Finite-time scaling for epidemic processes with power-law superspreading events

Carles Falcó [1,2,*] and Álvaro Corral [1,2,3]

[1] *Centre de Recerca Matemàtica, Edifici C, Campus Bellaterra, E-08193 Barcelona, Spain*

[2] *Departament de Matemàtiques, Facultat de Ciències, Universitat Autònoma de Barcelona, E-08193 Barcelona, Spain*

[3] *Complexity Science Hub Vienna, Josefstädter Straße 39, 1080 Vienna, Austria*

(Received 4 November 2021; revised 2 March 2022; accepted 31 May 2022; published 21 June 2022)

Epidemics unfold by means of a spreading process from each infected individual to a variable number of secondary cases. It has been claimed that the so-called superspreading events of the COVID-19 pandemic are governed by a power-law-tailed distribution of secondary cases, with no finite variance. Using a continuous-time branching process, we demonstrate that for such power-law-tailed superspreading, the survival probability of an outbreak as a function of both time and the basic reproductive number fulfills a "finite-time scaling" law (analogous to finite-size scaling) with universal-like characteristics only dependent on the power-law exponent. This clearly shows how the phase transition separating a subcritical and a supercritical phase emerges in the infinite-time limit (analogous to the thermodynamic limit). We also quantify the counterintuitive hazards posed by this superspreading. When the expected number of infected individuals is computed removing extinct outbreaks, we find a constant value in the subcritical phase and a superlinear power-law growth in the critical phase.

DOI: 10.1103/PhysRevE.105.064122

I. INTRODUCTION

The ongoing COVID-19 pandemic has raised considerable concern over the superspreading phenomenon. In the propagation of infectious diseases, superspreading refers to when a single infected individual triggers a large number of secondary cases. Superspreading has been previously proposed to occur with diseases such as SARS [1], MERS [2], measles [3], and the Ebola virus disease. The name of the disease is Ebola virus disease [4]. Naturally, understanding superspreading is crucial not only for identifying which events drive the propagation, but also for implementing effective contention measures [5–7].

Most definitions of superspreading have been rather vague or arbitrary. For instance, some authors may define a superspreading event if a single individual provokes the direct contagion of at least 10 other individuals (secondary cases) [5,8]. Lloyd-Smith *et al.* [3] associated the phenomenon to the presence of outliers in the distribution of secondary cases when these are modeled in terms of a Poisson distribution (with a mean given by the empirically found value of the basic reproductive number R_0). Thus, an excess of outliers would suggest that the Poisson distribution is not appropriate, and a negative binomial is introduced instead [3,9,10] (this comprises Poisson as a particular case and arises as a mixture of Poisson secondary cases with γ-distributed rates for different individuals).

In other instances, superspreading has been associated with the 20/80 rule [11], in which the top 20% of most infective individuals are the direct cause of a very large percentage of direct transmission (e.g., 80%). But note that the fulfillment of the 20/80 rule, providing a single pair of numbers, is not sufficient to characterize a probability distribution. In concrete terms, although for precise values of its parameters the negative binomial distribution fulfills the rule, other distributions may be tuned to fulfill the rule as well (for instance, the power law [12]). In summary, the common approaches to superspreading identify it with a distribution of secondary cases that has a large variance [13] (or at least larger than that of a Poisson distribution).

Recent empirical observations of SARS-CoV and SARS-CoV-2 transmission showed that superspreading makes the tail of the distribution of secondary cases in these diseases incompatible with an exponential tail (which characterizes the negative binomial). Instead, the decay is consistent with a power-law tail [14] with an exponent γ in between 2 and 3 for the probability mass function p_k [15–17]. A fundamental difference between both types of distributions is that power-law-tailed distributions (with such a value of γ) cannot be characterized by their variance, which diverges. In this context, the mean, R_0, is of limited applicability, as a standard error cannot be associated with it and variability becomes infinite, making the value of R_0 difficult to constrain empirically and making extremal superspreading events probable occurrences. In an abuse of language, we will refer to "power-law superspreading" when dealing with power-law tails with an exponent in the range $2 < \gamma < 3$. Although the empirical evidence supporting power-law superspreading could be weak, we can speculate that the power-law scenario makes sense in light of both our knowledge of human social behavior and the airborne transmission of COVID-19 [18] (airborne

*Present address: Mathematical Institute, University of Oxford, OX2 6GG Oxford, United Kingdom.

transmission can skyrocket the number of secondary cases in poorly ventilated indoor spaces).

In this article, we explore the mathematical consequences of a power-law distribution of secondary cases, as claimed for COVID-19 in Ref. [14]. In this way, we show that a simple branching-process model teaches important lessons to understand spreading in infectious diseases and its degree of universality, in particular regarding power-law superspreading. The advantage of using a simple model is that it can be very transparent to show universal features, whereas more complicated models may be of little practical use if their parameters cannot be precisely constrained [19].

Although most used epidemic models are of the compartmental type (or are based on compartments) [13,20–22], branching processes are well-known in the field [3,23–26], and they are more convenient when dealing with stochasticity (stochasticity is fundamental when there is large variability, and superspreading is all about large variability), and when it is required to count the number of individual cases.

Also, branching processes can approximate more complicated stochastic models [27], and they are closely related to well-studied epidemic models on random networks [28,29]. Indeed, the equivalence between epidemic percolation networks and branching processes has been considered before [30], and it is known that under certain conditions both models predict the same probability of epidemic, outbreak size distribution, and epidemic threshold—at least during the initial spread of the disease [31]. Recently, different types of branching-process models have been applied to study COVID-19 [32–35]. Further, network models with a power-law distributed number of connections have been extensively studied, e.g., in Refs. [21,36,37].

First, we introduce the well-known continuous-time branching process; then, we study it for the infinite-variance case using a rather general family of power-law-tailed distributions. We find that a finite-time scaling law provides a universal description of power-law superspreading in terms of a unique scaling function that is independent of model parameters. The finite-time scaling illustrates how a continuous phase transition separating a subcritical and a supercritical phase only emerges in the limit of infinite time (which plays the role of a thermodynamic limit). Next, we compare with the case of finite-variance spreading, with some counterintuitive results arising in the comparison. Finally, we identify and quantify the hazard potentially arising from power-law superspreading.

II. PRELIMINARIES

We consider the age-dependent branching process with exponential lifetimes, also known as the continuous-time branching process [38]. At $t = 0$ an initial element is created. After an exponentially distributed lifetime, the element generates a random number k of offspring elements and is removed from the population. The new elements evolve in the same stochastic way, each with an identical and independent exponential distribution of lifetimes, with rate λ, and an identical and independent distribution of offspring, given by the probability p_k (with $k = 0, 1, \ldots$).

The branching-process assumption takes for granted a well-mixed and infinite susceptible population (thus, one only needs to care about infected individuals), as well as totally independent secondary cases. There are no sources of heterogeneity other than those coming from the offspring distribution. Note that the time dynamics given by the exponential lifetimes is different from the way time progression is incorporated in network epidemic models [39,40]. The higher-order structure of social interactions has been taken into account in recent models [41].

In the epidemic-spread analogy, elements are infected individuals, offspring are secondary cases, and the removal of individuals at the end of their lifetime corresponds either to recovery or death. The total number of cases (secondary and beyond) triggered by the initial infected individual will constitute an epidemic outbreak. In usual approaches, the offspring distribution p_k can be given by a Poisson distribution, but, as we have mentioned, the negative binomial has been used to account for superspreading [3]. In contrast, as was recently proposed, we will consider p_k as a power-law-tailed distribution [14].

The offspring distribution is characterized by its probability generating function (PGF), $f(s) = \sum_{k=0}^{\infty} p_k s^k$. The mean (expected number of secondary cases, which is the basic reproductive number) is obtained as $R_0 = \langle k \rangle = f'(s)|_{s=1}$ (the prime denotes derivative). The key (random) variable is $Z(t)$, which counts the number of infected individuals at time t. Its PGF $F(s,t)$ obeys

$$\frac{\partial F(s,t)}{\partial t} = \lambda[f(F(s,t)) - F(s,t)], \tag{1}$$

with initial condition $F(s,0) = s$ (at $t = 0$ there is one single element) [38].

Derivation of $F(s,t)$ with respect to s and taking $s = 1$ yields $\mu(t) = \langle Z(t) \rangle$, the expected number of elements (infected individuals) at t, fulfilling $d\mu(t)/dt = \lambda(R_0 - 1)\mu(t)$, with initial condition $\mu(0) = 1$. We will refer to $\mu(t)$ as the mean instantaneous size of the outbreak, or just the size. Straightforward integration leads to $\mu(t) = e^{(R_0-1)\lambda t}$, which is decreasing if $R_0 < 1$ and increasing if $R_0 > 1$. The case $R_0 = 1$ corresponds to the critical point (see below). It is remarkable that the offspring distribution has null influence on $\mu(t)$, except for its mean value R_0. In other words, superspreading effects, no matter how they are defined (from negative binomials or from power laws), do not change the behavior, as long as R_0 takes the required value.

The reason for this is that the mean number of infections does not tell the whole story (only an averaged story). To proceed, we need to calculate $\eta(t)$, the probability that the outbreak is extinct at time t, i.e., the probability of $Z(t) = 0$. As $\eta(t) = F(0,t)$, we only need to take $s = 0$ in the equation for $F(s,t)$, which leads to

$$\frac{d\eta(t)}{dt} = \lambda[f(\eta(t)) - \eta(t)], \tag{2}$$

with $\eta(0) = 0$ [38]. As in the Galton-Watson (discrete-time) model [38,42,43], the equation has a stable fixed-point solution, η^*, fulfilling $\eta^* = f(\eta^*)$, and $f'(\eta^*) \leqslant 1$. Note that in the equation for $\eta(t)$, the offspring PGF appears explicitly.

III. POWER-LAW-TAILED OFFSPRING DISTRIBUTIONS

Although different definitions have been proposed [44], one can simply consider power-law-tailed (PLT) distributions as those that behave asymptotically as a power law, i.e., $p_k k^\gamma \xrightarrow[k\to\infty]{} c$ for $c > 0$ and an exponent $2 < \gamma < 3$ ensuring infinite variance. Power-law-tailed distributions defined in this way constitute a subclass of the so-called regularly varying distributions (or, roughly speaking, fat-tailed distributions, which, in extreme-value theory, correspond to the so-called Fréchet maximum domain of attraction [44]). See the Appendix B for other possibilities.

To find the expansion of the PGF $f_{\rm plt}(s)$ of p_k, we look at $f''_{\rm plt}(s) = \sum_{k=0}^{\infty} k(k-1)p_k s^{k-2}$. Note that $f''_{\rm plt}(s)$ is well-defined for $0 \leqslant s < 1$ and diverges as $s \to 1$ (divergence of the second moment), and also $k(k-1)p_k \sim k^{2-\gamma}$ for large k. By an Abelian theorem [45] (applicable when $\gamma - 2 < 1$), $f''_{\rm plt}(s)$ behaves as $c\Gamma(3-\gamma)/(1-s)^{3-\gamma}$ near $s = 1$. Integrating twice and using that $f_{\rm plt}(1) = 1$ and $f'_{\rm plt}(1) = R_0$, we can write $f_{\rm plt}(1-\epsilon) \approx 1 - R_0\epsilon + c\Gamma(1-\gamma)\epsilon^{\gamma-1}$ for small ϵ.

Now we are able to find the probability of extinction from Eq. (2) when this is close to 1. Let us introduce the survival probability of the outbreak at time t, which is $q(t) = 1 - \eta(t)$. Notice that the survival probability is the survivor function of the outbreak lifetime i.e., a complementary cumulative distribution function, but referring to outbreaks, not individuals, and thus $\eta(t)$ is the corresponding cumulative distribution function.

For long times, and close to the critical point (which separates sure extinction for $R_0 \leqslant 1$ from a small probability of survival for $R_0 > 1$), $\eta(t)$ will be close to 1 and $q(t)$ will be close to 0. So, we will be able to apply in Eq. (2) the previous expansion of the PGF around $\eta(t) = 1$ [i.e., $q(t) = 0$] to get

$$\frac{dq(t)}{dt} = \lambda[(R_0 - 1)q(t) - c\Gamma(1-\gamma)q(t)^{\gamma-1}],$$

disregarding terms $O(q(t)^2)$. As we cannot apply the original initial condition, because the equation is not valid for short times, we substitute it for $q(t_0) = q_0$, with q_0 unknown. The resulting solution is [46]

$$q(t) = \left(\frac{e^{(\gamma-2)(R_0-1)\lambda\Delta t}(R_0-1)/[c\Gamma(1-\gamma)]}{e^{(\gamma-2)(R_0-1)\lambda\Delta t} - 1 + q_0^{2-\gamma}(R_0-1)/[c\Gamma(1-\gamma)]} \right)^{\frac{1}{\gamma-2}} \tag{3}$$

with $\Delta t = t - t_0$.

IV. FINITE-TIME SCALING

Close to the critical point, the solution verifies a finite-time scaling law (analogous to finite-size scaling replacing system size by time [47,48]). Defining the rescaled variable

$$z = (\gamma - 2)(R_0 - 1)\lambda t \tag{4}$$

with $t - t_0 \simeq t$, and disregarding the last term in the denominator of Eq. (3) (which can be done close to the critical point, equivalent to long times when z is finite), we can write

$$q(t) = \left(\frac{1}{(\gamma-2)c\Gamma(1-\gamma)\lambda t} \right)^{\frac{1}{\gamma-2}} G_\gamma(z) \propto \frac{G_\gamma(z)}{t^{1/(\gamma-2)}} \tag{5}$$

with the γ-dependent scaling function

$$G_\gamma(z) = \left(\frac{ze^z}{e^z - 1} \right)^{\frac{1}{\gamma-2}}, \tag{6}$$

and where the dependence on the unknown initial condition has disappeared. Note that the new variable z, Eq. (4), absorbs in a rescaled way both the temporal dependence and the distance to the critical point (thus, in the forthcoming equations, time dependence is included both in t and z).

Therefore, for a fixed exponent γ, displaying $(c\lambda t)^{1/(\gamma-2)}q(t)$ versus z yields a unique z-dependent curve independent of λ, t, R_0, and any other parameter of the offspring distribution (as long as z is kept constant). Further, for different values of γ, displaying $[(\gamma-2)c\Gamma(1-\gamma)\lambda t]q(t)^{\gamma-2}$ versus z the curve becomes additionally independent of γ, and therefore "universal," with γ-independent scaling function $G(z) = [G_\gamma(z)]^{\gamma-2}$. The universal γ-independent scaling law is

$$q(t)^{\gamma-2} = \frac{1}{(\gamma-2)c\Gamma(1-\gamma)\lambda t} G(z) \propto \frac{1}{t} G(z). \tag{7}$$

The data collapses in Fig. 1, obtained from computer simulations, show how these finite-time scalings work.

The limiting behavior of the scaling function is

$$G_\gamma(z) \to \begin{cases} (|z|e^{-|z|})^{1/(\gamma-2)} & \text{for } z \to -\infty, \\ 1 & \text{for } z = 0, \\ z^{1/(\gamma-2)} & \text{for } z \to \infty \end{cases} \tag{8}$$

(it can be interesting to compare the resulting exponential decay for $q(t)$ in the subcritical phase with the empirical findings of Ref. [36]). Using this limiting behavior in the scaling law, the asymptotics of $q(t)$ (limit $t \to \infty$, close to the critical point) becomes

$$q(t) \xrightarrow[t\to\infty]{} \begin{cases} 0 & \text{for } R_0 < 1, \\ [(\gamma-2)c\Gamma(1-\gamma)\lambda t]^{-1/(\gamma-2)} & \text{for } R_0 = 1, \\ \{(R_0-1)/[c\Gamma(1-\gamma)]\}^{1/(\gamma-2)} & \text{for } R_0 > 1. \end{cases} \tag{9}$$

This change of behavior at the critical value $R_0 = 1$ can be understood as a phase transition (and $R_0 = 1$ as a critical point), with R_0 the control parameter and $\lim_{t\to\infty} q(t)$ the order parameter, and with the asymptotic limit playing the role of the thermodynamic limit (infinite-system-size limit). This shows how the phase transition emerges when $t \to \infty$. As $2 < \gamma < 3$, the order-parameter exponent $\beta = 1/(\gamma-2)$ is in the range $1 < \beta < \infty$ and the transition is not sharp but continuous with a continuous first derivative at $R_0 = 1$. So, the order of the transition is higher than second (in contrast to the finite-variance case; see below). The result for the critical phase is in agreement with Ref. [49] for a discrete-time branching process. An equivalent result to the one for the supercritical phase is known in the context of percolation in scale-free networks [50].

The close-to-critical regime, for which our results are valid, can be of great practical interest, as spontaneous changes in human behavior and implementation of contention measures usually lead to a decrease in the value of R [51,52] (the equivalent of R_0 when this changes its value).

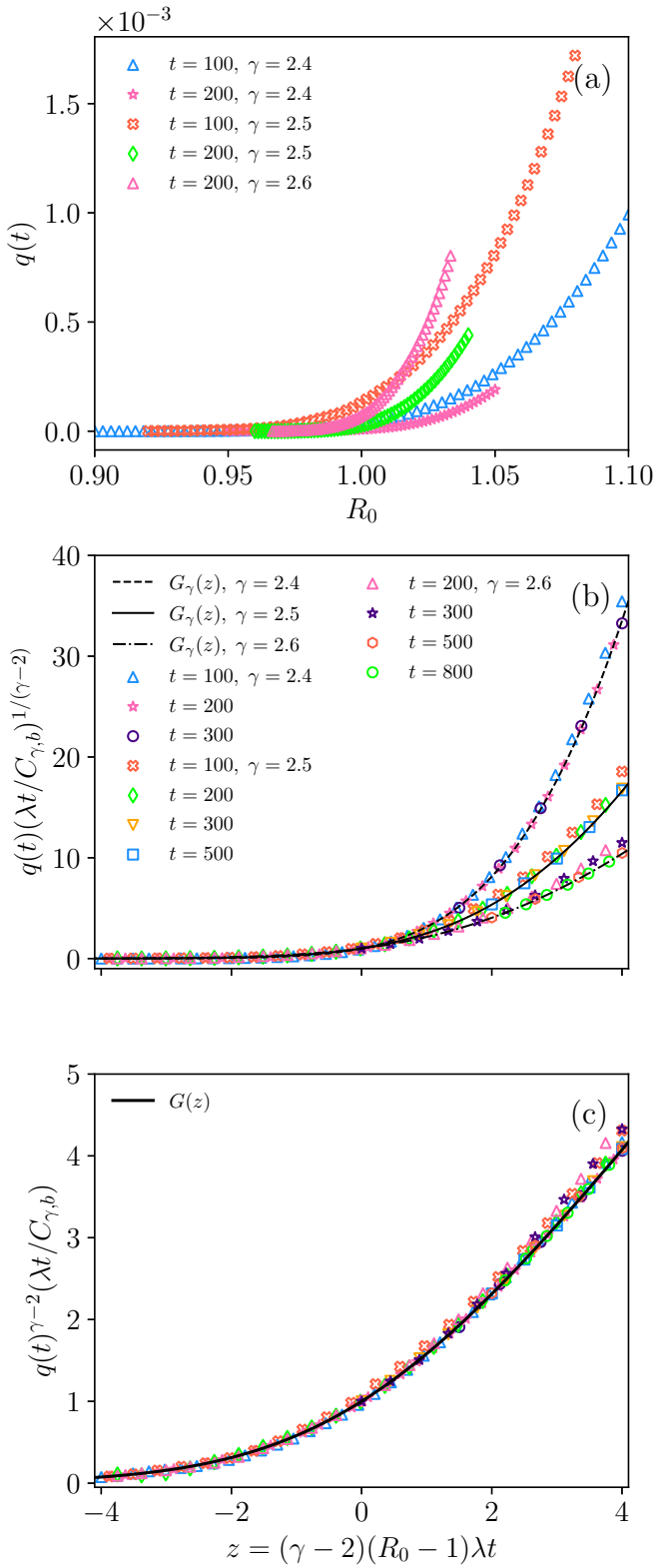

FIG. 1. Results of computer simulations of the continuous-time branching process with offspring distribution given by the shifted power-law distribution (see the Appendixes), for different values of b (to sweep R_0) and different values of γ and t, with $\lambda = 1$. (a) Survival probabilities $q(t)$ versus R_0. (b) Simple rescaling of $q(t)$ versus the rescaled distance to the critical point z. (c) General rescaling of $q(t)^{\gamma-2}$.

V. COMPARISON WITH THE FINITE-VARIANCE CASE

The result for the case of finite variance [53] (Poisson, negative binomial, etc., but also power-law tail with $\gamma > 3$) can be considered included in the previous expressions. Indeed, taking Eqs. (4), (6), (5), and (9) and replacing $\gamma - 2$ by 1 and $c\Gamma(1-\gamma)$ by $\sigma^2/2$, with σ^2 the variance of the offspring distribution in the critical point, one recovers the formulas for the finite-variance case [53].

Thus, the power-law behavior of the order parameter as a function of R_0 in the case of infinite variance, Eq. (9), translates into a linear function in the case of finite variance, i.e., $q(t) \xrightarrow[t\to\infty]{} 2(R_0 - 1)/\sigma^2$, for $R_0 > 1$. This highlights the importance of determining not only the mean R_0 of the offspring distribution, but also its variance (when it is finite [10]). The problem with using the Poisson distribution for offspring is that the variance is equal to R_0 and, close to the critical point, both are close to 1. But there is nothing special with regard to the negative binomial, apart from allowing a variance different from R_0; any distribution with the same variance and R_0 would lead not only to the same asymptotic solution for $q(t)$ but to the same finite-time scaling law [53]. In other words, superspreading with finite variance does not lead to any new phenomenology. It is only for power-law superspreading (with infinite variance) that superspreading becomes a new phenomenon, in the sense that new universality classes arise.

The previous simple expression for the limit of $q(t)$ in the finite-variance case [together with $q(t) \to 0$ for $R_0 \leqslant 1$] corresponds to the usual transcritical bifurcation [54]. Nevertheless, the power-law case with $2 < \gamma < 3$ also corresponds to a transcritical bifurcation, despite the fact that the behavior in the supercritical phase is not linear. Comparing, for the same values of R_0, the supercritical phases for finite and infinite variances, one can see that, sufficiently close to the critical point, the linear term is above the nonlinear one, and therefore the probability $q(t)$ that an outbreak does not become extinct is smaller if there is power-law superspreading (in fact, this probability is zero at first order in $R_0 - 1$, in comparison with the finite-variance case). Thus, power-law superspreading makes extinction of the outbreaks easier.

We can quantify the differences in the outbreak lifetimes t. This is a random variable with survivor function $q(t)$. Although we have only calculated the tail of $q(t)$, this is enough to characterize the expected lifetime $\langle t \rangle$ of an outbreak. From Eqs. (8) and (9), it is clear that in the infinite-variance case $\langle t \rangle$ is finite for $R_0 < 1$ [because $q(t)$ decays exponentially] and infinite for $R_0 > 1$ [because $q(t)$ does not tend to zero, and therefore it has a nonzero mass at infinity]. This is valid also for finite-variance offspring distributions.

The qualitative behavior at the critical point $R_0 = 1$ is different and counterintuitive. In both cases, we have critical slowing down (power-law decay in time), but in the finite-variance case $q(t) \sim 1/t$, which means that the power-law exponent of the density is 2 and $\langle t \rangle$ diverges, whereas for infinite variance, $q(t) \sim 1/t^{1/(\gamma-2)}$, leading to an exponent of the density larger than 2 and therefore to a finite mean value $\langle t \rangle$. In other words, in the critical phase, spreading with finite variance leads (despite the probability of extinction being 1) to never-ending outbreaks (in expected value, not in single

realizations), but infinite-variance superspreading reduces the expected lifetime to be finite.

In any case, power-law-tailed outbreak lifetimes (or total outbreak sizes [8,51,55,56]) are not an indication of power-law superspreading, as in the critical point power laws arise with any sort of spreading, whereas outside the critical point power-law lifetimes do not take place, regardless of the spreading. It is important then not to confuse these two different types of power laws. And of course, the occurrence of large outbreaks is not an indication of superspreading (they may arise even for the Poisson distribution if $R_0 \geqslant 0$).

VI. HAZARD FROM POWER-LAW SUPERSPREADING

We have seen that the expected number of infected individuals varies exponentially as $\mu(t) = e^{(R_0-1)\lambda t}$, independently of the spreading characteristics, but the survival probability of an outbreak is decreased when there is power-law superspreading. Which are the hazards coming from this, then? Obviously, $\mu(t)$ is not highly informative as it contains the contribution from outbreaks that have become extinct (and contribute with a value of zero, but are counted).

In a formula, $\mu(t) = q(t)I_{\rm sur}(t) + \eta(t) \times 0$, with $I_{\rm sur}(t)$ the expected number of infected individuals for outbreaks that are not extinct at time t [note that $I_{\rm sur}(t)$ is an average between infected individuals, in contrast to $\mu(t)$]; therefore, $I_{\rm sur}(t) = \mu(t)/q(t)$, and the decrease in $q(t)$ for power-law superspreading will yield an increase in $I_{\rm sur}(t)$. In concrete terms, substituting the scaling law for $q(t)$ [Eq. (5)] we get another finite-time scaling law,

$$I_{\rm sur}(t) = \left[c\Gamma(1-\gamma)(\gamma-2)\lambda t\left(\frac{e^z-1}{z}\right)\right]^{\frac{1}{\gamma-2}} \xrightarrow[t\to\infty]{} \begin{cases} \left(\frac{c\Gamma(1-\gamma)}{1-R_0}\right)^{1/(\gamma-2)} & \text{for } R_0 < 1, \\ [(\gamma-2)c\Gamma(1-\gamma)\lambda t]^{1/(\gamma-2)} & \text{for } R_0 = 1, \\ \left(\frac{c\Gamma(1-\gamma)}{R_0-1}\right)^{1/(\gamma-2)} e^{(R_0-1)\lambda t} & \text{for } R_0 > 1 \end{cases} \tag{10}$$

[the case of finite variance is recovered with the substitutions $c\Gamma(1-\gamma) \to \sigma^2/2$ and $\gamma - 2 \to 1$].

These results mean that, in the subcritical phase, the very few outbreaks that survive reach a fixed average (instantaneous) size $I_{\rm sur}$ (while they survive [57]), with a higher $I_{\rm sur}$ in the case of power-law superspreading (in comparison with the finite-variance case). In contrast, in the supercritical phase the nonextinct outbreaks grow exponentially (with a prefactor that can be very high when γ is close to 2). It is at the critical point that one finds an important qualitative difference between the finite-variance case and the power-law case: in the former case, the average size of the outbreaks that survive diverges linearly, but for power-law superspreading the growth is superlinear (as a power law of t with exponent larger than 1). We note then a trivial yet important observational bias, due to the fact that, at time t, we only see the outbreaks that have not become extinct. This is a dramatic realization of the survivorship bias (where survivorship refers to the outbreak, not to the individuals).

VII. DISCUSSION

We have made clear how a continuous-time branching process with power-law-tailed secondary cases (in correspondence with recent observational results describing superspreading in COVID-19 [14]) has properties that are qualitatively different from the case of finite variance. The latter constitutes a well-known mean-field universality class with order-parameter exponent $\beta = 1$ [58], whereas the power-law superspreading leads to a continuum of universality classes depending on the value of the secondary-case power-law exponent γ [59]. Further, we derive the existence of a finite-time scaling law describing the probability of outbreak survival as a function of R_0 and time, Eqs. (5) and (7), and we calculate the exact value of the scaling functions, Eq. (6). These scaling laws could be extended to random networks. We also show the peculiar behavior of $I_{\rm sur}(t)$, Eq. (10).

It would be desirable to apply our results to the COVID-19 pandemic in order to obtain the probability of outbreak extinction after some time as a function of R_0, for which the offspring power-law exponent γ has been estimated. However, in addition to the exponent γ, knowledge of the constant c in the asymptotic power-law formula is also fundamental (a relation between c and R_0 exists, but it is model-dependent). In other words, it is not enough to know the distribution of secondary cases for large k [14], but one needs to know the whole "population" to which those large outbreaks belong. Thus, concentrating only on large outbreaks is useless for the calculation of the survival probability.

ACKNOWLEDGMENTS

We appreciate feedback from M. Boguñá, L. Lacasa, and R. Pastor-Satorras. Support from projects PGC-FIS2018-099629-B-I00 from the Spanish MICINN, CEX2020-001084-M from the Spanish State Research Agency, through the Severo Ochoa and María de Maeztu Program for Centers and Units of Excellence in R&D, as well as the CERCA Programme/Generalitat de Catalunya, is acknowledged.

APPENDIX A: SHIFTED POWER-LAW OFFSPRING DISTRIBUTION

One particular case of a power-law-tailed distribution is given by the shifted power-law distribution. This is given by

$$p_k = \frac{1}{\zeta(\gamma, b)(b+k)^\gamma} \tag{A1}$$

for $k = 0, 1, \ldots$ with γ the exponent, b a location parameter, and $\zeta(\gamma, b)$ the Hurwitz zeta function, ensuring normalization. $\gamma > 2$ leads to a finite mean and $\gamma < 3$ leads to an infinite variance. This is the range we consider, together with $b > 0$. The mean of the distribution can be calculated directly to be $R_0 = \zeta(\gamma-1, b)/\zeta(\gamma, b) - b$. The case $b = 0$, truncated from below at $k = 1$, would lead to the standard discrete power law (straight in a log-log plot), whereas $b > 0$ leads to a shifted discrete power law. The shift does not alter the power-law behavior at the tail, i.e., $p_k \sim 1/k^\gamma$, for any $b > 0$ and large k.

Using the notation in the main text, in this case we have $c = 1/\zeta(\gamma, b)$. However, this simple case admits an

alternative but equivalent analysis. The PGF of the shifted power law (pl) is $f_{\rm pl}(s) = \Phi(s,\gamma,b)/\zeta(\gamma,b)$, with $\Phi(s,\gamma,b) = \sum_{k=0}^{\infty} s^k/(b+k)^\gamma$ the so-called Lerch transcendent and $\zeta(\gamma,b) = \Phi(1,\gamma,b)$. When $b=1$ one has $\zeta(\gamma,1) = \zeta(\gamma)$ (the Riemann zeta function) and $\Phi(s,\gamma,1) = \mathrm{Li}_\gamma(s)/s$, with $\mathrm{Li}_\gamma(s)$ the polylogarithm, arising in integrals that appear in the study of the Bose-Einstein condensation and whose asymptotic behavior near $s=1$ is well known. A generalization of this is possible [60], yielding

$$\Phi(s,\gamma,b) = \frac{1}{s^b}[\Gamma(1-\gamma)(-\ln s)^{\gamma-1} + \zeta(\gamma,b) + \zeta(\gamma-1,b)\ln s + O(\ln^2 s)],$$

valid for $|\ln s| < 2\pi$, with $b>0$ and γ a positive noninteger; $\Gamma()$ is the Gamma function. Since we are interested in s close to but smaller than 1, we can write $s = 1-\epsilon$, and then $s^{-b} \simeq 1 + b\epsilon$ and $\ln s \simeq -\epsilon$; thus

$$f_{\rm pl}(1-\epsilon) = \frac{\Phi(1-\epsilon,\gamma,b)}{\zeta(\gamma,b)} = 1 - R_0\epsilon + \frac{\Gamma(1-\gamma)}{\zeta(\gamma,b)}\epsilon^{\gamma-1} + O(\epsilon^2),$$

where we have assumed the range of interest, $2<\gamma<3$. The rest of the calculation is identical to that in the main text, just replacing c by $1/\zeta(\gamma,b)$.

APPENDIX B: ALTERNATIVE POWER-LAW TAIL

Our results also hold for distributions p_k that have similar behavior to power-law-tailed distributions, but they do not satisfy the condition $p_k k^\gamma \xrightarrow[k\to\infty]{} c$. For instance, one could consider distributions satisfying the alternative condition

$$\sum_{k=0}^{\infty} k^2\left(p_k - \frac{c}{(k+1)^\gamma}\right) < \infty \tag{B1}$$

for a given real positive constant c. Note that this condition is in fact different from the one used in the main text, since there are distributions satisfying this condition that do not tend in the limit to a power law; e.g., any distribution obeying $p_k \sim (1+(-1)^k/2)k^{-\gamma}$ for large k (the sum can be shown to converge using Leibniz's criterion for alternating series). Conversely, there are also distributions for which the limit of $p_k k^\gamma$ for $k\to\infty$ exists, but the condition above is not satisfied [e.g., when $p_k \sim k^{-\gamma}(1+1/\ln k)$ for large k, with $\gamma<3$].

To find the expansion of the PGF $f_{\rm plt}(s)$ of p_k when p_k verifies the condition given by Eq. (B1), we define $g(s) = f_{\rm plt}(s) - c\zeta(\gamma) f_{\rm pl1}(s)$ with $\zeta(\gamma)$ the Riemann zeta function and $f_{\rm pl1}(s)$ the PGF of the shifted power law with $b=1$. The condition above guarantees the existence of $g'(1)$ and $g''(1)$, and then $g(1) = 1 - c\zeta(\gamma)$ and $g'(1) = R_0 - c\zeta(\gamma)R_{\rm pl1}$, with R_0 the mean of p_k, and $R_{\rm pl1}$ the mean of the shifted power law. In this way, we can write $f_{\rm plt}(1-\epsilon) = g(1-\epsilon) + c\zeta(\gamma) f_{\rm pl1}(1-\epsilon) = 1 - R_0\epsilon + c\Gamma(1-\gamma)\epsilon^{\gamma-1} + O(\epsilon^2)$, which is identical to the PGF derived in the main text (although the range of validity is different).

[1] Z. Shen, F. Ning, W. Zhou, X. He, Ch. Lin, D. Chin, Z. Zhu, and A. Schuchat, Superspreading SARS events, Beijing, 2003, Emerg. Infect. Dis. **10**, 256 (2004).

[2] A. Kucharski and C. Althaus, The role of superspreading in Middle East respiratory syndrome coronavirus (MERS-CoV) transmission, Euro Surveillance **20**, 14 (2015).

[3] J. Lloyd-Smith, S. Schreiber, P. E. Kopp, and W. Getz, Superspreading and the effect of individual variation on disease emergence, Nature (London) **438**, 355 (2005).

[4] M. S. Y. Lau, B. D. Dalziel, S. Funk, A. McClelland, A. Tiffany, S. Riley, C. J. E. Metcalf, and B. T. Grenfell, Spatial and temporal dynamics of superspreading events in the 2014–2015 West Africa Ebola epidemic, Proc. Natl. Acad. Sci. (USA) **114**, 2337 (2017).

[5] D. Lewis, The superspreading problem, Nature (London) **590**, 544 (2021).

[6] B. F. Nielsen, L. Simonsen, and K. Sneppen, COVID-19 Superspreading Suggests Mitigation by Social Network Modulation, Phys. Rev. Lett. **126**, 118301 (2021).

[7] A. Vespignani, H. Tian, C. Dye, J. Lloyd-Smith, R. M. Eggo, M. Shrestha, S. V. Scarpino, B. Gutierrez, M. U. G. Kraemer, J. Wu, K. Leung, and G. M. Leung, Modelling COVID-19, Nat. Rev. Phys. **2**, 279 (2020).

[8] A. Kucharski, *The Rules of Contagion* (Basic Books, London, 2020).

[9] A. Endo, S. Abbott, A. Kucharski, and S. Funk, Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China, Wellcome Open Res. **5**, 67 (2020).

[10] L. Hébert-Dufresne, B. M. Althouse, S. V. Scarpino, and A. Allard, Beyond R_0: heterogeneity in secondary infections and probabilistic epidemic forecasting, J. R. Soc. Interface **17**, 20200393 (2020).

[11] A. Galvani and R. May, Epidemiology: Dimensions of superspreading, Nature (London) **438**, 293 (2005).

[12] M. E. J. Newman, Power laws, Pareto distributions and Zipf's law, Contemp. Phys. **46**, 323 (2005).

[13] A. L. Hill, The math behind epidemics, Phys. Today **73**, 28 (2020).

[14] F. Wong and J. J. Collins, Evidence that coronavirus superspreading is fat-tailed, Proc. Natl. Acad. Sci. (USA) **117**, 29416 (2020).

[15] For other datasets, better fits have been found by using Poisson mixtures as distributions of secondary cases [16].

[16] C. Kremer, A. Torneri, S. Boesmans, H. Meuwissen, S. Verdonschot, K. Vanden Driessche, C. Althaus, C. Faes, and N. Hens, Quantifying superspreading for COVID-19 using Poisson mixture distributions, Sci. Rep. **11**, 14107 (2021).

[17] Reference [14] reports an exponent between 1 and 2 because it refers to the exponent $\gamma-1$ of the complementary cumulative distribution function.

[18] T. Greenhalgh, J. L. Jimenez, K. A. Prather, Z. Tufekci, D. Fisman, and R. Schooley, Ten scientific reasons in support of airborne transmission of SARS-CoV-2, The Lancet **397**, 1603 (2021).

[19] 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).

[20] L. A. Meyers, Contact network epidemiology: Bond percolation applied to infectious disease prediction and control, Bull. Am. Math. Soc. **44**, 63 (2007).

[21] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, Epidemic processes in complex networks, Rev. Mod. Phys. **87**, 925 (2015).

[22] 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).

[23] M. Lipsitch, T. Cohen, B. Cooper, J. M. Robins, S. Ma, L. James, G. Gopalakrishna, S. K. Chew, C. C. Tan, M. H. Samore, D. Fisman, and M. Murray, Transmission dynamics and control of severe acute respiratory syndrome, Science **300**, 1966 (2003).

[24] D. Easley and J. Kleinberg, *Networks, Crowds, and Markets* (Cambridge University Press, Cambridge, 2010).

[25] J. C. Miller, A primer on the use of probability generating functions in infectious disease modeling, Infectious Disease Modell. **3**, 192 (2018).

[26] J. Pausch, R. Garcia-Millan, and G. Pruessner, Noise can lead to exponential epidemic spreading despite R_0 below one, arXiv:2109.00437.

[27] L. J. S. Allen, A primer on stochastic epidemic models: Formulation, numerical simulation, and analysis, Infectious Disease Modell. **2**, 128 (2017).

[28] E. Kenah and J. Robins, Second look at the spread of epidemics on networks, Phys. Rev. E **76**, 036113 (2007).

[29] S. Bhamidi, D. Nam, O. Nguyen, and A. Sly, Survival and extinction of epidemics on random graphs with general degree, Ann. Probab. **49**, 244 (2021).

[30] E. Kenah and J. Miller, Epidemic percolation networks, epidemic outcomes, and interventions, Interdisciplinary Perspectives on Infectious Diseases **2011**, 1 (2011).

[31] E. Kenah and J. M. Robins, Network-based analysis of stochastic SIR epidemic models with random and proportionate mixing, J. Theor. Biol. **249**, 706 (2007).

[32] J. Riou and C. L. Althaus, Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020, Euro Surveill. **25**, 2000058 (2020).

[33] A. L. Bertozzi, E. Franco, G. Mohler, M. B. Short, and D. Sledge, The challenges of modeling and forecasting the spread of COVID-19, Proc. Natl. Acad. Sci. (USA) **117**, 16732 (2020).

[34] J. Levesque, D. W. Maybury, and R. H. A. D. Shaw, A model of COVID-19 propagation based on a gamma subordinated negative binomial branching process, J. Theor. Biol. **512**, 110536 (2021).

[35] L. Zhang, H. Wang, Z. Liu, X. F. Liu, X. Feng, and Y. Wu, A heterogeneous branching process with immigration modeling for COVID-19 spreading in local communities in China, Complexity **2021**, 6686547 (2021).

[36] R. Pastor-Satorras and A. Vespignani, Epidemic Spreading in Scale-Free Networks, Phys. Rev. Lett. **86**, 3200 (2001).

[37] C. Castellano and R. Pastor-Satorras, Thresholds for Epidemic Spreading in Networks, Phys. Rev. Lett. **105**, 218701 (2010).

[38] M. Kimmel and D. E. Axelrod, *Branching Processes in Biology* (Springer-Verlag, New York, 2002).

[39] P.-A. Noël, B. Davoudi, R. C. Brunham, L. J. Dubé, and B. Pourbohloul, Time evolution of epidemic disease on finite and infinite networks, Phys. Rev. E **79**, 026101 (2009).

[40] A. J. Allen, M. C. Boudreau, N. J. Roberts, A. Allard, and L. Hébert-Dufresne, Predicting the diversity of early epidemic spread on networks, Phys. Rev. Res. **4**, 013123 (2022).

[41] G. St-Onge, H. Sun, A. Allard, L. Hébert-Dufresne, and G. Bianconi, Universal Nonlinear Infection Kernel from Heterogeneous Exposure on Higher-Order Networks, Phys. Rev. Lett. **127**, 158301 (2021).

[42] T. E. Harris, *The Theory of Branching Processes* (Springer, Berlin, 1963).

[43] A. Corral and F. Font-Clos, Criticality and self-organization in branching processes: application to natural hazards, in *Self-Organized Criticality Systems*, edited by M. Aschwanden (Open Academic, Berlin, 2013), pp. 183–228.

[44] I. Voitalov, P. van der Hoorn, R. van der Hofstad, and D. Krioukov, Scale-free networks well done, Phys. Rev. Res. **1**, 033034 (2019).

[45] G. H. Weiss, *Aspects and Applications of the Random Walk* (North Holland, Amsterdam, 1994).

[46] WolframAlpha, https://www.wolframalpha.com.

[47] R. Garcia-Millan, F. Font-Clos, and A. Corral, Finite-size scaling of survival probability in branching processes, Phys. Rev. E **91**, 042122 (2015).

[48] A. Corral, R. Garcia-Millan, and F. Font-Clos, Exact derivation of a finite-size scaling law and corrections to scaling in the geometric Galton-Watson process, PLoS ONE **11**, e0161586 (2016).

[49] A. Saichev, A. Helmstetter, and D. Sornette, Power-law distributions of offspring and generation numbers in branching models of earthquake triggering, Pure Appl. Geophys. **162**, 1113 (2005).

[50] R. Cohen, S. Havlin, and D. ben Avraham, Structural properties of scale-free networks, in *Handbook of Graphs and Networks*, edited by S. Bornholdt and H. G. Schuster (Wiley, Berlin, 2002), pp. 85–110.

[51] A. Corral, Tail of the distribution of fatalities in epidemics, Phys. Rev. E **103**, 022315 (2021).

[52] S. Manrubia and D. H. Zanette, Individual risk-aversion responses tune epidemics to critical transmissibility ($R = 1$), Proc. R. Soc. London **92**, 211667 (2022).

[53] N. Wei and G. Pruessner, Comment on "Finite-size scaling of survival probability in branching processes," Phys. Rev. E **94**, 066101 (2016).

[54] A. Corral, L. Alsedà, and J. Sardanyès, Finite-time scaling in local bifurcations, Sci. Rep. **8**, 11783 (2018).

[55] P. Cirillo and N. N. Taleb, Tail risk of contagious diseases, Nat. Phys. **16**, 606 (2020).

[56] A. Corral, Finite-size scaling versus dual random variables and shadow moments in the size distribution of epidemics, arXiv:2011.04316.

[57] The existence of a constant, nonzero value of I_{sur} could explain the observed persistence of computer viruses with low values of R_0; see Chap. 6 of Ref. [8]. In particular, this could be another way to look at the results of Ref. [36].

[58] R. Garcia-Millan, J. Pausch, B. Walter, and G. Pruessner, Field-theoretic approach to the universality of branching processes, Phys. Rev. E **98**, 062107 (2018).

[59] The situation is similar to the generalized central-limit theorem [61], for which there is a unique universality class (the Gaussian distribution) when the initial variance is finite, but infinite classes (the Lévy stable distributions) for diverging variance. Also, thinning processes show a similar separation but between finite and infinite means [62] (this sharp separation between finite and infinite first or second moments does not happen for extreme-value limit distributions [63]).

[60] A. Erdélyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Higher Transcendental Functions, Vol. 1, 1.11 The Function* $\Psi(z, s, v) = \sum_{n=0}^{\infty}(v+n)^{-s}z^n$ (Krieger, New York, 1981).

[61] J.-P. Bouchaud and A. Georges, Anomalous diffusion in disordered media: statistical mechanisms, models and physical applications, Phys. Rep. **195**, 127 (1990).

[62] A. Corral, Scaling in the timing of extreme events, Chaos Solitons Fractals **74**, 99 (2015).

[63] S. Coles, *An Introduction to Statistical Modeling of Extreme Values* (Springer, London, 2001).