On the spreading of epidemics and percolation theory

Armin Bunde, Shlomo Havlin and Josef Ludescher

1 Introduction

In a typical epidemic like COVID-19, infected individuals are sick and contagious over a certain time span τ before they get immunized either by recovering or dying. During this contagion period which for COVID-19 typically ranges between one and two weeks, an infected individual can infect other individuals. When more individuals are infected than recover, the net reproductive ratio R_0 is greater than one and the epidemic spreads, otherwise it tends to die out.

At the beginning of an epidemic, the number of newly infected individuals depends on the number of contacts each individual has during the contagion time τ and on the intensity and duration of the contacts. Not every contact transmits the virus and leads to an infection. Intense contacts lead to an infection with higher probability q, and q may even increase, on a larger time scale, due to mutations of the virus. Most contacts are within the family, between close friends, workmates, and schoolmates, but there are also contacts between individuals that do not know each other, in restaurants and theaters, in hospitals, as well as in buses, trains, airplanes. Since people are mobile, an infected individual can also infect others in different cities, countries and continents, allowing the disease to spread around the world.

Armin Bunde
Institut für Theoretische Physik, Justus-Liebig-Universität, D-35392 Giessen, Germany, e-mail: arminbunde00@gmail.com

Shlomo Havlin
Department of Physics, Bar-Ilan University, 52900 Ramat Gan, Israel e-mail: havlins@gmail.com

Josef Ludescher
Potsdam Institute for Climate Impact Research, D-14773 Potsdam, Germany e-mail: josef.ludescher@pik-potsdam.de

The number of contacts a person can have differs tremendously. Some people mostly stay at home and have little contacts to others, while others are highly mobile and meet dozens or even hundreds of people every day. Accordingly, the effect of different people on the spreading of an epidemic is very different: some have nearly no effects while others (the superspreaders) actually drive the spreading of the disease.

When trying to model the spreading of an epidemic, it is helpful when one knows in greater detail the structure of the network, i.e., the links between different people. The stable links between family, close friends, work mates and school mates, for example, trigger the spreading of the disease preferably in the neighborhood (village or city) of an infected individual, while random links between individuals in buses, trains, airplanes and large festivities trigger the spreading of the epidemics to distant places.

The models that aim to describe the spreading of an epidemic can be distinguished in the way the links between different people are treated. Most mathematical studies assume explicitely that populations are "fully mixed', i.e. an infective individual is equally likely to spread the disease to any other member of the population [1–3]. This assumption enables one to write down, in the limit of large population size N, nonlinear differential equations for the number of infective (I), susceptible (S) and recovered and immune or dead (R) individuals as a function of time [4–8].

The SIR equations can be solved in closed form mathematically and yield a useful insight into the development of an epidemic outbreak. The problem with this approach is that the non-homogeneous population topology, in particular the crucial role of the superspreaders in an epidemic outbreak, cannot be accounted for (see also the discussion in [10]). Also, in this approach the (Euclidean) distance between two individuals is irrelevant, in strong contrast to reality. In improvements of the SIR model, in order to obtain some kind of spatial confinement, one often divides a larger area (e.g. a country) into smaller areas. In each small area the individuals living there interact via the SIR model, while the contacts between these areas are considered separately by flux equations (see the Chapter by Dirk Brockmann in this book).

In the past two decades, motivated by the increasing interest in network theory and its applications in nature, society, and technology [11–13], the spreading of diseases in social networks came into the focus of the statistical physics community (see, e.g., [8, 10–31, 33, 34, 36–42]). Within a network approach, the susceptible, infective and recovered individuals form the nodes of a network. The disease can spread between linked nodes. The approach allows to study both the role of the superspreaders and the role of spatial constraints on the individuals [16, 29].

In this Chapter, we will mainly focus on network models that reveal the *characteristic* features of epidemic spreading. By borrowing ideas from *percolation theory* we discuss how epidemics are structured, in particular close to the transition from a localized to a spreading state. The models allow us to discuss, quite generally and partly in an analytical way, the role of the super-

spreaders in an epidemic outbreak, the immunization strategies one should use to bring the epidemic to an halt, and the effect of spatial constraints on the spreading of an epidemic.

Our Chapter is complementary to the chapter by Dirk Brockmann in this volume [8], where the spreading of a disease is considered on a global scale and different questions, like quantitative assessments of an epidemic outbreak or the reconstruction of the actual outbreak location, are being addressed.

First we focus briefly on the historical SIR model.

2 The SIR model

The SIR model is the most widely used epidemic model. As pointed out in [10], the basic ideas were first formulated (though never published) by Lowell Reed and Wade Hampton Frost about 100y ago. In the model, a population of N individuals is considered to have 3 states: infective (I), susceptible (S), and recovered (and this way immunized) or dead (R). Infective individuals (I) have contacts with randomly chosen individuals of all three states at an average rate β per unit time. They infect those that are susceptible (S), and recover and acquire immunity (or die) (R) at an average rate γ per unit time. By definition, the mean contagion time is $\tau = 1/\gamma$.

The rate equations of the SIR model are [4]:

$$\frac{dS}{dt} = -\beta \frac{I}{N} S, \quad \frac{dI}{dt} = (\beta \frac{S}{N} - \gamma) I, \quad \frac{dR}{dt} = \gamma I \tag{1}$$

Accordingly, in the first steps of an epidemic where $S \approx N$, the (mean) number of infective individuals $I(t)$ can be well approximated by

$$I(t) = I(0) e^{(\beta/\gamma - 1)t/\tau}, \tag{2}$$

and either increases exponentially when $\beta/\gamma > 1$ or decreases exponentially when $\beta/\gamma < 1$; Thus, the threshold $\beta/\gamma = 1$ separates a phase where the epidemic does not spread from a phase where it spreads.

It is convenient to split β/γ into the mean number of new contacts ($\equiv \langle k \rangle$) an infective individual has during the contagion time τ and the probability q that it infects them, i.e. $\beta/\gamma \equiv \langle k \rangle q$; the brackets $\langle ... \rangle$ denote an average over a distribution $P(k)$ of new contacts k. Within this notation, the epidemic only can spread when q is above the epidemic threshold

$$q_c = 1/\langle k \rangle. \tag{3}$$

When $q\langle k \rangle$ is above one, the epidemic grows as long as $dI(t)/dt \geq 0$. This requires

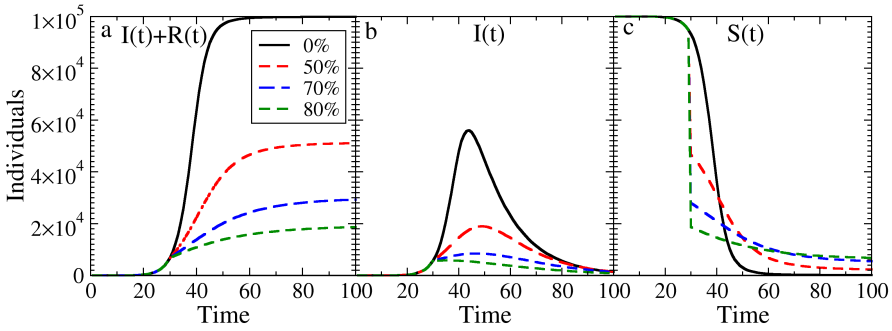

Fig. 1 (color online) Development of an epidemic within the SIR model (Eq. (1)), with and without vaccination. (a) shows the fraction of infected (infective and recovered) individuals $I_{\rm total}(t)/N = (I(t)+R(t))/N$ for $\beta/\gamma = 6$, without vaccination (full black line) and with vaccination of a certain fraction of the population (0.5, 0.7 and 0.8) at a time t_v, when 5 percent of the susceptible population has been infected. (b,c) The same as (a), but for the fractions of infective and susceptible individuals $I(t)$ and $S(t)$.

$$\frac{\beta}{\gamma}\frac{S(t)}{N} \equiv \langle k\rangle q \frac{S(t)}{N} > 1. \tag{4}$$

This condition can be interpreted in several ways:

1. When $\langle k\rangle q$ is reduced by some measures, for example by wearing face masks (which reduces q) or contact restrictions (which reduces $\langle k\rangle$), the epidemic slows down.
2. Since $S(t) = N-(I(t)+R(t))$, the epidemic will die out naturally, without any measures, when the fraction $I_{\rm total}/N \equiv (I(t)+R(t))/N$ of infected (infective plus recovered) individuals is larger than $1-1/(\langle k\rangle q)$; $1-1/(\langle k\rangle q)$ is referred to as herd immunity.
3. When at time t_v a fraction S_v/N of the susceptible individuals $S(t_0)$ has been vaccinated, the epidemic will stop growing when $I_{\rm total}/N + S_v/N$ reaches the herd immunity.

In the δ variant of Covid 19, $\langle k\rangle q$ is supposed to be between 6 and 7. Then the herd immunity is between $5/6 \approx 0.8333$ and $6/7 \approx 0.8571$. Accordingly, the SIR model predicts, for example, that the epidemic should stop growing in populations where 80 percent of the population has been vaccinated and 5 percent has been recovered. Figure 1 shows such a case.

The SIR equations can easily be modified by including death and birth processes as well as (time dependent) containment and quarantine measures [7]. Also dependencies on the locations of the considered individuals can be incorporated (see the chapters by Dirk Brockmann [8] and Jean-Philippe Platteau, Shlomo Weber, and Hans Wiesmeth [9] in this volume).

The advantage of the SIR equations and their simpler variants is that they are easy to handle and can be solved mathematically in closed form. The

disadvantage is that they do not reflect reality, since (i) they do not consider spatial constraints, and, more importantly, (ii) no distinction is made between individuals with many contacts (superspreaders that drive the epidemic) and those with few contacts. This has an enormous impact on the conditions for an epidemic outbreak and leads to false predictions of the herd immunity by the SIR model, as we show later. The model also does not allow any insight into the size distribution of epidemic events and the spatial structure of an epidemic outbreak, in particular near the epidemic threshold $\beta/\gamma = 1$.

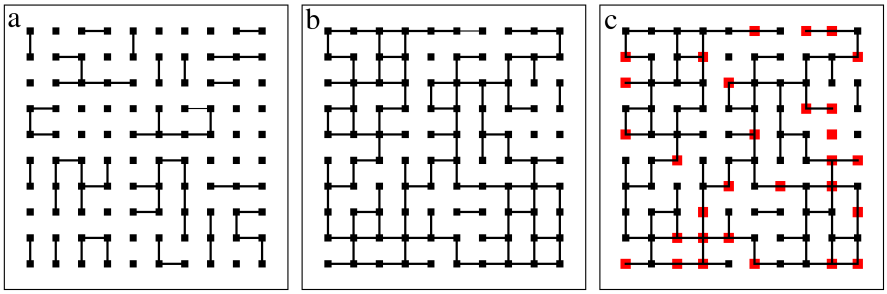

Fig. 2 (color online) Random resistor networks and epidemic spreading. The figure shows a random resistor network consisting of metallic nodes (black squares) and metallic bonds between nearest neighbor nodes. A fraction $1-q$ of bonds is cut. In (a), $1-q = 0.7$, in (b) $1-q = 0.4$. (c) is the same as (b), but in addition a fraction $1-p = 0.3$ of the metallic nodes is substituted by insulating nodes (red squares). While in (a) only small metallic clusters occur, there exists in (b) a metallic path that connects opposite faces of the network and electric current can flow. The added insulating nodes in (c) prevent an electric current to flow. In an epidemic, the nodes are individuals and the bonds specify the infectious contacts between them. Vaccinated individuals that cannot transmit the infection, are equivalent to the insulating nodes.

3 Epidemic models and percolation

Figure 2 shows a random resistor network on a square lattice. At each site of the lattice there is a metallic node (shown as black square). In the regular case, all neighboring nodes are connected by metallic wires and the network is an electric conductor. When we cut randomly a fraction $1 - q$ of the metallic wires (bonds), clusters of sites connected by metallic bonds are formed. When q is close to unity, all clusters are small and an electric current cannot flow between opposite sides of the network. When q is close to zero, there exists a spanning cluster (also called infinite cluster) that connects opposite sides of the network and an electric current can flow. A critical concentration q_c

(here $q_c = 1/2$) separates an insulating phase from a conducting phase. This problem is known as bond percolation [43, 44].

A possible application of bond percolation in chemistry is the polymerization process where small branching molecules can form large molecules by activating more and more bonds between them. If the activation probability q is above the critical threshold, a network of chemical bonds spanning the whole system can be formed, while below criticality only macromolecules of finite size can be generated. This process is called a sole-gel transition. An example of this gelation process is the boiling of an egg, which at room temperature is liquid and upon heating becomes a more solid-like gel. In addition to this prominent example, percolation aspects have been found useful to describe a large number of disordered systems in nature, ranging from thin films, supercooled water, and urban traffic to galactic structures (see, e.g., [44, 45] and references therein).

Here we are interested in the relation between random resistor networks and epidemic spreading. To this end, let us assume that all sites in the lattice represent individuals that can be infected by a disease. In its simplest form, the epidemic starts with one infective individual. During the first time unit (the contagion period τ), the infective individual can infect each of its neighbors with probability q before it gets immunized either by recovering or dying. Each of the infected neighbors will in turn infect, in the 2nd time unit between $t = \tau$ and 2τ, its susceptible neighbors with probability q and so on.

Another way of realizing this epidemic process is to start with a large network where the bonds between nearest neighbors are cut with probability $1 - q$ and then choose randomly one of the sites as starting point for an epidemic. The sites directly and indirectly connected to it specify how many individuals will become infected. It is obvious that the size distribution of the metallic clusters in a very large random resistor network is identical to the distribution of the clusters of infected individuals, and also the critical concentrations in both systems are the same. An epidemic outbreak occurs when $q > q_c$.

Next we assume that in the random resistor network, not all sites act as metallic nodes, but with a probability $1 - p$ the metallic nodes have been substituted by insulating nodes (see Fig. 1c). Only metallic nodes connected by metallic wires can conduct electric current. When $q = 1$, this kind of problem is called site percolation, and the critical concentration p_c denotes the threshold concentration of metallic nodes that is needed for obtaining a spanning metallic cluster. For the square lattice, p_c is about 0.5928 [46]. For arbitrary q, $p_c \equiv p_c(q)$ depends on q.

When considering the spreading of epidemics, the insulating nodes represent individuals that have been immunized by vaccination, and $p_c(q)$ separates a phase where the epidemic can spread from a phase where only finite clusters of infected individuals exist; $v_c(q) \equiv 1 - p_c(q)$ is then the minimum fraction of individuals that must be vaccinated in order to prevent the epidemic from spreading. This will be discussed in Fig. 5.

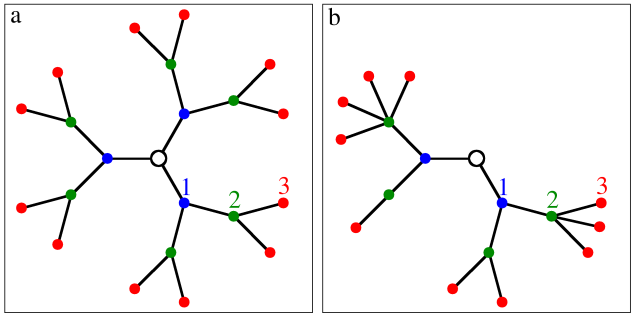

Fig. 3 (color online) Three shells of a Cayley tree, (a) with fixed number of links $k = 3$ emanating from each node, (b) with a distribution of k. The origin of the tree where the epidemic starts, is drawn as large open circle. The length of the bonds shown here is irrelevant.

First we discuss epidemics in tree-like structures (Cayley trees, also called Bethe lattices) where the relevant quantities can be obtained analytically.

4 Epidemics on tree-like structures

The Cayley tree is a structure without loops which is generated as follows [43, 44]: we start with a central site from which k branches emanate. At the end of each branch there is another site, and these k sites constitute the first shell of the Cayley tree. The branches constitute the links between the sites. From each site in the first shell $k - 1$ new branches grow out, this way generating $k(k - 1)$ sites in the second shell. This process is continued (see Fig. 2a for $k = 3$) and large Cayley tress with k links emanating from each site can be generated this way. For $k = 2$, the tree reduces to a one-dimensional chain. Figure 2b shows a Cayley tree where the number of links emanating from a node are drawn from a certain distribution. In this Section we will focus on the Caylee tree with k fixed, where most results of interest can be obtained analytically.

There are no loops in a Cayley tree, since any two sites are connected by only *one* path. The (Euclidean) distance r has no meaning here. The only relevant distance between two sites is the number of links on the shortest path between them (often called *topological* or *chemical* distance). For example, the chemical distance between the central site and a site on the ℓth shell is ℓ.

The ℓth shell contains $k(k - 1)^{\ell-1}$ sites, increasing exponentially with ℓ. In a d-dimensional lattice, with d finite, the number of sites at distance ℓ increases as ℓ^{d-1}. Since the exponential increase can be considered as a power-law behavior with an infinite exponent, the Cayley tree can be regarded as an infinite-dimensional lattice.

4.1 Spreading of a disease

Assume that the epidemic starts at the origin of the tree. In the first time unit, the infective individual at the origin infects each of the k individuals in shell 1 with probability q before it recovers. In the 2nd time unit, each of the infected individuals in shell 1 will in turn infect its $k-1$ neighbors in shell 2 with probability q and so on. Accordingly, when we consider many epidemic events, all starting in the origin of a Cayley tree, on the average

$$I_\ell = k(k-1)^{\ell-1}q^\ell = \frac{k}{k-1}[(k-1)q]^\ell, \quad \ell = t, \tag{5}$$

individuals, all located in shell ℓ, will be infected at the end of the ℓth time unit. For t approaching ∞, the number of infective individuals tends to zero exponentially for $q(k-1) < 1$, and diverges for $q(k-1) > 1$. Accordingly, for q above the epidemic threshold

$$q_c = \frac{1}{k-1}, \tag{6}$$

a typical epidemic outbreak ceases to be confined to a finite number of individuals, and expands to fill an extensive fraction of the tree. The epidemic threshold q_c is the same as in the SIR model, since in the Cayley tree the mean number of new contacts is $k-1$. Due to its tree structure, the increasing number $I_{\rm total}(t)$ of infected and recovered individuals is irrelevant here. Accordingly, the Cayley tree can only be used to describe the initial phase of an epidemic where the number of infected individuals increases exponentially.

The mean number of infected individuals $I_{\rm total}(t)$ (except the one at the center) is obtained by summing $I_\ell(t)$ over all shells ℓ between 1 and t. Since $\sum_{\ell=1}^{t} a^\ell = a(a^t-1)/(a-1)$ we obtain (with $k-1 = 1/q_c$)

$$I_{\rm total}(t) = \frac{q}{q-q_c}(1+q_c)\big((q/q_c)^t - 1\big). \tag{7}$$

Equation (7) is rigorous and holds for all $t \geq 1$. For q below q_c, $I_{\rm total}$ reaches the constant value $I_{\rm total}(\infty) = q(1+q_c)/(q_c-q)$ for large t, while well above q_c, $I_{\rm total}(t)$ increases exponentially, $I_{\rm total}(t) \sim (q/q_c)^t$.

4.2 Behavior at the epidemic threshold

For studying the behavior of $I_{\rm total}(t)$ close to q_c we follow percolation theory [43, 44]: by an epidemic, clusters of infected sites are generated with (on the average) I_ℓ sites at chemical distance ℓ from the origin. Accordingly, $I_\ell / \sum_{\ell=1}^{\infty} I_\ell$ can be interpreted as correlation function $g(\ell)$, with the correlation length ξ given by

$$\xi^2 = \frac{\sum_{\ell=1}^{\infty} \ell^2 I_\ell}{\sum_{\ell=1}^{\infty} I_\ell}. \tag{8}$$

ξ is the characteristic topological length of the cluster of infected individuals. Since $\sum_1^\infty \ell^2 a^\ell = -a(a+1)/(a-1)^3$ and $\sum_1^\infty a^\ell = -a/(a-1)$, we obtain

$$\xi^2 = q_c \frac{q_c + q}{(q_c - q)^2}, \quad q < q_c. \tag{9}$$

Thus the correlation length ξ diverges as $\xi \sim |q_c - q|^{-\nu}$, with $\nu = 1$ for the Cayley tree.

Close to q_c, where $\epsilon \equiv (q_c - q)/q_c$ tends to zero, $(q/q_c)^t \equiv (1-\epsilon)^t \cong \exp(-\epsilon t)$. This yields

$$I_{\rm total}(t) = t\ f(t/\xi) \tag{10}$$

with $f(x) = (1+q_c)[1 - \exp(-x\sqrt{2})]/(x\sqrt{2})$ for $q < q_c$ and $f(x) = (1+q_c)[\exp(x\sqrt{2}) - 1]/(x\sqrt{2})$ for $q > q_c$.

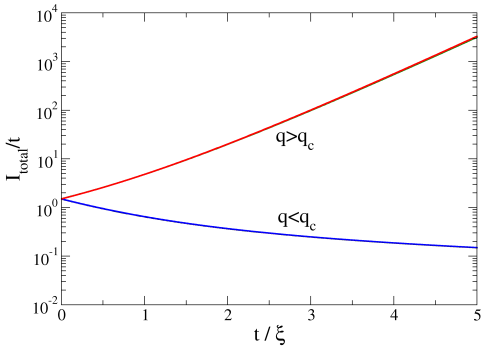

Fig. 4 (color online) Scaling behavior near the epidemic threshold in a Cayley tree. The figure shows $I_{\rm total}(t)/t$ for $|(q-q_c)/q_c| = 10^{-2}, 10^{-3}$, and 10^{-4}, as a function of t/ξ. The data collapse shows that $I_{\rm total}(t)/t$ does not depend separately on q and t, but only on the combined variable t/ξ with $\xi \propto |q - q_c|^{-\nu}$ and $\nu = 1$ here.

Figure 4 shows $I_{\rm total}/t$ as a function of t/ξ below and above criticality. The data collapse confirms the scaling behavior (10) which is characteristic for epidemic spreading in any kind of network with any distribution of links. Around the epidemic threshold, in the initial stages of the epidemic, $I_{\rm total}(t)$ always increases algebraically with t, in some cases with exponents differing from 1. The scaling function f describes the change of $I_{\rm total}$ when departing from the epidemic threshold. f does not depend separately on q and t, but solely on the combined variable $x = t/\xi(|q - q_c|)$, and its functional form is different below and above the epidemic threshold.

4.3 Cluster statistics

Below the epidemic threshold q_c, $I_{\rm total}(\infty)$ specifies the mean number of individuals $\langle s\rangle$ that will be infected during an epidemic event. By definition, $I_{\rm total}(\infty)$ results from an average over many events. In order to obtain the *distribution* of the cluster size s, we consider the probability $w_s(q)$ that the infective site at the origin of the tree has infected exactly $s-1$ other sites. In general, the probability of finding a finite cluster with s infected sites surrounded by t non-infected perimeter sites is $sq^s(1-q)^t$. On the Caylee tree, there exists a unique relation between s and t. A cluster of 1 site is surrounded by k, a cluster of 2 sites by $k+(k-2)$ perimeter sites. In general, a cluster of s sites has $k-2$ more perimeter sites than a cluster of $s-1$ sites, and thus $t(s)=k+(s-1)(k-2)=2+(k-2)s$. Hence, up to a constant normalization factor, $w_s(p)$ is given by

$$w_s(q) \propto g_s s q^s (1-q)^{2+(k-2)s} \equiv g_s s (1-q)^2 [q(1-q)^{k-2}]^s, \qquad (11)$$

where g_s is the number of configurations for an s-site cluster.

One can show [43, 44] that near q_c, to lowest order in ϵ, w_s also shows scaling behavior,

$$w_s(q) = s^{1-\tau} h(\epsilon s^\sigma) \qquad (12)$$

with $\tau = 5/2$, $\sigma = 1/2$, and $h(x) \propto \exp(-x^2/(2-2q_c))$ for the Cayley tree. Accordingly, at the epidemic threshold ($\epsilon = 0$) there is no characteristic size and the cluster size distribution decays by a power law. Below and above q_c, the correction to this power-law distribution only depends on the combined variable ϵs^σ and crosses over to an exponential decay for s above $\epsilon^{-1/\sigma}$.

Above q_c, there is a finite probability P_∞ for an epidemic outbreak where the epidemic spreads over the tree and a finite fraction P_∞ of individuals becomes infected. Accordingly, a finite fraction $1-P_\infty$ of the individuals will not be infected. Below q_c, $P_\infty \equiv 0$ by definition. One can show [43, 44] that above q_c, P_∞ increases by a power of ϵ,

$$P_\infty \sim \epsilon^\beta, \qquad (13)$$

with $\beta = (\tau-2)/\sigma$. Since for the Cayley tree $\tau = 5/2$ and $\sigma = 2$, we find $\beta = 1$ here, i.e., P_∞ increases linearly with increasing distance $q - q_c$ from the epidemic threshold q_c.

We like to stress that Eqs. (9), (10), (12), and (13) are important, since they are not only valid for epidemic spreading on a Cayley tree, but also on the other more complex networks that we will discuss below, with appropriate exponents ν, τ, σ, and β and appropriate scaling functions $h(x)$ below and above q_c.

Next we are interested in the time course of $s(t)$ during a typical epidemic event close to the epidemic threshold.

4.4 Large single epidemic events near the epidemic threshold

Since near q_c the mean number of infected sites $I_{\text{total}}(t)$ grows algebraically with time, we expect that also $s(t)$ will increase by a power law, $s(t) \propto t^{d_\ell}$, before reaching its asymptotic value. By using Eq. (12) one can show easily that close to q_c $I_{\text{total}}(t)$ and $s(t)$ are related by $I_{\text{total}}(t) \sim s(t)^{3-\tau}$ and thus

$$s(t) \sim t^2. \tag{14}$$

Since in epidemic spreading, time t is identical to the topological distance ℓ between the origin and shell ℓ of the tree, $d_\ell = 2$ is the topological dimension of the cluster of infected sites. While above q_c the dimension of the largest spreading cluster is infinity on large scales, the dimension of the cluster is finite at q_c. As a consequence, at the epidemic threshold, the total number of infected individuals form quite a dilute structure. The structure has self-similar properties as we we will see in Section 5.

The subexponential power law behavior (14) occurs when we are close to the epidemic threshold and is typical for critical phenomena at a critical point [47]. It is not unlikely that in a real epidemic, as a consequence of containment strategies, the epidemic is driven close to the epidemic threshold. In this case, we expect, according to Eq. (14), that the number of infected individuals grows quadratically with time t.

Interestingly, when inspecting data from China collected in February 2020, Maier and Brockmann [7] noticed that, unexpectedly, the epidemic did not take off exponentially, but rather by a power law with an exponent close to 2. To describe this effect they introduced a parsimonious SIR-type model that captures both quarantine of symptomatic infected individuals, as well as population-wide isolation practices in response to containment policies or behavioral changes. From our point of view, a power law behavior with a universal exponent 2 is quite natural when containment strategies drive the epidemic towards the epidemic threshold.

In addition, when there are strong constraints on the mobility of a population that makes long-distant contacts unlikely and also restricts the number of contacts an individual can have, the epidemic effectively spreads like on a network where practically only nearest neighbors are linked. In such a network, $s(t)$ also grows quadratically in time (see Sect. 5).

4.5 The effect of partial immunization

Next we consider the case that not all individuals on the Cayley tree (with k fixed) are susceptible, but a certain fraction $v = 1 - p$ of the population has been vaccinated and is immune. In this case, on the average only a fraction p

of the individuals in shell ℓ is susceptible, and the mean number I_ℓ of infected individuals during time step $t = \ell$ becomes

$$I_\ell = k(k-1)^{\ell-1}(pq)^\ell = \frac{k}{k-1}[(k-1)(qp)]^\ell, \ell = t. \tag{15}$$

Equation (15) is identical to Eq. (5), when q in (5) is substituted by $q_{\rm eff} = qp$. Accordingly, the epidemic stops growing, when a critical fraction $v_c(q) \equiv 1 - p_c(q)$ has been vaccinated. The condition $(k-1)qp < 1$ yields (with $q_c = p_c = 1/(k-1)$)

$$p_c(q) = \begin{cases} 1 & , q \le q_c \\ p_c/q & , \text{else.} \end{cases} \tag{16}$$

Accordingly, the fraction of immunized individuals v needs to exceed $v_c(q) = 1 - p_c(q) = 1 - 1/[(k-1)q]$ to prevent epidemic spreading. For example, when $(k-1)q = 6$ (as for the δ-variant of COVID-19), 83.33 percent of the population must be vaccinated, the same result as for the SIR model.

5 Epidemics on planar lattices

Next, in order to see the effects of spatial constraints on epidemic spreading and to learn more on the structure of an epidemic event near the epidemic threshold, we consider epidemics in 2 dimensional lattice structures, (a) in square lattices where $k = 4$ and (b) in triangular lattices where $k = 6$. Due to the spatial constraints that allow loops the critical thresholds q_c and p_c are no longer identical to $1/(k-1)$. For the square lattice, both $q_c = 1/2$ and $p_c \cong 0.593$ are well above $1/3$, and for the triangular lattice, both $q_c = 2\sin(\pi/18) \approx 0.347$ and $p_c = 1/2$ are well above $1/5$ [46].

To an excellent approximation [48], the critical lines follow the relation $\ln p/\ln p_c + \ln q/\ln q_c = 1$ which yields

$$p_c(q) = \begin{cases} 1 & , q \le q_c \\ p_c/q^a & , \text{else.} \end{cases} \tag{17}$$

with $a = \ln p_c/\ln q_c$. Equation (17) is exact for the Cayley tree where $a = 1$.

Figure 5 shows $v_c(q) = 1 - p_c(q)$ (a) for the square lattice and the corresponding Cayley tree with $k = 4$, and (b) for the triangular lattice and the corresponding Cayley tree with $k = 6$. Due to the strong spatial constraints in both lattices, epidemic spreading is less effective than in the Cayley trees, and less people need to be vaccinated.

Figure 6 illustrates the spreading of a typical epidemic in a 200×200 square lattice, for $v = 1 - p = 0$ and $q = 0.45$, 0.5, and 0.55. For $q = 0.45$, the epidemic cluster is quite small. At the epidemic threshold $q_c = 0.5$, the spreading cluster shows a quite fragile structure characterized by the

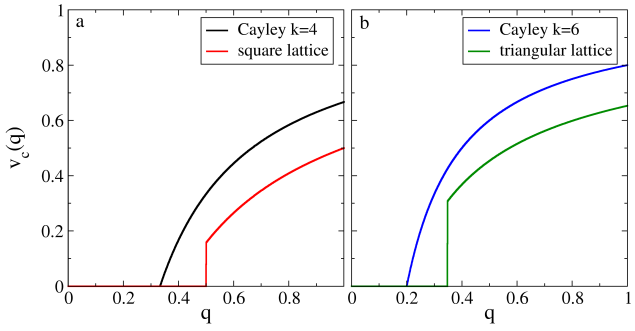

Fig. 5 (color online) Random immunization in loopless networks and lattice structures. The figure shows the fraction of individuals $v_c(q)$ that must be immunized to prevent epidemic spreading as a function of the infection probability q, for (a) the square lattice and the corresponding Cayley tree with $k = 4$ and (b) the triangular lattice and the corresponding Cayley tree with $k = 4$.

topological dimension $d_\ell \cong 1.678$ [44], being smaller than for the Cayley tree. The cluster is self-similar: A part of it, magnified, looks like the whole. A demonstration of the self-similarity is shown in [44]. At $q = 0.55$ already, the spreading cluster is compact on large scales with $d_\ell = 2$. Accordingly, above the epidemic threshold the epidemic grows quadratically in time, due to the strong spatial constraints (imposed by the lattice structure) that allow transmission of the infection solely between nearest neighbor sites.

The scaling laws Eqs. (10) and (12) that describe the cluster size distribution and the mean number of infected individuals in the Cayley tree hold also in lattice structures, only the exponents are different. Also the correlation length ξ and the size P_∞ of the epidemic outbreak above q_c follow power laws, but with different exponents than in Eqs. (9) and (13).

Next we discuss epidemic spreading on more general networks which are closer to reality since the number k of links emanating from a node can have a broad distribution. Spatial restrictions can also be easily introduced on these networks.

6 Epidemics on networks without spatial constraints

Perhaps the most popular random network is the Erdos-Renyi (ER) network, where N nodes are linked randomly by n links [49, 50]. To generate the ER network, one considers each pair of nodes (there are $m = N(N-1)/2$ pairs) and connects it with probability $p_0 = n/m$. For large N, the number of links k emanating from each node follows a Poissonian distribution, with mean $\langle k \rangle = \sum_k kP(k) = 2n/N$.

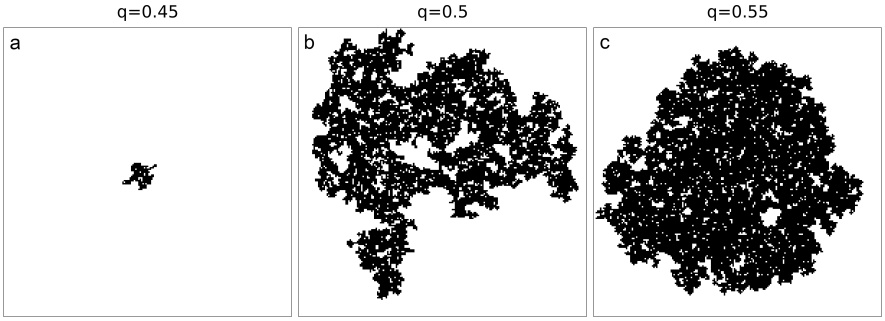

Fig. 6 (color online) Epidemic spreading on a square lattice, with infection probabilities $q = 0.45$ (a), 0.5 (b), and 0.55 (c). Right at the epidemic threshold $q_c = 0.5$, large clusters of infected individuals show a self-similar structure.

For better understanding the role of the superspreaders, we are interested in networks with a more general distribution $P(k)$. To generate such a network one first draws a sequence of N links $k_1, k_2, \ldots k_N$ from this distribution and assigns to each node $i, i = 1, 2, \ldots N$, the appropriate number k_i of links that will emerge from it. Then one connects randomly pairs of nodes until all links are used up. The network generated this way is chosen uniformly at random from the set of all graphs with the selected link sequence [10, 51].

Like in the Cayley tree, the Euclidean distance between 2 nodes is irrelevant here, only the topological distance ℓ matters. In contrast to the Cayley tree, loops can occur, but are irrelevant below the epidemic threshold since the overall fraction of loops in the network is smaller than S/N where S is the size of the largest existing cluster [16,37]. Thus we expect that the critical exponents are the same as for the Cayley tree.

The question is how the critical thresholds q_c and p_c can be obtained in these structures. It has been proven by Molloy and Reed [51] (see also [16]) that in random networks, a spanning cluster almost surely exists when

$$\sum_{k=1}^{N} k(k-2)P(k) \equiv \langle k^2 \rangle - 2\langle k \rangle > 0, \tag{18}$$

and almost surely all clusters are small when $\sum_k k(k-2)P(k) < 0$. Relation (18) can be used to determine p_c as follows.

When one removes randomly a fraction $(1-p)$ of nodes leaving a fraction p intact, the distribution changes into a distribution $P(p,k)$; by definition, $P(1,k) \equiv P(k)$. For a spanning cluster to exist, $P(p,k)$ must satisfy (18). It has been shown in [16], that $\sum_k kP(p,k) = p\sum_k kP(k)$ and $\sum_k k^2P(p,k) = p^2\sum_k k^2P(k) + p(1-p)\sum_k kP(k)$. Inserting this result into (18), yields [16]

$$p_c = \frac{\sum_{k=1}^{N} kP(k)}{\sum_{k=1}^{N} k(k-1)P(k)} \equiv \frac{1}{\langle k^2\rangle/\langle k\rangle - 1.} \tag{19}$$

Using a generating function approach [17] it was shown in [10,17] that also q_c is determined by (19), i.e., $q_c = p_c$. Since loops are irrelevant below q_c, we expect that $p_c(q)$ also satisfies (16). In the following, we focus on q_c.

For networks with fixed k, $\langle k^2\rangle = k^2$, and q_c reduces to $q_c = 1/(k-1)$). For the Poissonian distribution, $\langle k^2\rangle = \langle k\rangle^2 + \langle k\rangle$ and q_c becomes $1/\langle k\rangle$ [49,50].

6.1 Power law distribution $P(k)$ with exponential cutoff

In a real population, few people have many contacts while many people have few contacts. A relevant distribution which reflects this feature and is seen in a variety of real-world networks is a power-law [24] with an exponent $\beta \geq 0$ and an exponential cutoff around $k = \kappa$ [10],

$$P(k) = \begin{cases} 0 & , k \leq k_{\min} - 1 \\ k^{-\beta} e^{-k/\kappa} / \sum_{k=k_{\min}}^{\infty} k^{-\beta} e^{-k/\kappa} & , k \geq k_{\min} \end{cases} \tag{20}$$

A prominent example of a power-law distribution (with β close to 3) in social networks is the distribution of the number of sexual contacts of males and females [20].

Since $\sum_{k=1}^{\infty} k^{-\beta} x^k$ is identical to the βth polylogarithm of x, $\mathrm{Li}_\beta(x)$, the normalization factor in (20) can be written as (with $x = \exp(-1/\kappa)$)

$$\sum_{k=k_{\min}}^{\infty} k^{-\beta} x^k = \mathrm{Li}_\beta(x) - \sum_{k=1}^{k_{\min}-1} k^{-\beta} x^k. \tag{21}$$

For $\beta = 1, 0, -1, -2, -3 \ldots$, there exist simple formulas for $\mathrm{Li}_\beta(x)$, for example: $\mathrm{Li}_1(x) = -\ln(1-x)$, $\mathrm{Li}_0(x) = x/(1-x)$, $\mathrm{Li}_{-1}(x) = x/(1-x)^2$.

Using (20) and (21), $\langle k\rangle$ and $\langle k^2\rangle$ can be calculated straightforwardly. The calculations are particularly easy for $\beta = 1$ in the limit of large κ and yield (for $k_{\min} = 1$ and 2)

$$q_c \cong 1/\kappa, \quad k_{\min} = 1, 2 \tag{22}$$

and

$$\langle k\rangle = \begin{cases} \kappa/ln\kappa & , \quad k_{\min} = 1 \\ \kappa/(\ln\kappa - 1) & , \quad k_{\min} = 2 \end{cases} \tag{23}$$

Equation (22) is interesting since it shows explicitly that *not* the average number of links emanating from one node, $\langle k\rangle$, determines the epidemic threshold (as in the fully mixed SIR model (1)), but the typical maximum number of links κ emanating from a node.

Figure 7 shows $q_c(\langle k\rangle - 1)$ as a function of $\langle k\rangle$ for $\alpha = 1, 2, 3$ and $k_{\min} = 2$. In the fully mixed SIR model where each infective individual transmits the disease to the same average number of others as in the network model, $q_c(\langle k\rangle - 1) = 1$. The figure shows that for the distributions of k-values considered here, the epidemic thresholds obtained from (1) strongly overestimate the real thresholds.

A direct simulation of the epidemic outbreak in networks with $\beta = 2$, $k_{\min} = 1$ and $\kappa = 5, 10$, and 20 has been performed by Newman [10]. He found that the epidemic outbreak occurred at the same q_c as predicted by (20) and (21). Above q_c, the fraction of individuals P_∞ infected during the epidemic outbreak, increases linearly with increasing distance $q - q_c$ from the epidemic threshold, in the same way as in the Cayley tree.

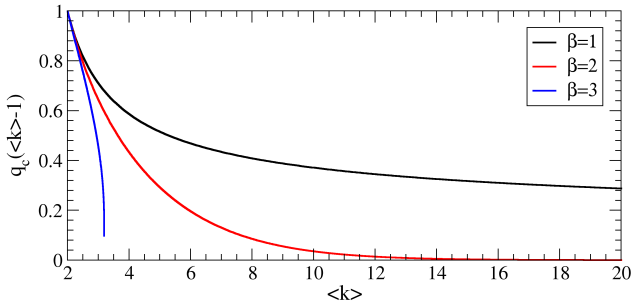

Fig. 7 (color online) Epidemic threshold q_c multiplied with $(\langle k\rangle - 1)$ as a function of $\langle k\rangle$ for $\alpha = 1$, 2, and 3 and $k_{\min} = 2$. The figure shows that the fully mixed SIR model where the effects of superspreaders are not taken into account and $q_c(\langle k\rangle - 1) = 1$, considerably overestimates the epidemic threshold and thus considerably underestimates the herd immunity.

6.2 Scale-free networks

In the limit of $\kappa \to \infty$, the distribution (20) reduces to

$$P(k) \propto k^{-\beta}, k = 1, 2, \cdots k_{\max}, \quad \beta > 1. \tag{24}$$

These scale-free networks have become very popular since they are believed to model a large number of networks in nature and technology [11]. Their maximum connectivity $k_{\max}$ can be estimated from $\int_{k_{\max}}^{\infty} P(k)dk = 1/N$ [16], yielding

$$k_{\max} \approx N^{1/(\beta-1)}. \tag{25}$$

For calculating $\langle k^2 \rangle$ and $\langle k \rangle$ for large N, one can substitute the relevant sums (see (18)) by an integral. This yields

$$\frac{\langle k^2 \rangle}{\langle k \rangle} = \frac{2-\beta}{3-\beta} \quad \frac{k_{\max}^{3-\beta} - 1}{k_{\max}^{2-\beta} - 1}, \tag{26}$$

and substituting (26) into (19) yields q_c. Accordingly, we can distinguish 3 β-regimes [16].

1. $\beta > 3$: $q_c = \beta - 3$ is finite.
2. $2 < \beta < 3$: $q_c \propto \frac{|3-\beta|}{|2-\beta|} N^{-\lambda}$ with $\lambda = (3-\beta)/(\beta-1) < 1$.
3. $1 < \beta < 2$: $q_c \propto \frac{|3-\beta|}{|2-\beta|} N^{-\lambda}$ with $\lambda = 1/(\beta-1) > 1$.

In regimes (2) and (3), q_c decreases with increasing number of nodes N. In both regimes a random immunization of the individuals is not useful since nearly every susceptible individual must be vaccinated to prevent the epidemic from spreading.

7 Networks with spatial constraints

Spatial constraints may play an important role in the spreading of an epidemic, when the individuals in a population are not fully mobile and thus cannot reach (and infect) any other individual with the same chance. In a highly mobile population, the spatial constraints are weak. Since the mobility of a population during an epidemic may be reduced by legal restrictions, it makes sense to study the influence of changing spatial constraints on the epidemic transition. For studies of the effects of legal traveling restrictions in Italy and China see [52, 53].

Typically, an infective individual has a larger preference to infect individuals close to him. In other words, two nodes in a social network are preferentially linked when the (Euclidean) distance r between them is small [38]. To model this, we place the nodes at the sites of a square lattice of length L, and connect pairs of nodes with a probability $p(r) \sim r^{-\delta}$ [39]. Including the normalization factor, $p(r)$ is

$$p(r) = \begin{cases} (2-\delta) L^{-(2-\delta)} r^{-\delta} & , \delta < 2 \\ (\delta - 2) r^{-\delta} & , \delta > 2 \ . \end{cases} \tag{27}$$

Figure 8 shows three Erdos-Renyi networks embedded in a square lattice, with $\langle k \rangle = 4$. In (a) the constraints are small ($\delta = 1$) and the underlying lattice seems irrelevant. In (c) the constraints are large ($\delta = 5$) and the network is nearly identical to a square lattice. (b) shows an intermediate case ($\delta = 3$) [55]. From $p(r)$ we obtain the mean link length $\langle r \rangle \propto \int_1^L dr r^{d-1} r p(r)$

and the maximum link length $r_{\max}$ defined by $\int_{k_{\max}}^{L} drr^{d-1}p(r) = 1/L^2$ where d is the dimension of the underlying lattice (here $d = 2$).

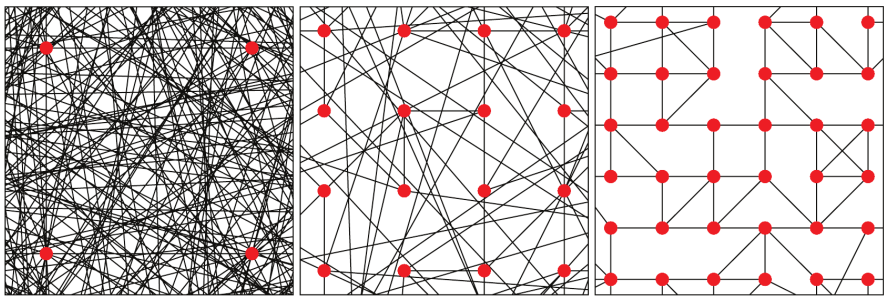

Fig. 8 (color online) Illustration of Erdos-Renyi networks with $\langle k \rangle = 4$ that are embedded in a square lattice, for (a) $\delta = 1$, (b) $\delta = 3$, and (c) $\delta = 5$. From [55].

For $\delta < 2$, both $\langle r \rangle$ and $r_{\max}$ are proportional to L. For δ between 2 and 4, $r_{\max}$ is still proportional to L but $\langle r \rangle$ is either proportional to $L^{3-\delta}$ (for δ between 2 and 3) or constant (for $\delta > 3$). For $\delta > 4$, $r_{\max}$ is proportional to $L^{2/(\delta-2)}$ such that $r_{\max}/L \to 0$ for $N \to \infty$ and $\langle r \rangle$ is constant. Thus we can distinguish 3 different δ-regimes:

1. $\delta < 2$: The spatial constraints are irrelevant. We expect the same values for the epidemic thresholds and the same critical behavior as for unconstrained networks.
2. $2 < \delta < 4$: Intermediate behavior: We expect that the epidemic thresholds increase with increasing constraint parameter δ.
3. $\delta > 4$: The spatial constraints are dominant and the epidemic threshold reaches its maximum value.

For the Erdos-Renyi networks demonstrated in Fig. 8, for example, we expect that $q_c = p_c \approx 1/4$ for $\delta < 2$ and $q_c \approx 0.5$, $p_c \approx 0.59$ for $\delta > 4$. In the δ-region in between, the critical points should increase continuously. Figure 9 confirms this behavior for p_c [40].

For scale-free networks, q_c and p_c should be unchanged for $\delta < 2$ and reach a constant value for $\delta > 4$. It has been shown in [55] that for $\beta < 3$, p_c still tends to zero with increasing system size N as $p_c \propto N^{-\tilde{\lambda}}$ but with an exponent $\tilde{\lambda}$ that decreases monotonically with δ, reaching 0 at $\delta = 4$.

Both examples show that the fraction of individuals v_c that need to be vaccinated, decreases when the spatial constraints are enhanced. Therefore, reducing the mobility is an effective weapon against an epidemic like COVID-19.

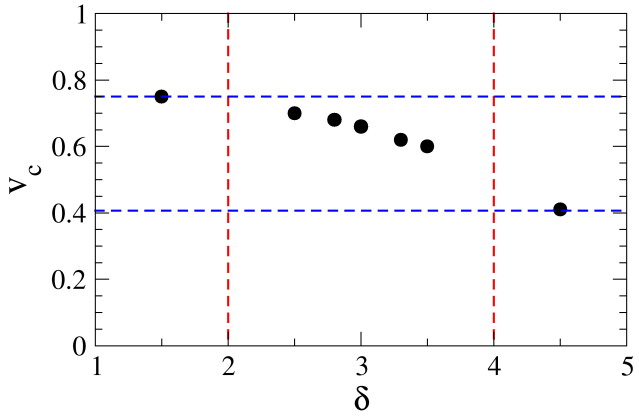

Fig. 9 (color online) Random immunization in the Erdos-Renyi networks shown in Fig. 8, under varying spatial constraints (Eq. (27)). The figure shows the minimum fraction of individuals $v_c = 1 - p_c$ that must be immunized to prevent epidemic spreading when the infection probability q equals 1. The vertical red lines at $\delta = 2$ and 4 separate the three regimes of weak, intermediate, and strong spatial constraints. The two blue lines mark the p_c-values for networks without spatial constraints (lower line) and for the square lattice (upper line), respectively. The data are from [34].

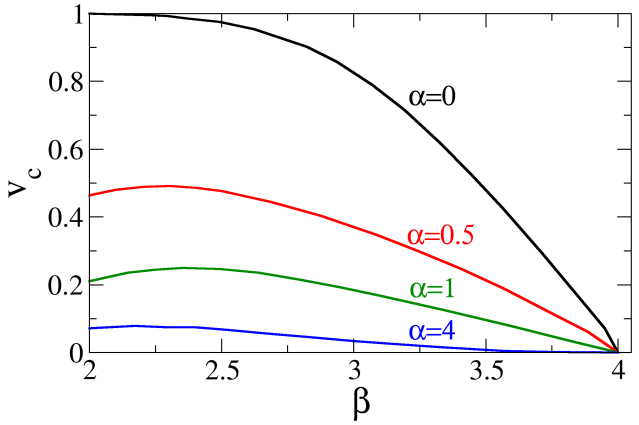

Fig. 10 (color online) Targeted immunization in scale-free networks with $N = 10^6$ nodes, with varying exponent β. The figure shows the minimum fraction of individuals $v_c = 1 - p_c$ that must be immunized to prevent epidemic spreading when the infection probability q equals 1, for different attack strategies with $\alpha = 0$ (random immunization), and $\alpha = 0.5, 1, 4$ (targeted immunization). The data are from [30].

8 Targeted immunization strategies

So far, we considered the effects of random vaccination on the spreading of an epidemic. We saw that when the number of contacts an individual has follows a broad distribution, nearly 100 percent of the population must be vaccinated in order to prevent an epidemic outbreak.

It has been shown by Cohen et al [18] that a far better strategy is a "targeted intentional attack" where first the superspreaders are tracked down rank-wise and then vaccinated. Following this strategy, only a comparatively small fraction of the population must be vaccinated. The problem is that not all superspreaders are known a priori and the question is how much do we need to know about the superspreaders to prevent epidemic outbreaks.

In a first attempt to deal with this problem, instead of deterministically choosing the individual with the largest number of links to be vaccinated, Gallos et al [30] introduced the probability

$$W(k_i) = \frac{k_i^\alpha}{\sum_{i=1}^{N} k_i^\alpha}, \quad 0 < \alpha < \infty, \tag{28}$$

for choosing an individual i with k_i contacts. The exponent α describes how effective the search in tracking down the superspreaders is. For $\alpha = 0$, the search is random and $W(k_i) = 1/N$. For $\alpha = \infty$, $W(k_{\max}) = 1$, i.e. always the best-connected superspreader is found.

To see, how effective this procedure is for general α, we consider again the extreme case $q = 1$, where an infective individual will infect all susceptible individuals contacted within the contagion time. Figure 10 shows the critical fraction of individuals v_c that must be vaccinated in order to prevent the epidemic from growing, as a function of β, for $\alpha = 0$, 0.5, 1, and 4, for networks of size $N = 10^6$. The data have been obtained both by numerical simulations and analytical calculations [30]. For $\beta = 4$, $p_c = 1$ and thus $v_c = 0$.

The black curve in Fig. 10 shows $v_c(\beta)$ for random vaccination, i.e., $\alpha = 0$. As expected, v_c approaches 1 for $\beta \to 2$. Remarkably, already for $\alpha = 0.5$, $v_c(\beta)$ is well below its value for random vaccination. Accordingly, even when the superspreaders are chosen with a rather small preference probability, only a finite fraction of individuals needs to be vaccinated. For $\alpha = 4$, $v_c(\beta)$ is the same as for the targeted attentional attack [18].

Efficient immunization strategies for the cases where the structure of the network is not known, have been also developed [27,32,35,56]. One approach is the acquaintance immunization [27] where one chooses individuals randomly and immunizes their friends. This can reduce the fraction of vaccinations required to stop the pandemic by a factor of 3. In another very efficient method called immunization with limited knowledge [56], one chooses randomly about 10 people at the time and immunizes the one that is most likely to have the highest number of contacts.

9 Conclusions

The COVID-19 epidemic is driven by contacts. When in a population each household is separated from the others for a certain period of time and the borders are closed, the epidemic will die out in this population. This "brute force" strategy is the most efficient strategy and has been used, for example, in China and New Zealand.

As suggested by the network models discussed here, a more refined alternative strategy is to focus on the superspreaders and superspreading events. A superspreader is everybody, who gets in close contact to a large number of people every day. Pupils and teachers are well-known superspreaders, and also medical doctors, healthcare workers, bus drivers, etc. A superspreading event is any event where a large number of people get close. Well-known superspreading events are mass events of all kinds like football events, large public concerts, theatrical performances, and worship services. Other superspreading events are large birthday and wedding parties and visits in crowded restaurants and bars, just to mention a few.

To contain a spreading epidemic and to prevent national health services to reach their limit, the first step is to deal with that fraction of superspreaders that is well known to the public and prevent them from transmitting the disease to a larger part of the population, either by vaccination or significant contact restrictions. The same holds for superspreading events: These events should either be prohibited or only take place on a considerably smaller, reduced-contact basis. As supplementary measures, mobility restrictions can be implemented. All these measures are particularly important and unavoidable as long as, like in COVID-19, vaccinations do not yet give full protection and recovered people can be infected again.

References

1. N.T.J. Bailey, *The Mathematical Theory of Infectious Diseases and its Applications* (Hafner Press, New York ,1975) [
2. R.M. Anderson, R.M. May, *Infectious Diseases of Humans* (Oxford University Press, Oxford, 1991)
3. H.W. Hethcote, SIAM Review 42, 599 (2000)
4. W.O. Kermack, A.G. McKendrick: Proc. R. Soc. Lond. A: Math. Phys. Eng. Sci. **115**, 700 (1927)
5. J.V. Noble, Nature **250**, 726 (1974)
6. M. Li, *An introduction to mathematical modeling of infectious diseases* (Springer, Berlin, 2018)
7. B. Maier, D. Brockmann, Science **368**, 742 (2020)
8. D. Brockmann, this volume
9. J.-P. Platteau, S. Weber, H. Wiesmeth, this volume

10. M.E.J. Newman, Phys. Rev. E **66**, 016128 (2002)
11. A.-L. Barabasi, *Linked: The New Science of Networks* (Perseus, Cambridge, MA, 2002)
12. M.E.J. Newman, *Networks* (Oxford University Press, Oxford, 2018)
13. L.M. Shekhtman, M.M. Danziger, S. Havlin, this volume
14. P. Grassberger, Math Biosci. **63**, 157 (1983)
15. C. Moore, M.E.J. Newman, Phys. Rev. E **61**, 5678 (2000)
16. R. Cohen, K. Erez, D. ben-Avraham, S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000)
17. D.S. Callaway, M.E.J. Newman, S.H. Strogatz, D.J. Watts, Phys. Rev. Lett. **85**, 5468 (2000)
18. R. Cohen, K. Erez, D. ben-Avraham, S. Havlin, Phys. Rev. Lett. **86**, 3682 (2001)
19. R. Pastor-Satorras, A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001)
20. F. Liljeros, C.R. Edling, L.A.N. Amaral, H.E. Stanley, Y. Aberg, Nature **411**, 907 (2001)
21. M. Kuperman, G. Abramson, Phys. Rev. Lett. **86**, 2909 (2001)
22. M.E.J. Newman, S.H. Strogatz, D. J. Watts, Phys. Rev. E **64**, 026118 (2001)
23. M.E.J. Newman, D.J. Watts, S.H. Strogatz, Proc. Natl. Acad. Sci. **99**, 2566 (2002)
24. R. Albert, A.-L. Barabasi, Rev. Mod. Phys. **74**, 47 (2002)
25. R. Pastor-Satorras, A. Vespignani, Phys. Rev. E **65**, 036104 (2002)
26. L. Sander, C.P. Warren, I.M. Sokolov, C. Simon, J. Koopman, Math. Biosci. **180**, 293 (2002)
27. R. Cohen, S. Havlin, D. ben Avraham, Phys. Rev. Lett. **91**, 247901 (2003)
28. S. Eubank, H. Guclu, V.S.A. Kumar, M.V. Marathe, A. Srinivasan, Z. Toroczkai, N. Wang, Nature **429**, 180 (2004)
29. N. Madar, T. Kalisky, R. Cohen, D. Ben-Avraham, S. Havlin, Eur. Phys. J. B **38**, 269 (2004)
30. L.K. Gallos, R. Cohen, P. Argyrakis, A. Bunde, S. Havlin, Phys. Rev. Lett. **94**, 188701 (2005)
31. V. Colizza, A. Barrat, M. Barthelemy, A. Vespignani, Proc. Natl. Acad. Sci. **103**, 2015 (2006)
32. L.K.Gallos, F. Liljeros, P. Argyrakis, A. Bunde, S. Havlin, Phys. Rev. E **75**, 045104 (R) (2007)
33. V. Colizza, A. Vespignani, J. Theor. Biol. **251**, 450 (2008)
34. K. Kosmidis, S. Havlin, A. Bunde, Europhys. Lett. **82**, 48005 (2008)
35. Y. Chen, G. Paul, S. Havlin, F. Liljeros, H.E. Stanley, Phys. Rev. Lett. **101**, 058701 (2008)
36. D. Balcan, V. Colizzac, B. Goncalves, H. Hu, J.J. Ramasco, A. Vespignani, Proc. Natl. Acad. Sci. **106**, 21484 (2009)
37. R. Cohen, S. Havlin, *Complex Networks* (Cambridge University Press, Cambridge, 2010)
38. M. Barthelemy, Phys. Rep. **499**, 1 (2010)
39. D. Li, K. Kosmidis, A. Bunde, S. Havlin, Nature Physics **7**, 481 (2011)
40. D. Li, G. Li, K. Kosmidis, H.E. Stanley, A. Bunde, S. Havlin, Europhys. Lett. **93** , 68004 (2011)
41. R. Pastor-Satorras, C. Castellano, P.V. Mieghem, A. Vespignani, Rev. Mod. Phys. **87**, 925 (2015)
42. B. Gross, S. Havlin, Applied Network Science **5** , 1 (2020)
43. D. Stauffer, A. Aharony, *Introduction to Percolation Theory* (Taylor & Francis, London, 1992)
44. A. Bunde, S. Havlin, *Fractals and Disordered Systems, 2nd ed.* (Springer, Berlin, 1995)
45. D. Li, B. Fu, Y. Wang, G. Lu, Y. Berezin, H.E. Stanley, S. Havlin, Proc. Natl. Acad. Sci. **112**, 669 (2015)

46. M.F. Sykes, J.W. Essam, J. Math. Phys. **5**, 1117 (1964)
47. S.K. Ma, *Modern Theory of Critical Phenomena* (Benjamin, Reading, 1976)
48. M. Yanuka, R. Englman, J. Phys. A: Math. Gen. **23**, L339 (1990)
49. P. Erdos, A. Renyi, Publ. Math. Debrecen **6**, 290 (1959)
50. P. Erdos, A. Renyi, Publ. Math. Inst. Hung. Acad. Sci. **5**, 17 (1960)
51. M. Molloy, B. Reed, Random. Struct. Algorithms **6**, 161 (1995)
52. B. Gross, Z. Zheng, S. Liu, X. Chen, A. Sela, J. Li, D. Li, S. Havlin, Europhys. Lett. **131**, 58003 (2020)
53. A. Smolynak, G. Bonaccorsi, A. Fiori, F. Pammoli, S. Havlin, Scientific Reports **11**, 21783 (2021)
54. T. Emmerich, A. Bunde, S. Havlin, Phys. Rev. E **86**, 046103 (2012).
55. T. Emmerich, A. Bunde, S. Havlin, G. Li, D. Li, Phys. Rev. E **87**, 032802 (2013)
56. Y. Liu, H. Sanhedrai, G.G. Dong, L.M. Shekhtman, F. Wang, S.V. Buldyrev, S. Havlin, National Science Review **8**, 229 (2021)