Bond Percolation in Small-World Graphs with Power-Law Distribution

Luca Becchetti
Sapienza University of Rome, Italy

Andrea Clementi
University of Rome Tor Vergata, Italy

Francesco Pasquale
University of Rome Tor Vergata, Italy

Luca Trevisan
Bocconi University, Milan, Italy

Isabella Ziccardi
Bocconi University, Milan, Italy

Abstract

Full-bond percolation with parameter p is the process in which, given a graph, for every edge independently, we keep the edge with probability p and delete it with probability $1-p$. Bond percolation is studied in parallel computing and network science to understand the resilience of distributed systems to random link failure and the spread of information in networks through unreliable links. Moreover, the full-bond percolation is equivalent to the *Reed-Frost process*, a network version of *SIR* epidemic spreading.

We consider *one-dimensional power-law small-world graphs* with parameter α obtained as the union of a cycle with additional long-range random edges: each pair of nodes $\{u, v\}$ at distance L on the cycle is connected by a long-range edge $\{u, v\}$, with probability proportional to $1/L^{\alpha}$. Our analysis determines three phases for the percolation subgraph G_p of the small-world graph, depending on the value of α.

- If $\alpha < 1$, there is a $p < 1$ such that, with high probability, there are $\Omega(n)$ nodes that are reachable in G_p from one another in $\mathcal{O}(\log n)$ hops;
- If $1 < \alpha < 2$, there is a $p < 1$ such that, with high probability, there are $\Omega(n)$ nodes that are reachable in G_p from one another in $\log^{\mathcal{O}(1)}(n)$ hops;
- If $\alpha > 2$, for every $p < 1$, with high probability all connected components of G_p have size $\mathcal{O}(\log n)$.

2012 ACM Subject Classification Theory of computation → Random network models

Keywords and phrases Information spreading, gossiping, epidemics, fault-tolerance, network self-organization and formation, complex systems, social and transportation networks

Digital Object Identifier 10.4230/LIPIcs.SAND.2023.3

Related Version *Full Version*: `https://arxiv.org/pdf/2205.08774.pdf`

Funding *Luca Trevisan and Isabella Ziccardi:* Supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 834861).
Luca Becchetti: Supported by the ERC Advanced Grant 788893 AMDROMA, EC H2020RIA project SoBigData++ (871042), PNRR MUR project PE0000013-FAIR, PNRR MUR project IR0000013-SoBigData.it.
Andrea Clementi: Partially supported by Spoke 1 "FutureHPC & BigData" of the *Italian Research Center on High-Performance Computing, Big Data and Quantum Computing (ICSC)* funded by *MUR Missione 4 Componente 2 Investimento 1.4: Potenziamento strutture di ricerca e creazione di "campioni nazionali" di R&S (M4C2-19) – Next Generation EU (NGEU).*

1 Introduction

Given a graph $G = (V, E)$ and a probability $p(e)$ associated to each edge $e \in E$, the *full-bond percolation* process[1] on G is the construction of a random subgraph $G_p = (V, E_p)$ of G, called the percolation graph, obtained by selecting each edge $e \in E$ to belong to E_p with probability $p(e)$, independently of the other edges. Often, the focus is on the *homogeneous* case in which all probabilities $p(e)$ are equal to the same parameter p. The main questions of interest in this case are, depending on the choice of G and p, what is the typical size of the connected components of G_p and the typical distances between reachable vertices.

The study of the percolation process originates from mathematical physics [14, 20, 23] and it has several applications in parallel and distributed computing and network science [1, 9, 13, 10, 12, 17]. For example, the study of network reliability in the presence of random link failures is equivalent to the study of the connectivity properties of the percolation graph of the network of links [12, 16].

The *independent cascade* is a process that models the spread of information and the influence of individual choices on others in social networks, and it is equivalent to a percolation process in a way that we explain next. In the independent cascade, we have a network $G = (V, E)$, an *influence* probability $p(e)$ associated to each edge, and a set $I_0 \subseteq V$ of network nodes that initially have a certain opinion[2]. The process evolves over time according to the following natural local rule: if a node u acquired the opinion at time t, the node v does not have the opinion, and the edge (u, v) exists in G, then node u will attempt to convince node v of the opinion, and it will succeed with probability $p(u, v)$. All nodes that were successfully convinced by at least one of their neighbors at time t will acquire the opinion, and will attempt to convince their neighbors at time $t + 1$ and so on. The independent cascade is studied in [13], where the problem of interest is to find the most "influential" initial set I_0. The resulting epidemic process is shown in [13] to be equivalent to percolation in the following sense: the distribution of nodes influenced by I_0 in the independent cascade process has the same distribution as the set of nodes reachable from I_0 in the percolation graph of G derived using the probabilities $p(e)$.

The *Reed-Frost* process is one of the simplest and cleanest models of *Susceptible-Infectious-Recovered* (*SIR*) epidemic spreading on networks [22, 24]. This process is identical to independent cascade with a fixed probability $p(e) = p$ for all edges e. We can interpret nodes that acquired the opinion in the previous step as *Infectious* nodes that can spread the disease/opinion, nodes that do not have the disease/opinion as nodes *Susceptible* to the infection, and nodes that acquired the disease/opinion two or more steps in the past as *Recovered* nodes that do not spread the disease any more and are immune to it. The probability p corresponds to the probability that a contact between an infectious person and a susceptible one causes a transmission of the disease from the former to the latter. The set I_0 is the initial set of infectious people at time 0. This process, being equivalent to independent cascade, is also equivalent to percolation [13]: the distribution of nodes that are infected and eventually recover in the Reed-Frost process is the same as the distribution of the set of nodes reachable from I_0 in the percolation graph of G derived using the probabilities $p(e)$. Furthermore, the distribution of nodes that are infectious at time t is precisely the distribution of nodes at distance t from I_0 in the percolation graph (see [8]).

[1] We simply write bond percolation when no confusion arises.
[2] Or hold a certain piece of information, or perform a certain action, these are all equivalent views that lead to the same distributed process.

We are interested in studying full-bond percolation (and hence reliability under random link failure, independent cascade, and Reed-Frost epidemic spreading) in one-dimensional small-world graphs with power-law distribution of edges. *Small-world graphs* are a collection of generative models of graphs, designed to capture properties of real-life social networks [10, 22] and communication networks [11]. In the model introduced by Watts and Strogatz [25], the network is obtained as a one-dimensional or two-dimensional lattice in which certain edges are re-routed to random destinations. In a refined model introduced by Kleinberg [7, 10], a possible edge between two nodes at distance L in the lattice exists with probability proportional to $L^{-\alpha}$, where the exponent α is a parameter of the model.

We study full bond percolation on the variant of Kleinberg's model defined below. This model has already been adopted in several previous papers [4, 5, 7, 10, 15, 22] to study bond percolation and epidemic processes, and to discover structural properties determining the performances of diffusion and navigations problems in real networks [26],

▶ **Definition 1** (1-D power-law small-world graphs). *For every $n \geqslant 3$ and $\alpha > 0$, an (undirected) random graph $G = (V, E)$ with $V = \{0, \ldots, n-1\}$ is sampled according to the distribution $\mathcal{SW}(n, \alpha)$ if $E = E_1 \cup E_2$, where: (V, E_1) is a cycle and its edges are called* ring-edges*, and, for each pair of non-adjacent vertices $u, v \in V$, the* bridge *$\{u, v\}$ is included in E_2 independently, with probability*

$$\mathbf{Pr}\left(\{u, v\} \in E_2\right) = \frac{1}{d(u, v)^{\alpha}} \cdot \frac{1}{C(\alpha, n)},$$

where $d(\cdot, \cdot)$ is the shortest-path distance in the ring and $C(\alpha, n)$ is the normalizing constant[3] $C(\alpha, n) = 2\sum_{x=2}^{n/2} x^{-\alpha}$.

The process of *long-range percolation* in the one-dimensional case is the variant of the percolation process applied to the generative model described above in which ring edges are preserved with probability one. The resulting percolation graphs are always connected, and the main question of interest is their diameter. Long-range percolation is well understood, and the one-dimensional case is studied in [4, 15, 25]. In particular, [4] provides bounds on the diameter and on the expansion of such graphs as a function of the power-law exponent α. Such results have been then sharpened and generalized to multi-dimensional boxes in [6]. The long-range percolation process, however, is not a realistic generative model for epidemiological processes, because even the most contagious diseases, including Ebola, do not have 100% probability of spreading through close contacts (see [18, 21, 22] for a discussion of this point).

Full-bond percolation in power-law small-world graphs has been studied in the case of infinite lattices, including the one-dimensional case that is the infinite analog of the model that we study in this paper. In the infinite case, the main questions of interest, which are studied in [5], are whether the percolation graph has an infinite connected component and, given two vertices, what is their typical distance in the percolation conditioned on them both being in the infinite component, as a function of their distance in the lattice. Although there are similarities, techniques developed to study infinite percolation graphs do not immediately apply to the finite case.

[3] Note that $C(n, \alpha)$ is not, strictly speaking, a constant, but rather a normalizing factor that depends on both α and n. It is always upper bounded by an absolute constant across the entire range of α, while it falls within an interval bounded by two constants when $\alpha > 1$. For the sake of conciseness, abusing terminology we write "constant" instead of normalizing factor.

1.1 Roadmap

In this paper, we study full-bond percolation in power-law small-world graphs. The rest of this paper is organized as follows. Section 2 provides an overview of our results and their main consequences. Section 3 introduces notation and some key preliminary notions, while Sections 4–6 provide the full analysis of full-bond percolation in one-dimensional power-law, small-world graphs. Each of Sections 4–6 includes a preliminary, informal description of the main techniques we adopt for that particular regime and a comparison with previous approaches.

Due to lack of space, the proofs of some technical results and a more extensive review of previous work is given in the full version of this paper [2].

2 Our Contribution

We analyze the bond percolation process over the small-world graph G sampled according to the distribution $\mathcal{SW}(n, \alpha)$. Consistently with previous results in long-range percolation models [4], our analysis shows that the process exhibits three different behaviors determined by different values of α. We formally state such results in the next three theorems.

▶ **Theorem 2** (Case $\alpha \in (2, +\infty)$)**.** *Let $\alpha > 2$ be a constant and $p < 1$ a percolation probability. Sample a graph $G = (V, E)$ from the $\mathcal{SW}(n, \alpha)$ distribution and let G_p be the percolation graph of G with percolation probability p. The following holds:*

1. *W.h.p., the connected components of G_p have size at most $\mathcal{O}(\log n)$;*
2. *For each node $v \in V$ and for any sufficiently large ℓ, with probability $1 - \Omega(\ell^{-(\alpha-2)/2})$, every node connected to v in G_p is at ring-distance no larger than $\mathcal{O}(\ell^2)$ from v.*

From an epidemiological point of view, this first regime is thus characterized by a negligible chance to observe an outbreak according to the Reed-Frost process, even in the presence of a large number (say some small root of n) of initially infected nodes (i.e. sources). In particular, the second claim of the theorem above strongly bounds the possible infected area of the ring graph.

The following case, determined by the range $1 < \alpha < 2$, shows the most interesting behavior.

▶ **Theorem 3** (Case $\alpha \in (1, 2)$)**.** *Let $\alpha \in (1, 2)$ be a constant and p a percolation probability. Sample a graph $G = (V, E)$ from the $\mathcal{SW}(n, \alpha)$ distribution and let G_p be the percolation graph of G with percolation probability p. Then, constants $\underline{p}, \overline{p} \in (0, 1)$ (with $\underline{p} \leqslant \overline{p}$) exist such that the following holds:*

1. *If $p > \overline{p}$, w.h.p., there exists a set of $\Omega(n)$ nodes that induces a connected sub-component in G_p with diameter $\mathcal{O}(\text{polylog}(n))$;*
2. *If $p < \underline{p}$, w.h.p. all the connected components of G_p have size $\mathcal{O}(\log n)$.*

The first claim above implies that, if p is sufficiently large (but still a constant smaller than 1), then, there is a good chance that few source nodes are able to infect a large (i.e. $\Omega(n)$) number of nodes and, importantly, this outbreak takes place at an almost exponential speed.

Finally, when $\alpha < 1$, we show the emergence of a behavior similar to that generated by one-dimensional small-world models with bridges selected according to the Erdős-Rényi distribution [3, 19].

▶ **Theorem 4** (Case $\alpha \in (0,1)$). *Let $\alpha \in (0,1)$ and p a percolation probability. Sample a graph $G = (V, E)$ from the $\mathcal{SW}(n, \alpha)$ distribution and let G_p be the percolation graph of G with percolation probability p. Then, constants $\overline{p}, \underline{p} \in (0,1)$ (with $\underline{p} \leqslant \overline{p}$) exist, such that the following holds:*

1. *If $p > \overline{p}$, w.h.p., there exists a set of $\Omega(n)$ nodes that induces a connected sub-component in G_p with diameter $\mathcal{O}(\log n)$. Moreover, for a sufficiently large $\beta > 0$, for any subset $S \subseteq V$ of size $|S| \geqslant \beta \log n$, the subset of nodes within distance $\mathcal{O}(\log n)$ in G_p from S has size $\Omega(n)$, w.h.p.*
2. *If $p < \underline{p}$, w.h.p. all the connected components of G_p have size $\mathcal{O}(\log n)$.*

In this last case, the presence of a sparse subset of relatively-long random bridges implies, above the probability threshold, that a few (i.e. $\Omega(\log n)$) sources w.h.p. generate a large outbreak at exponential speed.

We conclude this section by observing that three phases above are characterized by sharply different distributions of the typical length (measured according to the ring metric) of the bridges. To gauge the difference, consider the expectation, for a fixed vertex, of the sum of the lengths (in ring metric) of the bridges incident on the vertex, and call this expectation $BL_{\alpha,n}$.

When $\alpha < 1$, we have that $BL_{\alpha,n}$ is linear in n. When $1 < \alpha < 2$, then $BL_{\alpha,n}$ is of the form $O(n^{2-\alpha})$, going to infinity with n, but sublinearly in n. Finally, when $\alpha > 2$, $BL_{\alpha,n}$ is a constant that depends only on α and is independent of n. Nodes have, in expectation, only one bridge, so $BL_{\alpha,n}$ is an indication of how much we can advance on the ring by following one bridge.

When $\alpha < 1$, the bridges are basically as good as random edges, and we would expect a giant component to emerge even after full-bond percolation, if p is large enough. When $\alpha > 2$, the bridges behave like a constant number of ring-edges, and we would not expect a large component when $p < 1$.

The case $1 < \alpha < 2$ is the one for which it is hardest to build intuition, and the fact that the bridges typically have length of the form $n^{1-\Omega(1)}$ might suggest that it would take $n^{\Omega(1)}$ steps to reach antipodal nodes. Previous work on long-range percolation had established a polylogarithmic diameter bound in the model in which ring-edges are not subject to percolation. In that model, all pairs of nodes are reachable in a polylogarithmic number of steps even though the typical bridge has length $n^{1-\Omega(1)}$, and this suggests that the shortest path structure is such that a small number of long bridges is used by several shortest paths. One thus would expect such a structure to be sensitive to full-bond percolation, and indeed the proof of [4] relies on the deterministic presence of the ring-edges. Instead, we prove that, when $1 < \alpha < 2$, w.h.p, most pairs of nodes are reachable from one another in a polylogarithmic number of hops after the full-bond percolation process.

3 Model and Preliminaries

In this paper, we study bond percolation of graphs sampled from $\mathcal{SW}(n, \alpha)$, as formalized in Definition 1. In this section, we define notation, key notions and tools that will be used throughout the rest of this paper. Further notation used in the proofs of specific results is introduced wherever it is used.

For the sake of completeness, we begin with the formal definition of bond percolation.

▶ **Definition 5** (Bond percolation). *Given a graph $G = (V, E)$ and a real $p \in [0,1]$, its bond* percolation graph *G_p is the random subgraph obtained from G by removing each edge $e \in E$ independently, with probability $1 - p$.*

Considered any graph $G = (V, E)$ and $v \in V$, we denote by $\mathcal{N}_G(v)$ the neighborhood of v in G, while $\deg_G(v) = |\mathcal{N}_G(v)|$ denotes the degree of v in G. We omit the subscript when G is clear from context.

If $G = (V, E_1 \cup E_2)$ is sampled from $\mathcal{SW}(n, \alpha)$ as in Definition 1, we say that a bridge $\{u, v\} \in E_2$ has *length* $d(u, v)$ (where $d(\cdot, \cdot)$ is as in Definition 1). We also say that u is at *ring-distance* $d(u, v)$ from v (and viceversa). Considered a bridge $\{u, v\}$, we say that $\{u, v\}$ is on the *clockwise* (respectively, *counter-clockwise*) side of u if $d(u, v)$ corresponds to moving clockwise (respectively, counter-clockwise) along the cycle (V, E_1) from u to v.

Given a graph $G = (V, E)$ and $s \in V$, we denote by $\Gamma_G(s)$ the connected component containing s in G. We may omit G when clear from context, while with a slight abuse of notation, we simply write $\Gamma_p(s)$ for $\Gamma_{G_p}(s)$ when we refer to the percolation G_p of some graph G, which will always be understood from context. Given $G = (V, E)$ and $S \subseteq V$, $\mathrm{diam}_G(S)$ is equal to the diameter of the subgraph of G induced by S if this is connected, otherwise $\mathrm{diam}_G(S) = \infty$. With a slight abuse of notation, we write $\mathrm{diam}_p(S)$ when G is the percolation graph G_p. Given $G = (V, E)$ sampled from $\mathcal{SW}(n, \alpha)$ and any subgraph $H = (V, E')$ such that $E' \subseteq E$ (e.g., G_p), we associate a *ring-metric* to H, so that the *ring-distance* between u, v is simply $d(u, v)$ defined above on G.

In the sections that follow, unless stated otherwise, probabilities are always taken over both the randomness in the sampling of G from $\mathcal{SW}(n, \alpha)$ and over the randomness of the percolation. We further remark that our choice of the normalizing constant $C(\alpha, n)$ in Definition 1 entails $\mathbf{E}[\deg(v)] = 3$, while the following, preliminary fact follows from a straightforward application of Chernoff bound:

▷ Claim 6. Sample a graph $G = (V, E)$ from $\mathcal{SW}(n, \alpha)$. Then,

$$\mathbf{Pr}\left(\max_v \deg(v) \leqslant 4 \log n + 2\right) \geqslant 1 - \frac{1}{n}\,. \tag{1}$$

3.1 Galton-Watson Branching Processes

Our analysis of the percolation process in part relies on a reduction to the analysis of appropriately defined branching processes.

▶ **Definition 7** (Galton-Watson Branching Process). *Let W be a non-negative integer random variable, and let $\{W_{t,i}\}_{t \geqslant 1, i \geqslant 1}$ be an infinite sequence of independent identically distributed copies of W. The* Galton-Watson branching process *generated by the random variable W is the process $\{X_t\}_{t \geqslant 0}$ defined by $X_0 = 1$ and by the recursion*

$$X_t = \sum_{i=1}^{X_{t-1}} W_{t,i}\,.$$

All properties of the process $\{X_t\}_{t \geqslant 0}$ (in particular, population size and extinction probability) are captured by the equivalent process $\{B_t\}_{t \geqslant 0}$, recursively defined as follows:

$$B_t = \begin{cases} 1, & t = 0; \\ B_{t-1} + W_t - 1, & t > 0 \text{ and } B_{t-1} > 0; \\ 0, & t > 0 \text{ and } B_{t-1} = 0\,, \end{cases}$$

where $W_1, W_2, \ldots$ are an infinite sequence of independent and identically distributed copies of W.

In the remainder, when we refer to the Galton-Watson process generated by W, we always mean the process $\{B_t\}_{t \geqslant 0}$. In particular, we define $\sigma = \min\{t > 0 : B_t = 0\}$ (we set $\sigma = +\infty$ if no such t exists). Note that, for any $T < \sigma$, we have $B_T = \sum_{t=1}^{T} W_t - T$.

4 The case $\alpha > 2$

We recall that G_p is the percolation graph of G, sampled from the $\mathcal{SW}(n, \alpha)$ distribution. When $\alpha > 2$, we show that each component of the percolation graph G_p has w.h.p. at most $\mathcal{O}(\log n)$ nodes. To prove this fact, we need to cope with the complex "connectivity" of G_p yielded by the percolation of both ring-edges and the random bridges. To better analyze this structure, we introduce the notion of *ℓ-graph* $G_p^{(\ell)}$ of G_p, where ℓ is any fixed integer $\ell > 0$. This new graph $G_p^{(\ell)}$ is in turn a one-dimensional small-world graph of n/ℓ "supernodes". It is defined by any fixed partition of V of disjoint ring intervals, each of them formed by ℓ nodes: such n/ℓ intervals are the nodes, called *super-nodes*, of $G_p^{(\ell)}$. The set of edges of $G_p^{(\ell)}$ is formed by two types of random links: the *super-edges* that connect two adjacent super-nodes and the *super-bridges* connecting two non adjacent super-nodes of the n/ℓ-size ring (for the formal definition of $G_p^{(\ell)}$ see Definition 10). We then prove that, for any $\alpha > 2$ and any $p < 1$, each super-node of this graph has:

1. Constant probability to have no neighbors (for every value of ℓ);
2. Probability $\mathcal{O}(1/\ell^{\alpha-2})$ to be incident to a super-bridge.

We then design an appropriate BFS visit of $G_p^{(\ell)}$ (see Algorithm 1) that keeps the role of super-edges and super-bridges well-separated. In more detail, starting from a single super-node s, we show that the number of super-nodes explored at each iteration of the visit turns out to be dominated by a branching process having two distinct additive contributions: one generated by the new, visited percolated super-edges and the other generated by the new, visited percolated super-bridges. Special care is required to avoid too-rough redundancy in counting possible overlapping contributions from such two experiments. Then, thanks to Claims 1 and 2 above, we can prove that, for $\ell = \mathcal{O}(1)$ sufficiently large, w.h.p. this branching process ends after $\mathcal{O}(\log n)$ steps, and this proves that the connected component of s in $G_p^{(\ell)}$ has $\mathcal{O}(\log n)$ super-nodes. We also remark that our approach above also implies the following interesting result: with probability $1 - \mathcal{O}(1/\ell^{\alpha-2})$, all super-nodes connected to v are at ring distance at most ℓ from v. This implies that, for each node s, the nodes in the connected component of s in G_p are within ring distance $\mathcal{O}(\ell^2)$ from s, with the same probability.

We now proceed formally with the proof of Theorem 2, which is a consequence of the two lemmas below.

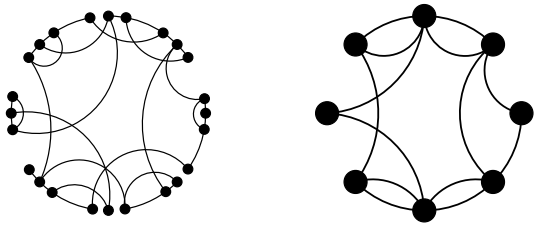
▶ **Lemma 8.** *Let $\alpha < 2$ be a constant and $p < 1$ a percolation probability. Sample a graph $G = (V, E)$ from the $\mathcal{SW}(n, \alpha)$ distribution and let G_p its percolation graph. For every $s \in V = \{0, \ldots, n-1\}$, the connected component $\Gamma_p(s)$ contains $\mathcal{O}(\log n)$ nodes with probability at least $1 - 1/n^2$.*

▶ **Lemma 9.** *Under the same hypotheses of Lemma 8, for every $s \in V = \{0, \ldots, n-1\}$ and for any sufficiently large ℓ,*

$$\mathbf{Pr}\left(\forall u \in \Gamma_p(s) : d(s,u) \leqslant 2\ell^2\right) \geqslant 1 - \frac{8}{(\alpha-2)\ell^{(\alpha-2)/2}}\,. \tag{2}$$

We remark that (2) implies that, for any increasing distance function $\ell = \ell(n) = \omega(1)$, every node in $\Gamma_p(s)$ has ring-distance from s not exceeding $2\ell^2$ with probability $1 - o(1)$.

The proof of Lemma 8 is given in Section 4.2, while the proof of Lemma 9 in Section 4.3. To prove these lemmas, we rely on the notion of ℓ-graph, defined in the following paragraph, together with supplementary notation that will be used in the remainder of this section.



(a) (b)

Figure 1 (a) A graph $G = (V, E)$ with a ring-metric. (b) The 3-graph of G.

4.1 ℓ-graphs

In what follows, we define a new graph on the vertex set $V = \{0, \dots, n-1\}$, starting from any one-dimensional small-world graph $G = (V, E_1 \cup E_2)$, where (V, E_1) is a cycle, which defines a ring-metric, and E_2 is an arbitrary subset of bridges. In the remainder, we always assume $n \geqslant 3$.

▶ **Definition 10** (ℓ-graph). *Let $H = (V, E_H)$ be a subgraph of a one-dimensional small-world graph G, where $E_H \subseteq E_1 \cup E_2$. For $\ell \geqslant 1$, consider an arbitrary partition of V into n/ℓ disjoint intervals of length ℓ, $\{I^{(1)}, \dots, I^{(n/\ell)}\}$ with respect to the ring metric induced by (V, E_1). The* ℓ-graph *associated to H is the graph $H^{(\ell)} = (V^{(\ell)}, E_H^{(\ell)})$, where $V^{(\ell)} = \{I^{(1)}, \dots, I^{(n/\ell)}\}$, and*

$$E_H^{(\ell)} = \left\{ \{I^{(h)}, I^{(k)}\} : \exists u \in I^{(h)}, v \in I^{(k)} \text{ s.t. } \{u, v\} \in E_H \right\}.$$

A generic element in $V^{(\ell)}$ thus corresponds to ℓ consecutive nodes in the ring (V, E_1), it is called *super-node*, and, depending on the context, is denoted by v or by its corresponding interval I_{v} in the partition $\{I^{(1)}, \dots, I^{(n/\ell)}\}$ of V. Fixed a partition, we denote with $E_1^{(\ell)}$ the set of links connecting two adjacent super-nodes in $V^{(\ell)}$ (that is, two adjacent intervals in (V, E_1)): notice that $(V^{(\ell)}, E_1^{(\ell)})$ is a ring of n/ℓ super-nodes, called *ℓ-ring*. Then, given a subgraph $H = (V, E_H)$ as in the definition above, the elements in $E_H^{(\ell)}$ can be partitioned into two subsets: the elements in $E_H^{(\ell)} \cap E_1^{(\ell)}$ are called *super-edges*, while all the remaining elements in $E_H^{(\ell)}$ are called *super-bridges*. $H^{(\ell)}$ can thus be seen as a subgraph of a one-dimensional small-world graph $G^{(\ell)}$ with n/ℓ super-nodes, formed by a ring $(V^{(\ell)}, E_1^{(\ell)})$, and an additional set of super-bridges. The example in Figure 1 summarizes the above definitions. Let $H = (V, E_H)$ as in Definition 10 and assume $u, v \in V$, with $u \in I^{(h)}$ and $v \in I^{(k)}$. Clearly, if $(u, v) \in E_H$ then $(I^{(h)}, I^{(k)}) \in E_H^{(\ell)}$. Moreover, the following fact straightforwardly holds:

▷ Claim 11. Let $H = (V, E_H)$ as in Definition 10 and let $s \in I^{(k)}$ for some $k \in \{1, \dots, n/\ell\}$. Then,

$$|\Gamma_H(s)| \leqslant \ell |\Gamma_{H^{(\ell)}}(I^{(k)})|.$$

The following, preliminary lemma is used in the proofs of Lemma 8 and Lemma 9.

▶ **Lemma 12.** *Assume the hypotheses of Lemma 8 (and Lemma 9), and let ${G_p}^{(\ell)} = ({V}^{(\ell)}, {E_p}^{(\ell)})$ be the ℓ-graph of G_p generated by any fixed interval partition of V. Then, for each $\mathrm{v} \in V^{(\ell)}$, we have:*

$$\mathbf{Pr}\left(\deg_{{G_p}^{(\ell)}}(\mathrm{v}) = 0\right) \geqslant (1-p)^2 e^{-2/(\alpha-2)}, \tag{3}$$

$$\mathbf{Pr}\left(\mathrm{v} \text{ is incident to a super-bridge in } {G_p}^{(\ell)}\right) \leqslant \frac{2}{(\alpha-2)\ell^{\alpha-2}}. \tag{4}$$

Proof. Consider any super-node $\mathrm{v} \in V^{(\ell)}$ and, for the rest of this proof, denote by w_1 and w_2 the two boundary nodes of I_{V}. Without loss of generality, we assume I_{V} includes the ring-edges that we traverse if we move on the cycle (V, E_1) from w_1 to w_2 counter-clockwise.

We first prove (3). The super-node v has no out-edges in ${G_p}^{(\ell)}$ if and only if i) each node in I_{V} has no bridge to a node in $V \setminus I_{\mathrm{V}}$ and ii) w_1 and w_2 share no-ring edges with nodes in $V \setminus I_{\mathrm{V}}$ in G_p. Condition i) above is equivalent to the following: for every $x = 1, \ldots, \ell$, the node $u \in I_v$ at distance $d(w_1, u) = x$ from w_1 has no bridge of length exceeding x on the clockwise side and of length exceeding $\ell - x$ on the counter-clockwise side. We have

$$\begin{aligned}
&\mathbf{Pr}\left(v \text{ has a bridge with length} > x \text{ in one side}\right) \leqslant \sum_{y=x+1}^{n/2} \frac{1}{C(\alpha,n)y^{\alpha}} \\
&\leqslant \int_{x+1}^{+\infty} \frac{1}{C(\alpha,n)(y-1)^{\alpha}} dy \leqslant \frac{1}{(\alpha-1)C(\alpha,n)x^{\alpha-1}} \leqslant \frac{1}{x^{\alpha-1}},
\end{aligned}$$

where the last inequality follows from the fact that $C(\alpha, n) \geqslant 1/2^{\alpha-1}$. From the inequality above, we have

$$\begin{aligned}
&\mathbf{Pr}\left(\deg_{{G_p}^{(\ell)}}(\mathrm{v}) = 0\right) \geqslant \left[(1-p) \cdot \prod_{x=1}^{\ell}\left(1 - \frac{1}{2x^{\alpha-1}}\right)\right]^2 \\
&\geqslant (1-p)^2 \cdot e^{-2\sum_{x=1}^{\ell} \frac{1}{x^{\alpha-1}}} \geqslant (1-p)^2 e^{-2/(\alpha-2)}.
\end{aligned}$$

We next prove (4). Let v_i a node in $V \setminus I_{\mathrm{V}}$ at ring-distance $i + \ell$ from I_{V}, i.e., such that $\min\{d(w_1, v_i), d(w_2, v_i)\} = i + \ell$. We have:

$$\begin{aligned}
&\mathbf{Pr}\left(v_i \text{ is not a neighbor of any node in } I_{\mathrm{V}} \text{ in } G_p\right) \\
&\geqslant \prod_{x=1}^{\ell}\left(1 - \frac{1}{C(\alpha,n)(x+\ell+i)^{\alpha}}\right) \geqslant e^{-\sum_{x=1}^{\ell} \frac{1}{(x+\ell+i)^{\alpha}}} \geqslant e^{-\frac{1}{(i+\ell)^{\alpha-1}}}.
\end{aligned}$$

Then, let ${(E_p)_2}^{(\ell)}$ denote the set of super-bridges in ${G_p}^{(\ell)}$, we use the above inequality to bound the expected number of super-bridges that are incident in v:

$$\begin{aligned}
&\mathbf{E}\left[\left|\mathcal{N}_{{G_p}^{(\ell)}}(\mathrm{v}) \cap {(E_p)_2}^{(\ell)}\right|\right] = 2\sum_{i=1}^{n/2-\ell} \mathbf{Pr}\left(v_i \text{ has a neighbor in } I_{\mathrm{V}}\right) \\
&\leqslant 2\sum_{i=1}^{+\infty} 1 - e^{-\frac{1}{(i+\ell)^{\alpha-1}}} \leqslant 2\sum_{i=1}^{+\infty} \frac{2}{(i+\ell)^{\alpha-1}} \leqslant \frac{2}{(\alpha-2)\ell^{\alpha-2}}.
\end{aligned}$$

Finally, the proof follows from

$$\mathbf{Pr}\left(\mathrm{v} \text{ has a super-bridge in } {G_p}^{(\ell)}\right) \leqslant \mathbf{E}\left[\left|\mathcal{N}_{{G_p}^{(\ell)}}(\mathrm{v}) \cap {(E_p)_2}^{(\ell)}\right|\right]. \qquad \blacktriangleleft$$

4.2 Proof of Lemma 8

To prove Lemma 8, we consider Algorithm 1 below that performs a BFS visit of the ℓ-graph of the percolation graph G_p. Then, in Lemma 14, we prove that this algorithm terminates after visiting (only) $\mathcal{O}(\log n)$ super-nodes, w.h.p., for a sufficiently large $\ell = \mathcal{O}(1)$. Together with Claim 11, this proves Lemma 8.

Algorithm 1 BFS visit of ${G_p}^{(\ell)}$.

1: **Input**: The ℓ-graph ${G_p}^{(\ell)} = (V^{(\ell)}, {E_p}^{(\ell)})$ of G_p; an initiator (super-node) $\mathrm{S} \subseteq V^{(\ell)}$
2: $Q = \{\mathrm{S}\}$
3: **while** $Q \neq \emptyset$ **do**
4: $\quad \mathrm{W} = \texttt{dequeue}(Q)$
5: $\quad \texttt{visited}(\mathrm{W}) = \texttt{True}$
6: $\quad$ **for** each X s.t. $\{\mathrm{X}, \mathrm{W}\}$ is a super-edge of ${E_p}^{(\ell)}$ and $\texttt{visited}(x) = \texttt{False}$ **do**
7: $\quad\quad \texttt{enqueue}(\mathrm{X}, Q)$
8: $\quad$ **for** each Y s.t. $\{\mathrm{Y}, \mathrm{W}\}$ is a super-bridge of ${E_p}^{(\ell)}$ and $\texttt{visited}(\mathrm{Y}) = \texttt{False}$ **do**
9: $\quad\quad \texttt{enqueue}(\mathrm{Y}, Q)$
10: $\quad\quad \mathrm{Y}_{\text{left}}$ = the super-node at distance 1 from Y on the ring at its left
11: $\quad\quad$ **if** $\texttt{visited}(\mathrm{Y}_{\text{left}}) = \texttt{False}$ and $\{\mathrm{Y}_{\text{left}}, \mathrm{Y}\} \in {E_p}^{(\ell)}$ **then**
12: $\quad\quad\quad \texttt{enqueue}(\mathrm{Y}_{\text{left}}, Q)$

▶ Remark 13. Note that, each time we add a super-node Y to queue Q, we also also add the super-node Y_{left} to its left on the ℓ-ring if Y_{left} is connected Y in $G_p^{(\ell)}$. So, in each while loop at line 3, for each super-node $\mathrm{W} \in Q$, at most one non-visited neighbor of W on the ℓ-ring will be added at line 7 since one of them has been already added to Q at the same while loop in which W has been added to Q (see line 10).

▶ **Lemma 14.** *Assume the hypotheses of Lemma 8 and fix any node $s \in V$. For any fixed interval partition*[4] *of the vertex set V, consider the ℓ-graph ${G_p}^{(\ell)}$ and the super-node $\mathrm{S} \in V^{(\ell)}$ such that $s \in I_S$. Then, a sufficiently large $\ell = \mathcal{O}(1)$ exists, depending only on p and α, such that Algorithm 1 terminates within $\mathcal{O}(\log n)$ iterations of the while loop in line 3, with probability at least $1 - 1/n^2$.*

Proof. For $t = 1, 2, \ldots$, let Q_t be the content of Q at the end of the t-th iteration of the while loop of Algorithm 1, and let W_t denote the subset of super-nodes that were added to Q during the t-th iteration. We have $|Q_0| = 1$ and

$$|Q_t| = \begin{cases} 0 \text{ if } |Q_{t-1}| = 0 \\ |Q_{t-1}| + |W_t| - 1 \text{ otherwise,} \end{cases}$$

Let X_t and Y_t denote the sets of super-nodes that were added to Q in the t-th iteration of the while loop, respectively at line 7 and at line 9. For $t \geqslant 2$, whenever a super-node is added to the queue, the queue also contains a super-node at distance 1 from it on the ℓ-ring, if the corresponding super-bridge is in $E_p^{(\ell)}$. We thus have the following formula $|W_t| = |X_t| + 2|Y_t|$, where $|X_t| \leqslant 1$ for $t \geqslant 2$ and $|X_1| \leqslant 2$. Let $\delta = 1 - p$, $\epsilon = 2 - \alpha$ and note that, from (3) and (4),

$$\mathbf{Pr}\left(|X_t| = 1\right) \leqslant 1 - \delta^2 e^{-2/\epsilon} \quad \text{and} \quad \mathbf{E}\left[|Y_t|\right] \leqslant \frac{2}{\epsilon \ell^{\epsilon}}. \tag{5}$$

Moreover, note that for every t, $|Q_t| = 0$, whenever $\sum_{i=1}^{t} |W_i| \leqslant t$. Then, we can write

$$\sum_{i=1}^{t} |W_i| - t = \sum_{i=1}^{t} |X_i| + 2 \sum_{i=1}^{t} |Y_i| - t\,,$$

[4] According to Definition 10.

where $|X_i|$ are $\{0,1\}$ random variables and, if w is the node extracted from Q in the i-th iteration of the while loop, $|Y_i|$ is the number of super-bridge neighbors of w, so that it can also be written as the sum of random variables in $\{0,1\}$. Note that $\{X_t\}_t$ and $\{Y_t\}_t$ are not independent random variables, because of the conditions that appear in lines 6, 8 and 11 of Algorithm 1. Still, it is easy to show that X_t's and Y_t's are dominated by independent copies of two random variables X and Y, such that $\mathbf{Pr}(|X|=1) = 1-\delta^2 e^{-2/\epsilon}$ and $\mathbf{E}[|Y|] \leqslant 1/(\epsilon\ell^\epsilon)$. Hence, Chernoff bound and (5) imply that, for a certain $t = \mathcal{O}(\log n)$ depending on δ and ϵ,

$$\mathbf{Pr}\left(\sum_{i=1}^{t}|X_i| \geqslant \left(1-\frac{\delta^2}{2}e^{-2/\epsilon}\right)t\right) \leqslant \frac{1}{2n^2}, \quad \text{and} \quad \mathbf{Pr}\left(\sum_{i=1}^{t}|Y_i| \geqslant \frac{2t}{\epsilon\ell^\epsilon}\right) \leqslant \frac{1}{2n^2}.$$

As a result, with probability at least $1-1/n^2$, we have:

$$\sum_{i=1}^{t}|X_i| + 2\sum_{i=1}^{t}|Y_i| - t \leqslant t\left(\frac{4}{\epsilon\ell^\epsilon} - \frac{\delta^2}{2}e^{-2/\epsilon}\right).$$

Hence, we can choose $\ell = \mathcal{O}(1)$ (depending only on ϵ and δ) large enough, so that with probability at least $1-1/n^2$, $\sum_{i=1}^{t}|W_i| < t$, so that $|Q_t| = 0$ for $t = \mathcal{O}(\log n)$. ◀

4.3 Proof of Lemma 9

Lemma 9 follows from Claim 11 and Lemma 15 below.

▶ **Lemma 15.** *Assume the hypotheses of Lemma 9. For sufficiently large ℓ depending only on p and α, consider the ℓ-graph ${G_p}^{(\ell)}$ and the super-node $s \in V^{(\ell)}$ such that $s \in I_s$. Then, with probability at least*

$$1 - \tfrac{8}{(\alpha-2)} \cdot \ell^{-(\alpha-2)/2},$$

all nodes in $\Gamma_{{G_p}^{(\ell)}}(s)$ are within ring-distance ℓ from s in the ring $(V^{(\ell)}, {E_1}^{(\ell)})$.

Proof. Without loss of generality, we assume s $= 0$ and let $V^{(\ell)} = \{0,\ldots,n/\ell\}$. Next, we consider nodes on the ring at increasing distance from s $= 0$ moving counter-clockwise, proving that nodes at distance exceeding $\ell^{(2-\alpha)/2}$ are not part of $\Gamma_{{G_p}^{(\ell)}}(\mathrm{s})$, the connected component of s. To this purpose, denote by K be the random variable indicating the ring-distance of the super-node v closest to s, such that v has no incident super-edges in ${G_p}^{(\ell)}$. From (3), and setting $\epsilon = 2-\alpha$ and $\delta = 1-p$, we have

$$\mathbf{Pr}\left(\text{v has no super-edges in } {G_p}^{(\ell)}\right) \geqslant \delta^2 e^{-2/\epsilon},$$

So, for $k = \ell^{\epsilon/2}$ and for ℓ large enough

$$\mathbf{Pr}(K > k) \leqslant (1-\delta^2 e^{-2/\epsilon})^k \leqslant \frac{1}{\ell^{\epsilon/2}}.$$

Moreover, denote by B_k be the event $B_k = \{\text{the } k \text{ nodes nearest to s have no super-bridges}\}$. From (4), from the independence of the edge percolation events and using a union bound, we have for the complementary event B_k^C that $\mathbf{Pr}(B_k^C) \leqslant k \cdot 2(\epsilon\ell^\epsilon)^{-1}$. Iterating the same argument for nodes on the clockwise side of s, if $k = \ell^{\epsilon/2}$ we have

$$\begin{aligned}&\mathbf{Pr}\left(\text{there is a node at distance} \leqslant k \text{ from s in } \Gamma_p^{(\ell)}(\mathrm{s})\right)\\ &\leqslant 2\mathbf{Pr}\left(\{K>k\} \cup B_k^C\right) \leqslant \frac{4}{\ell^{\epsilon/2}} + \frac{4}{\epsilon\ell^{\epsilon/2}},\end{aligned}$$

which completes the proof. ◀

5 The case $1 < \alpha < 2$

The aim of this section is to prove Theorem 3.

When $1 < \alpha < 2$, we make an analysis, based on a suitable inductive interval partitioning of the initial ring. In this case, our technique is inspired by the partitioning method used in [20] and [4]. However, though the high-level idea is similar, our analysis needs to address technical challenges that require major adjustments. In particular, the analysis of [20] only addresses and relies upon properties of infinite one-dimensional lattices with additional long-range links, applying Kolmogorov 0-1 law to the tail event of percolation. On the other hand, [4] borrows from and partly extends [20] to the finite case, proving a result similar to ours for finite one-dimensional lattices, but with important differences both in the underlying percolation model and in the nature of the results they obtain. As for the first point, in their model, ring-edges are deterministically present, while only long-range edges (what we call bridges in this paper) are affected by percolation. In contrast, in our model, *all* edges are percolated to obtain a realization of G_p. Consequently, the graph is deterministically connected in [4], while connectivity is something that only occurs with some probability in our setting. In particular, we face a subtler challenge, since we need to show the existence of a connected component that both spans a constant fraction of the node *and* has polylogarithmic diameter[5]. Moreover, their result for the diameter of the percolated graph only holds with probability tending to 1 in the limit, as the number of nodes grows to infinity, while we are able to show high probability.

We next provide a technical overview of our approach, highlighting the main points of the proof in which our percolation model requires a major departure from the approach of [4].

Similarly to [4], we consider a diverging sequence $\{N_k\}_k$ and for each k we consider in G_p an interval of length N_k consisting of adjacent nodes on the ring. Departing from [4], we then prove that each interval of size N_k contains a constant fraction $(1-\varepsilon_k)N_k$ of nodes, that induces a connected component of diameter D_k, with probability $1-\delta_k$, considering only edges *internal* to the interval and where ε_k and δ_k are suitable constants that only depend on k, p and α. To prove this fact, for each k, we consider an arbitrary interval I_k of size N_k and we proceed inductively as follows:

1. We divide I_k in N_k/N_{k-1} smaller intervals of size N_{k-1};
2. We assume inductively that, with probability $1-\delta_{k-1}$, each of these intervals contains $(1-\varepsilon_{k-1})N_{k-1}$ nodes that induces a connected component of diameter D_{k-1};
3. With concentration arguments, we prove that a constant fraction of the intervals in which we divide I_k have the above property, with probability that increases with k (we call these intervals *good*);
4. We consider all good intervals of the previous point, and we prove that all these intervals are connected to each other with probability increasing in k.

The above reasoning therefore implies, for some ε_k, the existence of a fraction $(1-\varepsilon_k)N_k$ of nodes in I_k, that with increasing probability in k, induces a connected component of diameter $D_k = 2D_k + 1$. More details about the sequences N_k, ε_k and δ_k can be found in the full proof. As mentioned earlier, our proof needs to specifically address percolation of both ring-edges and bridges. This in particular, means tackling points (2) - (4) above, which is a major challenge not present in [4] and is taken care of in the technical Lemma 18 and in

[5] As an example, we might have a connected component including a constant fraction of the nodes and linear diameter, containing a smaller connected component, still spanning a constant fraction of the nodes, but of polylogarithmic diameter.

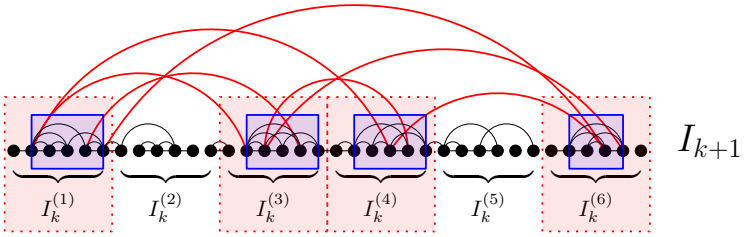


Figure 2 Visualization of the proof of Lemma 16. Red dotted boxes identify "good" intervals in H_{k+1} (intervals for which the event A_k holds); for each good box, the included blue box contains the subset of nodes that induce a connected subgraph of size at least $(1-\varepsilon_k)N_k$ nodes and diameter D_k. In the lemma, we have to prove the existence of at least a certain number of red boxes, and that blue boxes are all mutually connected (i.e. the existence of the red edges in the picture).

the proof of the main Lemma 16 itself. In particular, considering only edges internal to the interval is very important, since, in this way, the events denoting the connection of disjoint intervals are independent. The recurrence of the diameter derives from the fact that the path connecting two nodes u and v in I_k consists of the following three sub-paths:

1. The first part has length at most D_k, and it is a path in the aforementioned "good" interval of size $(1-\varepsilon_{k-1})N_{k-1}$ containing u;
2. The second part consists of a single edge, connecting the good interval containing u with the good interval of v;
3. The third has length at most D_k, and is a path of the good interval of size $(1-\varepsilon_{k-1})N_{k-1}$ containing v.

Finally (Lemma 16), we consider k such that $N_k = \Omega(n)$. For such k, we prove that

$$\delta_k = \mathcal{O}(1/n^{\varepsilon}), \quad \varepsilon_k \ll 1, \quad D_k = \text{polylog}(n),$$

so this implies the existence in G_p of $\Omega(n)$ nodes inducing a connected component with diameter $\text{polylog}(n)$, w.h.p.

We now proceed with the formal analysis which requires the two lemmas below.

▶ **Lemma 16.** *Under the hypotheses of Theorem 3, assume $G = (V, E)$ is sampled from $\mathcal{SW}(n, \alpha)$ and let G_p be its percolation graph. Then two constants $\overline{p} < 1$ and $\eta > 1$ exist such that, if $p > \overline{p}$, w.h.p. there is a set of $\Omega(n)$ nodes in G_p that induces a connected sub-component with diameter $\mathcal{O}(\log^{\eta} n)$.*

▶ **Lemma 17.** *Under the hypotheses of Theorem 3, assume $G = (V, E)$ is sampled from $\mathcal{SW}(n, \alpha)$ and let G_p its percolation graph. Then, a constant $\underline{p} > 0$ exists (in particular, $\underline{p} = 1/3$) such that, if $p < \underline{p}$, w.h.p. each connected component of G_p has size at most $\mathcal{O}(\log n)$.*

The proof of Lemma 16 is given in Section 5.1, while the proof of Lemma 17 follows an approach, based on Galton-Watson processes, similar to that of Lemma 8, and it is given in Section 5.2.

5.1 Proof of Lemma 16

The lemma that follows is our new key ingredient to apply an inductive approach similar to the one in [4] in the case of full-bond percolation.

▶ **Lemma 18.** *Assume the hypotheses of Lemma 16, let* $\beta = \frac{1}{2}\alpha(3-\alpha)$ *and*

$$N_k = e^{\beta^k} \quad \text{and} \quad C_k = e^{\beta^{k-1}(\beta-1)}.$$

Let I_k *be an arbitrary interval of* N_k *adjacent nodes in the cycle* (V, E_1)*,* D_k *any finite integer. Define the event* $A_k = \{\exists S \subseteq I_k : |S| \geqslant (1-\varepsilon_k)N_k \ \wedge diam_p(S) \leqslant D_k\}$*. Assume further that, for suitable, real constants* ε_k, δ_k *and* p_k *in* $(0,1)$*,*

$$\mathbf{Pr}\,(A_k) \geqslant 1-\delta_k, \quad \text{if } p \geqslant p_k.$$

Then, if we consider an interval I_{k+1} *of* N_{k+1} *adjacent nodes, it holds:*

$$\mathbf{Pr}\,(A_{k+1}) \geqslant 1-\delta_{k+1}, \quad \text{if } p \geqslant p_{k+1},$$

where

$$\delta_{k+1} = 2C_{k+1}^{-0.2}, \quad \varepsilon_{k+1} = \varepsilon_k + \delta_k + C_{k+1}^{-0.2},$$
$$D_{k+1} = 2D_k + 1, \quad p_{k+1} = \frac{0.9(\alpha-1)(2-\alpha)}{(1-\varepsilon_k)^2(4.2-2\alpha)}C(n,\alpha).$$

Proof. In words, A_k is the event that there exists a subset of the nodes in I_k that induces a connected component of G_p of size $\geqslant (1-\varepsilon_k)N_k$ nodes and diameter $\leqslant D_k$. Consider the interval I_{k+1} and divide it into C_{k+1} disjoint intervals of size N_k (note that $N_{k+1} = N_k \cdot C_{k+1}$). Denote by H_{k+1} the set of sub-intervals of I_{k+1} of size N_k for which event A_k holds. In particular, for every $i = 1, 2, \ldots, C_{k+1}$, we use the indicator variable $A_k^{(i)}$ to specify whether or not A_k holds for the i-th sub-interval, so that

$$|H_{k+1}| = \sum_{i=1}^{C_{k+1}} A_k^{(i)}.$$

where independence of the $A_k^{(i)}$'s follows from disjointness of the sub-intervals of I_{k+1}. For the sake of brevity, let B_{k+1} denote the event $\{|H_{k+1}| \geqslant (1-\delta_k)C_{k+1} - C_{k+1}^{0.8}\}$. Application of Chernoff's bound then implies

$$\mathbf{Pr}\,(B_{k+1}) \geqslant 1 - e^{-2C_{k+1}^{0.6}} \geqslant 1 - C_{k+1}^{-0.2}.$$

Note that by definition, each sub-intervals in H_{k+1} contains at least one subset (of the nodes) that induces a connected component in G_p with at least $(1-\varepsilon_k)N_k$ nodes and diameter at most D_k. Our goal is to show that, with some probability, these intervals are all connected to each other: this in turn implies the existence of a set of nodes that induces a larger connected component, with diameter at most $2D_k+1$ and containing at least $|H_{k+1}|(1-\varepsilon_k)N_k$ nodes, see see Figure 2 for a visual intuition of the proof. In the remainder of this proof, we denote by F_{k+1} the event that all connected components associated to intervals in H_{k+1} are mutually connected. In particular, we prove that $\mathbf{Pr}\,(F_{k+1} \mid B_{k+1}) \geqslant 1 - C_{k+1}^{-0.2}$, so that

$$\mathbf{Pr}\,(A_{k+1}) = \mathbf{Pr}\,(F_{k+1} \cap B_{k+1}) = \mathbf{Pr}\,(B_{k+1})\,\mathbf{Pr}\,(F_{k+1} \mid B_{k+1}) \geqslant 1 - 2C_{k+1}^{-0.2},$$

which implies that A_{k+1} holds with probability at least $1-\delta_{k+1}$, whenever we set

$$\delta_{k+1} = 2C_{k+1}^{-0.2} \quad \varepsilon_{k+1} = \varepsilon_k + \delta_k + C_{k+1}^{-0.2} \quad \text{and} \quad D_{k+1} = 2D_k + 1.$$

Now, we estimate the probability that, given B_{k+1}, F_{k+1} holds. Two nodes in I_{k+1} have distance at most N_{k+1} on the cycle. Moreover, if we consider two sub-intervals of I_{k+1} both belonging to H_{k+1}, the corresponding connected components contain at least $(1-\varepsilon_k)N_k$

nodes each, accounting for at least $(1-\varepsilon_k)^2 N_k^2$ pairs $\{u,v\}$, with u belonging to the first and v to second connected component. So, two given sub-intervals in H_{k+1}, they are not connected with probability at most

$$\left(1-\frac{p}{cN_{k+1}^{\alpha}}\right)^{(1-\varepsilon_k)^2 N_k^2} \leqslant \exp\left(-\frac{p}{c}(1-\varepsilon_k)^2\frac{e^{2\beta^k}}{e^{\alpha\beta^{k+1}}}\right) = \exp\left(-\frac{p}{c}(1-\varepsilon_k)^2 e^{\beta^k(2-\alpha\beta)}\right),$$

where in the remainder of this proof, we write c for $C(\alpha,n)$, for the sake of readability. If we consider all pairs of intervals in H_{k+1}, a simple union bound allows us to conclude that the probability that the intervals in H_{k+1} are not all mutually connected is at most

$$C_{k+1}^2 \exp\left(-\frac{p}{c}(1-\varepsilon_k)^2 e^{\beta^k(2-\alpha\beta)}\right) \leqslant \exp\left(\beta^k\left(2(\beta-1)-\frac{p}{c}(1-\varepsilon_k)^2(2(2-\alpha)+0.2)\right)\right),$$

where the last inequality follows from the definition of β. Finally, the quantity above can be upper bounded as follows

$$\exp\left(-\beta^k 0.1(\alpha-1)(2-\alpha)\right) = \exp(-0.2\beta^k(\beta-1)) = C_{k+1}^{-0.2},$$

whenever p satisfies $p \geqslant \frac{c(2-\alpha)(\alpha-1)0.9}{(1-\varepsilon_k)^2(2(2-\alpha)+0.2)}$. ◀

Now we are ready to prove Lemma 16.

Proof of Lemma 16. Let C_k and N_k defined as in the claim of Lemma 18. First, we consider the series $\sum_k C_k^{-0.2}$ and, since $C_k = e^{\beta^{k-1}(\beta-1)}$ with $\beta > 1$, we notice that the series is convergent, i.e. $\sum_{k=1}^{\infty} C_k^{-0.2} < +\infty$.[6] This means that the tail of the series converges to zero, and this implies that there exists a constant h such that

$$\sum_{k=h}^{+\infty} C_k^{-0.2} \leqslant \frac{1}{100}. \tag{6}$$

The constant h depends only on α. In particular, it increases as $\alpha \to 2^-$ and $\alpha \to 1^+$.

Now we consider I_h, an arbitrary interval of size $N_h = e^{\beta^h}$. Next, we let

$$\delta_h = 1 - p^{e^{\beta^h}}, \quad \varepsilon_h = 0, \quad D_h = e^{\beta^h}, \quad p_h = p,$$

and we consider the event A_h, defined as in the statement of Lemma 18. If no ring edge belonging to I_h is percolated, A_h is trivially true: this implies that, for every $p \geqslant p_h$, $\mathbf{Pr}(A_h) \geqslant 1-\delta_h$, where probability is over the edges with endpoints in I_h. If A_{h+1} and N_{h+1} are defined like in its statement, Lemma 18 then implies that, for an arbitrary interval I_{h+1} of size N_{h+1}, we have

$$\mathbf{Pr}(A_{h+1}) \geqslant 1-\delta_{h+1},$$

if $p \geqslant p_{h+1}$ and whenever we take

$$\delta_{h+1} = 2C_{h+1}^{-0.2}, \quad \varepsilon_{h+1} = 1 - p^{e^{\beta^h}} + C_{h+1}^{-0.2},$$
$$D_{h+1} = 2e^{\beta^h} + 1, \quad p_{h+1} = \frac{0.9c(\alpha-1)(2-\alpha)}{(4.2-2\alpha)},$$

where the probability is taken over the randomness of the edges with endpoints in I_{h+1}.[7]

[6] We did not try to optimize constants and the choice 0.2 for the exponent is not necessarily optimal.
[7] Recall that $c = C(\alpha,n)$ in the remainder of this proof.

If we iteratively apply Lemma 18, we thus have for each $k \geqslant 1$ such that $N_k \leqslant n$,

$$\mathbf{Pr}(A_k) \geqslant 1 - \delta_k,$$

where $\delta_k = 2C_k^{-0.2}$, $p_k = \frac{0.9c(\alpha-1)(2-\alpha)}{(1-\varepsilon_{k-1})^2(4.2-2\alpha)}$, and ε_k and D_k are defined by the recurrencies:

$$\begin{cases} \varepsilon_k = \varepsilon_{k-1} + \delta_{k-1} + C_k^{-0.2}, \text{ if } k > h \\ \varepsilon_h = 1 - p^{e^{\beta^h}} \end{cases} \quad \text{and} \quad \begin{cases} D_k = 2D_{k-1} + 1, \text{ if } k > h \\ D_h = e^{\beta^h}. \end{cases}$$

Now, we solve the recurrence for ε_k, obtaining $\varepsilon_k = 1 - p^{e^{\beta^h}} + 3\sum_{i=h}^{k} C_i^{-0.2}$, where, leveraging (6), we take $p > \sigma$, with $\sigma > 0$ such that $1 - \sigma^{e^{\beta^h}} = 1/50$. With this choice, for each $k \geqslant 1$ we obtain

$$\varepsilon_k \leqslant 1 - p^{e^{\beta^h}} + \frac{3}{100} \leqslant \frac{1}{20}.$$

Moreover, for each $i \leqslant k$, we have that $D_k \leqslant 2^{k-h+1} e^{\beta^h}$. If we take $m = \log_\beta(\log n)$, then $N_m = e^{\beta^m} = n$ and $C_m = n^{(\beta-1)/\beta}$. We also have:

$$D_m \leqslant \frac{e^{\beta^h}}{2^{h-1}} 2^{\log_\beta(\log(n))} = \frac{e^{\beta^h}}{2^{h-1}} (\log n)^{\log_\beta 2}.$$

Setting $\eta = \log_\beta 2$, we have $\eta > 1$, since $\beta < 2$. We thus have $D_m = O((\log^\eta n))$. Moreover, if $p \geqslant \sigma$

$$\delta_m \leqslant 2n^{-0.2(\beta-1)/\beta}, \ \varepsilon_m \leqslant \frac{1}{20}, \ p_m \geqslant \frac{0.9c(\alpha-1)(2-\alpha)}{(19/20)^2(4.2-2\alpha)}.$$

Finally, if

$$p \geqslant \max\left\{\sigma, \frac{0.9c(\alpha-1)(2-\alpha)}{(19/20)^2(4.2-2\alpha)}\right\} := \overline{p}$$

then,

$$\mathbf{Pr}(A) \geqslant 1 - \delta_m \geqslant 1 - 2n^{-0.2(\beta-1)/\beta},$$

where

$$A = \left\{\exists S \subseteq V : |S| > \frac{19}{20} n \wedge \operatorname{diam}_p(S) = O(\log^\eta(n))\right\},$$

i.e., w.h.p. G_p contains an induced subgraph of size at least $(19/20)n$ nodes and diameter $O(\log^\eta(n))$. ◀

5.2 Proof of Lemma 17

Let $\underline{p} = 1/3$ and consider an arbitrary node $s \in V$. We consider an execution of the BFS in Algorithm 2 with input the percolation subgraph $G_p = (V, E_p)$ of $G = (V, E)$ and the source s.

Algorithm 2 BFS visit of G_p.

1: **Input**: the subgraph $G_p = (V, E_p)$, a source $s \in V$
2: $Q = \{s\}$
3: $R = \emptyset$
4: **while** $Q \neq \emptyset$ **do**
5: $\quad w = \texttt{dequeue}(Q)$
6: $\quad R = R \cup \{w\}$
7: $\quad$ **for** each neighbor x of w in G_p such that $x \notin R$ **do**
8: $\quad\quad \texttt{enqueue}(x, Q)$

We consider the generic t-th iteration of the while loop at line 4 and we denote by W_t the number of nodes added to the queue Q at line 8. Moreover, B_t is the set of nodes that are in R in the t-th iteration. By its definition, B_t is a branching process described by the following recursion:

$$\begin{cases} B_t = B_{t-1} + W_t - 1 & \text{if } B_{t-1} \neq 0 \\ B_t = 0 & \text{if } B_{t-1} = 0 \\ B_0 = 1. \end{cases} \tag{7}$$

Note that, from Definition 1, each node $w \in V$ has expected degree $\mathbf{E}[\deg(w)] = 3$. So, since each edge in G is also in G_p with probability p, we have $\mathbf{E}[W_1] = 3p$ and $\mathbf{E}[W_t] \leqslant 3p$ for $t > 1$.[8] Since $p < \underline{p}$, there is a constant δ such that $p = (1-\delta)/3$ and, for each $t \geqslant 0$

$$\mathbf{E}[W_t] = 1 - \delta.$$

We consider the T-th iteration of the while loop, where $T = \gamma \log n$. Note that the random variables $W_1, \ldots, W_T$ are not independent as noted earlier but, as remarked in the proof of Lemma 14, it is easy to show that they are stochastically dominated by T independent random variables distributed as W_1. For the sake of simplicity, we abuse notation, by using $W_1, \ldots, W_T$ to denote the T independent copies of W_1 in the remainder of this proof. Now, if

$$\sum_{i=1}^{T} W_i - T < 0,$$

then $B_T = 0$. We notice that each W_i can be written as a sum of $n+2$ independent Bernoulli random variables.[9] Hence, applying Chernoff's bound to $W = \sum_{i=1}^{T} W_i$ we obtain:

$$\mathbf{Pr}\left(W > (1+\delta)\mathbf{E}[W]\right) \leqslant e^{-\frac{\delta^2}{2}\mathbf{E}[W]}.$$

Next, since $\mathbf{E}[W] < (1-\delta)T$,

$$\mathbf{Pr}\left(W > (1-\delta^2)T\right) \leqslant e^{-\frac{\delta^2(1-\delta)}{2}T} \leqslant \frac{1}{n^2},$$

where the last inequality follows if we take $\gamma \geqslant 4/(\delta^2(1-\delta))$. This allows us to conclude that

$$\mathbf{Pr}\left(B_T = 0\right) \geqslant 1 - \frac{1}{n^2},$$

[8] We have not strict equality for $t > 1$, which follows since the W_t's are not independent in general, since line 8 is only executed if $x \notin R$.

[9] Assume node v is visited in the t-iteration. Then we have 2 indicator variables for the 2 ring edges incident in v, plus n indicator variables, corresponding to n bridges potentially incident in v.

for some $T = \mathcal{O}(\log n)$, which in turn implies that with the above probability, the connected component of which v is part contains at most $\mathcal{O}(\log n)$ nodes. Finally, a union bound over all nodes in V concludes the proof.

6 The case $\alpha < 1$

In this section, we prove Theorem 4. When $\alpha < 1$, we first notice that every bridge-edges (u, v) is in G_p with probability at least pc/n, for some constant c. This reduces our problem to the analysis of the percolation graph H_p of a graph H, where H is the union of a ring and an Erdős-Rényi graph $\mathcal{G}_{n,q}$. [3] contains a detailed analysis of this case, and we use their results to prove that G_p, for any sufficiently large p, contains a connected component with diameter $\mathcal{O}(\log n)$.

We finally remark that, for every value α, when $p < 1/3$ the connected components of G_p have at most $\mathcal{O}(\log n)$ nodes. This fact easy follows from concentration techniques, and the fact that every node in G_p has expected degree at most 3.

The formal proof of Theorem 4 requires the two lemmas below.

▶ **Lemma 19.** *Under the hypotheses of Theorem 4, sample a $G = (V, E)$ from the $\mathcal{SW}(n, \alpha)$ distribution and let G_p be its percolation subgraph. Then, a constant $\overline{p} < 1$ exists such that, if $p > \overline{p}$ then, w.h.p., there is a set of $\Omega(n)$ nodes in G_p that induce a connected subgraph of diameter $\mathcal{O}(\log n)$.*

▶ **Lemma 20.** *Under the hypothesis of Theorem 4, sample a $G = (V, E)$ from the $\mathcal{SW}(n, \alpha)$ distribution and let G_p be the percolation subgraph of G. Then, a constant $\underline{p} > 0$ exists such that, if $p < \underline{p}$ then, w.h.p., each connected component of G_p has size at most $\mathcal{O}(\log n)$.*

A key observation to prove Lemma 19 is that, when $\alpha < 1$, the percolation graph of a graph sampled from $\mathcal{SW}(n, \alpha)$ stochastically dominates a cycle with additional Erdős-Rényi random edges. To prove Lemma 19 we thus use a previous result in [3] for this class of random graphs. In particular, we first give an equivalent formulation of Lemma 5.1 in [3], stating that, with constant probability the sequential BFS visit (Algorithm 2) reaches $\Omega(\log n)$ nodes within $\mathcal{O}(\log n)$ rounds. We then consider a *parallel* BFS visit (see Algorithm 4 in Section 6.1) and show an equivalent formulation of Lemma 5.2 in [3], stating that, w.h.p., the parallel BFS visit starting with $\Omega(\log n)$ nodes reaches a constant fraction of nodes within $\mathcal{O}(\log n)$ rounds.

The full proof of Lemma 19 is given in Section 6.1, while the proof of Lemma 20 is omitted, since it proceeds along the very same lines as the proof of Lemma 17, which does not depends on the value of α.

6.1 Proof of Lemma 19

We first prove the following fact.

▷ Claim 21. Under the hypotheses of Lemma 19, a constant $c \in (0, 1)$ (depending only on α) exists such that, for any $u, v \in V$,

$$\mathbf{Pr}\left((u, v) \in E_p\right) \geqslant \frac{pc}{n}.$$

Proof. The normalizing constant $C(\alpha, n)$ (see Definition 1) can be upper bounded as follows

$$C(\alpha, n) = 2\sum_{x=2}^{n/2} \frac{1}{x^\alpha} \leqslant \frac{2}{2^\alpha} + \int_2^{n/2} \frac{2}{x^\alpha} dx \leqslant 2^{1-\alpha} + \frac{2}{1-\alpha}\left(\frac{n}{2}\right)^{1-\alpha}.$$

Hence, a constant $c \in (0,1)$ (depending only on α) exists such that

$$\begin{aligned}\mathbf{Pr}\left((u,v) \in E_p\right) &= \frac{p}{C(\alpha,n)d(u,v)^\alpha} \geqslant \frac{p}{C(\alpha,n)(n/2)^\alpha} \\ &\geqslant \frac{p}{2^{1-\alpha}(n/2)^\alpha + n/(1-\alpha)} \geqslant \frac{p \cdot c}{n}.\end{aligned}$$

◁

The above fact proves that, when $\alpha < 1$, the percolation subgraph G_p of a graph sampled from $\mathcal{SW}(n,\alpha)$ stochastically dominates a cycle with additional Erdős-Rényi random edges. To prove Lemma 19 we thus use a previous result in [3] on such class of random graphs. In particular, we first give an equivalent formulation of Lemma 5.1 in [3], that states that, with constant probability the sequential BFS visit (Algorithm 3) reaches $\Omega(\log n)$ nodes within $\mathcal{O}(\log n)$ rounds. We then consider a *parallel*-BFS visit (Algorithm 4) and give an equivalent formulation of Lemma 5.2 in [3], that states that, w.h.p., the parallel-BFS visit starting with $\Omega(\log n)$ nodes reaches a constant fraction of nodes within $\mathcal{O}(\log n)$ rounds.

We also introduce a slightly different version of the BFS visit of Algorithm 4, where we have also a set R_0 of removed nodes in input.

Algorithm 3 BFS visit of G_p.

Input: the subgraph $G_p = (V, E_p)$, an initiator $s \subseteq V$, a set of removed nodes $R_0 \subseteq V$.
$Q = \{s\}$
$R = R_0$
while $Q \neq \emptyset$ **do**
    $w = \texttt{dequeue}(Q)$
    $R = R \cup \{w\}$
    for each neighbor x of w in G_p such that $x \notin R$ **do**
        $\texttt{enqueue}(x, Q)$

▶ **Lemma 22.** *Under the hypothesis of Lemma 19, let $v \in V$ be a node and $c \in (0,1)$ be a constant as in Claim 21. For every $\beta > 0$, $\varepsilon > 0$, and percolation probability $p > \frac{\sqrt{c^2+6c+1}-c-1}{2c} + \varepsilon$, there are positive parameters k and γ (depending only on ε, c and p) such that the following holds: the BFS visit (Algorithm 3) with input G_p, v, and a set R_0 with $|R_0| \leqslant \log^4 n$, with probability γ, a time $\tau_1 = \mathcal{O}(\log n)$ exists such that*

$$|(R \setminus R_0) \cup Q| \geqslant n/k \qquad OR \qquad |Q| \geqslant \beta \log n\,.$$

Algorithm 4 Parallel BFS visit of G_p.

Input: the subgraph $G_p = (V, E_p)$, a set of initiators $I_0 \subseteq V$, a set of removed nodes $R_0 \subseteq V$
$Q = I_0$
$R = R_0$
while $Q \neq \emptyset$ **do**
    $A = R \cup Q$
    $X = \text{neighbors}(Q)$
    $Q' = Q$
    $Q = \emptyset$
    while $Q' \neq \emptyset$ **do**
        $w = \texttt{dequeue}(Q')$
        $R = R \cup \{w\}$
        for each $x \in X$ **do**
            $\texttt{enqueue}(x, Q)$

▶ **Lemma 23.** *Under the hypothesis of Lemma 19, let $c \in (0,1)$ be a constant as in Claim 21. For every $\varepsilon > 0$ and percolation probability $p > \frac{\sqrt{c^2+6c+1}-c-1}{2c} + \varepsilon$, there are positive parameters k, β (depending only on c, p and ε) such that the following holds: For any pair of sets $I_0, R_0 \subseteq V$, with $|I_0| \geqslant \beta \log n$ and $|R_0| \leqslant \log^4 n$, in the parallel BFS-visit (Algorithm 4) with input G_p, I_0, and R_0, with probability at least $1 - 1/n$, a time $\tau_2 = \mathcal{O}(\log n)$ exists such that*

$$|(R \setminus R_0) \cup Q| \geqslant n/k.$$

Now we are ready to prove Lemma 19.

Proof of Lemma 19. Let c be as in Claim 21 and $\overline{p} = \frac{\sqrt{c^2+6c+1}-c-1}{2c}$. Let $p > \overline{p} + \varepsilon$, for an arbitrarily-small constant $\varepsilon > 0$, and let $\beta > 0$ be the constant in Lemma 23 and k be the constant in Lemmas 22 and 23. We consider the following process, where we initialize $R_0 = \emptyset$ and $\tau_1 = \mathcal{O}(\log n)$ is as in Lemma 22.

1. Consider a node $v \in V \setminus R_0$.
2. From v, perform a sequential-BFS visit (Algorithm 2), with input G_p, v, and R_0, for τ_1 while loops and add to R_0 the sets Q and R as they are at the end of the τ_1-th iteration of the while loop ($R_0 = R_0 \cup R \cup Q$).
3. If $|Q| \geqslant \beta \log n$ or $|Q \cup R| \geqslant n/k$, interrupt the process.
4. Restart from 1.

Let $\gamma > 0$ be the constant in Lemma 22. We prove that the process above terminates within $\sigma = \log_{1-\gamma}(n)$ iterations, w.h.p.

First we notice that, at each iteration of the process, the set R_0 grows, w.h.p., at most of size $\mathcal{O}(\log^2 n)$, since each node in G_p has degree at most $\mathcal{O}(\log n)$, w.h.p. (Claim 6) and so, in τ_1 iteration of the parallel-BFS, at most $\mathcal{O}(\log^2 n)$ nodes will be reached by v. This implies that, at each iteration $i \leqslant \sigma$ of the process $|R_0| = \mathcal{O}(\log^3 n)$, w.h.p.

From Lemma 22 it follows that,a sequential-BFS with in input G_p, v and R_0 with $|R_0| = \mathcal{O}(\log^3 n)$ is such that, at the end of τ_1-th iteration

$$\mathbf{Pr}\left(|Q| \geqslant \beta \log n \text{ or } |R \cup Q| \geqslant n/k\right) \geqslant \gamma > 0\,.$$

Therefore, the probability that the process exceeds σ iterations is at most $(1-\gamma)^\sigma \leqslant 1/n$.

So, w.h.p., a node v exists such that the sequential-BFS starting from v, after $\mathcal{O}(\log n)$ steps, satisfies at least one of the two conditions: i) $|Q| \geqslant \beta \log n$ or ii) $|Q \cup R| \geqslant n/k$.

If ii) holds, the lemma is proven. Indeed, w.h.p. we have the existence of a node v such that there is a set of $\Omega(n)$ nodes at distance at most $\mathcal{O}(\log n)$ from v.

If i) holds, it suffices to perform a sequential-BFS (Algorithm 2) with in input G_p, $I_0 = Q$ and R_0 and apply Lemma 23 to claim that such BFS reaches at least $\Omega(n)$ nodes in $\mathcal{O}(\log n)$ steps. ◀

References

1 Noga Alon, Benjamini Itai, and Stacey Alan. Percolation on finite graphs and isoperimetric inequalities. *Annals of Probability*, 32:1727–1745, 2004.

2 Luca Becchetti, Andrea Clementi, Francesco Pasquale, Luca Trevisan, and Isabella Ziccardi. Bond percolation in small-world graphs with power-law distribution. *arXiv preprint*, 2022. arXiv:2205.08774.

3 Luca Becchetti, Andrea E. F. Clementi, Riccardo Denni, Francesco Pasquale, Luca Trevisan, and Isabella Ziccardi. Percolation and epidemic processes in one-dimensional small-world networks - (extended abstract). In *Proceedings of LATIN 2022: the 15th Latin American Symposium*, volume 13568 of *Lecture Notes in Computer Science*, pages 476–492. Springer, 2022. doi:10.1007/978-3-031-20624-5_29.

4 Itai Benjamini and Noam Berger. The diameter of long-range percolation clusters on finite cycles. *Random Struct. Algorithms*, 19(2):102–111, 2001. doi:10.1002/rsa.1022.

5 Marek Biskup. On the scaling of the chemical distance in long-range percolation models. *The Annals of Probability*, 32(4):2938–2977, 2004. doi:10.1214/009117904000000577.

6 Marek Biskup. Graph diameter in long-range percolation. *Random Structures & Algorithms*, 39(2):210–227, 2011. doi:10.1002/rsa.20349.

7 Duncan S. Callaway, John E. Hopcroft, Jon Kleinberg, Mark E. J. Newman, and Steven H. Strogatz. Are randomly grown graphs really random? *Phys. Rev. E*, 64:041902, September 2001. doi:10.1103/PhysRevE.64.041902.

8 Wei Chen, Laks Lakshmanan, and Carlos Castillo. Information and influence propagation in social networks. *Synthesis Lectures on Data Management*, 5:1–177, October 2013. doi:10.2200/S00527ED1V01Y201308DTM037.

9 Hyeongrak Choi, Mihir Pant, Saikat Guha, and Dirk Englund. Percolation-based architecture for cluster state creation using photon-mediated entanglement between atomic memories. *npj Quantum Information*, 5(1):104, 2019.

10 Easley David and Kleinberg Jon. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, USA, 2010.

11 Michele Garetto, Weibo Gong, and Donald F. Towsley. Modeling malware spreading dynamics. In *Proceedings IEEE INFOCOM 2003, The 22nd Annual Joint Conference of the IEEE Computer and Communications Societies, San Franciso, CA, USA, March 30 - April 3, 2003*, pages 1869–1879. IEEE Computer Society, 2003. doi:10.1109/INFCOM.2003.1209209.

12 Anna R. Karlin, Greg Nelson, and Hisao Tamaki. On the fault tolerance of the butterfly. In *Proceedings of the twenty-sixth annual ACM symposium on Theory of Computing*, pages 125–133, 1994.

13 David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. *Theory of Computing*, 11(4):105–147, 2015. doi:10.4086/toc.2015.v011a004.

14 Harry Kesten. The critical probability of bond percolation on the square lattice equals 1/2. *Communications in mathematical physics*, 74(1):41–59, 1980.

15 Jon Kleinberg. Navigation in a small world. *Nature*, 406:845, 2000.

16 Alexander Kott and Igor Linkov. *Cyber resilience of systems and networks*. Springer, 2019.

17 Rémi Lemonnier, Kevin Seaman, and Nicolas Vayatis. Tight bounds for influence in diffusion networks and application to bond percolation and epidemiology. In *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'14, pages 846–854, Cambridge, MA, USA, 2014. MIT Press.

18 Cristopher Moore and Mark E. J. Newman. Epidemics and percolation in small-world networks. *Physical Review E*, 61(5):5678, 2000.

19 Cristopher Moore and Mark E. J. Newman. Exact solution of site and bond percolation on small-world networks. *Phys. Rev. E*, 62:7059–7064, November 2000. doi:10.1103/PhysRevE.62.7059.

20 Charles M. Newman and Lawrence S. Schulman. One dimensional $1/|j-i|^s$ percolation models: The existence of a transition for $s \leqq 2$. *Communications in Mathematical Physics*, 104(4):547–571, 1986.

21 Mark E. J. Newman and Duncan J. Watts. Scaling and percolation in the small-world network model. *Physical review E*, 60(6):7332, 1999.

22 Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, and Alessandro Vespignani. Epidemic processes in complex networks. *Rev. Mod. Phys.*, 87:925–979, August 2015. doi:10.1103/RevModPhys.87.925.

23 Vinod K.S. Shante and Scott Kirkpatrick. An introduction to percolation theory. *Advances in Physics*, 20(85):325–357, 1971. doi:10.1080/00018737100101261.

24 Wei Wang, Ming Tang, H Eugene Stanley, and Lidia A Braunstein. Unification of theoretical approaches for epidemic spreading on complex networks. *Reports on Progress in Physics*, 80(3):036603, February 2017. doi:10.1088/1361-6633/aa5398.

25 Duncan J. Watts and Steven H. Strogatz. Collective dynamics of "small-world" networks. *Nature*, 393(6684):440–442, 1998.

26 Tongfeng Weng, Michael Small, Jie Zhang, and Pan Hui. Lévy walk navigation in complex networks: A distinct relation between optimal transport exponent and network dimension. *Scientific Reports*, 5(1):17309, 2015.