

Outbreak statistics and scaling laws for externally driven epidemics

Sarabjeet Singh*

Theoretical and Applied Mechanics, Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, New York 14853, USA

Christopher R. Myers†

Laboratory of Atomic and Solid State Physics, and Institute of Biotechnology, Cornell University, Ithaca, New York 14853, USA

(Received 31 December 2013; published 3 April 2014)

Power-law scalings are ubiquitous to physical phenomena undergoing a continuous phase transition. The classic susceptible-infectious-recovered (SIR) model of epidemics is one such example where the scaling behavior near a critical point has been studied extensively. In this system the distribution of outbreak sizes scales as $P(n) \sim n^{-3/2}$ at the critical point as the system size N becomes infinite. The finite-size scaling laws for the outbreak size and duration are also well understood and characterized. In this work, we report scaling laws for a model with SIR structure coupled with a constant force of infection per susceptible, akin to a “reservoir forcing”. We find that the statistics of outbreaks in this system fundamentally differ from those in a simple SIR model. Instead of fixed exponents, all scaling laws exhibit tunable exponents parameterized by the dimensionless rate of external forcing. As the external driving rate approaches a critical value, the scale of the average outbreak size converges to that of the maximal size, and above the critical point, the scaling laws bifurcate into two regimes. Whereas a simple SIR process can only exhibit outbreaks of size $O(N^{1/3})$ and $O(N)$ depending on whether the system is at or above the epidemic threshold, a driven SIR process can exhibit a richer spectrum of outbreak sizes that scale as $O(N^\xi)$, where $\xi \in (0, 1] \setminus \{2/3\}$ and $O((N/\ln N)^{2/3})$ at the multicritical point.

DOI: [10.1103/PhysRevE.89.042108](https://doi.org/10.1103/PhysRevE.89.042108)

PACS number(s): 02.50.Ey, 05.70.Fh, 87.10.Ca

I. INTRODUCTION

Epidemic models have proven to be extremely useful in understanding the spread of infectious diseases, rumors, computer viruses, and fads [1,2]. These models constitute a broader category of models describing physical processes that exhibit a second-order phase transition at a critical threshold [3]. As is characteristic of such transitions, epidemic models exhibit power-law scaling in various statistics characterizing infectious outbreaks at the critical threshold. The classic susceptible-infectious-recovered (SIR) model [Eq. (1)] has been the most widely studied epidemic model in the literature [2,4]. In this model, a small number of infected hosts start an “outbreak” in a susceptible pool. An infected host continues to infect susceptible hosts before becoming recovered. The model can be specified using the following rate equations where the rates represent probabilities per unit time of each “reaction” taking place:

$$\begin{aligned} (S, I, R) &\xrightarrow{\alpha SI/N} (S-1, I+1, R) \\ (S, I, R) &\xrightarrow{I} (S, I-1, R+1). \end{aligned} \quad (1)$$

Note that the rates reported in Eq. (1) are rescaled by the rate of recovery, without loss of generality.

The SIR model has an epidemic threshold ($\alpha = 1$ in our case), below which all outbreaks are small [with size $o(N)$] and above which some outbreaks are large [with size $O(N)$] [4]. At the critical threshold, the distribution of outbreak sizes shows the universal scaling of $P(n) \sim n^{-3/2}$, which is invariant to changes in the microscopic details of the model [2,5]. The

size of the average and the maximal outbreaks scale as $N^{1/3}$ and $N^{2/3}$ at the critical point, respectively [2]. Extensions of the SIR that include multiple stages exhibit the scaling $P(n) \sim n^{-(1+2^{-p})}$ for outbreak sizes where p is the number of stages that an infected host crosses before being recovered [5]. The scaling exponents for the average and the maximal outbreak sizes in this multistage SIR are functions of both p and 2^{-p} , introducing a discrete degree of variability in the scaling depending on the number of stages involved.

A different extension to the simple SIR includes an external force of infection action on each susceptible as follows:

$$\begin{aligned} (S, I, R) &\xrightarrow{\alpha S(I+\nu)/N} (S-1, I+1, R) \\ (S, I, R) &\xrightarrow{I} (S, I-1, R+1). \end{aligned} \quad (2)$$

In this system, each susceptible experiences an additional force of infection with rate $\alpha\nu/N$, where ν is a dimensionless parameter reflecting the external driving rate. Such a model describes infection dynamics where a pathogen that is sustained in a reservoir repeatedly jumps to susceptible hosts [6,7], as might be applicable to the study of cross-species infections such as zoonotic diseases that jump from animals to humans. By construction, the model allows for reintroduction of infection after an outbreak has died out so long as there remain any susceptible host. In the subcritical case ($\alpha < 1$), the process alternates between periods of highly stochastic externally forced outbreaks and periods of no activity [see Fig. 1(a)]. While the dynamics of this type of model have been examined previously, the calculation of the distributions of outbreak sizes and durations has surprisingly received no attention. These statistics are important for several reasons. From a theoretical perspective, we demonstrate here that outbreak statistics qualitatively differ for the externally driven

*ss2365@cornell.edu

†c.myers@cornell.edu

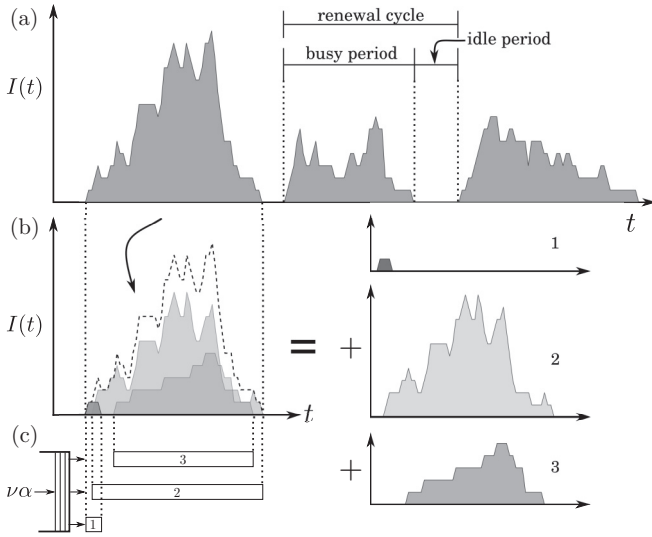


FIG. 1. (a) A realization of the externally driven SIR process with three outbreaks. (b) The decomposition of an outbreak into its constituent micro-outbreaks. Each micro-outbreak is associated with a single imported infection. (c) The $M/G/\infty$ queue where a micro-outbreak is analogous to a customer being serviced at a station. In this realization, the busy period starts when the first customer enters service and ends when the second customer leaves service. The number of customers served during a busy period corresponds to the number of micro-outbreaks that constitute an outbreak. The statistics of the composite outbreaks depends strongly on the driving rate ν , which is our primary focus here.

system than for the simple SIR. From a practical point of view, reservoir-driven zoonotic outbreaks are known to be sporadic [6], and time-series data for outbreaks exhibit active and nonactive phases over long periods of time. The information on statistics of individual outbreaks allows one to assess from given data whether the rate $\nu\alpha$ of external introduction and the rate α of infectious contact are constant over a given period of time or varying from outbreak to outbreak. In this work, we solve for the distribution of outbreak sizes $P(n)$ for this system in the limit of $N \rightarrow \infty$. From the analytical distribution, we distill scaling laws for all quantities of interest for both infinite and finite population systems at the epidemic threshold of $\alpha = 1$ with varying values of ν . In previous related work, we have calculated similar sorts of outbreak statistics for a different type of externally driven SIR system [8], arising from the coupling of an epidemic outbreak across two populations (e.g., animals and humans). The case of constant external forcing considered here would be more applicable to situations where infection is at an endemic equilibrium in the reservoir.

II. INFINITE POPULATION

In the limit of infinite system size, the simple SIR process converges in distribution to a linear birth-death (BD) process whose analysis has provided crucial insights into the full nonlinear process. Similarly, the distribution of the driven SIR process converges to a linear birth-death-immigration (BDI) process [Eq. (3)] as $N \rightarrow \infty$, which has been analyzed

extensively in the literature [9,10],

$$\begin{aligned} (I, R) &\xrightarrow{\alpha(I+1)} (I+1, R) \\ (I, R) &\xrightarrow{I} (I-1, R+1). \end{aligned} \quad (3)$$

Note that the BDI process with $\nu = 0$ is identical to the BD process. Of particular interest here are the subcritical and critical cases ($\alpha \leq 1$) where outbreaks occur sporadically from imported infections that arrive with rate $\nu\alpha$ but go extinct with probability 1. The time-dependent distribution of the number of infected hosts in the strictly subcritical case ($\alpha < 1$) which starts with no infection is given by a negative binomial distribution [10],

$$\mathbb{P}[I(t) = n] = \binom{n + \nu - 1}{n} (1 - \alpha)^\nu \alpha^n \frac{[1 - e^{-(\alpha-1)t}]^n}{[1 - \alpha e^{-(\alpha-1)t}]^{n+\nu}}, \quad (4)$$

which is succinctly expressed using a probability generating function (PGF),

$$A(x; t) = \sum_{n=0}^{\infty} \mathbb{P}[I(t) = n] x^n = \left[\frac{1 - \alpha}{1 - \alpha x + \alpha e^{-(\alpha-1)t} (x - 1)} \right]^\nu. \quad (5)$$

Due to an infinite susceptible pool and repeated introductions, the epidemic never goes extinct in the subcritical process and the number of currently infectious hosts converges to a limiting distribution as $t \rightarrow \infty$ [10],

$$A(x; \infty) = \sum_{n=0}^{\infty} \mathbb{P}[I(\infty) = n] x^n = \left(\frac{1 - \alpha}{1 - \alpha x} \right)^\nu. \quad (6)$$

The limiting subcritical BDI process can be interpreted as a renewal process where one idle period ($I = 0$) and one busy period ($I > 0$) together form a renewal cycle [9]. We shall define an outbreak in the BDI process to be synonymous with the busy period of the renewal cycle. To obtain the distribution of outbreak sizes, we first draw the analogy between the BDI process and the $M/G/\infty$ queue, as has been reported in the literature [9,11]. The notation $M/G/\infty$ describes a queueing process where customers arrive at an infinite server station according to a Poisson process and enter into service immediately. The service time at a server has a general distribution that is specified. In the notation, M stands for Markovian arrival process, G stands for the general service time distribution, and ∞ stands for the infinite number of servers [12]. The busy period for the queue is defined as the time period when at least one customer is still in service.

Each imported infection can be imagined to be an arrival in the infinite server queue. The service through a single server is then analogous to a ‘‘micro-outbreak’’ in the BDI process, i.e., the chain of infections originating from a single imported infection. A micro-outbreak is then mathematically equivalent to a BD process [Eq. (3) with $\nu = 0$] with a single infectious host at the beginning. Thus, the distribution of service times at a single server in the queue system is the same as the distribution of outbreak durations in a BD process, whose closed form solution is available [10]. Finally, the busy period of a BDI process is mathematically equivalent to the busy period of

an $M/G/\infty$ queue whose statistics can be calculated using established methods in queueing theory [13]. The intuition for infinite servers comes from the fact that because outbreaks are occurring in an infinite susceptible pool, there is no constraint on how many individual micro-outbreaks can be initiated on overlapping time scales. See Fig. 1 for illustration of the preceding concepts. Before embarking on new calculations, we first report some results from the literature that we can make use of. For instance, from Eq. (6) the probability that the limiting BDI process is in the idle state ($I = 0$) is given by

$$q_0 = (1 - \alpha)^\nu. \quad (7)$$

This is equal to the probability that the equivalent $M/G/\infty$ queue is in the idle state. From queueing theory [14], we know the following:

$$\text{Mean no. of customers served in busy period} = 1/q_0. \quad (8)$$

The number of customers served during the busy period of the queue correspond to the number of imported infections in a single outbreak. The total outbreak size can be obtained by integrating over all the micro-outbreaks emanating from imported infections. Heuristically, we know that the average outbreak size for the BD process is given by $(1 - \alpha)^{-1}$. Using this result, we can guess that the average outbreak size for the BDI process would scale as

$$\langle n \rangle \sim \frac{1}{q_0(1 - \alpha)} = (1 - \alpha)^{-(1+\nu)}. \quad (9)$$

Note that the probability of the process being in the idle state [Eq. (7)] is always greater than 0 as long as $\alpha < 1$ and ν is finite. Thus, a subcritical BDI process can never be driven into a perpetual busy period, and, accordingly, there does not exist any critical driving rate.

The average duration of the outbreak can be derived using the theory of renewal processes (see Ref. [9] for a more rigorous derivation). Arrivals in the analogous $M/G/\infty$ queue form a Poisson process with rate $\nu\alpha$. A busy period begins when an arrival takes place at the end of an idle period. Thus, the renewal cycle (idle period + busy period) is a thinned Poisson process which occurs with rate $\nu\alpha q_0$ (the original rate multiplied by the probability that the arrival occurs when the cycle is in the idle state). The average duration of a renewal cycle is $1/(\nu\alpha q_0)$ and the average duration of a busy period is a fraction $1 - q_0$ of the cycle duration. Combining these results, the average duration of an outbreak in the BDI process is given by

$$\langle t \rangle = \frac{1 - q_0}{\nu\alpha q_0} = \frac{(1 - \alpha)^{-\nu} - 1}{\nu\alpha}. \quad (10)$$

The critical BDI process ($\alpha = 1$) is an interesting analog to the critical BD process for which some results can be derived using the PGF in Eq. (5). For instance, the distribution of the number of infectious hosts as a function of time (with no infection at time 0) is generated by

$$K(x, t) = [1 - (x - 1)t]^{-\nu}, \quad (11a)$$

which does not have a steady-state solution. The average number of infectious hosts grows linearly with time,

$$\langle I(t) \rangle = \nu t, \quad (11b)$$

and the probability of the process being in the idle state decays with time,

$$p_0(t) = (1 + t)^{-\nu}. \quad (11c)$$

As expected, the average outbreak size [Eq. (9)] and duration [Eq. (10)] diverge at the critical threshold.

The existing generating functions for the BDI process [Eqs. (5) and (11a)] do not describe the busy period of the process in isolation. One can only query the distribution of the number of infected (or recovered) hosts at time t without conditioning on whether the process is busy or idle and without any knowledge of how many outbreaks have occurred before t . For calculating the statistics of a single outbreak, integrating the time-dependent generating functions unconditionally would be incorrect. Instead, one must integrate over the duration of a single outbreak, which corresponds to the busy period of the analogous $M/G/\infty$ queue. The calculation for the number of customers served in the busy period exists for the $M/G/\infty$ queue [13] that we shall adopt for our purpose. The calculation presented here is done for arbitrary values of α and ν assuming that the outbreak sizes are finite. The first ingredient in this calculation are the statistics of a BD process, which are summarized in the following PGF. Let

$$F(x, y; t) = \sum_{m, n} \mathbb{P}[I(t) = m, R(t) = n] x^m y^n \quad (12)$$

be the PGF for the joint distribution of infectious and removed hosts in a BD process with birth rate α and death rate set to 1 that starts with one infectious individual at time 0. From Ref. [10] the exact solution of the PGF is given by

$$F(x, y; t) = \frac{\Lambda_0(\Lambda_1 - x) + \Lambda_1(x - \Lambda_0)e^{-\alpha(\Lambda_1 - \Lambda_0)t}}{(\Lambda_1 - x) + (x - \Lambda_0)e^{-\alpha(\Lambda_1 - \Lambda_0)t}}, \quad (13)$$

where $\Lambda_0(y)$ and $\Lambda_1(y)$ are roots of the following quadratic equation such that $0 < \Lambda_0 < 1 < \Lambda_1$.

$$\alpha w^2 - (\alpha + 1)w + y = 0. \quad (14)$$

The joint distribution of the duration T and the size $R(T)$ of an outbreak can be summarized using $F(0, y; t)$, i.e.,

$$F(0, y; t) = \sum_{n \geq 1} \mathbb{P}[T \leq t, R(T) = n] y^n. \quad (15)$$

The trick that yields the desired result is to use the PGF $F(0, y; t)$ in place of the service time distribution for calculating the number of customers served in a busy period of $M/G/\infty$ queue (see Appendix A for details). The intuition comes from the fact that the outbreak duration and size are correlated random variables, and integrating the joint distribution preserves the correlation. Once we substitute $F(0, y; t)$ and simplify the integration, we obtain the following PGF [Eq. (16a)] for the joint distribution of the number of imported infections and outbreak size during the busy period of the BDI process:

$$G(x, y) = 1 - \frac{1}{\nu} \frac{\Lambda_1 z^\alpha (1 - z)^b}{\int_z^1 r^{\alpha-1} (1 - r)^{b-1} dr}, \quad (16a)$$

where

$$z = 1 - \frac{\Lambda_0}{\Lambda_1}, \quad a = 1 - \nu x, \quad b = \nu \left(\frac{1 - \Lambda_0 x}{\Lambda_1 - \Lambda_0} \right), \quad (16b)$$

$$\Lambda_0, \Lambda_1 = \frac{(\alpha + 1) \mp \sqrt{(\alpha + 1)^2 - 4\alpha y}}{2\alpha}. \quad (16c)$$

The marginal distribution of outbreak sizes is generated by $G(1, y)$. Let $H(y)$ be the PGF for the marginal distribution at the critical threshold $\alpha = 1$. This simplifies some of the terms in the PGF as follows:

$$H(y) = 1 - \frac{1}{\nu} \frac{\Lambda_1 z^a (1 - z)^b}{\int_z^1 r^{a-1} (1 - r)^{b-1} dr}, \quad (17a)$$

where

$$z = 1 - \frac{\Lambda_0}{\Lambda_1}, \quad a = 1 - \nu, \quad b = \frac{\nu}{2}, \quad (17b)$$

$$\Lambda_0, \Lambda_1 = 1 \mp \sqrt{1 - y}. \quad (17c)$$

The integral in the denominator of Eq. (17a) can be solved explicitly for $\nu \in \mathbb{Z}_{>0}$. For an arbitrary ν , the integral can be represented as the difference between the beta function $B(a, b)$ and incomplete beta function $B(a, b; z)$. Let this integral be denoted by $J(a, b; z)$.

$$J(a, b; z) = B(a, b) - B(a, b; z). \quad (18)$$

The asymptotic form for $P(n)$ —the probability of having an outbreak of size n —can be obtained by the singularity expansion of the PGF $H(y)$ around $y = 1$ (or $z = 0$). For $\nu < 1$ (which implies $a, b > 0$), the incomplete beta function can be approximated by

$$B(a, b; z) \sim \frac{z^a (1 - z)^b}{a} \quad (19)$$

in the limit of $z \rightarrow 0$ [15]. The PGF $H(y)$ simplifies as

$$\begin{aligned} H(y) &\sim 1 - \frac{1}{\nu B(a, b)} \frac{\Lambda_1 z^a (1 - z)^b}{\left(1 - \frac{z^a (1 - z)^b}{a B(a, b)}\right)} \\ &= 1 - \frac{1}{\nu B(a, b)} \Lambda_1 z^a (1 - z)^b \left[1 + \frac{z^a (1 - z)^b}{a B(a, b)} + \dots \right] \\ &= 1 - \frac{2^{1-\nu} (1 - y)^{(1-\nu)/2} (y)^{\nu/2}}{\nu B(a, b)} + \dots, \end{aligned} \quad (20)$$

where the simplification in the last step follows from substituting for z from Eq. (17b). From the leading-order term $(1 - y)^{(1-\nu)/2}$, we can assert the following asymptotic form for $P(n)$ as described in Ref. [16]:

$$P(n) \sim n^{-(3-\nu)/2}. \quad (21)$$

The scaling law is verified in Fig. 2. As expected, the power law becomes more and more flat with increasing ν . The pronounced bump in the simulations is a finite-size effect due to the clustering of outbreaks that would have continued to exhibit the power-law scaling if the system size was infinite [17]. It can be verified that all moments of the distribution diverge for any value of $\nu \in [0, 1)$.

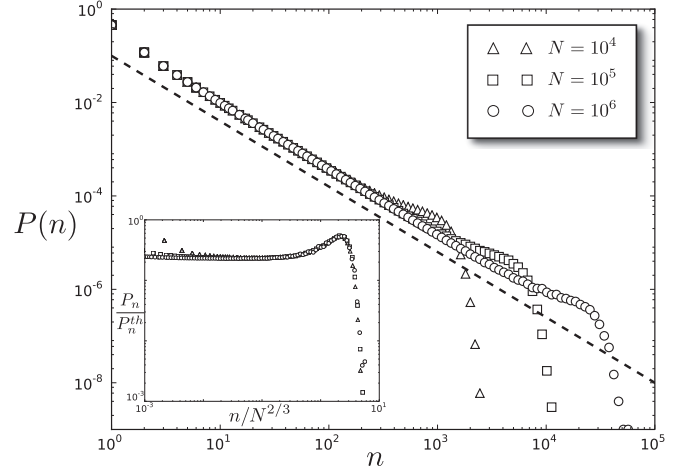


FIG. 2. The probability of having an outbreak of size n empirically calculated from 10^9 realizations of the process for different values of N and $\alpha = 1, \nu = 0.2$. The dashed line shows the slope of the analytical scaling predicted from theory [Eq. (21)], ignoring any constant prefactors. Inset: The collapse of outbreak sizes when scaled by $N^{2/3}$. The bump near the exponential cutoff represents the probability mass associated with outbreaks that would have continued along the power law in an infinite-size system but are clustered due to finite-size effects.

At $\nu = 1$, the PGF $H(y)$ in Eq. (17a) can be simplified further into a closed form solution,

$$H(y) = 1 + \frac{\sqrt{y}}{\ln\left(\frac{\sqrt{1-y}}{1+\sqrt{y}}\right)}, \quad (22)$$

whose singularity analysis around $y = 1$ yields the following asymptotic form for $P(n)$:

$$P(n) \sim \frac{1}{n \ln^2 4n} \left[\frac{1}{2} - \frac{\gamma}{\ln 4n} + O\left(\frac{1}{\ln^2 4n}\right) \right], \quad (23)$$

where $\gamma = 0.5772\dots$ is the Euler-Mascheroni constant. As in the case of $\nu < 1$, all moments of the distribution diverge in this case as well. See Fig. 3 for comparison with stochastic simulations.

The case of $\nu > 1$ requires a careful analysis of the function $J(a, b; z)$ because the parameter a becomes negative in this regime. Since $J(a, b; z)$ is the difference of beta and incomplete beta functions, the following identity holds:

$$J(a, b; z) = \frac{(a + b)}{a} J(a + 1, b; z) - \frac{z^a (1 - z)^b}{a}. \quad (24)$$

Consider the case where $\nu \in (1, 2)$, which implies that $0 < a + 1 < 1$. In this case, as $z \rightarrow 0$,

$$J(a + 1, b; z) \sim B(a + 1, b)$$

$$J(a, b; z) \sim \frac{(a + b)}{a} B(a + 1, b; z) - \frac{z^a (1 - z)^b}{a} \quad (25)$$

and the PGF $H(y)$ simplifies as follows:

$$\begin{aligned} H(y) &\sim 1 + \frac{a}{\nu} \frac{\Lambda_1}{[1 - (a + b)B(a + 1, b)z^{-a}(1 - z)^{-b}]} \\ &= 1 + \frac{a\Lambda_1}{\nu} [1 + (a + b)B(a + 1, b)z^{-a}(1 - z)^{-b} + \dots]. \end{aligned} \quad (26)$$

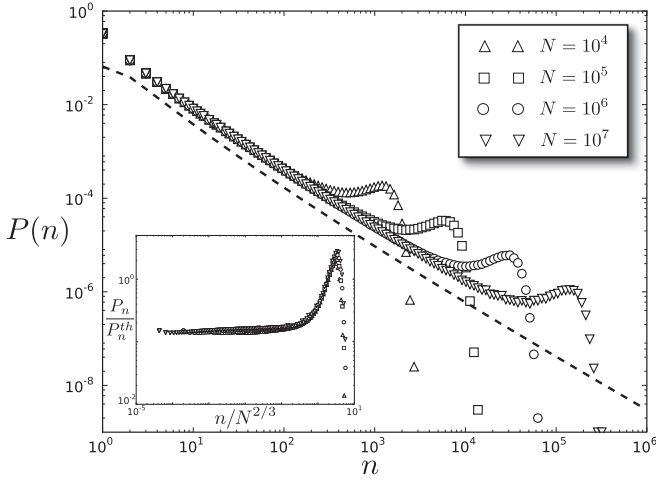


FIG. 3. The probability of having an outbreak of size n for $\alpha = 1, \nu = 1$. Dashed line plotted at an offset represents the analytical scaling [Eq. (23)]. Inset: The collapse of outbreak sizes when scaled by $N^{2/3}$.

Substituting for z , we obtain a series expansion in fractional powers of $\sqrt{1-y}$. The leading term in the expansion is of the order $(1-y)^{(\nu-1)/2}$, which provides the following asymptotic form for $P(n)$:

$$P(n) \sim n^{-(\nu+1)/2}. \quad (27)$$

Similarly, by binning ν in $\{(2,3), (3,4), \dots\}$ and applying the property given by Eq. (24) iteratively, we obtain the asymptotics as in Eq. (27). The same is true when integral values are chosen for ν , in which case the PGF can be simplified further. For instance, substituting $\nu = 2$ in Eq. (17a) simplifies the PGF to

$$H(y) = \frac{1 - \sqrt{1-y}}{2}, \quad (28)$$

which is same as the PGF for the BD process with a prefactor of $1/2$. Asymptotic analysis reveals a scaling exponent of $3/2$, consistent with Eq. (27). The agreement of Eq. (27) with stochastic simulations is demonstrated in Fig. 4.

Equation (27) suggests that the outbreak size distribution falls off more steeply with increasing ν . This seems counterintuitive at first because one would expect that, with increasing ν , there should be a greater probability of larger outbreaks, leading to more slowly decaying distribution. The resolution of this puzzle can be found by looking at the total probability mass contained in the generating function. It can be verified that for $\nu \leq 1$, $H(y) \rightarrow 1$ as $y \rightarrow 1$, i.e., the distribution is proper. But for $\nu > 1$, the distribution becomes defective such that

$$\lim_{y \rightarrow 1} H(y) = \nu^{-1}. \quad (29)$$

The remaining probability mass $1 - \nu^{-1}$ is associated with the infinite-sized outbreak. This effect can be seen in stochastic simulations (Fig. 4) where the outbreaks not accounted by the power-law cluster in the bump of the distribution. More

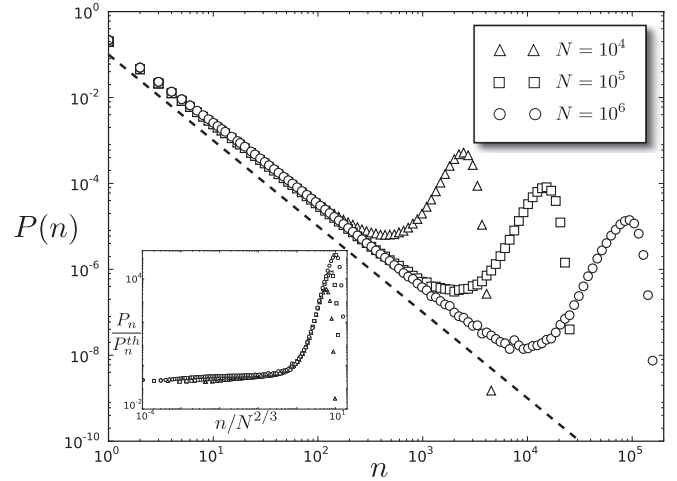


FIG. 4. The probability of having an outbreak of size n for $\alpha = 1, \nu = 3$. Dashed line represents the analytical scaling [Eq. (27)] at an offset. The finite-size bump is more pronounced because it has accumulated outbreaks of two distinct scales. The bump starts at the scale of $N^{2/3}$ consistent with what is observed for $\nu \leq 1$. However, the exponential cutoff which marks the end of the bump occurs at a different scale as evident from the lack of scaling collapse for large n in the inset. The presence of two scales in the bump is discussed in detail in Sec. III A and illustrated in Fig. 7.

formally, we have the following:

$$\lim_{n \rightarrow \infty} \mathbb{P}[R(\infty) > n] \sim \begin{cases} O(n^{-(1-\nu)/2}) & \nu < 1, \\ O((\ln n)^{-1}) & \nu = 1, \\ 1 - \nu^{-1} + O(n^{-(\nu-1)/2}) & \nu > 1. \end{cases} \quad (30)$$

Thus, for $\nu > 1$, the probability of having an outbreak size exceeding any arbitrary scale converges to a constant value of $1 - \nu^{-1}$, whereas the same probability diminishes with n for the case of $\nu \leq 1$. For $\nu > 1$, the distribution represented by the generating function $H(y)$ excludes the infinite-sized outbreak. Thus, the statistics are conditional on a finite-sized outbreak. The distribution falls more steeply with increasing ν because more and more outbreaks escape to infinity with probability $1 - \nu^{-1}$. Nevertheless, the statistics of the power-law regime are interesting to analyze even if they represent part of the distribution. For instance, the k^{th} moment of the distribution is finite only if $\nu > 2k + 1$ and diverges otherwise,

$$\langle n^k \rangle \sim \begin{cases} \frac{1}{\nu - (2k + 1)} & \text{if } \nu > 2k + 1, \\ \infty & \text{otherwise.} \end{cases} \quad (31)$$

Finally, we summarize the asymptotic statistics calculated in the preceding section

$$P(n) \sim \begin{cases} n^{-(3-\nu)/2} & \nu < 1, \\ \frac{1}{n \ln^2 4n} \left[\frac{1}{2} - \frac{\gamma}{\ln 4n} + O\left(\frac{1}{\ln^2 4n}\right) \right] & \nu = 1, \\ n^{-(\nu+1)/2} & \nu > 1. \end{cases} \quad (32)$$

We can now put the results in some perspective. The external driving can be thought of as a ‘‘coupling agent’’ that combines

an increasing number of micro-outbreaks into a single outbreak as ν is increased. When $\alpha = 1$ and ν is above 1, the external driving binds an infinite number of micro-outbreaks into one contiguous outbreak with probability $1 - \nu^{-1}$. Qualitatively, the BDI process can also be interpreted as a two-state Markov chain that switches between the idle period and the busy period. In this interpretation the idle period is positive recurrent if $\alpha < 1$ (busy period ends with probability 1 and in finite time), null recurrent if $\alpha = 1, \nu \leq 1$ (busy period ends with probability 1 but the expected duration is ∞), and transient if $\alpha = 1, \nu > 1$ (busy period can persist indefinitely). The case of $\alpha > 1$ is trivial since a supercritical process can grow exponentially even without the external forcing. The idle period is thus transient in this case.

III. FINITE POPULATION

A. Outbreak size

For a finite-sized system [Eq. (2)], we first establish the scaling of the “maximal” outbreak size echoing the analysis in Refs. [2,5]. Let there be a maximal size M , such that the outbreak cannot exceed this size due to depletion in the susceptible pool. For $\nu < 1$, the algebraic distribution in Eq. (21) gives an estimate for the average outbreak size,

$$\langle n \rangle = \sum_{n \leq M} n P(n) \sim \sum_{n \leq M} n^{-(1-\nu)/2} \sim M^{(1+\nu)/2}. \quad (33)$$

In a finite-sized system, the effective rate of infectious contact per infected host is reduced to $\alpha_* = 1 - M/N$ due to depletion. From Eq. (9), we obtain a second estimate for the scaling of the average outbreak size,

$$\langle n \rangle \sim (1 - \alpha_*)^{-(1+\nu)} = (N/M)^{1+\nu}. \quad (34)$$

Equating the two estimates we obtain the following scaling laws for $\nu < 1$:

$$M \sim N^{2/3} \quad \text{and} \quad \langle n \rangle \sim N^{(1+\nu)/3}. \quad (35)$$

The scaling of M is verified in Fig. 2 (inset) and that of $\langle n \rangle$ in Fig. 5. For $\nu = 1$, we use the expression for $P(n)$ in Eq. (23) and obtain M as the solution of the following implicit equation:

$$\left(\frac{N}{M}\right)^2 = \frac{M}{2 \ln^2 4M} + O\left(\frac{M}{\ln^3 4M}\right), \quad (36)$$

whose solution to a first-order approximation leads to the following scaling laws:

$$M \sim (N \ln N)^{2/3} \quad \text{and} \quad \langle n \rangle \sim \left(\frac{N}{\ln^2 N}\right)^{2/3}. \quad (37)$$

However, numerical results obtained from stochastic simulations reveal slightly different scaling laws,

$$M \sim (N^2 \ln N)^{1/3}, \quad \text{and} \quad \langle n \rangle \sim \left(\frac{N}{\ln N}\right)^{2/3}, \quad (38)$$

that differ from theory by a factor of $(\ln N)^{1/3}$ in M and $(\ln N)^{-2/3}$ in $\langle n \rangle$. The empirical scaling law can be obtained if Eq. (36) is replaced with the following:

$$\left(\frac{N}{M}\right)^2 \sim \frac{M}{\ln M}. \quad (39)$$

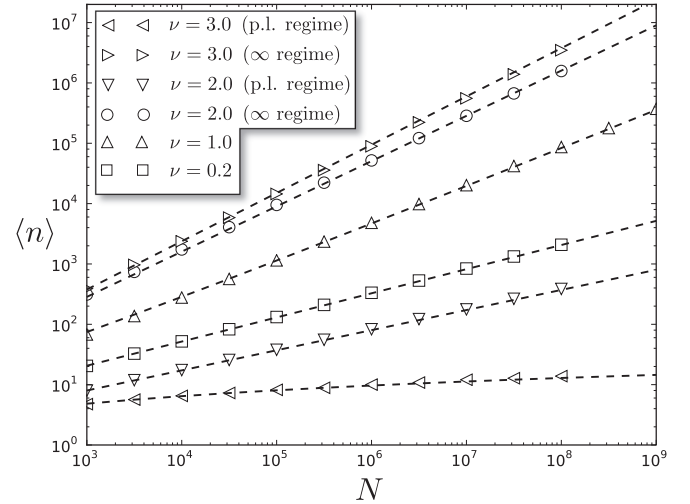


FIG. 5. Finite-size scaling for average outbreak size ($\nu = \{0.2, 1, 2, 3\}$). The statistics of the power-law regime and the infinite regime were calculated separately using $N^{2/3}$ as the separation boundary. Note that the scaling for $\nu = 3$ is purely logarithmic in the power-law regime. Dashed lines represent the scaling laws predicted from theory [Eq. (44)].

Although the power-law part of the scaling—that is, the term $N^{2/3}$ —is consistent between both the empirically observed [Eq. (38)] and the theoretically calculated [Eq. (37)] scaling, we are unable to resolve the logarithmic corrections and pose their solution as an open problem. The agreement of the scaling laws [Eq. (38)] with results from stochastic simulations is shown in Figs. 5 and 6. Henceforth, we shall refer only to the empirical scaling law for $\nu = 1$ where the logarithmic corrections are important.

The case of $\nu > 1$ requires careful consideration. The analysis on the infinite-sized system revealed that outbreaks

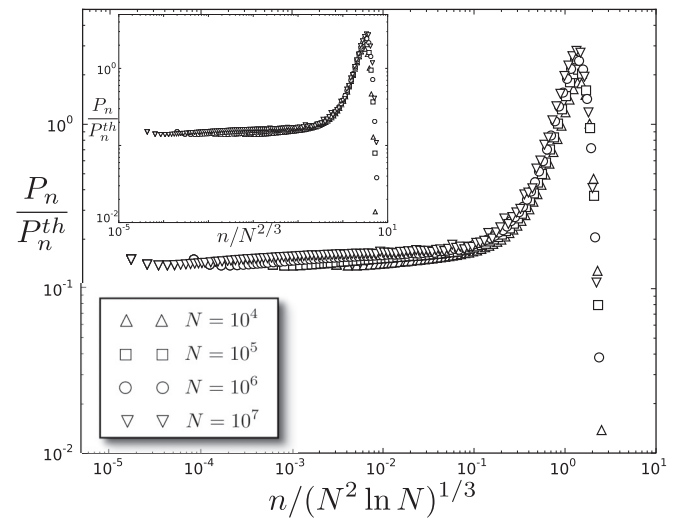


FIG. 6. Scaling collapse at $M \sim (N^2 \ln N)^{1/3}$ for $\nu = 1$. The y axis is scaled by the theoretical scaling law of Eq. (23). Note that the scaling collapse is distinct from the one shown in inset (done at the scale of $N^{2/3}$). While the power laws collapse on top of each other at $N^{2/3}$, the exponential cutoffs collapse at $(N^2 \ln N)^{1/3}$. This separation of scales is more pronounced for $\nu > 1$ (see Fig. 7).

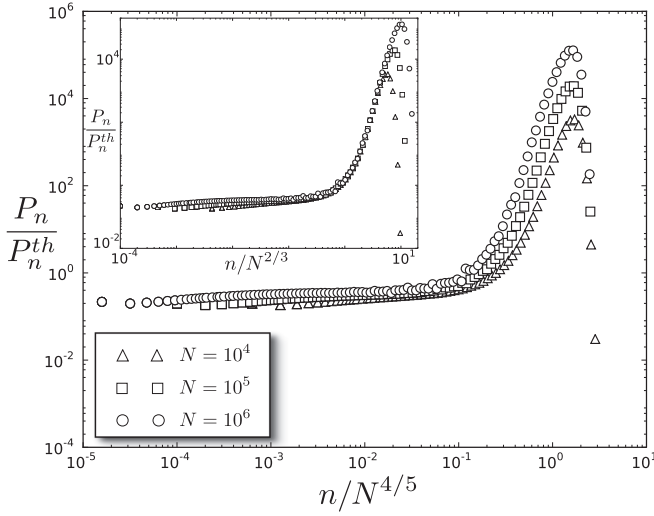


FIG. 7. Scaling of the outbreak size distribution by $M \sim N^{(v+1)/(v+2)}$ for $\nu = 3$. The y axis is scaled by the theoretical scaling law of Eq. (27). Similarly to Fig. 6, the scaling collapse is distinct from the one shown in inset (outbreaks scaled by $N^{2/3}$). The power-law regime exhibits a scaling collapse at $N^{2/3}$, while the exponential cutoffs collapse at $N^{4/5}$.

occur according to a power-law distribution [Eq. (27)] with probability ν^{-1} or are infinite in size with probability $1 - \nu^{-1}$. Henceforth, we shall label these as the “power-law regime” and the “infinite regime”, respectively. The average outbreak size in the power-law regime diverges when $\nu < 3$ [see Eq. (31)]. For finite systems, we expect that both the infinite regime and the power-law regime would admit two different scaling laws for average outbreak size and duration. The power-law regime admits a positive exponent for the scaling law only for $\nu \in (1, 3)$. Let $\langle n \rangle_\infty$ be the average outbreak size conditioned on the outbreak being in the infinite regime. From Eq. (33), note that as $\nu \rightarrow 1$, the average outbreak size $\langle n \rangle$ approaches M in scale. For $\nu = 1$, we found empirically that $\langle n \rangle \sim M / \ln M$ [see Eq. (39)]. For $\nu > 1$, intuition suggests that in the infinite regime $\langle n \rangle_\infty \sim M$, i.e., all outbreaks will be clustered at one scale. Using Eq. (34), we obtain the following scaling relationship:

$$\langle n \rangle_\infty \sim M \sim N^{(v+1)/(v+2)}. \quad (40)$$

Figure 7 shows the scaling collapse in agreement with Eq. (40).

The exponent of the scaling law in Eq. (40) is an increasing function of ν that lies in the interval $(2/3, 1)$ for all $\nu > 1$. The lower bound of $2/3$ is consistent with the fact that the scaling law for $\nu < 1$ has $2/3$ as the upper bound [Eq. (35)] and that the same exponent shows up at $\nu = 1$ albeit with logarithmic factors [Eq. (38)]. But the above scaling law will hold with probability $1 - \nu^{-1}$ that corresponds to the infinite regime. Let $\langle n \rangle_{pl}$ be the average outbreak size in the power-law regime. In a finite-sized system, there will be another scale L up to which the power-law regime holds, and any outbreak exceeding that enters the infinite regime. Using Eq. (27), we can estimate the scaling in the power-law regime,

$$\langle n \rangle_{pl} \sim \sum_{n \leq L} n^{(1-\nu)/2} \sim \begin{cases} L^{(3-\nu)/2} & \nu \in (1, \infty) \setminus \{3\}, \\ \ln L & \nu = 3. \end{cases} \quad (41)$$

L can be deduced by noting that in the limit of $\nu \rightarrow 1$, the scaling law [Eq. (41)] has to approach $N^{2/3}$ in order for the exponent to be consistent with the scaling laws for $\nu \leq 1$ [Eq. (35) and (38)]. This is true only when L scales as the following:

$$L \sim N^{2/3}, \quad (42)$$

and thus we arrive at the following scaling laws:

$$\langle n \rangle_{pl} \sim \begin{cases} N^{1-\nu/3} & \nu \in (1, \infty) \setminus \{3\}, \\ \ln N & \nu = 3. \end{cases} \quad (43)$$

The scale of $N^{2/3}$ as being the upper bound of all power laws in $P(n)$ is confirmed in simulations (see the insets in Figs. 2, 3, and 4). The derivation in Eq. (43) assumes that L only depends on N and not ν . Intuitively, L is the scale at which the power-law regime is impacted by the finiteness of the system and thus should only depend on N . Other model parameters only determine how fast or slow the process approaches that scale. We now summarize the finite-size scaling laws for the average outbreak size as follows:

$$\langle n \rangle \sim \begin{cases} N^{(1+\nu)/3} & \nu < 1, \\ \left(\frac{N}{\ln N} \right)^{2/3} & \nu = 1, \\ \left. \begin{cases} N^{1-\nu/3} & \text{w.p. } \nu^{-1} \\ N^{(v+1)/(v+2)} & \text{w.p. } 1 - \nu^{-1} \end{cases} \right\} & \nu \in (1, \infty) \setminus \{3\}, \\ \left. \begin{cases} \ln N & \text{w.p. } 1/3 \\ N^{4/5} & \text{w.p. } 2/3 \end{cases} \right\} & \nu = 3, \end{cases} \quad (44)$$

where w.p. is an abbreviation for “with probability”. The agreement of these results with stochastic simulations is shown in Fig. 5. Similarly, the summary table for the maximal outbreak size is shown below:

$$M \sim \begin{cases} N^{2/3} & \nu < 1, \\ (N^2 \ln N)^{1/3} & \nu = 1, \\ N^{(v+1)/(v+2)} & \nu > 1. \end{cases} \quad (45)$$

The critical point of $\nu = 1$ separates the scaling behavior of M into one as being a power law with fixed exponent of $2/3$ and the other as a power law with continuously varying exponent.

For $\nu > 1$, the scaling exponent of the average outbreak size bifurcates at the value of $2/3$; the two different exponents move in opposite directions with increasing ν [compare $1 - \nu/3$ with $(\nu + 1)/(\nu + 2)$, both of which start off from the value $2/3$ as $\nu \rightarrow 1+$]. A crucial insight from these results is that the average outbreak size scales as N^ξ , where $\xi \in (0, 1) \setminus \{2/3\}$ (at $\xi = 2/3$, logarithmic corrections are present). The scaling law is always sublinear as long as $\alpha = 1$, i.e., there cannot occur an outbreak that scales as $O(N)$ no matter how strongly the system is driven. Only for $\alpha > 1$ would there be an $O(N)$ outbreak with ν having no qualitative bearing on the statistics. This is because the multiplicative nature of the supercritical BD process always dominates the constant rate of growth from external driving.

Using the above results, we can calculate the scaling window for the scaling laws, i.e., the distance from the threshold

boundary within which the scaling laws are applicable [2]. The scaling window is a characteristic of the finite system size and shrinks to 0 in the limit of $N \rightarrow \infty$. For finite N , the system need not be right at the critical threshold $\alpha = 1$ for the scaling laws to be valid. Using Eqs. (34) and (44) we obtain

$$|\alpha - 1| \sim \begin{cases} N^{-1/3} & \nu < 1, \\ \left(\frac{N}{\ln N}\right)^{-1/3} & \nu = 1, \\ \begin{cases} N^{-1/3} & \text{power-law regime} \\ N^{-1/(\nu+2)} & \text{infinite regime} \end{cases} & \nu > 1. \end{cases} \quad (46)$$

For a fixed N , the infinite regime has the largest window that grows with ν .

B. Outbreak duration

With the effective transmission rate $\alpha_* = 1 - M/N$ below 1, the scaling behavior for outbreak durations can be obtained by using Eq. (10),

$$\langle t \rangle = \frac{(1 - \alpha_*)^{-\nu} - 1}{\nu \alpha_*} \sim \begin{cases} \ln(N/M) & \nu = 0, \\ (N/M)^\nu & \nu > 0. \end{cases} \quad (47)$$

For $\nu \leq 1$, we arrive at the following using Eq. (45):

$$\langle t \rangle \sim \begin{cases} \ln N & \nu = 0, \\ N^{\nu/3} & 0 < \nu < 1, \\ \left(\frac{N}{\ln N}\right)^{1/3} & \nu = 1. \end{cases} \quad (48)$$

For $\nu > 1$, we have the bifurcation of behavior into the power-law regime and the infinite regime. Since we already know the scale of M in the infinite regime [from Eq. (45)], we obtain

$$\langle t \rangle_\infty \sim N^{\nu/(\nu+2)}, \quad \nu > 1. \quad (49)$$

In the power-law regime, we resort to the survival function for calculating the scaling for the average outbreak duration (see Appendix B) and obtain the following:

$$\langle t \rangle_{pl} \sim \begin{cases} T_c^{2-\nu} & \nu \in (1, \infty) \setminus \{2\}, \\ \ln T_c & \nu = 2, \end{cases} \quad (50)$$

where T_c is the cutoff time scale for the power-law regime. For $\nu > 1$, we know that the cutoff length scale for the power law is $L \sim N^{2/3}$. We now estimate the relationship between L and T_c . From Eq. (11b), the mean number of infectious hosts increases linearly with time. Thus, the outbreak size grows quadratically with time,

$$\frac{dR}{dt} \sim t, \quad R \sim t^2, \quad (51)$$

and this gives the relationship between T_c and L as

$$T_c \sim \sqrt{L}. \quad (52)$$

The same scaling relationship was noted in Ref. [2] for the simple SIR. Using Eqs. (50), (52), and (42), we obtain

$$T_c \sim N^{1/3}, \quad \langle t \rangle_{pl} \sim \begin{cases} N^{(2-\nu)/3} & \nu \in (1, \infty) \setminus \{2\}, \\ \ln N & \nu = 2. \end{cases} \quad (53)$$

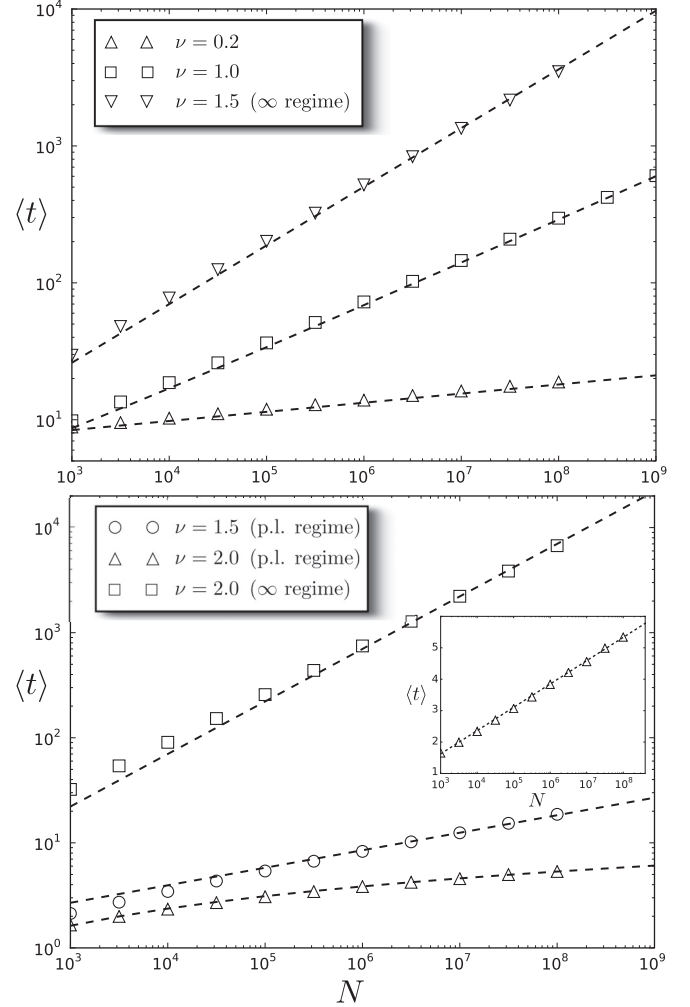


FIG. 8. Finite-size scaling for average outbreak duration ($\nu = 0.2, 1, 1.5, 2$). The plots are split into two figures for clarity. The statistics of the power-law regime and the infinite regime were calculated separately using $N^{2/3}$ as the separation boundary. Inset shows the scaling behavior for $\nu = 2$ on a log-linear plot. Dashed lines represent scaling laws predicted from theory [Eq. (54)].

The summary of the scaling laws for the average outbreak duration is given below and the agreement with stochastic simulations is shown in Fig. 8,

$$\langle t \rangle \sim \begin{cases} \ln N & \nu = 0, \\ N^{\nu/3} & \nu < 1, \\ \left(\frac{N}{\ln N}\right)^{1/3} & \nu = 1, \\ \begin{cases} N^{(2-\nu)/3} & \text{w.p. } \nu^{-1} \\ N^{\nu/(\nu+2)} & \text{w.p. } 1 - \nu^{-1} \end{cases} & \nu \in (1, \infty) \setminus \{2\}, \\ \begin{cases} \ln N & \text{w.p. } 1/2 \\ N^{1/2} & \text{w.p. } 1/2 \end{cases} & \nu = 2. \end{cases} \quad (54)$$

On comparing the scaling laws for the average outbreak size [Eq. (44)] and duration [Eq. (54)], we note that in all power-law regimes ($\nu \notin \{0, 1, 2, 3\}$), the following relationship

holds:

$$\frac{\langle n \rangle}{\langle t \rangle} \sim N^{1/3}. \quad (55)$$

For $\nu \in \{0, 1, 2, 3\}$, the relationship is not too far off either with the presence of logarithmic factors,

$$\frac{\langle n \rangle}{\langle t \rangle} \sim \begin{cases} \frac{N^{1/3}}{\ln N} & \nu = 0, \\ \left(\frac{N}{\ln N}\right)^{1/3} & \nu = 1, \\ \frac{N^{1/3}}{\ln N} & \nu = 2, \\ N^{1/3} \ln N & \nu = 3. \end{cases} \quad (56)$$

In the infinite regime, however (only for $\nu > 1$),

$$\frac{\langle n \rangle_\infty}{\langle t \rangle_\infty} \sim N^{1/(\nu+2)}, \quad (57)$$

the two scales converge for fixed N and increasing ν . This result shows that there is a universality in the power-law characteristics regardless of whether ν is below or above the critical value of 1, and it ties with the universality of $N^{2/3}$ as the characteristic scale to which all power laws in $P(n)$ extend. The presence of an ‘‘infinite regime’’ does not preclude this universality.

C. Convergence near critical points

The finite-size scaling laws come with a caveat: that the system size should be large enough or the parameter ν should be far away from critical points to avoid any interference from the logarithmic factors [see Eq. (44) and (54)]. For instance, if $\nu = 1 \pm \epsilon$, the logarithmic factor present in the scaling laws for $\nu = 1$ interferes with the scaling laws for $\nu < 1$ and $\nu > 1$ if $\epsilon \ll 1$. Since all scaling laws in the infinite regime have a monotonically increasing exponent, they would not be subjected to any interference near the critical points. Similarly to Eq. (46), we can calculate heuristically, the ‘‘interference window’’ for ν within which scaling power laws will be muddled via interference from logarithmic factors. Near $\nu = 1$, interference would occur if the two estimates of $\langle n \rangle$ at and above $\nu = 1$ are similar in scale, i.e.,

$$\left(\frac{N}{\ln N}\right)^{2/3} \sim N^{1-\nu/3}, \quad (58)$$

which gives the interference window as

$$|\nu - 1| \sim \frac{\ln(\ln N)}{\ln N}. \quad (59)$$

The window is a slowly decreasing function of N . The same functional form is obtained if we compare the scale of $\langle n \rangle$ at and below $\nu = 1$, as well as near all other critical points for both $\langle n \rangle$ and $\langle t \rangle$, i.e.,

$$|\nu - \nu_c| \sim \frac{\ln(\ln N)}{\ln N}. \quad (60)$$

Thus, for moderate values of N , the scaling laws for the power-law regime are likely to suffer from interference from logarithmic factors unless ν is far away from its critical values.

IV. DISCUSSION

In this work, we have solved for the statistical properties of the externally forced SIR model through rigorous analysis of the relevant stochastic process. By invoking the analogy between the BDI process and the $M/G/\infty$ queue, we were able to leverage existing results in building the theory for the process. The external driving acts as a binding agent for micro-outbreaks and is especially significant when $\alpha = 1$. In this case, $\nu = 1$ emerges as a second critical point in the process separating a state of recurring outbreaks from one with a single perpetual outbreak. Although power-law characteristics at the critical point were expected, the tunability of the power law by the external forcing with a precise functional form is a nontrivial result that was revealed through calculations. The finite-size scaling laws exhibit a continuum of scaling exponents governed by the driving rate that has some important implications for understanding reservoir-driven epidemics. This work also elucidates the universality of the scale of the maximal outbreak size and of the ratio of average outbreak size and duration when the distribution of sizes $P(n)$ follows a power law.

Our results provide a framework for interpreting time-series data from reservoir-driven outbreaks where the time scale of primary infections (direct reservoir transmission) and secondary infections (transmission among hosts) are comparable, and where it is not feasible to conduct field studies necessary to distinguish among them. In cases of sufficiently weak reservoir forcing, individual chains of secondary transmission can be explained by simple SIR statistics. But we demonstrate here that if the system is near the critical threshold, which is typical of emerging infectious diseases, the statistics of the process depend strongly on the reservoir forcing. Similarly, if fine-scale data were available that allowed one to resolve each micro-outbreak separately, then the simple SIR process is sufficient to describe the data. But, typically, such fine-scale data are difficult to collect on the time scales of outbreaks, and practitioners often have to contend with coarse-scale data on composite outbreaks, which is precisely where our theory and results serve a strong purpose.

ACKNOWLEDGMENTS

The authors thank David Schneider, Jason Hindes, and Oleg Kogan for helpful discussions. This work was supported by the Science & Technology Directorate, Department of Homeland Security, via interagency Agreement No. HSHQDC-10-X-00138.

APPENDIX A: DERIVATION OF GENERATING FUNCTION

From Ref. [13], the number of customers served in the busy period of an $M/G/\infty$ queue with arrival rate λ and service time distribution $U(s)$ is generated by the following PGF:

$$G(x) = 1 - \frac{1}{\lambda Q(x)}$$

$$Q(x) = \int_0^\infty \exp\left[-\lambda t + \lambda x \int_0^t U(s) ds\right] dt. \quad (A1)$$

The number of customers served corresponds to the number of micro-outbreaks in the BDI process that occur on overlapping time scales. The distribution $U(s)$ corresponds to the duration of a micro-outbreak, i.e., the BD process. The joint distribution of duration T and size $R(T)$ of an outbreak in the BD process is generated by

$$\begin{aligned} F(0, y; s) &= \sum_{n \geq 1} \mathbb{P}[T \leq s, R(T) = n] y^n \\ &= \frac{\Lambda_0 \Lambda_1 (1 - e^{-\alpha(\Lambda_1 - \Lambda_0)s})}{\Lambda_1 - \Lambda_0 e^{-\alpha(\Lambda_1 - \Lambda_0)s}}, \end{aligned} \quad (\text{A2})$$

where $\Lambda_0(y)$ and $\Lambda_1(y)$ are roots of the following quadratic equation such that $0 < \Lambda_0 < 1 < \Lambda_1$,

$$\alpha w^2 - (\alpha + 1)w + y = 0. \quad (\text{A3})$$

Substituting $\nu\alpha$ for λ and $F(0, y; s)$ for $U(s)$ in Eq. (A1) and simplifying the integral, we obtain the PGF for the joint distribution of number of micro-outbreaks and outbreak size in the BDI process,

$$G(x, y) = 1 - \frac{1}{\nu} \frac{\Lambda_1 z^\alpha (1 - z)^b}{\int_z^1 r^{a-1} (1 - r)^{b-1} dr}, \quad (\text{A4})$$

where

$$\begin{aligned} z &= 1 - \frac{\Lambda_0}{\Lambda_1}, \quad a = 1 - \nu x, \quad b = \nu \left(\frac{1 - \Lambda_0 x}{\Lambda_1 - \Lambda_0} \right) \\ \Lambda_0, \Lambda_1 &= \frac{(\alpha + 1) \mp \sqrt{(\alpha + 1)^2 - 4\alpha y}}{2\alpha}. \end{aligned}$$

APPENDIX B: SURVIVAL FUNCTION FOR $\nu > 1$

To calculate the survival function for $\alpha = 1, \nu > 1$, we repeat the calculation of the previous section but limit the integration in Eq. (A1) to a finite (but large) t rather than ∞ . In doing so, the PGF $G(x, y; t)$ reflects the distribution for those outbreaks that end before time t ,

$$G(x, y; t) = 1 - \frac{1}{\nu} \frac{\Lambda_1 z_0^\alpha (1 - z_0)^b}{\int_{z_0}^{z_t} r^{a-1} (1 - r)^{b-1} dr}, \quad (\text{B1})$$

where

$$z_t = 1 - \frac{\Lambda_0}{\Lambda_1} e^{-(\Lambda_1 - \Lambda_0)t} \quad (\text{B2})$$

and α is set to 1 for Λ_0 and Λ_1 . The distribution function for the duration T is the total probability contained in the PGF, i.e.,

$$\mathbb{P}[T < t] = \lim_{(x, y) \rightarrow (1, 1)} G(x, y; t). \quad (\text{B3})$$

In the limit $(x, y) \rightarrow (1, 1)$,

$$a \rightarrow 1 - \nu, \quad b \rightarrow \frac{\nu}{2}, \quad \Lambda_0, \Lambda_1 \rightarrow 1,$$

$$z_t \rightarrow 2(1 + t)\sqrt{1 - y},$$

$$\int_{z_0}^{z_t} r^{a-1} (1 - r)^{b-1} dr \rightarrow \frac{z_t^a - z_0^a}{a}.$$

Taking the limit and simplifying the expression, we obtain

$$\mathbb{P}[T < t] \sim \frac{\nu^{-1} - (1 + t)^{1-\nu}}{1 - (1 + t)^{1-\nu}}. \quad (\text{B4})$$

As $t \rightarrow \infty$, the probability converges to ν^{-1} , which is the probability that the outbreak has a finite size. With probability $1 - \nu^{-1}$ the outbreak persists indefinitely. The survival function $P(t)$ is thus defined for finite-size outbreaks,

$$\begin{aligned} P(t) &= \mathbb{P}[t < T < \infty] \sim \frac{1 - \nu^{-1}}{(1 + t)^{\nu-1} - 1} \\ &\sim \frac{1}{t^{\nu-1}} \quad \text{for large } t. \end{aligned} \quad (\text{B5})$$

The average duration can be calculated as

$$\langle t \rangle_{pl} \sim - \int_0^{T_c} t \frac{dP}{dt} dt \sim \begin{cases} T_c^{2-\nu} & \nu \in (1, \infty) \setminus \{2\}, \\ \ln T_c & \nu = 2, \end{cases} \quad (\text{B6})$$

where T_c is a cutoff time scale in the BDI process with finite system size. We would like to note that the technique of using the survival function is also applicable when $\nu < 1$ and it yields the desired scaling laws [Eq. (54)] when applied.

-
- [1] H. Andersson and T. Britton, *Stochastic Epidemic Models and Their Statistical Analysis*, Vol. 4 (Springer, New York, 2000).
- [2] E. Ben-Naim and P. Krapivsky, *Eur. Phys. J. B* **85**, 1 (2012).
- [3] J. P. Sethna, *Statistical Mechanics: Entropy, Order Parameters, and Complexity* (Oxford University Press, Oxford, 2006).
- [4] F. Brauer, P. Van den Driessche, J. Wu, and L. Allen, *Mathematical Epidemiology* (Springer, Berlin, 2008).
- [5] T. Antal and P. Krapivsky, *J. Stat. Mech.: Theor. Exp.* (2012) P07018.
- [6] J. Lloyd-Smith, D. George, K. Pepin, V. Pitzer, J. Pulliam, A. Dobson, P. Hudson, and B. Grenfell, *Science* **326**, 1362 (2009).
- [7] P. Rohani, R. Breban, D. E. Stallknecht, and J. M. Drake, *Proc. Natl. Acad. Sci. USA* **106**, 10365 (2009).
- [8] S. Singh, D. J. Schneider, and C. R. Myers, *Phys. Rev. E* **89**, 032702 (2014).
- [9] A. M. Zubkov, *Theory Probab. Appl.* **17**, 174 (1972).
- [10] N. Bailey, *The Elements of Stochastic Processes with Applications to the Natural Sciences* (Wiley-Interscience, New York, 1990).
- [11] S. Ong, *Metrika* **43**, 221 (1996).
- [12] D. G. Kendall *et al.*, *Ann. Math. Stat.* **24**, 338 (1953).
- [13] D. Shanbhag, *J. Appl. Probab.* **3**, 274 (1966).
- [14] J. Virtamo, Lecture Notes, Helsinki University of Technology (2005).
- [15] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in FORTRAN 77: Volume 1 of Fortran Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1992).
- [16] P. Flajolet and R. Sedgewick, *Analytic Combinatorics* (Cambridge University Press, Cambridge, 2009).
- [17] K. Christensen and N. R. Moloney, *Complexity and Criticality*, Vol. 1 (Imperial College Press, London, 2005).