J. Phys. A: Math. Theor. 55 (2022) 224009 (12pp)

https://doi.org/10.1088/1751-8121/ac6716

A branching random-walk model of disease outbreaks and the percolation backbone

Paulo Murilo C de Oliveira^{1,2,*}, Daniel A Stariolo^{1,2,*} and Jeferson J Arenzon^{2,3,*}

¹ Instituto de Física, Universidade Federal Fluminense, Av. Litorânea s/n, 24210-340 Boa Viagem, Niterói, RJ, Brazil

² Instituto Nacional de Ciência e Tecnologia—Sistemas Complexos, Rio de Janeiro RJ, Brazil

³ Instituto de Física, Universidade Federal do Rio Grande do Sul, CP 15051, 91501-970 Porto Alegre RS, Brazil

E-mail: pmco@if.uff.br, danielstariolo@id.uff.br and arenzon@if.ufrgs.br

Received 15 December 2021, revised 9 March 2022 Accepted for publication 13 April 2022 Published 12 May 2022



Abstract

The size and shape of the region affected by an outbreak is relevant to understand the dynamics of a disease and help to organize future actions to mitigate similar events. A simple extension of the SIR model is considered, where agents diffuse on a regular lattice and the disease may be transmitted when an infected and a susceptible agents are nearest neighbors. We study the geometric properties of both the connected cluster of sites visited by infected agents (outbreak cluster) and the set of clusters with sites that have not been visited. By changing the density of agents, our results show that there is a mixed-order (hybrid) transition separating a phase where the outbreak cluster is finite from the other one where it percolates through the system. Moreover, the outbreak cluster seems to have the same exponents of the backbone of the critical cluster of the ordinary percolation while the clusters with unvisited sites have a size distribution with a Fisher exponent $\tau < 2$.

Keywords: branching random-walk, SIR model, backbone percolation

(Some figures may appear in colour only in the online journal)

*Authors to whom any correspondence should be addressed.

1751-8121/22/224009+12\$33.00 © 2022 IOP Publishing Ltd Printed in the UK

1. Introduction

An epidemic outbreak may be controlled once herd immunity develops because most of the population has been either vaccinated or infected by the contagious agent and recovered. Amidst those individuals that no longer can get infected, groups of yet susceptible ones may remain. The sizes of these groups and the distances and connecting routes between them will determine the consequences of a future outbreak. Thus, along with the temporal course of an epidemic, for future prevention and mitigation of similar events it is also important to model and characterize both the spatial extent [1, 2] and the geometric properties of the regions that had been affected. Moreover, the interest in spreading processes goes beyond the propagation of diseases, since epidemiological models consider general mechanisms that have been applied, for example, in studies of how opinions, gossips, or fake news propagate through the complex networks of social contact [3-6].

A cornerstone of many theoretical studies in epidemiology is the SIR model and its many variants (see references [7-11] for recent reviews). Infectious agents (I) may either transmit the disease, with a given rate, to susceptible ones (S) through direct contact or, spontaneously, get removed (R) from the process by dying or recovering. This compartmental model has an absorbing state where the infected individuals get extinct and the disease can no longer propagate. Depending on the parameters, there is a continuous transition between a phase where the outbreak is controlled fast, leaving only a finite number of agents that have been infected, and another phase where most of the agents became infected and, once recovered, occupy a macroscopic fraction of the lattice. The nature of this transition, belonging to the percolation universality class, has been largely studied [12–14]. Several modifications have been introduced in the original SIR model (further compartments, non-permanent immunity, different networks and lattices, disorder, vaccination, etc) and both the asymptotic and dynamical properties of the contagion process in these scenarios have been studied, along with the possibility of changing the universality class of the transition. In some cases, the transition changes from continuous to discontinuous, or even to a hybrid (or mixed order) one, where the discontinuity in the order parameter is accompanied by critical fluctuations [15, 16]. Such hybrid transitions have been observed in models with cooperative spreading where multiple strains (or multiple contagion steps) are involved [17-22].

Analogously to the SIR model, the system we consider in this paper is a stochastic branching process, where the agents are random walkers with volume exclusion [23-25] and the infection proceeds through close encounters between infected agents and susceptible ones. Once the outbreak is over, we study the spatial extent of the infected region, i.e., the asymptotic set of all sites that have been visited by an infected agent before recovery [1, 2]. Depending on the parameters of the model, there is a phase transition connected with the size of this region. The main purpose of this paper is to characterize this transition and determine which is the universality class this model belongs to. Besides studying the cluster of sites visited by infected agents, it is also interesting to study those that were not. As the number of agents increase, facilitating the propagation of the disease, the set of unvisited sites appears fragmented in several independent regions. The geometric properties of these clusters are also interesting, in particular, the area distribution close to the percolation threshold and the characterization of how homogeneous these areas are.

While exploring the properties of this simple model, we unveiled its connection with the backbone of the percolating cluster of ordinary percolation. Despite its simplifying assumptions, it nonetheless provides a good basis for the study of the spatial extension of an epidemic spread. In section 2 we describe the model and the quantities we considered to characterize its

behavior. In section 3 we show and analyze our results, obtained through extensive numerical simulations. The discussion and some conclusions are presented in section 4.

2. Model and observables

We consider the model introduced in reference [26] where $N \leq L^2$ agents are initially distributed at random and without superposition on an $L \times L$ square lattice with periodic boundary conditions. As will be discussed at the end of this section, it is possible [27, 28] to transform the results obtained with a constant, discrete N to a continuous variable p. We choose an initial condition in which only one agent is infected (I) while the other N-1 are susceptible (S). Starting from this initial configuration, at each Monte Carlo step, each agent sequentially chooses one of its nearest neighbor sites at random. If the later is vacant, the agent moves to that site. Otherwise, in case the agent is currently infected and the tossed neighbor is a susceptible, then the later also becomes infected with probability p_{inf} . Each infected agent remains in the I state during t_{rec} time steps (kept constant), after which it becomes recovered (R) and unable to further propagate the disease to other agents. Notice that because of the excluded volume, recovered agents indeed help hindering the transmission of the disease. We here adopt $p_{inf} = 0.5$ and $t_{rec} = 20$ and discuss, in the conclusions, on the generality of the results. Notice that having $t_{rec} > 1$ makes the model non Markovian [29–31]. The above steps are repeated until no infected agent remains in the system, what defines the total time T of the spreading process. The model also considers that the timescale of any demographic process is much longer than the contagion one, and the population is kept fixed. We also consider that displacements and contagion are local processes despite the complex way humans are networked and the possibility of long range travels.

Here we focus on several properties of the many absorbing states attained at the end of the spreading process. Of primary interest is the single cluster, of size M, that includes all sites that have been visited by the infected agents. An example of such a cluster is shown in figure 1 for p slightly above the percolation threshold p_c (to be determined later). The sample averaged fraction of sites belonging to the infected cluster is given by the strength $m = \langle M \rangle / L^2$, which acts as an order parameter for the propagation of the disease. It vanishes in the thermodynamic limit, $L \to \infty$, if the population density is smaller than a critical value p_c , while it is finite for $p > p_c$. Thus, the epidemic threshold is associated with this cluster becoming macroscopic, i.e., occupying a finite fraction of the system when $L \to \infty$. Although we did not attempt to measure the basic reproduction number \mathcal{R}_0 , i.e., the number of secondary infections that a single infected individual can generate before it dies or becomes cured [9, 32, 33], the outbreak cluster must be related to it [34]. Indeed, the presence of a percolating cluster above the threshold is the difference between an infection that dies out and the onset of an epidemic.

The size fluctuations, $\chi = L^{-2}(\langle M^2 \rangle - \langle M \rangle^2)$, and the associated Binder cumulant [35], $U = 1 - \langle M^4 \rangle / 3 \langle M^2 \rangle^2$, are computed as well. The latter has the interesting property of having null anomalous dimension, i.e., in the thermodynamic limit U is a step function with two constant trays, one for $p < p_c$ (whose height depends on the nature of the transition) and another one at U = 2/3 for $p > p_c$. The value that U assumes exactly at p_c is an indication of the order of the transition. For continuous transitions this isolated value is intermediary between both trays. On the other hand, for discontinuous transitions the cumulant at p_c assumes a non trivial value [36–39] below both trays.

Two other step functions with null anomalous dimension are also useful in this context: the wrapping probability W around the torus and the connection probability Y [40–42]. Above p_c , the outbreak cluster percolates in a finite system, i.e., it wraps around the system and touches the opposite boundaries. Thus, W is the fraction of samples that present a wrapping cluster.



Figure 1. Final configuration of the branching process for L = 200 and p = 0.31 (slightly above the percolation threshold) showing the single cluster formed by the sites visited by infected agents (blue). The remaining sites, on the other hand, are distributed among several clusters of varying sizes (white regions) that make the blue cluster fractal at the transition.

Between two parallel lines a distance L/2 apart from each other, Y is 1 if both lines cross the outbreak cluster and 0 otherwise. Since one can scan L different pairs of parallel lines for each configuration (L/2 along each of the torus main directions), this observable can be computed with more precision than W.

Besides the single, connected cluster associated with the outbreak, the complementary set of the unvisited sites (figure 1, white sites) is also of interest. Those domains, as they may become trapped inside the outbreak cluster, may have many different sizes in the same sample. We measure the cluster size distribution close to the percolation threshold. The number of distinct cluster sizes of a given configuration, H, indicates how heterogeneous this distribution is and has been recently subject of study in different systems [43–49]. In the limit $p \ll 1$, the unvisited sites form a large domain and $H \simeq 1$. In the other limit, $p \simeq 1$, a few isolated holes of approximately unitary size remain in the visited cluster and, being mostly of the same size, once again $H \simeq 1$. As p moves to intermediate values, H increases because the structure of the outbreak cluster becomes more complex and a set of interior holes develops, with different sizes. As a consequence of the distribution of cluster sizes being a power-law close to the threshold, $n(s) \sim s^{-\tau}$ (τ is the Fisher exponent), H develops a peak that grows as $H_{peak} \sim L^{2/\tau}$.

Finally, following references [27, 28], if all L^2 values of a given observable X_N have been measured for a constant, discrete N, a transformation to a continuous variable p is obtained by

$$X(p) = \sum_{N=1}^{L^2} {\binom{L^2}{N} p^N (1-p)^{L^2 - N} X_N}.$$
(1)

This procedure is equivalent to the traditional transformation from the microcanonical to the canonical ensemble. Notice that although in principle all the values $N \in [1, L^2]$ should be considered in the above sum, because the coefficients of X_N are highly peaked, it is the neighboring region to the specific value of N when $N/L^2 \sim p$ that contributes the most. Of course, it is also possible to populate the initial system with a probability p of occupying each site, the results



Figure 2. The infected agents visit, on average, $\langle M \rangle$ sites during the outbreak, forming a single cluster that occupies a fraction $m \equiv \langle M \rangle / L^2$ of the system. There is a threshold p_c (indicated by the vertical dashed line), separating a region $(p < p_c)$ where this cluster is finite from another one, $p > p_c$, with a giant, percolating cluster. As the system size increases, the curves become steeper. Inset: difference between the order parameter m and the jump $m_0 \simeq 0.3070$ as a function of 1/L, at the estimated critical point $p_c = 0.3086$. The best fit (thin solid line) gives $\beta / \nu \simeq 0.362$. The thick solid line is L^{2-d_B} , where $d_B \simeq 1.643$ is the backbone fractal dimension (see text).

being consistent. A large number of samples has to be considered, nonetheless, in order to achieve the desired precision. Indeed, for the averages shown here, no less than 10^5 samples have been used.

3. Results

3.1. Outbreak cluster

The set of evidences presented in this section points to a hybrid transition between a phase with a finite, non-percolating cluster of sites visited by the infected agents and another one with a giant, percolating cluster. In a hybrid, or mixed-order transition, the order parameter has a finite jump at p_c . Nonetheless, criticality remains after discounting the size of the jump from $m, m - m_o \sim L^{-\beta/\nu}$, corresponding to a critical cluster of mass

$$M = m_0 L^2 + m_1 L^{d_{\rm F}},\tag{2}$$

where d_F , the cluster fractal dimension, obeys the hyperscaling relation $d_F = d - \beta/\nu$ with the exponent β defined by $m - m_o \sim (p - p_c)^{\beta}$. Thus, the critical cluster has a compact component (first term) along with a fractal part (second term). In standard second order transitions, $m_o \rightarrow 0$ and the compact region is missing. Discontinuous transitions miss the fractal term as $\beta \rightarrow 0$. In figure 2, the order parameter m is shown as a function of p for different system sizes L (the critical point, $p_c \simeq 0.3086$, that will be more precisely determined later, is shown as a vertical dashed line). As L increases, the curves become steeper. The behavior of $m - m_o$ as a function of 1/L at the critical point p_c is shown in the inset along with a power-law fit (thin solid line). From this, we obtain the exponent $\beta/\nu \simeq 0.362$ and the jump, $m_o \simeq 0.3070$ (no sensible change is found in the exponent when using different values of this constant). For reasons that will be discussed later, we also show (thick solid line) that this behavior is consistent with



Figure 3. The connection probability Y(p) for several values of *L*. Top inset: data collapse onto a universal curve *g* using $Y = g[(p - p_c)L^{1/\nu}]$. Bottom inset: corrections to finite size scaling. The black curve shows the fitting using equation (3) for the adopted value $\lambda = 0.7943$, which leads to $a_0 = 0$. The critical point is then $p_c = 0.3086$. Extra sizes were considered, from L = 19 up to 363.

the cluster having the same fractal dimension as the random percolation backbone, i.e., L^{2-d_B} , where $d_F = d_B \simeq 1.643$. The backbone is the subset of the critical cluster without the dangling ends, i.e., those sites that are not relevant for the transport properties through the cluster [50].

The connection probability Y(p) is shown in figure 3 for different system sizes. From these data we obtain the exponent ν since, because of its null anomalous dimension, the vertical scale of Y does not change and the scaling is given by $Y = g[(p - p_c)L^{1/\nu}]$, where g(x) is a universal function. The scaling variable is $(p - p_c)L^{1/\nu}$ and, therefore, $1/\nu$ is the leading exponent for the scaling transformation along the horizontal axis. In the top inset of figure 3 we observe that the collapse is very good with $\nu = 2$ and $p_c \simeq 0.3086$, ruling out the possibility that the transition is in the random percolation universality class. This precise value for p_c will be determined in the sequence using the data for Y(p).

Following the method of references [51, 52], one determines the point p_L where Y crosses a constant value λ , i.e., $Y(p_L) = \lambda$, for each L. Considering corrections to scaling up to the third order, this sequence of crossing points is fitted with the series expansion

$$p_L(\lambda) = p_c + \frac{1}{L^{1/\nu}} \sum_{n=0}^{3} \frac{a_n(\lambda)}{L^n},$$
(3)

with parameters p_c and $a_n = a_n(\lambda)$. Since $\nu > 1$, the limit $1/L \rightarrow 0$ is problematic if $a_0 \neq 0$. In this case, the curve p_L versus 1/L will touch the vertical axis at p_c as expected, but with a diverging derivative, what obviously weakens the accuracy in determining p_c . However, carefully choosing the value of λ , it is possible to obtain $a_0 \simeq 0$ and a null slope where p_L touches the vertical axis. This is the case of the so-called Pinson number adopted for W in the percolation problem [27], whose value is exactly known from conformal invariance arguments, and corresponds to the value that the step function assumes at the isolated point p_c in the thermodynamic limit. In our case, one needs to find the proper λ value by the fitting procedure itself and approaching $a_0 = 0$, as in references [51, 52]. Fitting equation (3) with $\nu = 2$, one can obtain the coefficients of the corrections to scaling and evaluate, with good precision, the location of the critical point. Indeed, $\lambda \simeq 0.7943$ tunes $a_0 \simeq 0$, leading to $p_c \simeq 0.3086$ as shown in the



Figure 4. The Binder cumulant for different system sizes. As *L* increases, the minimum of *U* moves to the right, approaching the transition point p_c . Inset: *U* versus $(p - p_c)L^{1/\nu}$ showing that, for sufficiently large sizes, there is a very good collapse with $p_c \simeq 0.306$ and $\nu = 2$.

bottom inset of figure 3. The estimated uncertainty is located at the last digit which does not sensibly change even when discarding the larger lattice sizes.

Further evidence that the transition is not continuous is given by the Binder cumulant U, as shown in figure 4 (main panel) for several lattice sizes. The behaviour is different from the typical one for continuous transitions. It shows a region of negative values with a minimum on the left of p_c . For small values of p, instead of forming a flat plateau at U = 2/3, it converges to a value below 2/3, probably due to the finite, increasing values of M (indeed, even for $p \rightarrow 0$, when $1 \le M \le t_{rec}$, we get U slightly below 2/3). Moreover, increasing L, the location of the minimum shifts towards the critical point p_c and seems to converge to a non trivial, negative value. This is an indication that, being a zero anomalous dimension quantity, in the thermodynamic limit U assumes an isolated value at p_c . The scaling behavior of U is shown in the inset of figure 4. In the critical region $U = f[(p - p_c)L^{1/\nu}]$ (f is a universal function) and a very good collapse is obtained with $\nu = 2$ and $p_c \simeq 0.306$ for the largest sizes. This slightly different value for p_c is probably due to the smaller sizes used to collapse U.

The fluctuations on the size of the outbreak cluster also have a non-trivial exponent, $\chi \sim |p - p_c|^{-\gamma}$. The main panel of figure 5 shows that the susceptibility χ , as a function of p, develops an increasing peak that moves toward p_c . In the bottom inset, we plot the height at $p_c \simeq 0.3086$, $\chi(p_c)$, versus L^{-1} in a log-log scale. From the best fit (thin solid line) we obtain that its anomalous dimension is $\gamma/\nu \simeq 1.91$. The top inset shows the collapse using this value for γ and $\nu = 2$. Although deviations are present for the small sizes, for sufficiently large systems (the largest two), the collapse is very good. The total time T for all infected agents to get removed also has a peak that moves toward p_c , but its height increases linearly, $T \sim L$ (not shown). As will be shown in the next section, the distribution of cluster sizes for those sites *not* visited by the infected ones is a power-law at p_c , $n(s) \sim s^{-\tau}$, with an exponent that is clearly smaller than 2. It was argued in reference [53] (see also references [53–56]) that for $\tau < 2$, $\gamma = 1/\sigma$ and $\nu = 1/\sigma d$. From these equations we obtain that $\gamma/\nu = d = 2$, i.e., $\gamma = 4$, what is shown as a thick straight line in the bottom inset of figure 5 for comparison. Albeit our result is close to this value, there is still a clear difference.



Figure 5. Behavior of the susceptibility χ as a function of p for different values of L, showing the presence of a pronounced peak close to the transition. (Top inset) Data collapse of χ around the transition using $\gamma/\nu \simeq 1.91$, $\nu = 2$ and $p_c = 0.3086$. Although finite size corrections appear for the smaller sizes, the two largest sizes are well superposed. (Bottom inset) Power-law increase of χ at $p = p_c$, $\chi(p_c) \sim L^{\gamma/\nu}$ where, from the fit (thin solid line), $\gamma/\nu \simeq 1.91$. The thick solid line shows the comparison with $\gamma/\nu = 2$ (see text). In this inset, extra sizes were considered, from L = 19 up to 363.



Figure 6. (Main panel) Heterogeneity *H* as a function of the occupation *p* for several system sizes. As in the ordinary percolation, for sufficiently large *L*, *H* has a peak that approaches the threshold and grows as $H^* \sim L^{1.11}$. (Inset) Collapse of *H* with $\tau \simeq 1.8$, $p_c \simeq 0.308$ and $\nu_H = (\tau - 1)/\nu \tau$ with $\nu \simeq 2$. We do not consider the smallest sizes because of the strong finite size effects.

3.2. Uninfected regions

We now consider the geometrical properties of the unvisited sites close to the percolation threshold, occupying the space left by the outbreak cluster. Although the latter is a single cluster, the sites never visited by an infected agent may form many disjoint clusters, as seen in figure 1. For each sampled configuration, the number of different sizes with at least one cluster present gives, once averaged, the size diversity (or heterogeneity) H [43–49]. The results are shown in figure 6 for several values of L. As expected, it presents a growing peak, of height



Figure 7. Cluster size distribution for the unvisited sites measured at the peak of H (see figure 6) for several system sizes. For comparison, we show (solid line) the $s^{-1.8}$ behavior of H^* , that is also consistent with the distribution of holes in the percolation backbone [56]. Inset: the flat region corresponds to the power-law behavior of n(s).

 H^* , at a concentration p^* close to the threshold. Similarly to the ordinary percolation [43], these peaks are located in the region where the outbreak cluster percolates while the largest uninfected cluster is still building up (in our case, $p > p_c$). The position of the peak moves to the right if one only considers the smaller sizes. However, for intermediate sizes it starts moving toward the previously estimated threshold p_c . This is an indication that the position of the peak of *H*, for the sizes we are able to simulate, does not provide a reliable extrapolation for the threshold. This pre-scaling, strong finite-size effect prevents a more precise estimate of the Fisher exponent τ , that is related with the height of the peak and with the power-law tail of the size distribution at the threshold.

Fitting the height of the peaks in figure 6, we obtain that $H^* \sim L^{1.11}$. Considering that $H^* \sim L^{d/\tau}$ [43, 47], it implies that the Fisher exponent is smaller than 2, $\tau \simeq 1.8$. This indeed provides a good collapse for the height, as shown in the inset of figure 6. All peaks can be centered with $p_c \simeq 0.308$, a value that is consistent with the previous estimates. As discussed in references [43, 47], the width of the critical region scales with L^{1/ν_H} instead of $L^{1/\nu}$, where $\nu_H = \tau \nu/(\tau - 1)$. With $\nu \simeq 2$, as can be seen in the inset of figure 6, the collapse is excellent.

Figure 7 shows the size distribution n(s) for the uninfected regions when the number of particles (not p) is kept constant and corresponds to the point where H has a peak, i.e., $N \simeq p^*L^2$. Since $p^* > p_c$, this peak occurs in the region where the outbreak cluster has a large probability of percolating. Thus, the complementary set of unvisited sites can percolate in this region only when the outbreak is halted very early, in the first steps of the dynamics. Once this happens, the percolating cluster will occupy most of the lattice and contribute to the peak located close to $s \sim L^2$, in the tail of the distribution. The initial part of the distribution is a power-law, $n(s) \sim s^{-\tau}$ whose width increases with L because p^* slowly approaches p_c . The exponent is consistent with our previous estimate obtained from the height of H at $p^*, \tau \simeq 1.8$, as can be seen in the figure (solid line). This value is compatible with the exponent of the size distribution of holes in the backbone of the random site percolation cluster [56], with the no-enclave percolation (NEP) model [53] and with the clusters formed by sites not visited by a random walk [57]. Moreover, for large values of s, there is a peak corresponding to a percolating unvisited cluster. In the NEP model, clusters that are fully surrounded by larger clusters are absorbed into the latter. A similar effect occurs in our model while the many infected agents

are randomly walking and visiting most of the sites around it. Both mechanisms decrease the number of enclaves in the large cluster.

4. Conclusions

We studied, from a statistical mechanics point of view, the equilibrium properties of a simple model for a disease outbreak. In a population of N mobile agents on a square lattice, a single initially infected agent may transmit the disease to nearest neighbors. After a given period of time, t_{rec} , an infected agent becomes unable to further transmit the disease and gets recovered. We define the outbreak cluster as the set of connected sites visited by the infected agents starting from the patient zero. For low densities, there is a large probability of the outbreak being halted at the early steps of the dynamics. In this case, the outbreak cluster would contain only a few sites. As N increases, the outbreak cluster also gets larger and, above a given threshold, it percolates through the lattice. Since infected agents become recovered after $t_{\rm rec}$ steps, the interior of the outbreak cluster may not get completely filled and many holes of different sizes may be present. We focus here on the geometrical properties of the single outbreak cluster and of the set of holes. Since there is not a simple way to make our algorithm incremental, we cannot use the full power of the Newman–Ziff algorithm [27, 28]. Nonetheless, the measures are obtained with a constant N and then reweighted following their procedure. In this way, our control variable is p, a continuous parameter equivalent to the density, or the probability of having an agent in each site. The results show a hybrid percolation transition at $p_c \simeq 0.308\,65$ (for the set of parameters chosen here), along with estimates for the critical exponents that strongly indicate that the generated outbreak cluster is in the same universality class of the random percolation backbone.

The backbone consists of a set of blobs connected by single links [58]. In the model we considered here, when a site gets infected, the next sites visited by its random-walk are included in the outbreak cluster until it gets removed. Each one of these regions is similar to a blob. When, through contact, a susceptible agent gets infected, it starts a new blob. The ensemble of blobs compose both the backbone and the outbreak cluster. The set of sites belonging to the backbone is usually identified among the larger set of the percolation cluster. Our model also provides a way of building a cluster whose critical properties are the same as the backbone of ordinary percolation, although a few dangling ends may remain.

Close to the threshold, the cluster size distribution of the holes is a power-law whose exponent τ is smaller than 2. Our result is compatible with the value obtained from the holes in the percolation backbone, whose τ is given by the Mandelbrot hyperscaling relation

$$au = 1 + \frac{d_{\rm B}}{2} \simeq 1.822,$$
(4)

where $d_{\rm B} \simeq 1.643$ is the fractal dimension of the backbone [58]. Once $d_{\rm B}$ (or τ) is known, other exponents can be determined. For example, $\beta/\nu = d - d_{\rm B} \simeq 0.357$ and $\gamma/\nu = d = 2$. These exponents are very close to the values we obtained for our model, with the exception of γ where we obtained $\gamma \simeq 1.91$ instead of 2.

Whether the conclusions drawn from the chosen parameters change with different sets of values is still to be verified. In particular it would be interesting to check whether the critical exponents are sensitive to such parameters, i.e., whether the universality class remains the same or a richer phase diagram is generated. For example, if t_{rec} , the time an agent remains infectious, were larger, each infected agent would have more time to diffuse and fill more holes in the outbreak cluster. However, the whole cluster would grow larger as more agents will get infected. The question is whether the outbreak cluster will get more compact or, for sufficiently

large systems, its fractal dimension will remain unchanged. A similar question applies for a smaller value of t_{rec} . In this case, the contribution of each infected agent to the outbreak cluster will be less compact, perhaps approaching the self-avoiding random walk case. The more realistic case in which t_{rec} may be different for each agent, and even vary significantly, can also be addressed [59]. In addition, it would be important to consider larger system sizes to confirm and extend the above results. Although constrained by the excluded volume condition that prevents more than one individual in the same site, the mobility of the agents induce some local shuffling and an effective longer interaction. Thus, it would be interesting to approach this problem analytically and check how well a mean field approximation would describe the results presented here. Also, the model considered here can be studied on a non-regular network in order to check whether the transition remains of mixed-order [60].

Acknowledgments

We dedicate this paper to our friend Bob Ziff. Work partially supported by the Brazilian agencies FAPERJ, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and CAPES.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

ORCID iDs

Daniel A Stariolo b https://orcid.org/0000-0001-6378-5448 Jeferson J Arenzon b https://orcid.org/0000-0001-8373-6475

References

- [1] Dumonteil E, Majumdar S N, Rosso A and Zoia A 2013 Proc. Natl Acad. Sci. USA 110 4239
- [2] Bordeu I, Amarteifio S, Garcia-Millan R, Walter B, Wei N and Pruessner G 2019 Sci. Rep. 9 15590
- [3] Castellano C, Fortunato S and Loreto V 2009 Rev. Mod. Phys. 81 591
- [4] Baronchelli A and Soc R 2018 R. Soc. Open Sci. 5 172189
- [5] Amaral M A and Arenzon J J 2018 Europhys. Lett. 124 18007
- [6] Amaral M A, Dantas W G and Arenzon J J 2020 Phys. Rev. E 101 062418
- [7] Pastor-Satorras R, Castellano C, Van Mieghem P and Vespignani A 2015 Rev. Mod. Phys. 87 925
- [8] Sun G-Q, Jusup M, Jin Z, Wang Y and Wang Z 2016 Phys. Life Rev. 19 43
- [9] Brauer F 2017 Infect. Dis. Mod. 2 113
- [10] de Arruda G F, Rodrigues F A and Moreno Y 2018 Phys. Rep. 756 1
- [11] Mata A S 2021 Chaos **31** 012101
- [12] Cardy J L and Grassberger P 1985 J. Phys. A: Math. Gen. 18 L267
- [13] Tomé T and Ziff R M 2010 Phys. Rev. E 82 051921
- [14] de Souza D R, Tomé T and Ziff R M 2011 J. Stat. Mech. P03006
- [15] Thouless D J 1969 Phys. Rev. 187 732
- [16] Cardy J L 1981 J. Phys. A: Math. Gen. 14 1407
- [17] Dodds P S and Watts D J 2004 Phys. Rev. Lett. 92 218701
- [18] Bizhani G, Paczuski M and Grassberger P 2012 Phys. Rev. E 86 011128
- [19] Cai W, Chen L, Ghanbarnejad F and Grassberger P 2015 Nat. Phys. 11 936

- [20] Bassler K E, Liu W, Schmittmann B and Zia R K P 2015 Phys. Rev. E 91 042102
- [21] Janssen H-K and Stenull O 2016 Europhys. Lett. 113 26005
- [22] Lee D, Choi W, Kertész J and Kahng B 2017 Sci. Rep. 7 5723
- [23] Boccara N and Cheong K 1992 J. Phys. A: Math. Gen. 25 2447
- [24] Zhang D-M et al 2005 Phys. Scr. 73 73
- [25] da Silva R and Fernandes H A 2015 J. Stat. Mech. P06011
- [26] de Oliveira P M C 2021 Physica A 561 125287
- [27] Newman M E J and Ziff R M 2000 Phys. Rev. Lett. 85 4104
- [28] Newman M E J and Ziff R M 2001 Phys. Rev. E 64 016706
- [29] Van Mieghem P and van de Bovenkamp R 2013 Phys. Rev. Lett. 110 108701
- [30] Kiss I Z, Röst G and Vizi Z 2015 Phys. Rev. Lett. 115 078701
- [31] Feng M, Cai S-M, Tang M and Lai Y-C 2019 Nat. Commun. 10 3748
- [32] Alves D, Haas V J and Caliri A 2003 J. Biol. Phys. 29 63
- [33] Brauer F and Castillo-Chavez C 2012 Mathematical Models in Population Biology and Epidemiology (Texts in Applied Mathematics) (Berlin: Springer)
- [34] Ziff R M 2021 Physica A 568 125723
- [35] Binder K 1981 Phys. Rev. Lett. 47 693
- [36] Challa M S S, Landau D P and Binder K 1986 Phys. Rev. B 34 1841
- [37] Lee J and Kosterlitz J M 1991 Phys. Rev. B 43 3265
- [38] Selke W and Shchur L N 2005 J. Phys. A: Math. Gen. 38 L739
- [39] Malakis A, Fytas N G and Gülpinar G 2014 Phys. Rev. E 89 042103
- [40] de Oliveira P M C 1992 Europhys. Lett. 20 621
- [41] de Oliveira P M C 1994 *Physica* A **205** 101
- [42] de Figueiredo Neto J M, de Oliveira S M M and de Oliveira P M C 1994 Physica A 206 463
- [43] Noh J D, Lee H K and Park H 2011 Phys. Rev. E 84 010101
- [44] Lee H K, Kim B J and Park H 2011 Phys. Rev. E 84 020101
- [45] Lv J-P, Yang X and Deng Y 2012 Phys. Rev. E 86 022105
- [46] Jo W S, Yi S D, Baek S K and Kim B J 2012 Phys. Rev. E 86 032103
- [47] de la Rocha A R, de Oliveira P M C and Arenzon J J 2015 Phys. Rev. E 91 042113
- [48] de Azevedo-Lopes A, de la Rocha A R, de Oliveira P M C and Arenzon J J 2020 Phys. Rev. E 101 012108
- [49] Mazzarisi O, de Azevedo-Lopes A, Arenzon J J and Corberi F 2021 Phys. Rev. Lett. 127 128301
- [50] Stauffer D and Aharony A 1994 Introduction to Percolation Theory (London: Taylor and Francis)
- [51] de Oliveira P M C, Nóbrega R A and Stauffer D 2003 Braz. J. Phys. 33 616
- [52] de Oliveira P M C, Nóbrega R A and Stauffer D 2004 J. Phys. A: Math. Gen. 37 3743
- [53] Sheinman M, Sharma A, Alvarado J, Koenderink G H and MacKintosh F C 2015 Phys. Rev. Lett. 114 098104
- [54] Sheinman M, Sharma A and MacKintosh F C 2016 Phys. Rev. Lett. 116 189802
- [55] Pruessner G and Lee C F 2016 Phys. Rev. Lett. 116 189801
- [56] Hu H, Ziff R M and Deng Y 2016 Phys. Rev. Lett. 117 185701
- [57] Federbush A and Kantor Y 2021 Phys. Rev. E 103 032137
- [58] Herrmann H J and Stanley H E 1984 Phys. Rev. Lett. 53 1121
- [59] Sun H, Kryven I and Bianconi G 2022 arXiv:2201.10773
- [60] Watts D J 2002 Proc. Natl Acad. Sci. USA 99 5766