

Effects of confinement and vaccination on an epidemic outburst: A statistical mechanics approachÓscar Toledano ¹, Begoña Mula ², Silvia N. Santalla ³, Javier Rodríguez-Laguna ² and Óscar Gálvez ¹¹*Departamento de Física Interdisciplinar, Facultad de Ciencias, UNED, Las Rozas, 28232 Madrid, Spain*²*Departamento de Física Fundamental, Facultad de Ciencias, UNED, 28040 Madrid, Spain*³*Departamento de Física & Grupo Interdisciplinar de Sistemas Complejos, Universidad Carlos III de Madrid, 28911 Leganés, Spain*

(Received 3 March 2021; revised 28 July 2021; accepted 13 September 2021; published 28 September 2021)

This work describes a simple agent model for the spread of an epidemic outburst, with special emphasis on mobility and geographical considerations, which we characterize via statistical mechanics and numerical simulations. As the mobility is decreased, a percolation phase transition is found separating a free-propagation phase in which the outburst spreads without finding spatial barriers and a localized phase in which the outburst dies off. Interestingly, the number of infected agents is subject to maximal fluctuations at the transition point, building upon the unpredictability of the evolution of an epidemic outburst. Our model also lends itself to testing vaccination schedules. Indeed, it has been suggested that if a vaccine is available but scarce it is convenient to carefully select the vaccination program to maximize the chances of halting the outburst. We discuss and evaluate several schemes, with special interest on how the percolation transition point can be shifted, allowing for higher mobility without epidemiological impact.

DOI: [10.1103/PhysRevE.104.034310](https://doi.org/10.1103/PhysRevE.104.034310)**I. INTRODUCTION**

Epidemic containment has been a crucial problem for humankind throughout our history, which has become of paramount importance since the COVID-19 outbreak late in 2019. The efforts to understand the spread of infectious diseases have attracted a large variety of professionals from all scientific fields, ranging from biology to sociology (see recent studies from different fields [1–3]). Mathematical modeling has also provided very relevant tools to analyze the stream of data concerning the infected population and has substantially contributed to policy design (see, e.g., [4–6]). Among the different approximations employed, agent-based models have been extensively used to provide policy recommendations even in the COVID-19 case [7], for which simulations based on stratified population dynamics were carried out [8]. Multiscale approaches are known to improve our ability to explain the geographical expansion of the disease, and they have been also used, along with data-driven simulations, to analyze the epidemic of COVID-19 in Brazil [9]. Epidemic waves have also been considered [9,10], employing stratified population dynamics and nonautonomous dynamics, where mitigation effects are subsequently imposed and relaxed.

During the COVID-19 epidemic and due to the scarce amount of vaccination doses in the first stages of the vaccination campaign, immunization schedules have also attracted attention of the modeling community with the aim of stifling the expansion of an epidemic burst by acting upon a number of individuals substantially below the percolation threshold [11], e.g., through the search of certain types of motifs in the contact network.

However, predictions about the evolution of an epidemic burst are known to be difficult and unreliable, because the

uncertainty in the initial data propagates exponentially [12]. Yet, on occasions the inherently unpredictability of the models can be turned in our favor. As mentioned in previous works, the fluctuations in the number of infected people during the evolution of an epidemic burst can provide useful information regarding our ability to stifle an ongoing epidemic [13].

In this work we propose a very simple agent model [14] of the susceptible-infected-recovered (SIR) type [15,16]. Agents follow a random walk in the vicinity of their homes, which are randomly distributed on a square lattice, wandering up to a maximum distance r , which may vary from agent to agent. When r is very short, agents are effectively confined at home, and an infectious outbreak will very likely die off. As this distance is increased, outbreaks have larger chances of spreading throughout the system until a percolation phase transition is reached [17] in which the presence of an infinite cluster becomes certain. Further increases of mobility will have a very limited impact on the spread of the infectious disease. Interestingly, predictions about the future evolution of the epidemic are more difficult near the transition point. Indeed, above the percolation threshold the mean-field theory associated with the SIR equations provide a very accurate prediction of the evolution of the model, while below the transition point typical outbursts will not propagate beyond a certain correlation length. But as the agent mobility approaches the percolation value, the precise geographical origin of the outbreak becomes crucial to predict the outcome. Even though an infinite connected cluster exists, the probability that the initial infected agent will be part of it is minimal at that point, leading to a maximal uncertainty. Thus we show that the fluctuations in the number of infected people become a very useful observable in order to pinpoint the phase transition.

Of course, our model is far too simple to be taken decisively in order to provide policy recommendations, which should be always undertaken with the assessment of experts from several fields. Yet our results suggest that lock-down measures increase their effectiveness very sharply once the percolation threshold has been crossed. Finding the minimal lock-down measures that lead to the phase transition is therefore of paramount importance and a very difficult task. We cannot provide a complete recipe to establish the location of that threshold in practice, but we provide an interesting proxy as was noted before [13]: fluctuations in the infection reach.

If an effective vaccine is available but scarce, a vaccination schedule becomes unavoidable, i.e., the sanitary authorities should consider which agents must be immunized first. According to their social role, some of these agents can be selected in order to minimize the spread of the infection [18–20]. Of course, many individuals should receive the vaccine because of other considerations, such as age or health condition, and we will not discuss these very relevant aspects. Our model lends itself very easily to the evaluation of the efficiency of a vaccination program. In it individuals are only distinguished through their relations: some homes are relatively isolated, so their inhabitants are not likely to spread the disease. Yet, a naive approach would consider that individuals with many connections should be the first candidates to be immunized. We will show that this criterion is not optimal. Indeed, some agents with few connections act as natural *bridges* between different clusters. Thus if they receive the vaccine the clusters will become isolated and the outburst will be halted.

This article is organized as follows. Section II discusses in detail our agent model and how our numerical simulations are performed. In Sec. III we describe our results regarding the percolation phase transition as the agent mobility is decreased, the secondary cases, and the effects of a finite probability of recovery. Our theoretical framework is discussed in Sec. IV. The efficiency of several vaccination schemes is discussed in Sec. V, especially in connection with the shift of the percolation threshold. Our conclusions and ideas for future work are discussed in Sec. VI.

II. THE CONFINED-SIR MODEL

We propose a simple model to characterize the effects of partial confinement during an epidemic expansion, which we call the *confined-SIR* model. Let us consider N agents moving on an $L \times L$ square lattice with periodic boundary conditions. Agent i possesses a *home*, determined by a fixed lattice point \vec{H}_i . Agents can move freely within their *wandering circles*, centered at \vec{H}_i and with radius r_i , which we will assume equal— $r_i = r$ for all i . Time advances in discrete steps, $\Delta t = 1$, in arbitrary units. At each time step, every agent performs a random step. If the step takes the agent out of the wandering circle a new step is attempted. Homes are distributed randomly over the whole lattice.

Agents can be classified into three compartments: susceptible (S), infected (I), and recovered (R). Susceptible agents get infected with probability $n\beta$ when they share a cell with n infected agents. Infected agents may get recovered at each time step, with probability γ . In some cases an effective SIR model

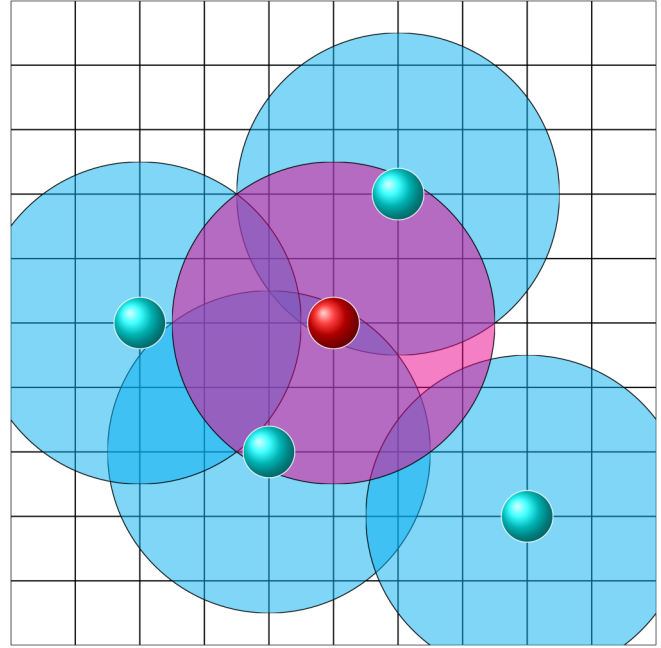


FIG. 1. Illustration of the confined-SIR model. The red agent at the center is infected and can move freely within its wandering circle, shown in magenta. Blue agents are susceptible and can move within their wandering circles, shown in cyan. Notice that the magenta circle intersects all cyan circles, thus providing a finite probability of infection for all those agents.

can be written down and solved analytically. An illustration is provided in Fig. 1.

Let A_i stand for the number of lattice points in the wandering circle of agent i , while A_{ij} will stand for the number of lattice points in the intersection between the wandering circles of agents i and j . The probability that agent i and agent j will collide at time t can be estimated as

$$C_{ij} = \frac{A_{ij}}{A_i A_j}. \quad (1)$$

Thus if agent $i \in S$ and $j \in I$, the probability per unit time that agent i will get infected by j becomes $P_{ij}^I = \beta C_{ij}$. Thus we can find the total probability per unit time that an agent i will get infected by just summing over all possible infection sources:

$$P_i^I = \beta \sum_{j \in I} C_{ij} = \beta \sum_{j \in I} \frac{A_{ij}}{A_i A_j}. \quad (2)$$

We will be mostly interested in the continuum limit, in which $L \rightarrow \infty$, $N \rightarrow \infty$ but the density $\rho = N/L^2$ is constant, in order to avoid lattice effects. In that limit, $A_i \approx \pi r^2$ and A_{ij} corresponds to the overlap between two circles whose centers are a distance d_{ij} apart, which is given by

$$A_{ij} \approx 2r^2 (\arccos(u_{ij}) - u_{ij} \sqrt{1 - u_{ij}^2}) \equiv 2r^2 f(u_{ij}), \quad (3)$$

where $u_{ij} = d_{ij}/(2r) \in [0, 1]$ is the dimensionless relative distance. Thus we have

$$P_i^I \approx \beta \sum_{j \in I} \frac{2f(u_{ij})}{\pi^2 r^2} \approx \beta \rho \frac{2}{\pi^2} \left[\int_0^1 du u f(u) \right] \equiv K \beta \rho, \quad (4)$$

showing that the effective infection probability does not depend on the wandering radius r .

The dynamics of the model can be approximated by a Markov chain. Let us consider that, at time t , we know the epidemiological status of each agent, the location of their home \vec{H}_i , and the wandering radius r . The evolution of the expected number of infected agents can now be found,

$$\langle I(t+1) \rangle = (1 - \gamma \langle I(t) \rangle) + \beta \left\langle \sum_{i \in S} \sum_{j \in I} C_{ij} \right\rangle. \quad (5)$$

Notice that C_{ij} are constants, but the right-hand side evolves due to the migration of agents between sets I and S . In the totally mixing limit, $r \sim L$ and all agents wander around the whole region. Therefore all $C_{ij} \sim 1/L^2$. Thus, in a mean-field approximation Eq. (5) becomes

$$\langle I(t+1) \rangle \approx (1 - \gamma) \langle I(t) \rangle + \frac{\beta}{L^2} \langle S(t) \rangle \langle I(t) \rangle, \quad (6)$$

and we obtain the usual SIR differential equations. Introducing fractional variables, $\mathcal{S} = \langle S \rangle / N$, $\mathcal{I} = \langle I \rangle / N$ and $\mathcal{R} = \langle R \rangle / N$, we reach

$$\begin{aligned} \dot{\mathcal{S}} &\approx -\beta \rho \mathcal{S} \mathcal{I}, \\ \dot{\mathcal{I}} &\approx -\gamma \mathcal{I} + \beta \rho \mathcal{S}(t) \mathcal{I}, \\ \dot{\mathcal{R}} &\approx \gamma \mathcal{I}. \end{aligned} \quad (7)$$

For early times, $\mathcal{S} \approx 1$ so the infection either expands or vanishes, depending on the value of the basic reproductive number, which in this case can be defined as

$$R_0 \equiv \frac{\beta \rho}{\gamma}. \quad (8)$$

Our main focus will be the probability distribution for the total size of the outburst, characterized by the number of affected agents, A , that have suffered the infection at any moment in the long run, and its deviation σ_A . Of course, there are other relevant observables, such as the maximum number of infected agents, but they are not considered in this work.

Our model presents a very clear graph structure. Indeed, two agents i and j may infect each other if and only if $C_{ij} > 0$. This condition determines an effective graph \mathcal{G} , such as the one shown in Fig. 2, in which the homes are represented as nodes and the links denote the pairs of agents that are able to meet. Both networks are obtained for $N = 100$ agents on 50×50 lattices, using $r = 2$ (top) and $r = 5$ (bottom). It can be readily seen that the network for $r = 2$ contains many disconnected clusters, while for $r = 5$ it contains a large cluster spanning most of the nodes.

These networks are examples of *random geometric graphs* (RGGs) [21–24], in which N points are randomly scattered on a $L \times L$ square and joined with edges whenever their distance is less than r . RGG have proved to be a very convenient framework for the analysis of epidemic outbursts [25], typically providing infected nodes with a certain probability per unit time of extending the infection to their neighbors. Yet our confined-SIR model presents a substantial difference with those models. Indeed, the graph structure is equivalent to that of an RGG model, but the actual distances are still relevant in our case. Since our infected agents perform a random walk,

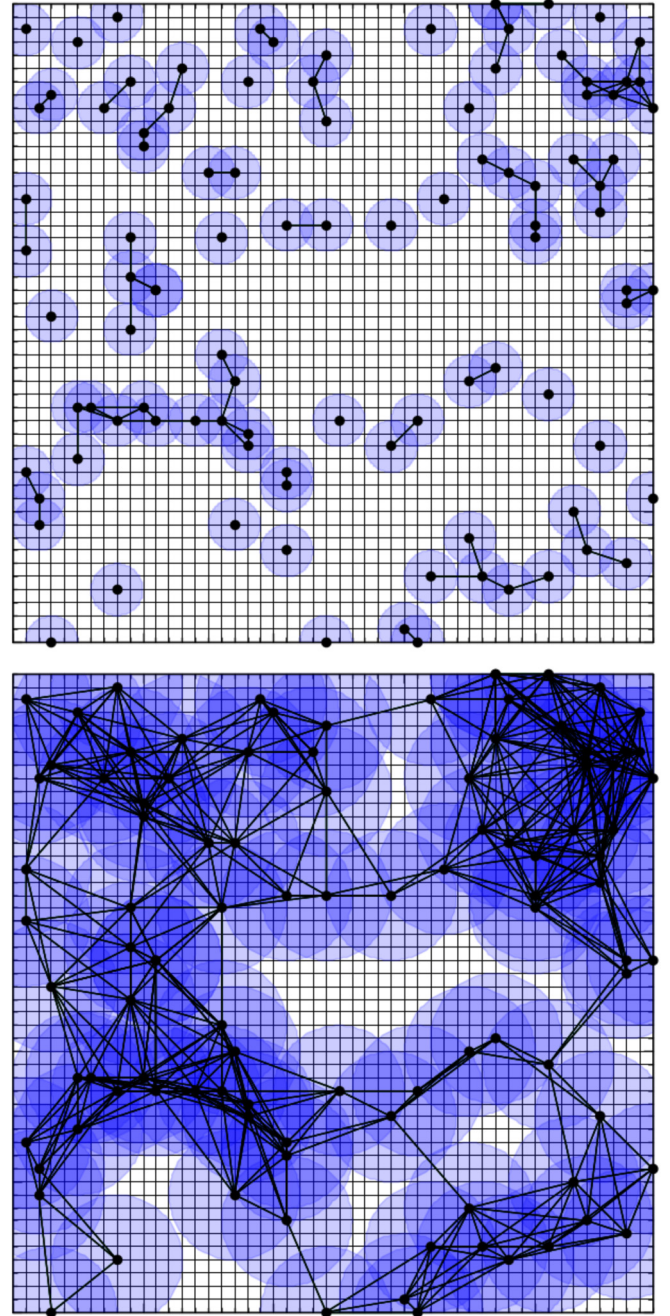


FIG. 2. Effective networks obtained from a confined-SIR system using $N = 100$ agents on a 50×50 lattice. Nodes are associated with the homes of the agents, while links are drawn if their respective dwellers are able to meet during their wandering. Top: using $r = 2$; bottom: using $r = 5$.

their collisions into susceptible agents typically take place one by one, instead of through multiple infections. This difference has been shown to be relevant in the literature [26].

III. NUMERICAL SIMULATIONS

We have run numerical simulations of our confined-SIR model on a square lattice, placing N random homes and using the same value of the wandering radius r for all agents. Un-

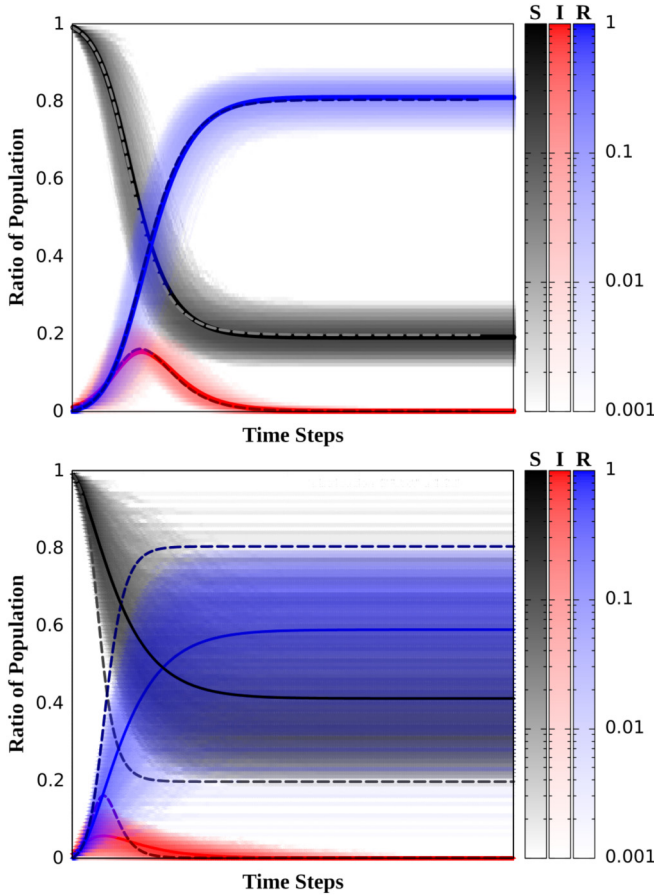


FIG. 3. Fraction of susceptible, infected, and recovered agents as a function of time, depicted with continuous black, red, and blue lines, respectively, shown as a time-dependent histogram through the color gradation (more intense corresponding to larger probabilities), along with their average values shown in continuous lines and the theoretical predictions of the SIR model in dashed lines. In both panels we have used $\beta = 0.02$, $\gamma = 5 \times 10^{-4}$, $L = 141$, and $N = 1000$. The top panel shows the results for a wandering radius $r = 40.3$ and the lower panel for $r = 8.9$.

less otherwise specified, we always average over $N_S = 5000$ samples, with fixed values for $\beta = 1$ and $\gamma = 0$, $N = 1000$ agents, and $\rho = 5 \times 10^{-6}$.

For large values of the wandering radius r we expect the epidemic burst to follow the mean-field approximation provided by the SIR equations for perfect mixing, Eq. (7). We have checked that conjecture numerically in Fig. 3 (top), where we show the probability distribution for the fractions S , I , and R as a function of time, along with their expected values for the theoretical prediction. However, the results for smaller values of r are very different, as we can see in Fig. 3 (bottom). Indeed, in this case the average values of S , I , and R differ notably from the theoretical predictions, because the mixing assumption is inadequate in this regime.

Thus, we are led to identify two different regimes. For large values of r we have nearly perfect mixing, and for low r values the system is naturally divided into isolated clusters. Both are separated by a *percolation transition*, which we characterize in the next section.

A. Percolation phase transition

The theory of bond percolation has been one of the foremost paradigms of statistical mechanics for more than 40 years [27–30], with applications to magnetism [31], wireless communications [32], ecological competition [33], or sequence alignment in molecular biology [34]. Recently a very relevant connection was described between the geodesics in strongly disordered networks and bond percolation [35]. Application of percolation theory to epidemics has been carried out by previous authors, such as those of Refs. [36,37], or more recently, Refs. [38,39].

Bond percolation on a fixed lattice is characterized by a single parameter p , the probability that a given bond will be present. Above a certain threshold value, $p > p_c$, the probability that the system will contain an infinite connected cluster reaches 1, with $p_c = 1/2$ for the square lattice. Our system presents a strong similarity with bond percolation but with another observable playing the role of the bond probability p . Let us consider N agents on an $L \times L$ lattice. The average distance between homes, \bar{r} , can be estimated as

$$\bar{r} = \frac{L}{\sqrt{N}}. \tag{9}$$

Now let us notice that the wandering radius r is only meaningful when compared to this average distance between homes. Thus we introduce a *mobility parameter*,

$$\varepsilon = \frac{r}{\bar{r}} = \frac{r\sqrt{N}}{L}. \tag{10}$$

We will readily show that the mobility parameter ε is the only relevant variable to determine the geometry of our system.

Critical points, such as the percolation transition, typically lead to large fluctuations and can be obtained by considering the deviation of the number of affected agents, σ_A . This magnitude is shown in Fig. 4 as a function of the mobility parameter for different lattice sizes (top) and densities (bottom). The validity of this strategy for characterizing the transition point has already been probed in the literature [13,40]. Indeed, we can observe that the fluctuations in the size of the outburst present a maximum for a certain value of ε which only depends on the ratio between the recovery and infection probabilities, γ/β .

Let us provide evidence that this critical value of the mobility parameter, ε_c , corresponds to the position of the percolation phase transition. In the vicinity of the percolation phase transition, many observables show critical behavior in the form of power laws. The most salient of those is given by the average size of a cluster, s . Below the transition we have

$$\langle s \rangle \sim |p - p_c|^{-\eta}, \tag{11}$$

with $\eta = 43/18 \approx 2.39$ for a square lattice [17]. Figure 5 shows the average cluster size of our effective networks as a function of the mobility parameter ε for different system sizes and populations. We can observe that for all system sizes considered, the average size of the cluster diverges as we approach a critical value ε_c , with an exponent which slightly differs from the value obtained in the square lattice, $\eta \approx 2.35$, which seems to be robust under changes in the lattice size and the density.

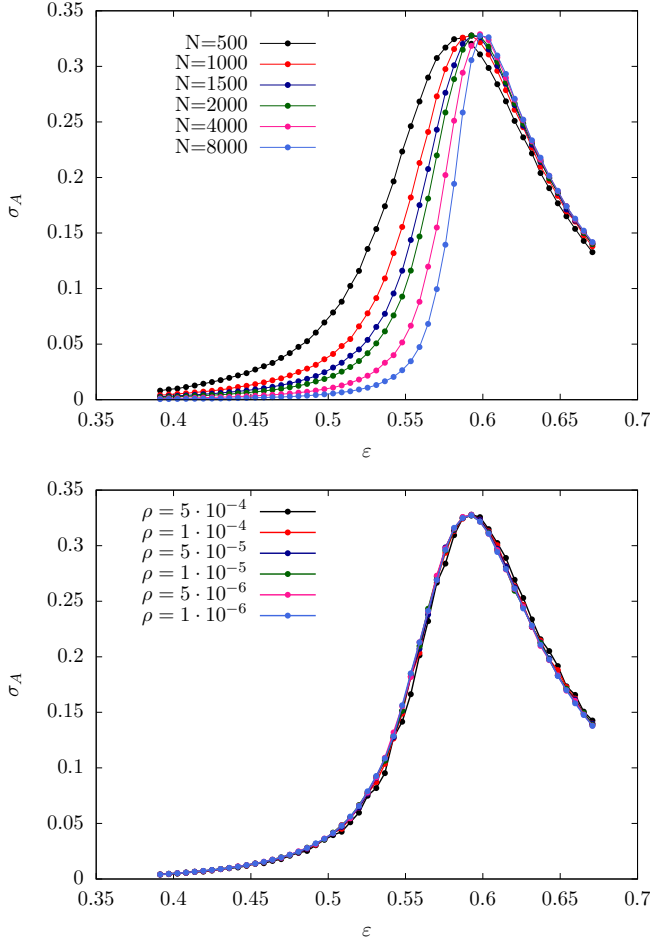


FIG. 4. Top: Deviation of the long-term fraction of affected agents σ_A as a function of the mobility parameter ε for different numbers of agents and a fixed density $\rho = N/L^2 = 5 \times 10^{-6}$, $\beta = 1$, and $\gamma = 0$. Bottom: Same observable as a function of the mobility parameter ε using $N = 1000$ agents for several values of the agent density ρ . Notice the collapse of all curves.

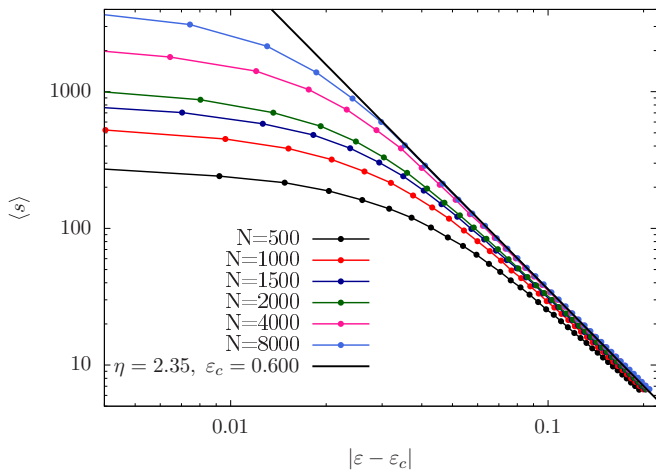


FIG. 5. Average cluster size as a function of $\varepsilon - \varepsilon_c$, using $\rho = 5 \times 10^{-6}$, $\beta = 1$, and $\gamma = 0$. Notice the power-law behavior, following Eq. (11), with an exponent $\eta \approx 2.35$, close to the theoretical prediction $\eta = 43/18$.

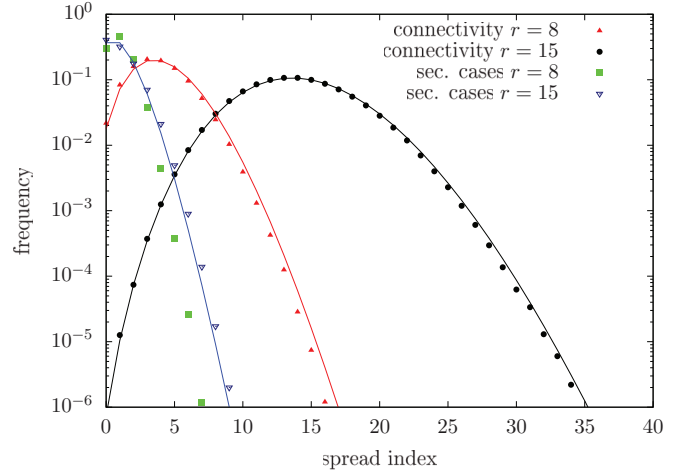


FIG. 6. Histogram of secondary cases in the confined-SIR model for $N = 1000$, $L = 447$, using $r = 8$ and $r = 15$, corresponding to $\varepsilon = 0.566$ and $\varepsilon = 1.060$. For comparison, we also show the connectivity histogram in the associated RGG and the fit to a Poisson law.

B. Secondary cases

Epidemic phenomena are characterized by the number of secondary cases, defined as the number of susceptible agents infected by a single infected node. The average value of this number is related to R_0 , defined in Eq. (8). Yet the deviation of that number has also proved to be a valuable tool to understand the evolution of an epidemic outburst. Indeed, given the exponential nature of the expansion, large fluctuations in the spread rate will naturally dominate the long-term evolution of the disease [41,42].

It is thus relevant to wonder whether the spread of the infection within our model gives rise to large deviations in the number of secondary cases, which we will characterize using its full histogram. In RGG models, a proxy for that value is the number of connections for each node. Figure 6 shows the full histogram for the number of neighbors of any given node, both before and after percolation, i.e., for $r = 8$ and 15 using $N = 1000$ and $L = 447$. The data fit perfectly to a Poisson distribution, corresponding to the theoretical expectation. Furthermore, the figure also shows the number of actual secondary cases, which is much more peaked, showing that the deviation is always low.

The reason for this difference between RGG and our model lies in the statistical properties of time sequences of infection events involving a common infected agent. In an RGG setup infection times are uncorrelated Poisson variates, but in our model those times are not independent. Indeed, an infected agent may require a minimal time to infect two different neighbours, given by the relative positions of the overlapping areas between the three wandering circles involved.

C. Nonzero recovery probability

Let us discuss the effect of a nonzero recovery probability, $\gamma > 0$. In that case, infected agents may recover before they can propagate the disease, and thus we are led to compare two natural times: $\tau_{ij} \approx (\beta C_{ij})^{-1}$ is the expected time be-

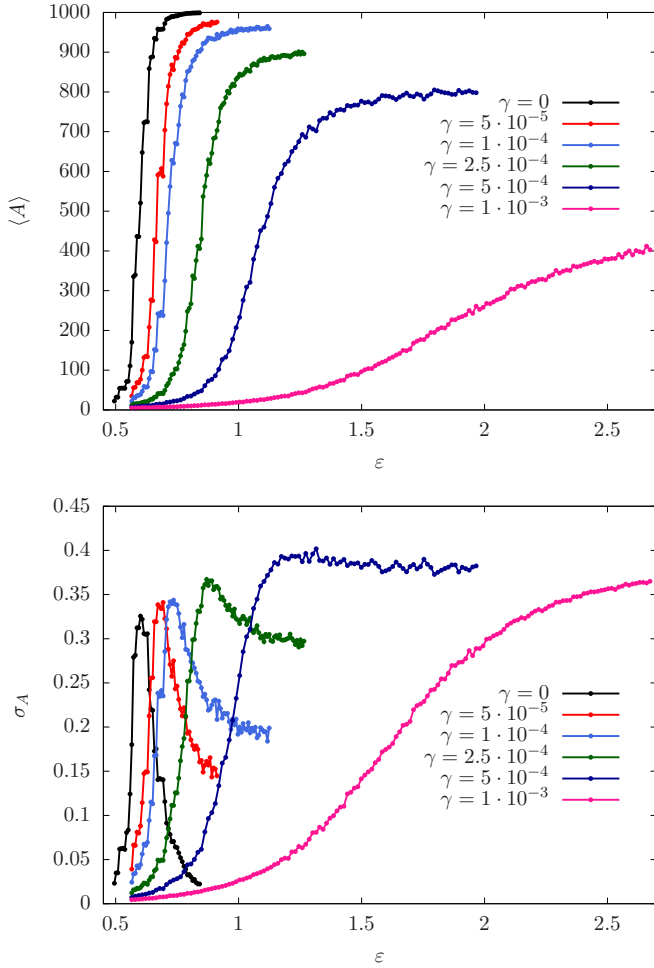


FIG. 7. Top: average number of affected agents $\langle A \rangle$ as a function of the mobility parameter for different values of the γ parameter. Bottom: Deviation of the same magnitude, σ_A . Simulations were performed with $N = 1000$, $N_S = 5000$ samples, $\beta = 1$, and $\rho = 5 \times 10^{-3}$.

fore the infection may propagate from agent i to agent j (or viceversa), while $\tau_R \approx \gamma^{-1}$ corresponds to the expected time before recovery. Thus if $\tau_{ij} \gg \tau_R$, we may assume that the infection will not be able to propagate from agent i to agent j , while in the opposite case, $\tau_R \gg \tau_{ij}$, we may neglect the possibility of recovery. We may claim that a finite recovery probability provides an effective cutoff for the local infection probabilities, thus removing weak links from the graph.

The top panel of Fig. 7 shows the average number of affected nodes $\langle A \rangle$ as a function of the mobility parameter as we increase the recovery probability per unit time γ . It shows a continuous decrease, as expected. The bottom panel of Fig. 7, on the other hand, shows the deviation in the number of affected nodes σ_A , showing that the fluctuations at the maximum do not depend strongly on γ . In addition, the value of the mobility parameter for which this maximum takes place grows rather fast with γ . Thus we are led to claim that the recovery probability per unit time strongly affects the value of the percolation threshold. Also, we can see in the bottom panel of Fig. 7 that for low γ the fluctuations show a peaked maximum, while it seems to reach a plateau for higher values of γ .

IV. THEORETICAL FRAMEWORK

Let us provide a theoretical framework in order to explain the nature of the phase transition observed in the system. A percolation transition of geometrical origin is combined with an epidemic spread transition that takes place when $R_0 > 1$ in a well-mixed system.

First, let us consider the percolation transition on an RGG of size $L \times L$ with N node and radius r . The probability that a single node will be isolated is $1 - \pi(2r/L)^2$ [21]. The expected number of isolated nodes will be approximately given by

$$\mu \approx N \left[1 - \pi \left(\frac{2r}{L} \right)^2 \right]^{N-1} \approx N \exp(-4\pi\varepsilon^2). \quad (12)$$

The average connectivity of a node will be $C = 4\pi\varepsilon^2$, thus showing that the number of isolated nodes will be given by $\mu \sim N e^{-C}$ [23]. Interestingly, this magnitude drops quickly as r increases at the percolation transition for large N . More rigorously, it has been proved that the longest edge of the minimal spanning tree (MST), ℓ , scales like

$$\left(\frac{r_c}{L} \right)^2 \sim \langle \ell^2 \rangle \sim \frac{\log N}{\pi N} \quad (13)$$

for large N [22]. The longest edge of the MST is a good approximation for the minimal value of r for which a giant component will exist within our graph, thus signaling the percolation transition. The validity of Eq. (13) can be checked in the top panel of Fig. 8, which shows our estimate for the critical radius r_c when $\gamma = 0$ using several values of N , along with two different fits,

$$\frac{r_c}{L} = \sqrt{\frac{K}{N^\chi}}, \quad (14)$$

which yields an exponent $\chi = 1.1$ or the slightly more accurate fit to

$$\frac{r_c}{L} = \sqrt{K \frac{\log(N)}{N^\chi}}, \quad (15)$$

which yields $\chi = 0.98$. The inset shows the results for the longest edge of the MST of graphs of the same type, along with both types of fits.

Now let us shift our attention to the characterization of the epidemic spread transition on generic networks [25,43]. The standard approach to locate the spread transition on a random graph is to equate the expected number of infections and the expected number of recoveries in the unit time, which is usually expressed in the mean-field equation,

$$\tilde{\beta} \langle k \rangle = \gamma, \quad (16)$$

where $\langle k \rangle$ is the average degree of the network and $\tilde{\beta} = K\beta\rho$ is the average infection rate for neighboring nodes, given by Eq. (4). It has been shown that $\langle k \rangle$ can be replaced with the largest eigenvalue of the adjacency matrix Λ_{\max} , providing better results [44]. Even more, we may take into account the presence of correlations [45]. Yet Eq. (16) is good enough as a first approximation.

In our case we will consider $\tilde{\beta} \langle k \rangle$ to be the average value of the total infection rate, given by Eq. (4), but *restricting*

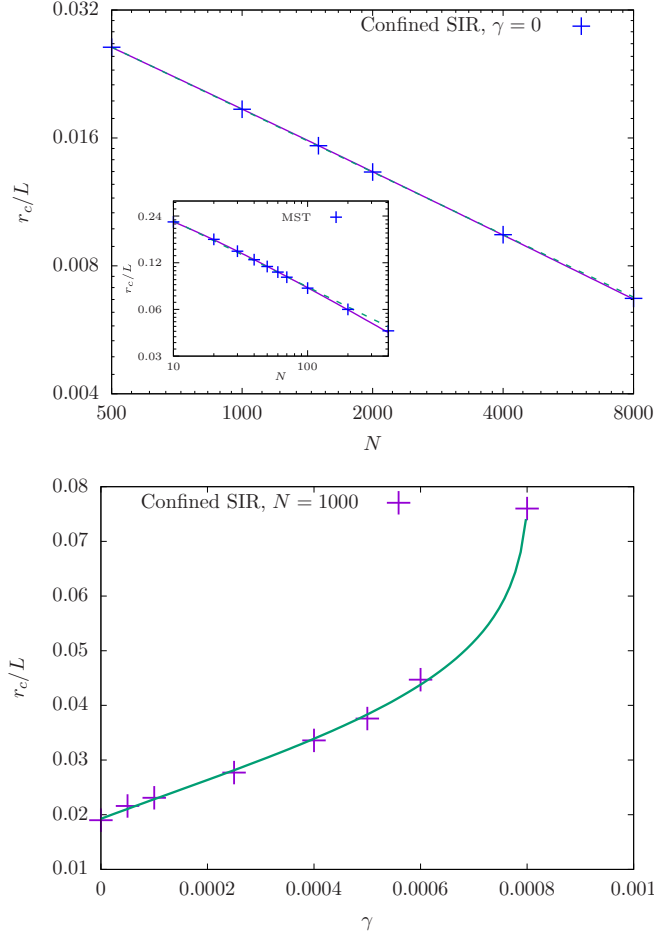


FIG. 8. Top: Transition radius as a function of the size system N for $\gamma = 0$. The dashed line is a fit to expression (14), while the slightly more accurate continuous line shows a fit to expression (15). Bottom: Transition radius as a function of γ , for $N = 1000$, showing a fit to expression (18). The transition value of r_c/L is approximately equivalent to $\langle \ell \rangle$, as shown in (13).

ourselves to nodes belonging to the largest connected component. The probability that a node will belong to this giant component may be approximated as $p_C \approx 1 - \exp(-4\pi\varepsilon^2)$, as we can readily check from Eq. (12). Thus we can write down a fundamental equation defining the critical mobility ε_c ,

$$\gamma_0(1 - e^{\kappa_1 - (\kappa_2\varepsilon)^2}) = \gamma, \quad (17)$$

where $\gamma_0 = K\rho\beta$, and κ_1 and κ_2 are fitting parameters. Solving for ε_c we obtain

$$\varepsilon_c = \kappa_2 \sqrt{\kappa_1 - \log\left(\frac{1}{1 - \gamma/\gamma_0}\right)}, \quad (18)$$

where $\gamma_0 = K\rho\beta$. Equation (18) shows that $r_c \rightarrow \infty$ as $\gamma \rightarrow \gamma_0$, showing that the infection will never spread beyond that point. The bottom panel of Fig. 8 shows the critical radius obtained for $N = 1000$ as a function of γ , along with a fit to Eq. (18), with κ_1 , κ_2 , and γ_0 as fitting parameters. Interestingly, the fit is able to guess a suitable value of $\gamma_0 \approx 0.0008$, despite the fact that the provided values were far from it.

V. VACCINATION SCHEDULES

Let us consider the possibility of providing immunity through vaccination to a (small) fraction of the population, f_v , for the case of zero recovery rate, $\gamma = 0$. In this section we will attempt to answer the following question: How do we select the agents that will receive the vaccine if our only aim is to minimize the size of a future epidemic outburst? Notice that, in practice, many other issues must be considered in this situation, such as the health conditions or the age of the patients.

The simplest vaccination schedule is merely to randomly select the individuals. Of course, we do not expect this schedule to be very efficient. We have considered several observables which can be employed to determine how useful it will be to provide the vaccine to a certain agent. The simplest one is the *total degree*, defined as the number of first neighbors in the effective graph. Naively, we might sort the agents by their degrees and immunize the first $f_v N$. Yet it is more efficient to *recompute* the degree of each agent after each selection, and we will do so unless otherwise specified.

The most promising observable is, nonetheless, the *betweenness centrality* (BC) [46] associated to each agent, BC_i , defined as the fraction of the total number of geodesics which go through agent i . This measure is global and takes $O(N^4)$ steps to compute using Dijkstra's algorithm to evaluate the geodesics, or worse when we recalculate BC_i after each vaccination [47]. In previous studies vaccination schemes based on immunizing the highest BC links have proved to be very efficient (see, e.g., [18]). Considering the high computational cost of the BC, we are led to propose a cheaper alternative, the *local betweenness centrality* (LBC) [48], which is defined for each site as its BC corresponding to a subgraph restricted to itself and its nearest neighbors. In intuitive terms, the LBC is high for a node that connects neighbors which are otherwise disconnected among themselves and is naturally related to the clustering coefficient.

Thus we have considered the following five vaccination schedules:

- (0) No vaccination, considered the base case.
- (1) Select randomly.
- (2) Select the agents with highest *degree* (HD).
- (3) Select the agents with highest *local betweenness centrality* (LBC).
- (4) Select the agents with highest *betweenness centrality* (BC).

The vaccination programs effectively change the topological properties of the network whenever the removed agents are not selected at random. Thus in Fig. 9 we can observe a specific example of a network where different vaccination schemes have been performed over the same amount of agents. In all cases, let us focus on the cluster structure. For random vaccination, Fig. 9(a), the large clusters remain untouched. For a degree-based vaccination scheme, Fig. 9(b), we can see that links have been removed from the *core* of the clusters but the clusters themselves remain connected. Indeed, immunizing the individuals with a large number of connections seems to have a low impact on the network structure in our case. The reason is that high-degree agents tend to be neighbors of other high-degree agents. Panel (c) of Fig. 9, on

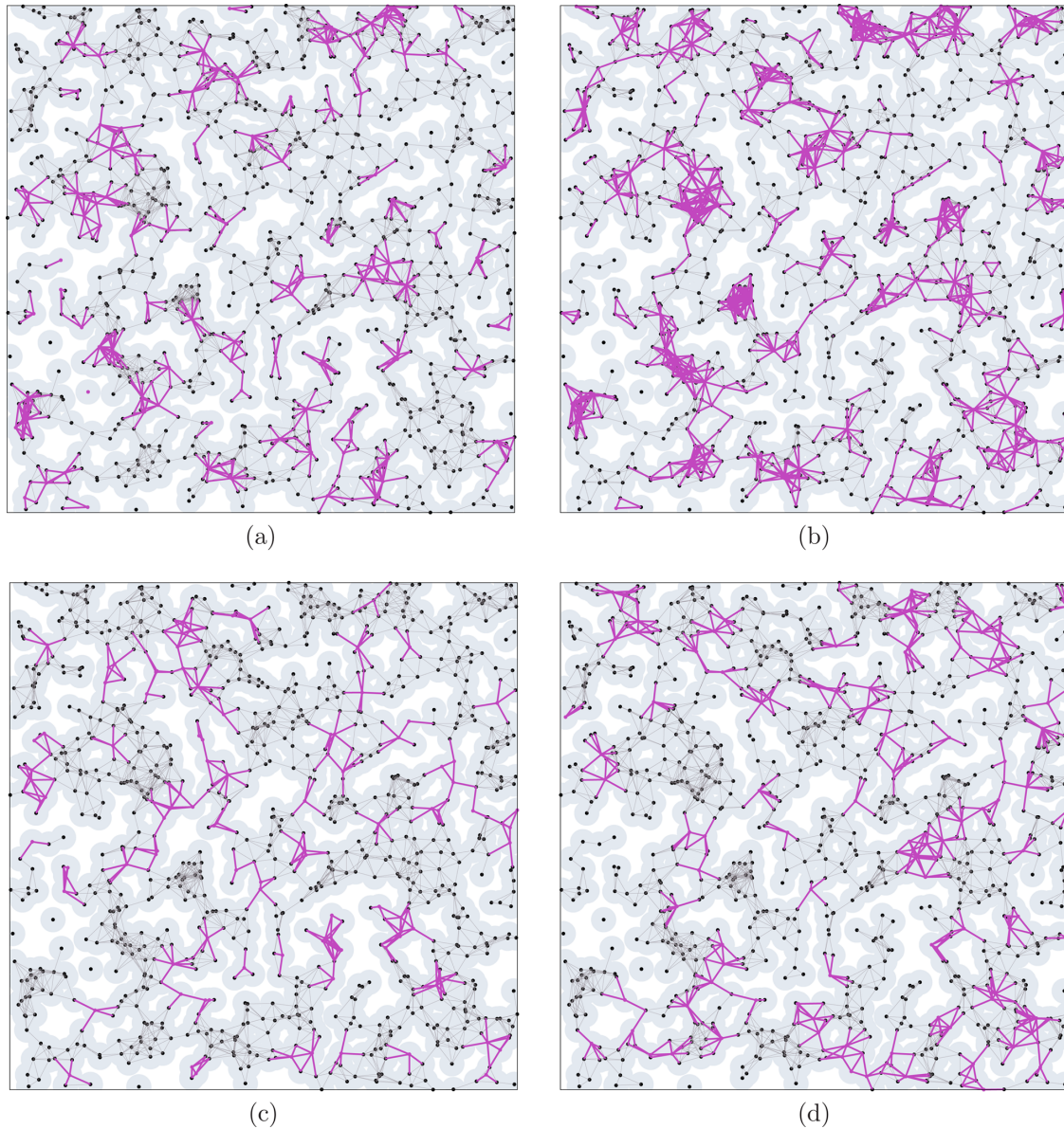


FIG. 9. Effect of different vaccination schedules on a fixed network: (a) random vaccination, (b) highest degree, (c) highest local betweenness centrality (LBC), and (d) highest betweenness centrality (BC). Black dots denote susceptible agents, and they are joined by black links. Immunized agents are colored purple, same as the dead links. The parameters for all cases are $N = 1000$, $\varepsilon = 0.89$, and a 10% vaccination fraction. Videos of the numerical simulations for these vaccinations schemes can be seen in [49].

the other hand, shows that removing agents with a large LBC breaks up some clusters but some large clusters still remain active, leading to a likely propagation of an epidemic outburst to a substantial fraction of the population. Figure 9(d) shows the resulting network when the agents with a largest BC have been removed, and we can readily see that all large clusters have indeed disappeared.

In order to describe the quantitative effect of the different vaccination schedules, the top panel of Fig. 10 depicts the expected value of the long-term number of affected agents as a function of the mobility parameter. The no-vaccination case, marked with the black curve, reaches the total number of agents, $N = 1000$, with the transition at around $\varepsilon_c \approx 0.6$. The effect of the random vaccination scheme is analogous to that

obtained by performing a reduction of the population density with a factor $1 - f_v$, and thus the mobility parameter at the transition point will follow the relation $\varepsilon_c = \varepsilon_c^{(0)}(1 - f_v)^{-1/2}$, where $\varepsilon_c^{(0)}$ stands for the critical value of the mobility parameter when no agents are vaccinated. Interestingly, random vaccination and highest degree vaccination provide similar outcomes, with a slightly higher critical value of the mobility parameter for the random case. Indeed, this shows that highest degree vaccination is not effective at all. Highest LBC vaccination, on the other hand, results in a substantial reduction in the number of affected agents. Yet the most effective vaccination schedule is, with a large difference, the one based on the highest BC (yellow line). This result is in agreement with previous studies [18–20]. Moreover, all schemes reach a similar

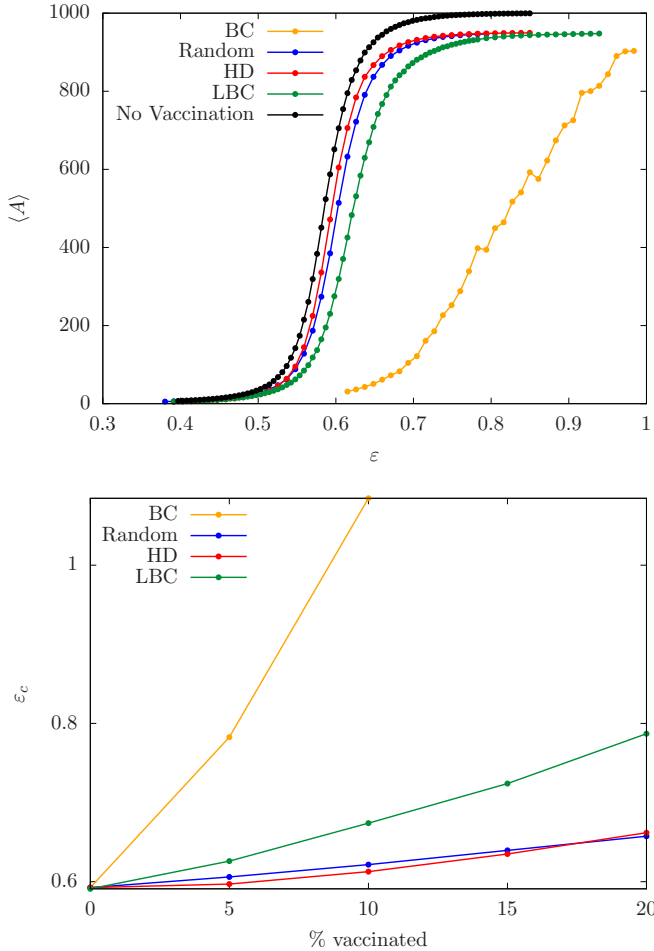


FIG. 10. Top: Average number of affected agents for each vaccination schedule as a function of the mobility parameter ϵ with a vaccination fraction of 5%. Bottom: Mobility parameter marking the percolation phase transition ϵ_c for each vaccination schedule as a function of the vaccination fraction.

value of the total number of affected agents for large mobility. Thus we conclude that an effective vaccination scheme will substantially increase the mobility threshold under which an epidemic outburst will die off. Yet the effects of vaccination can be substantially reduced above that mobility threshold.

How effective are the vaccination schedules in shifting the percolation transition point? In order to answer that question we have traced the percolation threshold ϵ_c as a function of the vaccination fraction f_v in the bottom panel of Fig. 10. The base value, for no vaccination, is $\epsilon_c \approx 0.6$ (as shown in Fig. 5). Highest BC vaccination results in a sharp increase of the percolation transition, reaching $\epsilon_c \approx 1.1$ for $\approx 10\%$ vaccination fraction. The highest LBC vaccination performs substantially worse but still provides a significant improvement of the epidemiological situation for scarce vaccines. Random and high degree vaccinations perform similarly, and none of them are quite effective.

VI. CONCLUSIONS AND FURTHER WORK

In this work we have presented a very simple agent model for epidemic expansion, the confined-SIR model, in which the

effects of partial confinements and vaccinations can be easily tested. We have developed a theoretical framework to characterize the confined-SIR model, which combines results from percolation theory in random geometric graphs with those of mean-field theory for epidemic expansion. It is relevant to notice that our model is far too simple to lead to policy recommendations without further insight from experts from different fields, ranging from virology to sociology. The main shortcoming of our model when applied to human epidemics is that human mobility is not geographically restricted in the same way as in our model. Indeed, even under lock down essential workers must attend their workplaces, which might be far away from their homes. Yet, we expect that some of our conclusions can be interesting for researchers in epidemic expansion and lead, in combination with other insights, to sensible policy recommendations which will help alleviate the effects of present and future epidemic outbursts.

Our first conclusion is that the effects of the confinement measures vary strongly when they cross the percolation threshold. Thus it is of paramount importance to design lockdown measures so that mobility is restricted sufficiently below the percolation threshold in order to limit the expansion of the epidemic. The determination of the percolation threshold is not easy in practice, but a good hint is provided by the fluctuations in the outburst sizes. Above the transition point, most outbursts reach a huge fraction of the population, and below it, they will only affect a few individuals. Yet, near the transition the number of affected agents may vary enormously depending on the location of the initial infected agents. In addition, the effect of increasing the recovery probability in our model causes a decrease of the number of final infected agents, since it provides an efficient cutoff for the infection probability between pairs of agents which seldom meet.

Our second conclusion is that in order to determine an efficient vaccination schedule, disregarding other healthcare considerations, *bridge* individuals should be especially targeted for vaccination, i.e., individuals which move between different clusters. Their immunization will lead to an effective confinement of an epidemic outburst to its initial cluster, thus creating effective firewalls between them. Bridge individuals can be detected via *betweenness centrality* (BC), which is a global and computationally expensive measure, or through easier proxies, such as the individual *local betweenness centrality* (LBC), which addresses the question, *Are your friends friends among them?* Individuals whose friends form a clique are not good candidates for vaccination, but individuals whose friends do not know each other are. Our results in this respect are mixed: vaccinating the highest BC agents is extremely efficient, but finding those agents is very hard, both in the simulations and in real life. Vaccinating the highest LBC agents makes more sense: choose the individuals whose contacts are not in contact among themselves, such as retail salespeople. Yet, more research is required in order to find optimal vaccination schedules.

Confined SIR models seem to be an appropriate tool in order to improve our intuition regarding the effectiveness of different strategies to stifle an epidemic outburst and to find relevant observables in order to characterize the current situation, allowing us to make meaningful predictions. For

example, in this work we have emphasized the very interesting role provided by the fluctuations in the maximal number of infected agents. Indeed, a very promising line of research is the statistical analysis of these fluctuations during real epidemic outbursts, such as COVID-19. These analysis present a very interesting challenge: fluctuations should be compared *ceteris paribus*, i.e., removing major differences between the different geographical areas and times.

ACKNOWLEDGMENTS

We would like to acknowledge J. E. Alvarellos, P. Córdoba-Torres, R. Cuerno, E. Korutcheva, and S. Ferreira for very useful discussions. This work was funded by Instituto de Salud Carlos III (Spain) through Grant No. COV20/01081, and the Spanish government through Grants No. PGC2018-094763-B-I00, No. PID2019-105182GB-I00, and No. PID2019-107514GB-I00.

-
- [1] B. Hu, H. Guo, P. Zhou and Zheng-Li Shi, Characteristics of SARS-CoV-2 and COVID-19, *Nat. Rev. Microbiol.* **19**, 141 (2020).
- [2] L. Wynants, B. V. Calster, G. S. Collins, R. D. Riley, G. Heinze, E. Schuit, M. M. J. Bonten, D. L. Dahly, J. A. Damen, T. P. A. Debray *et al.*, Prediction models for diagnosis and prognosis of COVID-19: Systematic review and critical appraisal, *BMJ* **369**, m1328 (2020).
- [3] L. J. Thomas, P. Huang, F. Yin, X. I. Luo, Z. W. Almquist, J. R. Hipp and C. T. Butts, Spatial heterogeneity can lead to substantial local variations in COVID-19 timing and severity, *Proc. Natl. Acad. Sci. USA* **117**, 24180 (2020).
- [4] X. Li, J. Guo, C. Gao, L. Zhang and Z. Zhang, A hybrid strategy for network immunization, *Chaos, Solitons Fractals* **106**, 214 (2018).
- [5] L. Hébert-Dufresne, A. Allard, J.-G. Young, and L. J. Dube, Global efficiency of local immunization on complex networks, *Sci. Rep.* **3**, 2171 (2013).
- [6] S. Riley, K. Eames, V. Isham, D. Mollison, and P. Trapman, Five challenges for spatial epidemic models, *Epidemics* **10**, 68 (2015).
- [7] A. Aleta, D. Martín-Corral, A. Pastore y Piontti, M. Ajelli, M. Litvinova, M. Chinazzi, N. E. Dean, M. E. Halloran, I. M. Longini Jr., S. Merler *et al.*, Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19, *Nat. Hum. Behav.* **4**, 964 (2020).
- [8] A. Arenas, W. Cota, J. Gómez-Gardeñes, S. Gómez, C. Granell, J. T. Matamalas, D. Soriano-Paños, and B. Steinegger, Modeling the Spatiotemporal Epidemic Spreading of COVID-19 and the Impact of Mobility and Social Distancing Interventions, *Phys. Rev. X* **10**, 041055 (2020).
- [9] G. S. Costa, W. Cota, and S. C. Ferreira, Outbreak diversity in epidemic waves propagating through distinct geographical scales, *Phys. Rev. Research* **2**, 043306 (2020).
- [10] G. A. Muñoz-Fernández, J. M. Seoane, and J. B. Seoane-Sepúlveda, A SIR-type model describing the successive waves of COVID-19, *Chaos, Solitons Fractals* **144**, 110682 (2021).
- [11] G. S. Costa and S. C. Ferreira, Nonmassive immunization to contain spreading on complex networks, *Phys. Rev. E* **101**, 022311 (2020).
- [12] M. Castro, S. Ares, J. A. Cuesta, and S. Manrubia, The turning point and end of an expanding epidemic cannot be precisely forecast, *Proc. Natl. Acad. Sci. USA* **117**, 26190 (2020).
- [13] P. Shu, W. Wang, M. Tang, and Y. Do, Numerical identification of epidemic thresholds for susceptible-infected-recovered model on finite-size networks, *Chaos* **25**, 063104 (2015).
- [14] O. Toledano, B. Mula, S. N. Santalla, O. Gálvez, and J. Rodríguez-Laguna, Confined SIR model for epidemic expansion, software repository, https://github.com/jvrlag/confined_sir (2021).
- [15] W. O. Kermack and A. G. McKendrick, A contribution to the mathematical theory of epidemics, *Proc. R. Soc. Lond. A* **115**, 700 (1927).
- [16] M. Y. Li, *An Introduction of Mathematical Modeling of Infectious Diseases* (Springer, New York, 2018).
- [17] D. Stauffer and A. Aharony, *An Introduction to Percolation Theory* (Taylor and Francis, London, 2003).
- [18] C. M. Schneider, T. Mihaljev, S. Havlin, and H. J. Herrmann, Suppressing epidemics with a limited amount of immunization units, *Phys. Rev. E* **84**, 061911 (2011).
- [19] J. T. Matamalas, A. Arenas, and S. Gómez, Effective approach to epidemic containment using link equations in complex networks, *Sci. Adv.* **4**, eaau4212 (2018).
- [20] Y. Yang, A. McKhann, S. Chen, G. Harling, and J.-P. Onnela, Efficient vaccination strategies for epidemic control using network information, *Epidemics* **27**, 115 (2019).
- [21] M. Penrose, *Random Geometric Graphs* (Oxford University Press, Oxford, England, 2003).
- [22] M. D. Penrose, The longest edge of the random minimal spanning tree, *Ann. Appl. Prob.* **7**, 340 (1997).
- [23] J. Dall and M. Christensen, Random geometric graphs, *Phys. Rev. E* **66**, 016121 (2002).
- [24] J. Díaz, D. Mitsche, and X. Pérez-Giménez, Large connectivity for dynamic random geometric graphs, *IEEE Trans. Mob. Comm.* **8**, 821 (2009).
- [25] E. Estrada, S. Meloni, M. Sheerin, and Y. Moreno, Epidemic spreading in random rectangular networks, *Phys. Rev. E* **94**, 052316 (2016).
- [26] S. Gómez, A. Arenas, J. Borge-Holthoefer, S. Meloni, and Y. Moreno, Discrete-time Markov chain approach to contact-based disease in complex networks, *Europhys. Lett.* **89**, 38009 (2010).
- [27] J. M. Hammersley and D. J. A. Welsh, First-passage percolation, subadditive processes, stochastic networks and generalized renewal theory, in *Bernoulli, Bayes, Laplace Anniversary Volume*, edited by J. Neyman and L. M. LeCam (Springer-Verlag, Berlin, 1965), p. 61.
- [28] C. D. Howard, Models of first passage percolation, in *Probability on Discrete Structures*, edited by H. Kesten (Springer, New York, 2004), pp. 125–173.
- [29] A. Auffinger, M. Damron, and J. Hanson, *50 Years of FPP*, University Lecture Series 68 (American Mathematical Society, Providence, RI, 2017).

- [30] P. Córdoba-Torres, S. N. Santalla, R. Cuerno, and J. Rodríguez-Laguna, Kardar-Parisi-Zhang universality in first passage percolation: The role of geodesic degeneracy, *J. Stat. Mech.* (2018) 063212.
- [31] D. B. Abraham, L. Fontes, C. M. Newman, and M. S. T. Piza, Surface deconstruction and roughening in the multizigzag model of wetting, *Phys. Rev. E* **52**, R1257(R) (1995).
- [32] S. Beyme and C. Leung, A stochastic process model of the hop count distribution in wireless sensor networks, *Ad Hoc Networks* **17**, 60 (2014).
- [33] G. Kordzakhia and S. P. Lalley, A two-species competition model on \mathbb{Z}^d , *Stoch. Proc. App.* **115**, 781 (2005).
- [34] R. Bundschuh and T. Hwa, An analytic study of the phase transition line in local sequence alignment with gaps, *Discr. Appl. Math.* **104**, 113 (2000).
- [35] D. Villarrubia, I. Á. Domenech, S. N. Santalla, J. Rodríguez-Laguna, and P. Córdoba-Torres, First-passage percolation under extreme disorder: From bond-percolation to KPZ universality, *Phys. Rev. E* **101**, 062124 (2020).
- [36] L. M. Sander, C. P. Warren and I. M. Sokolov, Epidemics, disorder, and percolation, *Physica A* **325**, 1 (2003).
- [37] J. C. Miller, Percolation and epidemics in random clustered networks, *Phys. Rev. E* **80**, 020901(R) (2009).
- [38] G. Oliveira, Early epidemic spread, percolation and Covid-19, *J. Math. Biol.* **81**, 1143 (2020).
- [39] F. Croccolo and H. E. Roman, Spreading of infections on random graphs: A percolation-type model for COVID-19, *Chaos, Solitons Fractals* **139**, 110077 (2020).
- [40] L. L. Ramírez-Ramírez and M. E. Thompson, Applications of the variance of final outbreak size for spreading in networks, *Method. Comput. Appl. Prob.* **16**, 839 (2014).
- [41] D. C. Adam, P. Wu, J. Y. Wong, E. H. Y. Lau, T. K. Tsang, S. Cauchemez, G. M. Leung, and B. J. Cowling, Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong, *Nat. Med.* **26**, 1714 (2020).
- [42] D. Lewis, Superspreading drives the COVID pandemic—and could help to tame it, *Nature (London)* **590**, 544 (2021).
- [43] W. Wang, M. Tang, H. E. Stanley, and L. A. Braunstein, Unification of theoretical approaches for epidemic spreading on complex networks, *Rep. Prog. Phys.* **80**, 036603 (2017).
- [44] V. M. Preciado and A. Jadbabaie, Spectral analysis of virus spreading in random geometric networks, in *Proceedings of the 48th IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference* (IEEE, Piscataway, NJ, 2009), pp. 4802–4807.
- [45] D. H. Silva, F. A. Rodrigues, and S. C. Ferreira, High prevalence regimes in the pair-quenched mean-field theory for the susceptible-infected-susceptible model on networks, *Phys. Rev. E* **102**, 012313 (2020).
- [46] L. C. Freeman, Centrality in social networks conceptual clarification, *Social Networks* **1**, 215 (1979).
- [47] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms* (MIT Press, Cambridge, MA, 1990).
- [48] H. Mahyar, R. Hasheminezhad, E. Ghalebi, A. Nazemian, R. Grosu, A. Movaghar, and H. R. Rabiee, Identifying central nodes for information flow in social networks using compressive sensing, *Social Netw. Anal. Mining* **8**, 33 (2018).
- [49] O. Toledano, B. Mula, S. N. Santalla, O. Gálvez, and J. Rodríguez-Laguna, 2021, Epidemic Modelling UNED website, <https://sites.google.com/view/epidemic-modelling-uned/home>.