q_n' = \pi\mu$, ν-a.s. Let $X_{\kappa,n} := \frac{1}{n}\sum_{j=0}^{\kappa}\sum_{k=0}^{\kappa} j N_{j,k}^{\pi}(n)$ 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 $\widetilde{\kappa}(\epsilon), \tilde{m}(\epsilon)$ such that for all $\kappa > \widetilde{\kappa}$ and $n > \tilde{m}$

$$\frac{1}{n} \sum_{\substack{(j,k)=(0,0),\\ j>\kappa \text{ or } k>\kappa}}^{(d_{\max},d_{\max})} jN_{j,k}(n) < \epsilon. \tag{20}$$

Thus, for $\kappa > \widetilde{\kappa}$, we have $X_{\kappa,n} \leq Q_n' \leq X_{\kappa,n} + \epsilon$. This implies that if

$$\tilde{X}_\kappa := \lim_{n\to\infty} X_{\kappa,n} = \sum_{j=0}^{\kappa} \sum_{k=0}^{\kappa} jp_{j,k}^{\text{bond}}, \ \ \nu - \text{a.s.}, \tag{21}$$

then $\lim\limits_{n\to\infty} Q_n' = Q$ holds ν- almost surely as well [16, (3.5)]. Thus, the goal is to prove Eq. (21). Applying [25, Lemma 6.8], we can show that this limit holds ν- almost surely, if for any $\epsilon > 0$,

$$\sum_{n=1}^{\infty} \mathbb{P}\left[\left|X_{\kappa,n} - \tilde{X}_\kappa\right| > \epsilon\right] < \infty. \tag{22}$$

We show this analogously to the proof of Eq. (18). By the definition of $X_{\kappa,n}$, $\tilde{X}_\kappa$, and $p_{j,k}^{\text{bond}}$, for all $\epsilon > 0$, there exists $\widetilde{N}$ such that for all $n > \widetilde{N}$, $\left|\mathbb{E}\left[X_{\kappa,n}\right] - \tilde{X}_\kappa\right| < \frac{\epsilon}{2}$. Combining this with the reverse triangle inequality, we find for any $\epsilon > 0$,

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

Remark that $\left|X_{\kappa,n} - \mathbb{E}\left[X_{\kappa,n}\right]\right| = \frac{1}{n}\sum_{j=0}^{\kappa}\sum_{k=0}^{\kappa} j\left(N_{j,k}^{\pi}(n) - \mathbb{E}\left[N_{j,k}^{\pi}(n)\right]\right)$. This implies that for $\epsilon' = \frac{\epsilon}{2\sum_{j\leq\kappa} j}$,

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

Using Eq. (19), we find $\mathbb{P}\left[\left|X_{\kappa,n} - \mathbb{E}\left[X_{\kappa,n}\right]\right| > \frac{\epsilon}{2}\right] \leq \sum_{j\leq\kappa,k\leq\kappa} o\left(\frac{1}{n^3}\right) \leq o\left(\frac{1}{n^{\frac{25}{9}}}\right)$. Here we used the fact that $d_{\max} = O\left(n^{1/9}\right)$ and that $N_{j,k}^{\pi}(n) = \mathbb{E}\left[N_{j,k}^{\pi}(n)\right] = 0$ when $j > d_{\max}$ or $k > d_{\max}$ or both. This proves Eq. (22) and hence proves that Q_n' converges ν- almost surely to Q. Similar derivations hold for the first mixed moment and the second moments. That is, all the moments of interest and the distribution itself converge simultaneously ν- almost surely for an element of $\mathcal{D}$. Thus, we have shown that $\left(\mathbf{d}_\pi^n\right)_{n\in\mathbb{N}}$ is ν- almost surely feasible. □

Let $E \subset \mathcal{D}$ be the event over which the degree progression is feasible. Since any element of $\mathcal{D}$ is ν- almost surely feasible, we have $\nu(E) = 1$. For any $\left(\mathbf{d}_\pi^n\right)_{n\in\mathbb{N}} \in E$, we may apply Theorem 3.11 to a sequence of random multigraphs $\left(\tilde{G}_{\mathbf{d}^n}^{\pi}\right)_{n\in\mathbb{N}}$ arising from uniformly random configurations. Lemma 4.2 states that if we condition $\mathbf{D}_\pi^n = \mathbf{d}_\pi^n$, Theorem 3.11 is applicable for all n. We will now fix $\left(\mathbf{d}_\pi^n\right)_{n\in\mathbb{N}} \in E$ and apply Theorem 3.11 to $\left(\tilde{G}_{\mathbf{d}^n}^{\pi}\right)_{n\in\mathbb{N}}$, distinguishing two cases for π: $\pi < \hat{\pi}^{\text{bond}}$ and $\pi > \hat{\pi}^{\text{bond}}$, with $\hat{\pi}^{\text{bond}} = \frac{\mu}{\mu_{11}}$.

Let $\pi < \hat{\pi}^{\text{bond}}$ and pick $\epsilon \in (0,1)$. 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 SCC contains no more than ϵn vertices. As $\frac{\mu_{11}^\pi}{\mu^\pi} = \pi\frac{\mu_{11}}{\mu} < 1$, Theorem 3.11 states that for all ϵ:

$$\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}_{\mathbf{d}^n}^\pi \in \mathcal{A}_\epsilon\left(\mathbf{d}_\pi^n\right) \mid \mathbf{D}_\pi^n = \mathbf{d}_\pi^n\right] = 1. \tag{23}$$

Next consider $\pi > \hat{\pi}^{\text{bond}}$. Define $\mathcal{B}_\theta\left(\mathbf{d}_\pi^n\right)$ to be the set of all graphs whose largest SCC has size in $((\theta-\varepsilon)n, (\theta+\varepsilon)n)$, with $\theta \in (0,1)$ and any vanishing ε. Since $\frac{\mu_{11}^\pi}{\mu^\pi} = \pi\frac{\mu_{11}}{\mu} > 1$, Theorem 3.11 states that there exists a unique $\theta = c^{\text{bond}}$ such that

$$\lim_{n\to\infty} \mathbb{P}\left[\tilde{G}_{\mathbf{d}^n}^\pi \in \mathcal{B}_\theta\left(\mathbf{d}_\pi^n\right) \mid \mathbf{D}_\pi^n = \mathbf{d}_\pi^n\right] = 1. \tag{24}$$

Moreover, the theorem determines that $c^{\text{bond}} := 1 - U_\pi^{\text{bond}}(x^*,1) - U_\pi^{\text{bond}}(1,y^*) + U_\pi^{\text{bond}}(x^*,y^*)$, where x^*, y^* are the smallest positive solutions of

$$\begin{aligned} x^* &= U_\pi^-(x^*), \\ y^* &= U_\pi^+(y^*), \end{aligned} \tag{25}$$

and functions U_π^+, U_π^-, and U_π^{bond} are as defined in Eq. (16).

To finalize the proof for Theorems 2.2 and 2.3, we need to supplement Eqs. (23) and (24) 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 the percolated graphs. As $\nu(E) = 1$, an analogous argument to that of Fountoulakis [16, p. 348] is applied to show that

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

Second, Theorems 2.2 and 2.3 make assertions about uniformly random *simple* graphs instead of random multigraphs. Lemma 3.8 can also be stated for the graph $\tilde{G}_{\mathbf{d}^n}^\pi$ conditioning on $G_{\mathbf{d}^n}$ (the graph to which percolation is applied) being simple. Applying this variant of Lemma 3.8 to the above limits, we deduce an equivalent statement as above for G_n^π:

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

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

4.2. Site percolation

The proof of Theorems 2.2 and 2.3 for site percolation has a similar structure as for bond percolation, and we will refer to Section 4.1 where applicable. The proof again is split into three steps. First, in Section 4.2.1, 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 4.2.2. The proof is completed in Section 4.2.3 by combining the first two steps with the results of Section 4.1.

As mentioned in Section 2, deleting a vertex means that we remove all edges adjacent to this vertex. In the setting of the configuration model, this translates into removing all stubs attached to a deleted vertex. Let us call these stubs $(W^{-,r}, W^{+,r})$. Additionally, the matches of $(W^{-,r}, W^{+,r})$ are also removed. The situation is further complicated by the fact that a match of a removed stub itself may or may not be attached to a deleted vertex. That is to say, for a deleted edge, both endpoints may be deleted or only one of them. To avoid double counting, we introduce $(W^{-,m}, W^{+,m})$ containing all the genuine 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 denoted by $(W^{-,\pi}, W^{+,\pi})$ as before, allowing us to write the full partition of stubs as

$$W^{-} = W^{-,\pi} \cup W^{-,r} \cup W^{-,m} \quad \text{and} \quad W^{+} = W^{+,\pi} \cup W^{+,r} \cup W^{+,m}. \tag{26}$$

Hence, the abbreviation is as follows: π – surviving stubs, r – stubs on removed vertices, and m – matches of stubs on removed vertices. These abbreviations will be used throughout the rest of the section.

4.2.1. A site-percolated configuration is a uniformly random configuration

We show in Lemma 4.6 that conditional on the stubs that are removed by the site percolation, the matching on the surviving stubs is uniformly random. This lemma, in turn, allows us to formulate Lemma 4.7 stating that the configuration after percolation is a uniformly random configuration conditional on its degree sequence.

Lemma 4.6. *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 Eq. (26), 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. □

Lemma 4.7. *Apply site percolation to a uniformly random configuration on (W^-, W^+). Conditional on $\boldsymbol{D}^n_\pi = \boldsymbol{d}^n_\pi$, any configuration on $\left(W^-_{\boldsymbol{d}^n_\pi}, W^+_{\boldsymbol{d}^n_\pi}\right)$ is equally likely.*

Proof. The proof of this lemma is identical to the proof of Lemma 4.2, replacing Lemma 4.1 with Lemma 4.6. □

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

The next step in the proof of Theorem 2.2 for the case of site percolation is proving the existence of the limit for $\frac{N^\pi_{j,k}(n)}{n}$, the fraction of vertices in the site-percolated graph with in-degree j and out-degree k. We then show that the value of this limit, distribution $p^{\text{site}}_{j,k}$, is closely related to the corresponding quantity for bond percolation. After this, we determine $\hat{\pi}^{\text{site}}$ and show in Section 4.2.3 that this is indeed the threshold for site percolation.

We introduce the following notation for the sizes of the subsets

$$\begin{aligned} s^- &:= |W^{-,\pi} \cup W^{-,m}|, \\ s^+ &:= |W^{+,\pi} \cup W^{+,m}|, \\ r^- &:= |W^{-,m}|, \\ r^+ &:= |W^{+,m}|. \end{aligned}$$

Note that $|W^{-,\pi}| = s^- - r^- = s^+ - r^+ = |W^{+,\pi}|$ and the remaining configuration on $(W^{-,\pi}, W^{+,\pi})$ forms a directed multigraph. As before, $N_{d^-,d^+}(n)$ is the number of vertices with degree (d^-, d^+) before percolation. Let $N^{\pi,r}_{d^-,d^+}(n)$ be the number of vertices of original degree (d^-, d^+) that survive site percolation. Then, $N_{d^-,d^+}(n) - N^{\pi,r}_{d^-,d^+}(n)$ is the number of removed vertices of original degree (d^-, d^+). Since each vertex is independently removed with probability $1 - \pi$,

$$\mathbb{E}\left[N^{\pi,r}_{d^-,d^+}(n)\right] = \pi N_{d^-,d^+}(n), \tag{27}$$

$$\mathbb{E}\left[N_{d^-,d^+}(n) - N^{\pi,r}_{d^-,d^+}(n)\right] = (1-\pi) N_{d^-,d^+}(n). \tag{28}$$

We proceed by formulating two lemmas showing that, with high probability, s^-, s^+ remain in some bounded shrinking interval I'_n and r^-, r^+ in I_n.

Lemma 4.8. *Let* $I'_n := \left[m\pi - n^{2/3} \ln n, m\pi + n^{2/3} \ln n\right]$, *then*

$$\mathbb{P}\left[s^-, s^+ \in I'_n\right] = 1 - \mathrm{e}^{-\Omega\left(\ln^2 n\right)}$$

holds separately for, respectively, s^- *and* s^+.

Proof. By using Eq. (27), we obtain

$$\mathbb{E}\left[s^-\right] = \sum_{d^-=0}^{d_{\max}} \sum_{d^+=0}^{d_{\max}} \pi d^- N_{d^-,d^+}(n) = m\pi$$

and

$$\mathbb{E}\left[s^+\right] = \sum_{d^-=0}^{d_{\max}} \sum_{d^+=0}^{d_{\max}} \pi d^+ N_{d^-,d^+}(n) = m\pi.$$

Using Hoeffding's inequality and the bound on the maximum degree, $d_{\max} \le n^{1/9}$ and, we obtain

$$\mathbb{P}\left[\left|s^- - \mathbb{E}\left[s^-\right]\right| > n^{2/3} \ln n\right] \le \mathrm{e}^{-\Omega\left(\ln^2 n\right)} \text{ and } \mathbb{P}\left[\left|s^+ - \mathbb{E}\left[s^+\right]\right| > n^{2/3} \ln n\right] \le \mathrm{e}^{-\Omega\left(\ln^2 n\right)}, \tag{29}$$

which proves the claim. □

Lemma 4.9. *Let* $I_n := \left[m\pi(1-\pi) - n^{2/3} \ln^2 n, m\pi(1-\pi) + n^{2/3} \ln^2 n\right]$, *then*

$$\mathbb{P}\left[r^+, r^- \in I_n \mid s^-, s^+ \in I'_n\right] = 1 - \mathrm{e}^{-\Omega\left(\ln^2 n\right)}$$

holds separately for, respectively, r^- *and* r^+.

Proof. We present the proof for r^-. The proof for r^+ is then identical up to switching the roles of in-stubs and out-stubs. Since we consider a uniformly random configuration on (W^-, W^+), the probability that a given 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 $s^-, s^+ \in I'_n$. Since r^- equals the number of in-stubs in $W^- \setminus W^{-,r}$ with a match in $W^{+,r}$, this implies that

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

To complete the proof, we will now show that

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

This is realized by applying Theorem 3.12 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 kth pair of a configuration in A by (x, y_A) and the kth 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^- changes by at most the symmetric difference of the two matchings, Theorem 3.12 states that

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

as $m \le n d_{\max} \le n^{10/9}$. □

Lemma 4.10. *Let $N^{\pi}_{j,k}(n)$ be the number of vertices in the site-percolated graph (or configuration) with in-degree j and out-degree k. The following limit exists*

$$p^{\text{site}}_{j,k} := \lim_{n\to\infty} \frac{\mathbb{E}\left[N^{\pi}_{j,k}(n)\right]}{n}, \text{for } j, k \in \mathbb{N}_0 \tag{30}$$

and

$$p^{\text{site}}_{j,k} = \begin{cases} \pi p^{\text{bond}}_{j,k}, & (j,k) \ne (0,0), \\ \pi p^{\text{bond}}_{0,0} + 1 - \pi, & (j,k) = (0,0), \end{cases} \tag{31}$$

where $p^{\text{bond}}_{j,k}$ is defined in Lemma 4.4.

Proof. If $j, k > d_{\max}$, then $p^{\text{site}}_{j,k} = N^{\pi}_{j,k}(n) = N_{j,k}(n) = 0$.

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

$$\mathbb{E}\left[N^{\pi}_{0,0}(n)\right] = \sum_{d^-=0}^{d_{\max}} \sum_{d^+=0}^{d_{\max}} \left((1-\pi) N_{d^-,d^+}(n) + \pi P_{0,0}\left(d^-, d^+\right) N_{d^-,d^+}(n)\right), \tag{32}$$

and otherwise,

$$\mathbb{E}\left[N^{\pi}_{j,k}(n)\right] = \sum_{d^-=j}^{d_{\max}} \sum_{d^+=k}^{d_{\max}} \pi P_{j,k}\left(d^-, d^+\right) N_{d^-,d^+}(n). \tag{33}$$

Hence, we need to derive the expression for $P_{j,k}(d^-, d^+)$. Let us determine $P_{j,k}(d^-, d^+, s^-, s^+, r^-, r^+)$, the probability $P_{j,k}(d^-, d^+)$ conditional on the values s^-, s^+, r^-, r^+. Site percolation combines the independent random processes of removing 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 of r^- (respectively, r^+), each subset of $W^- \setminus W^{-,r}$ (respectively, $W^+ \setminus W^{+,r}$) of this size is equally likely to be $W^{-,m}$ (respectively, $W^{+,m}$). Hence,

$$P_{j,k}\left(d^-, d^+, r^-, r^+, s^-, s^+\right) = \binom{d^-}{d^- - j}\binom{d^+}{d^+ - k}\frac{\binom{s^- - d^-}{r^- - d^- + j}}{\binom{s^-}{r^-}}\frac{\binom{s^+ - d^+}{r^+ - d^+ + k}}{\binom{s^+}{r^+}}. \tag{34}$$

Let intervals I_n, I_n' are as defined in Lemmas 4.8 and 4.9. We have that, uniformly in $r^-, r^+ \in I_n$ and $s^-, s^+ \in I_n'$,

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

Plugging this equation into Eq. (34) gives

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

uniformly for all $s^-, s^+ \in I_n'$ and $r^-, r^+ \in I_n$. It turns out that a violation of the latter condition is unlikely, as $\mathbb{P}\left[s^- \notin I_n' \vee s^+ \notin I_n' \vee r^- \notin I_n \vee r^+ \notin I_n\right]$ is small, which in turn helps to bound the value of $\mathbb{E}\left[N_{j,k}^{\pi}(n)\right]$. By applying Hoeffding's inequality, we have

$$\mathbb{P}\left[|N_{d^-,d^+}^{\pi,r}(n) - \mathbb{E}\left[N_{d^+,d^-}^{\pi,r}(n)\right]| > \sqrt{n}\,\ln n\right] < \mathrm{e}^{-\Omega\left(\ln^2 n\right)}, \tag{35}$$

and in combination with Eq. (27), this gives

$$N_{d^-,d^+}^{\pi,r}(n) \in I_n''(d^-, d^+) := \left[\pi N_{d^-,d^+}(n) - \sqrt{n}\,\ln n, \pi N_{d^-,d^+}(n) + \sqrt{n}\,\ln n\right],$$

with probability $1 - \mathrm{e}^{-\Omega\left(\ln^2 n\right)}$. Combining this with Lemmas 4.8 and 4.9 gives

$$\begin{aligned}
&\mathbb{P}\left[s^- \notin I_n' \vee s^+ \notin I_n' \vee r^- \notin I_n \vee r^+ \notin I_n \vee N_{d^-,d^+}^{\pi,r}(n) \notin I_n''\left(d^-, d^+\right)\right] \\
\le&\mathbb{P}\left[s^- \notin I_n'\right] + \mathbb{P}\left[s^+ \notin I_n'\right] + \mathbb{P}\left[r^- \notin I_n\right] + \mathbb{P}\left[r^+ \notin I_n\right] + \mathbb{P}\left[N_{d^-,d^+}^{\pi,r}(n) \notin I_n''\left(d^-, d^+\right)\right] \\
=&o\left(\frac{1}{n^3}\right) + \mathbb{P}\left[r^- \notin I_n\right] + \mathbb{P}\left[r^+ \notin I_n\right],
\end{aligned}$$

where $\mathbb{P}\left[r^- \notin I_n\right] = o\left(n^{-3}\right)$ and $\mathbb{P}\left[r^+ \notin I_n\right] = o\left(n^{-3}\right)$. Therefore,

$$\mathbb{P}\left[s^- \notin I_n' \vee s^+ \notin I_n' \vee r^- \notin I_n \vee r^+ \notin I_n \vee N_{d^-,d^+}^{\pi,r}(n) \notin I_n''\left(d^-, d^+\right)\right] = o\left(n^{-3}\right). \tag{36}$$

This allows us to determine a lower and upper bound for the value $\mathbb{E}\left[N_{j,k}^{\pi}(n)\right]$. Since $N_{d^-,d^+}^{\pi,r}(n) \le N_{d^-,d^+}(n)$ 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 $n > N$

$$\sum_{\substack{(d^-,d^+)=(0,0)\\ d^- \ge \kappa+1 \text{ or } d^+ \ge \kappa+1}}^{(d_{\max},d_{\max})} P_{j,k}\left(d^-, d^+\right) N_{d^-,d^+}^{\pi,r}(n) \le \sum_{\substack{(d^-,d^+)=(0,0)\\ d^- \ge \kappa+1 \text{ or } d^+ \ge \kappa+1}}^{(d_{\max},d_{\max})} N_{d^-,d^+}(n) \le \epsilon n. \tag{37}$$

In combination with Eq. (33), this implies for $(j,k) \neq (0,0)$

$$\sum_{\substack{d^-=j,\\ d^+=k}}^{\kappa} P_{j,k}\left(d^-,d^+\right) N^{\pi,r}_{d^-,d^+}(n) \leq \mathbb{E}\left[N^{\pi}_{j,k}(n)\right] \leq \sum_{\substack{d^-=j,\\ d^+=k}}^{\kappa} P_{j,k}\left(d^-,d^+\right) N^{\pi,r}_{d^-,d^+}(n) + \epsilon n. \tag{38}$$

Using Eq. (36) on the LHS of the above equation, we find

$$\begin{aligned}
\mathbb{E}\left[N^{\pi}_{j,k}(n)\right] &\geq \sum_{\substack{d^-=j,\\ d^+=k}}^{\kappa} \sum_{\substack{\tilde{r}^-\in I'_n,\\ \tilde{r}^+\in I'_n}} \sum_{\substack{\tilde{s}^-\in I_n,\\ \tilde{s}^+\in I_n}} \sum_{\tilde{d}_{d^-,d^+}\in I''_n(d^-,d^+)} \tilde{d}_{d^-,d^+} P_{j,k}\left(d^-,d^+,\tilde{r}^-,\tilde{r}^+,\tilde{s}^-,\tilde{s}^+\right) \\
&\quad \times \mathbb{P}\left[r^-=\tilde{r}^-, r^+=\tilde{r}^+, s^-=\tilde{s}^-, s^+=\tilde{s}^+, N^{\pi,r}_{d^-,d^+}(n)=\tilde{d}_{d^-,d^+}\right] + o\left(\frac{1}{n^2}\right) \\
&= \sum_{\substack{d^-=j,\\ d^+=k}}^{\kappa} \sum_{\tilde{d}_{d^-,d^+}\in I''_n(d^-,d^+)} \tilde{d}_{d^-,d^+} \binom{d^-}{d^--j}\binom{d^+}{d^+-k} \pi^{j+k}(1-\pi)^{d^-+d^+-j-k} \\
&\quad \times \mathbb{P}\left[N^{\pi,r}_{d^-,d^+}(n)=\tilde{d}_{d^-,d^+}\right]\left(1+\mathcal{O}\left(\frac{\ln^2 n}{n^{1/3}}\right)\right) + o\left(\frac{1}{n^2}\right)
\end{aligned}$$

Since Eq. (35) and $N^{\pi,r}_{d^-,d^+}(n) \leq n$ gives

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

we obtain the lower bound

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

In a similar fashion, using the RHS of Eq. (38) gives the upper bound

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

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

$$p^{\text{site}}_{j,k} := \lim_{n\to\infty}\frac{\mathbb{E}\left[N^{\pi}_{j,k}(n)\right]}{n} = \pi\sum_{d^-=j}^{\infty}\sum_{d^+=k}^{\infty} p_{d^-,d^+}\binom{d^-}{j}\binom{d^+}{k}\pi^{j+k}(1-\pi)^{d^--j+d^+-k}. \tag{39}$$

For $(j,k) = (0,0)$, we need to use Eq. (32) instead of Eq. (33). Since $N^{\pi,r}_{d^-,d^+}(n) \le N_{d^-,d^+}(n)$ 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 $n > N$:

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

Thus, the equivalent of Eq. (38) for $(j,k) = (0,0)$ becomes

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

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

$$\lim_{n\to\infty}\frac{\mathbb{E}\left[N^{\pi}_{0,0}(n)\right]}{n} = (1-\pi) + \pi\sum_{\substack{d^-=j,\\ d^+=k}}^{\infty} p_{d^-,d^+}\binom{d^-}{j}\binom{d^+}{k}\pi^{j+k}(1-\pi)^{d^--j+d^+-k} = p^{\text{site}}_{0,0}. \tag{40}$$

Comparing Eqs. (39) and (40) with Eq. (9), we obtain Eq. (31). □

From Eq. (31), we can see that $p^{\text{site}}_{j,k}$ is a probability distribution and has moments

$$\mu^{\pi,\text{site}} := \mu^{\pi,\text{site}}_{10} = \mu^{\pi,\text{site}}_{01} = \pi\mu^{\pi,\text{bond}} = \pi^2\mu \quad \text{and} \quad \mu^{\pi,\text{site}}_{11} = \pi\mu^{\pi,\text{bond}}_{11} = \pi^3\mu_{11}. \tag{41}$$

Additionally, the generating function $U^{\text{site}}_{\pi}(x,y)$ for $p^{\text{site}}_{j,k}$ is given by

$$U^{\text{site}}_{\pi}(x,y) := 1-\pi+\pi U(1-\pi+\pi x, 1-\pi+\pi y), \tag{42}$$

and auxiliary functions U^+_π, U^-_π are the same as in Eq. (16). It is left to determine $\hat{\pi}^{\text{site}}$. Theorem 3.11 states that the percolation threshold is $\pi = \hat{\pi}^{\text{site}}$ such that $\sum_{j,k=0}^{\infty} jkp^{\text{site}}_{j,k} = \sum_{j,k=0}^{\infty} jp^{\text{site}}_{j,k}$. Combining this with Eq. (41), we find that the percolation thresholds for site and bond percolation are equal:

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

This can be explained by noting that the expected degree distribution after site percolation is a rescaled version of the degree distribution after bond percolation, exept for $(0,0)$. Hence, one expects the 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.

4.2.3. *Determining π^{site}_c and c^{site}*

To finalize the proof of Theorems 2.2 and 2.3 for the case of site percolation, it remains to show that $\pi^{\text{site}}_c = \hat{\pi}^{\text{site}}$ as well as to determine c^{site}, the fraction of vertices in GSCC. This is done by following a similar exposition as in Section 4.1.3, which covers bond percolation. We therefore focus on explaining the necessary adjustments to be made to apply the derivations in Section 4.1.3 to site percolation.

First, we need to replace the distribution $p^{\text{bond}}_{j,k}$ with $p^{\text{site}}_{j,k}$. Lemma 4.7 proves the equivalent statement for site percolation to that of Lemma 4.2 for bond percolation. However, the counterpart of Eq. (22)

requires a different proof. Conditional on a certain realization of $(W^{-,r}, W^{+,r})$ and the values $s^-, s^+ \in I'_n$, $r^-, r^+ \in I_n$, the value of $N^{\pi}_{j,k}(n)$ is determined by the random choice of $(W^{-,m}, W^{+,m})$. By changing one element of $(W^{-,m}, W^{+,m})$, the value of $N^{\pi}_{j,k}(n)$ changes by at most 2. Thus, Theorem 3.13 can be applied to obtain

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

Using Lemmas 4.8 and 4.9, it follows that

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

and, since κ is bounded, this completes the proof of the counterpart of Eq. (22).

The last change is related to the fact that Theorem 3.11 is now applied to a feasible degree progression with $p^{\text{site}}_{j,k}$ as degree distribution instead of $p^{\text{bond}}_{j,k}$. In Section 4.2.2, we found that $\hat{\pi}^{\text{site}} = \frac{\mu}{\mu_{11}}$. Furthermore, in analogy to calculations in Section 4.1.3, we choose x^*, y^* to be the solution of Eq. (25) to find that

$$c^{\text{site}} := 1 - U^{\text{site}}_{\pi}(x^*, 1) - U^{\text{site}}_{\pi}(1, y^*) + U^{\text{site}}_{\pi}(x^*, y^*) = \pi c^{\text{bond}}.$$

which completes the proof of Theorems 2.2 and 2.3 for site percolation.

Acknowledgement. The authors are grateful to Rik Versendaal for corrections and improvements to earlier drafts.

References

[1] Amini, H. (2010). Bootstrap percolation and diffusion in random graphs with given vertex degrees. *The Electronic Journal of Combinatorics 17*(1): R25.

[2] Bloznelis, M., Götze, F., & Jaworski, J. (2012). Birth of a strongly connected giant in an inhomogeneous random digraph. *Journal of Applied Probability 49*(3): 601–611.

[3] Bollobás, B., Borgs, C., Chayes, J., & Riordan, O. (2010). Percolation on dense graph sequences. *The Annals of Probability 38*(1): 150–183.

[4] Bollobás, B. & Riordan, O. (2006). *Percolation*, Cambridge: Cambridge University Press.

[5] Cai, X.S., Caputo, P., Perarnau, G., & Quattropani, M. (2021). Rankings in directed configuration models with heavy tailed in-degrees. preprint, arXiv:2104.08389.

[6] Cai, X.S. & Perarnau, G. (2020). Minimum stationary values of sparse random directed graphs. preprint, arXiv:2010.07246.

[7] Cai, X.S. & Perarnau, G. (2021). The giant component of the directed configuration model revisited. *ALEA Latin American Journal of Probability and Mathematical Statistics 18*(1): 1517–1528.

[8] Cai, X.S. & Perarnau, G. (2023). The diameter of the directed configuration model. *Annales de l'Institut Henri Poincare (B) Probabilites et statistiques 59*(1): 244–270.

[9] Chen, N. & Olvera-Cravioto, M. (2013). Directed random graphs with given degree distributions. *Stochastic Systems 3*(1): 147–186.

[10] Cooper, C. & Frieze, A. (2004). 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.

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

[12] Coulson, M. (2021). The strong component structure of the barely subcritical directed configuration model. preprint, arXiv:2105.11434.

[13] Coulson, M. & Perarnau, G (2021). Weak components of the directed configuration model. In Nešetřil, Jaroslav and Rué Juanjo (eds.) *Extended Abstracts EuroComb 2021*, 14, Trends in Mathematics, Barcelona: Birkhäuser Springer, pp. 682–687.

[14] Donderwinkel, S. & Xie, Z. (2021). Universality for the directed configuration model: metric space convergence of the strongly connected components at criticality. Preprint, arXiv:2105.11434.

[15] Durrett, R. (2007). *Random graph dynamics*. Cambridge: Cambridge University Press.
[16] Fountoulakis, N. (2007). Percolation on sparse random graphs with given degree sequence. *Internet Mathematics 4*(4): 329–356.
[17] Goldschmidt, C. & Stephenson, R. (2019). The scaling limit of a critical random directed graph. preprint, arXiv:1905.05397.
[18] Graf, A. (2016). On the strongly connected components of random directed gaphs with given degree sequences, Waterloo, CA: University of Waterloo
[19] Janson, S. (2009). On percolation in random graphs with given vertex degrees. *Electronic Journal of Probability 14*(1): 86–118.
[20] LaMar, M.D. (2011). 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)*: Phoenix. IEEE, pp. 3348–3359.
[21] McDiarmid, C. (1989). On the method of bounded differences. *Surveys in Combinatorics 141*(1): 148–188.
[22] Molloy, M. & Reed, B. (1995). A critical point for random graphs with a given degree sequence. *Random Structures & Algorithms 6*(2–3): 161–180.
[23] Molloy, M. & Reed, B. (1998). The size of the giant component of a random graph with a given degree sequence. *Combinatorics, Probability and Computing 7*(3): 295–305.
[24] Penrose, M.D. (2016). The strong giant in a random digraph. *Journal of Applied Probability 53*(1): 57–70.
[25] Petrov, V.V. (1995). *Limit theorems of probability theory: Sequences of independent random variables*. Oxford: Clarendon Press.
[26] Pittel, B. & Poole, D. (2016). Asymptotic distribution of the numbers of vertices and arcs of the giant strong component in sparse random digraphs. *Random Structures & Algorithms 49*(1): 3–64.
[27] van der Hofstad, R. (2016). *Random graphs and complex networks*, Vol. 1. Cambridge: Cambridge University Press.
[28] van Ieperen, F. (2020). Percolation in random directed graphs with an arbitrary degree distribution. Thesis, Utrecht University.
[29] Łuczak, T. & Cohen, J.E. (1992). Giant components in three-parameter random directed graphs. *Advances in Applied Probability 24*(4): 845–857.
[30] Łuczak, T. & Seierstad, T.G. (2009). The critical behavior of random digraphs. *Random Structures & Algorithms 35*(3): 271–293.

Cite this article: van Ieperen F. and Kryven I. (2024). Percolation in simple directed random graphs with a given degree distribution. *Probability in the Engineering and Informational Sciences* 38(2): 268–289. https://doi.org/10.1017/S0269964823000128