EPIDEMIC SIZE IN THE SIS MODEL
OF ENDEMIC INFECTIONS

DAVID A. KESSLER,* *Bar-Ilan University*

Abstract

We study the susceptible-infected-susceptible model of the spread of an endemic infection. We calculate an exact expression for the mean number of transmissions for all values of the population size (N) and the infectivity. We derive the large-N asymptotic behavior for the infectivitiy below, above, and in the critical region. We obtain an analytical expression for the probability distribution of the number of transmissions, n, in the critical region. We show that this distribution has an $n^{-3/2}$ singularity for small n and decays exponentially for large n. The exponent decreases with the distance from the threshold, diverging to ∞ far below and approaching 0 far above.

Keywords: Infection; epidemic; SIS

2000 Mathematics Subject Classification: Primary 92D30
Secondary 60J70; 60J27

1. Overview

The susceptible-infected-susceptible (SIS) model of Weiss and Dishon [20] is one of the simplest models of endemic infections. The model describes the evolution of an infection in a fixed population, with no restriction on the possibility of reinfection of a previously infected and now recovered individual. This is in contrast to the venerable susceptible-infection-recovered (SIR) model [11], where reinfection is not permitted. Both models exhibit a threshold value of the infectivity, below which the infection immediately dies out. Below the threshold, where only a tiny fraction of the population is impacted, the two models have essentially equivalent statistical properties. Above the threshold, in the SIR model the infection is self-limiting, since in a fixed population the number of potential new victims, the susceptible pool, is monotonically decreasing in size. The SIS model, on the other hand, describes an endemic infection which can (above the threshold) persist indefinitely, at least at the deterministic level. Thus, the statistics of infection size in the two models above the threshold are very different.

The statistics of the mean time to extinction in the SIS model have been much studied, starting with the original paper of Weiss and Dishon [20], continuing with Kryscio and Lefévre [14] and Anderson and Djehiche [2], among others, and most recently, with Doering *et al.* [6]. In these latter papers the large population limit of the mean extinction time was investigated. The mean extinction time goes from a logarithmic dependence on the population size, N, below the threshold, to a $\sqrt{N}$ dependence exactly at the threshold, to an exponential dependence above the threshold. In this paper we will focus on the mean *number* of transmissions till extinction. This is a more pertinent method of characterizing the epidemic and the threshold transition. We will derive an exact formula for this quantity for general N and infectivity, and then examine its large-N asymptotics. As we will see, above the threshold, the mean

epidemic size is directly related to the mean epidemic duration. However, at and below the threshold, these quantities are quite different. Furthermore, the number of infection events is directly relevant when considering the probability of a mutation of the pathogen, as mutations are most probable during the exponential growth phase following a new infection [3]. These mutations are implicated in the conversion of a sub-threshold, weakly transmittable pathogen into a super-threshold variety capable of inducing a major epidemic.

Of particular interest will be the critical regime separating the above and below threshold cases. As already noted in [17], for a range of infectivities of width $1/\sqrt{N}$ around the threshold, there is a crossover region that interpolates between the above and below threshold cases. The existence of a large-N scaling theory in this region was recently proved in [7]. We will see this crossover region and its characteristic scaling arising naturally from our general result for the mean infection size.

After this treatment, dealing exclusively with the perhaps most biologically relevant case of a single initial infection, we extend our results to an arbitrary number of initial infections, again deriving an exact formula and then examining the large-N asymptotics. In the crossover regime we will have to distinguish the cases when the number of initial infections is small, comparable to, or much larger than $\sqrt{N}$.

From looking only at the mean number of infections we move on to consider the entire probability distribution for the number of infections. We first briefly discuss the above and below threshold cases, and then focus on the critical threshold regime. In the particular case of the infectivity exactly at the threshold value, the entire probability distribution for the appropriate scaling variable (the number of infections divided by N) can be explicitly displayed. In general, we can express the probability distribution as an inverse Laplace transform. This is sufficient to calculate the limiting behavior of the distribution for small and large epidemics, and to recover our expression for the mean in the critical regime. We then conclude with a few observations.

2. Preliminaries

We begin with a description of the SIS model. The N individuals in the population are divided into two subclasses: the susceptible pool, of size S, and the infected (and infectious) class, of size I, where $N = S + I$. The disease is transmitted from an infected individual to a susceptible individual with rate α/N, so that

$$(S, I) \overset{\alpha SI/N}{\to} (S - 1, I + 1).$$

Infected individuals recover with rate β, reverting back to susceptibles:

$$(S, I) \overset{\beta I}{\to} (S + 1, I - 1).$$

Of primary interest is the case where initially $S = N - 1$ and $I = 1$, so that the outbreak is sparked by a single infected individual. The outbreak terminates when the last infected individual recovers and I returns to 0.

This stochastic process is traditionally approximated (for large populations) by the rate equations

$$\dot{S} = -\frac{\alpha}{N}SI + \beta I, \qquad \dot{I} = \frac{\alpha}{N}SI - \beta I.$$

Using the conservation of N, we obtain

$$\dot{I} = (\alpha - \beta)I - \frac{\alpha}{N}I^2,$$

which is a logistic-type equation. We see that there is a transitition at $R_0 \equiv \alpha/\beta = 1$, where R_0 is equal to the mean number of primary infections caused in a large population of susceptibles by an infected individual. It is clear that, if $R_0 < 1$, the $I = 0$ state is stable, whereas, for $R_0 > 1$, the rate equation predicts a stable equilibrium state at

$$S_* = \left(\frac{\beta}{\alpha}\right)N \quad \text{and} \quad I_* = \left(1 - \frac{\beta}{\alpha}\right)N.$$

Thus, at the classical level, $R_0 = 1$ marks the threshold between an infection that becomes endemic and those that fail to spread.

3. Mean number of infections

In the original Weiss–Dishon paper [20], an exact expression for the mean time to extinction, starting from the completely infected state, was derived. The generalization of this to an arbitrary number of initially infected individuals was given in [15] and rediscovered in [6]. However, the mean *number* of infections is the quantity of primary interest in an infection model. We can focus on this quantity if we eliminate time, considering only the transitions between states. We characterize the system by the number, T, of transitions the system has undergone. In each transition the number of infected individuals either rises or falls by 1, so that I undergoes a kind of random walk. The probability of an upward transition is

$$p^+ = \frac{R_0 S}{R_0 S + N} = \frac{R_0(N - I)}{R_0(N - I) + N},$$

whereas the probability of a downward transition is $p^- = 1 - p^+$. These probabilities are unequal and depend on I, so that the walk is biased, with a 'space'-dependent drift. (From here on, we will refer to T as Time, with the lower-case word 'time' retaining its usual meaning, and trust that this will not lead to confusion.) It is easy to see that, at the point of extinction, the total number of infections, including the initial n_o infected individuals, is just $n = (T_{\text{ext}} + n_o)/2$. The number of induced infections is of course n_o smaller. Since the results of Weiss and Dishon [20], Leigh [15], and Doering *et al.* [6] for the mean time to extinction apply to a general one-step random walk, we can apply them directly to calculate the mean number of infections. Specializing to the case where we initially have exactly one infected individual, we have, for the mean extinction Time, τ_1,

$$\tau_1 = \sum_{j=1}^{N} \frac{1}{p_j^-} \prod_{k=1}^{j-1} \frac{p_k^+}{p_k^-}.$$

Here we have explicitly indicated the dependence of the transition probabilities on I via a subscript. Substituting these probabilities into the expression for τ_1, we find that

$$\tau_1 = \sum_{j=1}^{N} \left(1 + R_0 - \frac{R_0 j}{N}\right) \frac{R_0^{j-1}(N-1)!}{N^{j-1}(N-j)!}.$$

Reordering the sum, we can rewrite this as

$$\tau_1 = \left(\frac{R_0}{N}\right)^{N-1} (N-1)! \sum_{n=0}^{N-1} \left(1 + \frac{R_0 n}{N}\right) \frac{1}{n!} \left(\frac{N}{R_0}\right)^n.$$

We can do better, since the second term in the sum is the same as the first, except for the last index, so τ_1 simplifies to

$$\tau_1 = \left(\frac{R_0}{N}\right)^{N-1} (N-1)! \left(2 \sum_{n=0}^{N-1} \frac{1}{n!} \left(\frac{N}{R_0}\right)^n - \frac{1}{(N-1)!} \left(\frac{N}{R_0}\right)^{N-1}\right)$$

$$= 2\left(\frac{R_0}{N}\right)^{N-1} (N-1)! \sum_{n=0}^{N-1} \frac{1}{n!} \left(\frac{N}{R_0}\right)^n - 1,$$

giving for $\bar{n}_1$, the mean number of infections starting from an infected individual,

$$\bar{n}_1 = \left(\frac{R_0}{N}\right)^{N-1} (N-1)! \sum_{n=0}^{N-1} \frac{1}{n!} \left(\frac{N}{R_0}\right)^n.$$

We recognize the sum as the first N terms of the Taylor expansion of the exponential $\exp(N/R_0)$. The behavior of the sum depends on whether R_0 is above or below 1. This follows from the fact that the terms in the Taylor expansion of e^x increase until $n = x$, and then decrease. For large x, in fact, the behavior of the terms with n is a Gaussian peaked at $n = x$. The behavior of the sum is then determined by whether the last term of the sum at $n = N - 1$ occurs before or after the peak at $n = N/R_0$, i.e. whether R_0 is above or below 1.

For R_0 above 1, the summed terms extend past the peak, which dominates the sum, and so, up to exponentially small corrections, the sum is just the exponential. Furthermore, the prefactor can be approximated via Stirling's formula, giving

$$\bar{n}_1 \approx \frac{\sqrt{2\pi N}}{R_0} \exp\left(N\left(\ln(R_0) + \frac{1}{R_0} - 1\right)\right). \tag{1}$$

Thus, as expected, the mean number of infected cases grows exponentially large with N, with the exponent going to 0 as R_0 approaches 1. Furthermore, the exponent is the same as for the mean first passage time (here actual time), as calculated in [6], and is equal to the action for the semiclassical extinction trajectory [8], [12]. This is because, above the threshold, the system remains an exponentially long time in the classically stable state. We plot $\bar{n}_1$ versus R_0 in Figure 1, together with the large-N asymptotic formula, (1). We see that the mean number of infections quickly grows to astronomical proportions as R_0 increases away from 1. To see the approach to the large-N result, we show in the left-hand part of Figure 1 the ratio of the exact results for $N = 100$ and 400 to the large-N asymptotic formula. We see that the approximate formula works excellently, except in the vicinity of the transition point, $R_0 = 1$, and improves with increasing N.

For R_0 below 1, the sum is cut off while the terms are still increasing with n, and so the largest terms in the sum are the last ones, which approximate a geometric series:

$$\bar{n}_1 \approx \sum_{k=0}^{\infty} R_0^k = \frac{1}{1 - R_0},$$

which of course is the same as in the SIR model, since the number of infected individuals is so small that no one gets a multiple infection. This infinite-N answer is compared to the finite-N results of Figure 2, where we see that it works well as long as we are sufficiently below $R_0 = 1$, and the range of agreement increases with N. As opposed to the supercritical case, here there
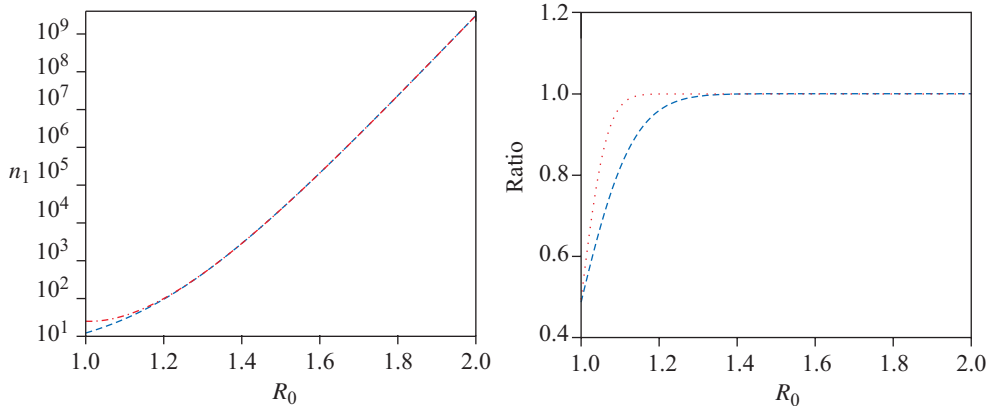
FIGURE 1: Supercritical regime ($R_0 > 1$). *Left*: mean total number of infections starting from one infected individual, $\bar{n}_1$, versus the infectivity parameter, R_0. The case of population size $N = 100$ (*dashed line*) is shown, together with the large-N asymptotic formula, (1) (*dash–dot line*). *Right*: ratio of the exact mean number of infections, $\bar{n}_1$, to that given by the large-N asymptotic formula, (1), versus the infectivity parameter, R_0. The cases of population size $N = 100$ (*dashed line*) and 400 (*dotted line*) are shown.
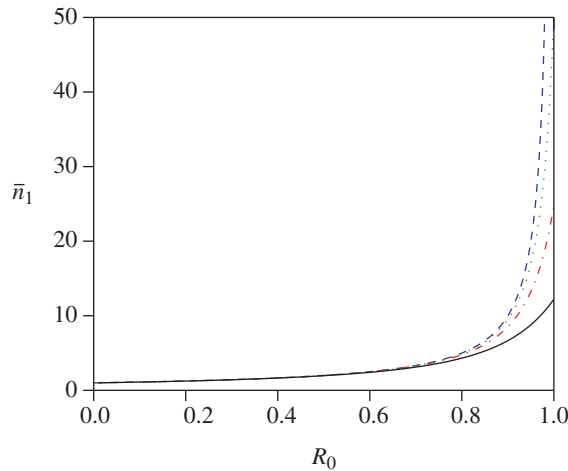

FIGURE 2: Subcritical regime ($R_0 < 1$). Mean total number of infections when starting with one infected individual, $\bar{n}_1$, versus the infectivity parameter, R_0, together with the prediction for an infinite population. Shown are the cases of population size $N = 100$ (*solid line*), 400 (*dash–dot line*), 1600 (*dotted line*), and ∞ (*dashed line*).

is in general no simple relation between the mean number of infections and the mean (actual) time to extinction, which is given by

$$\bar{t}_1 = \sum_{k=1}^{N} \frac{1}{\beta k} \left(\frac{R_0}{N} \right)^{k-1} \frac{(N-1)!}{(N-k)!} \approx \sum_{k=1}^{\infty} \frac{1}{\beta k} R_0^{k-1} = -\frac{1}{\beta R_0} \ln(1 - R_0).$$

This is of course due to the fact that in the subcritical case the number of infections is not strongly peaked about some value, as it was in the supercritical case.
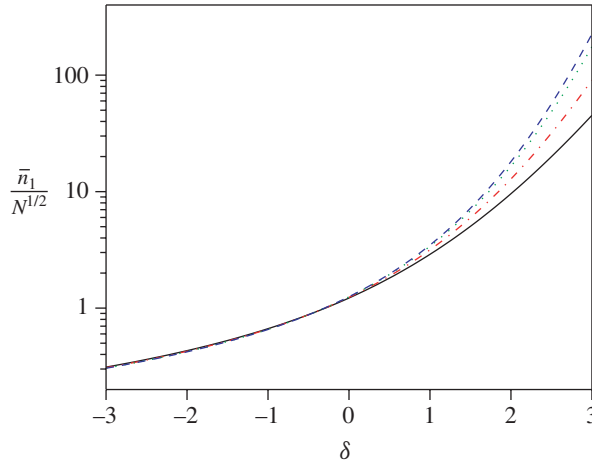
FIGURE 3: Critical regime ($R_0 = 1 + \delta/\sqrt{N}$). Scaled mean total number of infections, $\bar{n}_1/\sqrt{N}$, versus the scaled infectivity parameter, δ, when starting with one infected individual. The cases of population size $N = 100$ (*solid line*), 400 (*dash–dot line*), and 6400 (*dotted line*) are shown, along with the large-N asymptotic result, (2) (*dashed line*).

For R_0 near 1, $R_0 = 1 + \delta N^{-1/2}$, with δ of order 1, there is a transition region. The dominant terms in the sum are again the last ones, which have a Gaussian character, with a maximum at $N/R_0 \approx N$, so that

$$
\begin{aligned}
\bar{n}_1 &= \sum_{k=0}^{N-1} \left(\frac{R_0}{N}\right)^k \frac{(N-1)!}{(N-1-k)!} \\
&= \sum_{k=0}^{N-1} R_0^k \prod_{j=1}^{k} \left(1 - \frac{j}{N}\right) \\
&\approx \sum_{k=0}^{N-1} \exp\left(k\left(\frac{\delta}{\sqrt{N}}\right) - \frac{k^2}{2N}\right) \\
&\approx \int_0^\infty dk \, \exp\left(k\left(\frac{\delta}{\sqrt{N}}\right) - \frac{k^2}{2N}\right) \\
&= \sqrt{\frac{\pi N}{2}} \exp\left(\frac{\delta^2}{2}\right) \left(1 + \text{erf}\left(\frac{\delta\sqrt{2}}{2}\right)\right).
\end{aligned}
\tag{2}
$$

This clearly reproduces the subcritical and supercritical results in the limit of large negative and positive δ, respectively. This formula is plotted in Figure 3 along with data for $N = 100$, 400, and 6400. We see that the finite-N data converge to the infinite-N prediction, with the finite-N effects being larger at larger δ.

4. Mean number of infections: general initial condition

These results are easily generalized to the case of n_o initially infected individuals, again starting from the corresponding mean first passage Time. We obtain $\bar{n}_{n_o} = \bar{n}_1 + \sum_{k=2}^{n_o} \sigma_k$, where
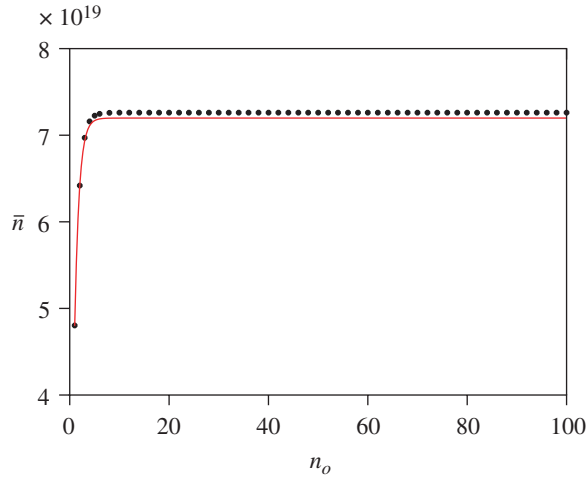
FIGURE 4: Exact calculation of $\bar{n}$ as a function of n_o for the supercritical case, $R_0 = 3$ (*filled circles*), together with the analytical approximation, (3) (*solid line*). Here $N = 100$.

$$\sigma_k \equiv \left(\frac{R_0}{N}\right)^{N-k} (N-k)! \sum_{j=0}^{N-k} \frac{1}{j!}\left(\frac{N}{R_0}\right)^j.$$

Again, these results are instructive in the various limits. For the above threshold case, $\sigma_k \approx R_0^{1-k}\bar{n}_1$, so that

$$\bar{n}_{n_o} \approx \left(1 + \frac{1}{R_0} + \frac{1}{R_0^2} + \cdots + \frac{1}{R_0^{n_o-1}}\right)\bar{n}_1$$

$$\approx (1 - R_0^{-n_o})\left[\frac{\sqrt{2\pi N}}{R_0 - 1}\exp\left(N\left(\ln(R_0) + \frac{1}{R_0} - 1\right)\right)\right]. \tag{3}$$

The prefactor is recognized as the probability of a biased random walk starting at n_o to survive forever. Thus, the mean number of infections is equal to the mean number of infections starting in the macroscopically infected metastable state multiplied by the probability of surviving long enough to reach this state. The expression in square brackets, the mean number of infections starting in the metastable state, $\hat{n}_{(R_0-1)/R_0}$, is itself simply related to the mean first passage time for this initial state calculated in [6]. If we account for the average time for a transition in the metastable state, $2(R_0 - 1)/N R_0$, we can easily obtain from the above the average Time to extinction, which is twice the average number of infections. This is because the overwhelming majority of the time is spent in the vicinity of the metastable state. In Figure 4 we present the exact results for $\bar{n}_{n_o}$ for the case in which $N = 100$ and $R_0 = 3$ together with our approximation, (3). We see that the agreement is quite satisfactory.

For R_0 below threshold,

$$\sigma_k \approx \sum_{j=0}^{\infty}\left(\left(1 - \frac{k}{N}\right)R_0\right)^j = \frac{1}{1 - (1 - k/N)R_0}.$$

Thus, for $n_o \ll N$, all the σs are approximately equal to $\bar{n}_1$, so that $\bar{n}_{n_o} \approx n_o\bar{n}_1$. This is clear, as the individual seeds act essentially independently, since they impact an infinitesimal fraction of
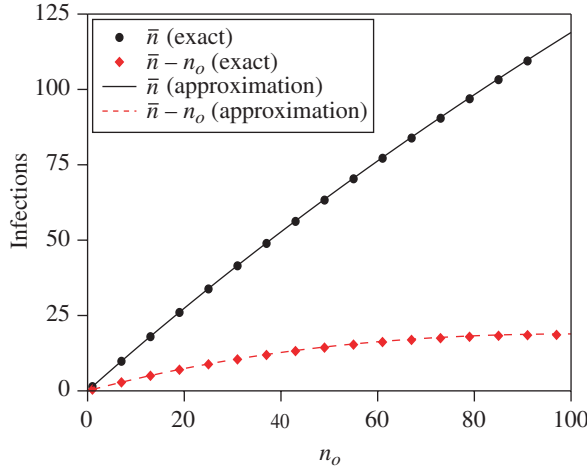
FIGURE 5: Exact calculation of $\bar{n}$ as a function of n_o for the subcritical case, $R_0 = 0.3$, together with the analytical approximation, (4). Also shown is the average number of *induced* infections, $\bar{n} - n_o$. Here $N = 100$.

the total population. This is in sharp contrast to the above threshold case, where $\bar{n}_{n_o}$ converges exponentially quickly over an $\mathcal{O}(1)$ range of n_o. For larger n_o of order N, we obtain

$$\bar{n}_{n_o} \approx \int_0^{n_o} \sigma_k \, dk = \frac{N}{R_0} \ln\left(\frac{1 - (1 - n_o/N)R_0}{1 - R_0}\right). \tag{4}$$

For the subcritical case, it is also interesting to consider the number of *induced* infections, since here most of the infections are just those of the initial state. For small n_o, the number of induced infections is approximately $n_o R_0/(1 - R_0)$, again proportional to n_o. For $n_o = N$, the number of induced infections is $(N/R_0) \ln(1/(1 - R_0)) - N$, which, for small R_0, is approximately $N R_0/2$. Thus, the interference between different initial seeds reduces the number of induced infections roughly by half in this case. The interference effect is of course even more dramatic for larger R_0. As R_0 approaches unity, the number of induced infections diverges only logarithmically for $n_o = N$, as opposed to the $1/(1 - R_0)$ divergence for small n_o. Of course, for R_0 even larger, in the supercritical regime, as we have seen, the interference effect is almost total, as increasing n_o beyond 10 or so has essentially no effect. The subcritical case is demonstrated in Figure 5 for the case in which $R_0 = 0.3$.

In the critical regime, things are of course a bit more complicated. Again, we first compute the σ_k:

$$\begin{aligned}
\sigma_k &= \sum_{j=0}^{N-k} R_0^j \prod_{\ell=0}^{j-1}\left(1 - \frac{k+l}{N}\right) \\
&\approx \sum_{j=0}^{N-k} \exp\left(j\left(\frac{\delta}{\sqrt{N}} - \frac{k}{N}\right) - \frac{j^2}{2N}\right) \\
&\approx \sqrt{\frac{\pi N}{2}} \exp\left(\frac{\delta - k/\sqrt{N}^2}{2}\right)\left(1 + \mathrm{erf}\left(\frac{\sqrt{2}}{2}\left(\delta - \frac{k}{\sqrt{N}}\right)\right)\right).
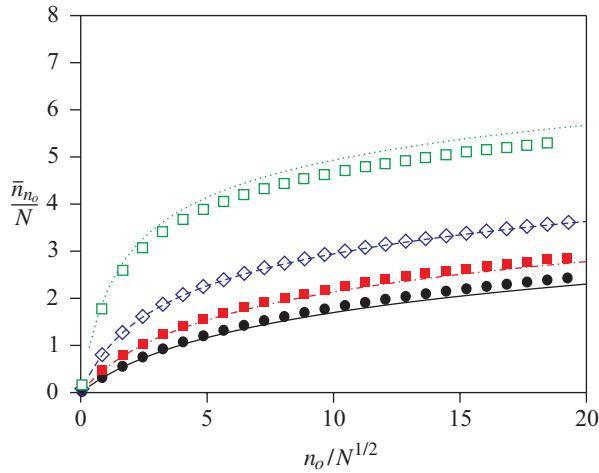\end{aligned}$$

FIGURE 6: Scaled average infection size $\bar{n}_{n_o}/N$ as a function of the scaled initial number of infected individuals, $n_o/\sqrt{N}$, for $N = 400$ and $\delta = -2$ (*filled circles*), -1 (*filled squares*), 0 (*open diamonds*), and 1 (*open squares*) ($R_0 = 0.9$, 0.95, 1, and 1.05, respectively). Also shown is the analytical approximation for these four cases: $\delta = -2$ (*solid line*), $\delta = -1$ (*dash–dot line*), $\delta = 0$ (*dashed line*), and $\delta = 1$ (*dotted line*).

We now have to integrate this with respect to k, whose typical scale is $\mathcal{O}(\sqrt{N})$:

$$\bar{n}_{n_o} = N\sqrt{\frac{\pi}{2}} \int_0^{n_o/\sqrt{N}} dx \, \exp\left(\frac{(\delta - x)^2}{2}\right)\left(1 + \text{erf}\left(\frac{\sqrt{2}}{2}\left(\delta - x\right)\right)\right). \quad (5)$$

One immediate result is that $\bar{n}_{n_o}$ starts out as $\mathcal{O}(\sqrt{N})$ for small n_o of order unity, but, for n_o of order $\mathcal{O}(\sqrt{N})$, the average number of infections rises to $\mathcal{O}(N)$. In Figure 6 we present the exact results for $\bar{n}_{n_o}$ in the critical region for $N = 400$ versus our scaling prediction, (5). We see that the scaling results are perfect for the exactly critical case, and the further we are from criticality, the larger the finite-N effects. Furthermore, the finite-N effects are larger for positive δ than for negative δ. Also, the larger the initial infection size, the larger the finite-N effects.

To better understand our scaling formula, we consider in turn the cases in which δ is large and negative, δ is of order unity, and δ is large and positive. In the former case, the argument of the erf is large and negative, and so

$$\bar{n}_{n_o} \approx \sqrt{N} \int_0^{n_o} dk \, \frac{1}{|\delta| + k/\sqrt{N}} = N \ln\left(1 + \frac{n_o}{|\delta|\sqrt{N}}\right).$$

Thus, $\bar{n}_{n_o}$ crosses over from a linear behavior at n_o of order unity to a logarithmic growth when n_o is of order $\sqrt{N}$. For δ of order unity, we note that $\bar{n}_{n_o}$ is given by a double sum, a sum of σ_k over k, each of which is in turn a sum over j. It is convenient to reverse the order of summation. The sums can then be approximated by integrals, and we integrate first with respect to k and then evaluate the integral over j. This gives the formula

$$\bar{n}_{n_o} = N \int_0^\infty dj \, \exp\left(-\frac{j^2}{2N}\right) e^{\delta j} \frac{1 - \exp(-jn_o/N)}{j}.$$

Again, for n_o small compared to $\sqrt{N}$, the answer is proportional to n_o. For large n_o, this can

be approximated as follows:

$$\bar{n}_{n_o} \approx \lim_{\varepsilon \to 0^+} N \left(\int_0^\infty \mathrm{d}j \exp\left(-\frac{j^2}{2N}\right) j^{\varepsilon-1} e^{j\delta} \right.$$
$$\left. - \int_0^\infty \mathrm{d}j \exp\left(-\frac{jn_o}{N}\right) j^{\varepsilon-1} (1 + j\delta + \cdots) \left(1 - \frac{j^2}{2} + \cdots\right)\right)$$
$$\approx N \left(\frac{1}{2}\gamma + \ln\left(n_o\sqrt{\frac{2}{N}}\right) + A(\delta) - \frac{\delta\sqrt{N}}{n_o} - \frac{N(\delta^2 - 1)}{2n_o^2}\right), \tag{6}$$

where $\gamma = 0.5772\ldots$ is the Euler constant and $A(\delta)$ is given by

$$A(\delta) \equiv \sqrt{\frac{\pi}{2}} \int_0^\delta \exp\left(\frac{t^2}{2}\right) \mathrm{d}t + \int_0^\infty \frac{\cosh k\delta - 1}{k} \exp\left(-\frac{k^2}{2}\right) \mathrm{d}k.$$

Here A is a monotonically increasing function of δ with $A(0) = 0$. For large negative δ, $A(\delta) \approx -\ln(\delta) - \gamma/2 - \ln(2)/2$, reproducing our previous result. For large positive δ, A grows quickly, $A(\delta) \approx \sqrt{2\pi} \exp(\delta^2/2)$. Thus, for all δ, $\bar{n}$ grows logarithmically in n_o for sufficiently large $n_o/\sqrt{N}$. However, since A grows so rapidly with δ, for large positive δ, this behavior is not readily visible in practice, since n_o can be no bigger than N. In Figure 7 we show $\bar{n}_{n_o}$ for different δs. We see that the large n_o approximation works well for $n_o/\sqrt{N}$ larger than 1 or so. For large positive δ, the argument of the erf is large and positive, and so the behavior for fixed $n_o/\sqrt{N}$ is most relevant. Then

$$\bar{n}_{n_o} \approx \sqrt{2\pi N} \exp\left(\frac{\delta^2}{2}\right) \int_0^{n_o} \mathrm{d}k e^{-\delta k/\sqrt{N}} = \frac{N\sqrt{2\pi} \exp(\delta^2/2)(1 - \exp(-\delta n_0/\sqrt{N}))}{\delta}. \tag{7}$$

The problem with this expression is that it is not at all accurate until δ is fairly large, around 6 or so. For such large δs, the concept of a critical region does not really apply until large Ns.
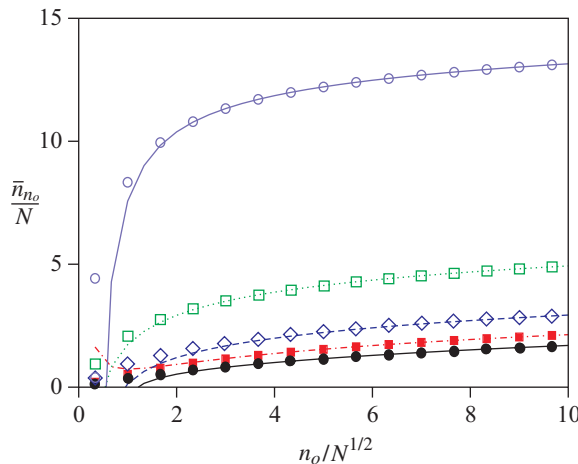

FIGURE 7: Calculation of the scaling form of the scaled mean epidemic size $\bar{n}_{n_o}/N$ as a function of the scaled initial epidemic size, $n_o/\sqrt{N}$, in the critical regime for intermediate $\delta = -2$ (*filled circles*), -1 (*filled squares*), 0 (*open diamonds*), 1 (*open suqares*), and 2 (*open circles*). Also plotted is the large n_o analytic approximation, (6), for $\delta = -2$ (*lower solid line*), -1 (*dash–dot line*), 0 (*dashed line*), 1 (*dotted line*), and 2 (*upper solid line*).
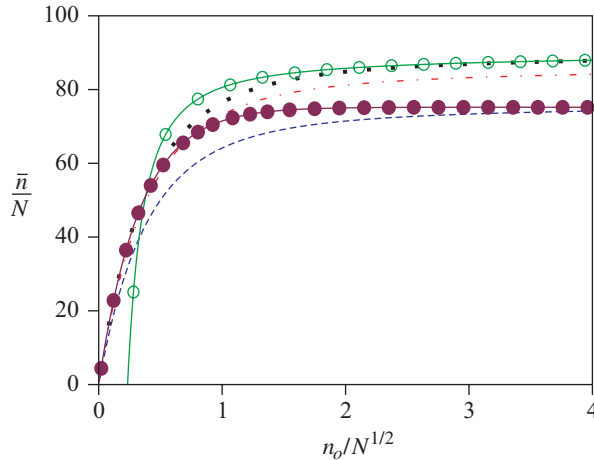
FIGURE 8: Scaling average epidemic size $\bar{n}/N$ as a function of the scaled initial epidemic size n_o for $N = 10\,000$ (*dashed line*) and $160\,000$ (*dash–dot line*) for the above critical case, $\delta = 3$. Also shown is the analytic critical scaling form, (5) (*dotted line*), and its large-n_o limit, (6) (*solid line with open circles*), and large-δ limit, (7) (*solid line with filled circles*).

For example, the next order correction to the action is of order $\delta^3/N^{1/2}$, which is only small for $N \sim \delta^6$. We can see this in Figure 8, where we examine the convergence of the finite-N results to the critical scaling result for $\delta = 3$. Only for $N = 160\,000$ is $\bar{n}$ approaching its limiting scaling form. This form is well approximated by our large n_o formula, (6), for $n_o/\sqrt{N} > 2$ and by the large δ formula, (7), for $n_o/\sqrt{N} < \frac{1}{2}$.

5. Distribution of the total number of infections

We now turn to investigate the probability distribution $P(n)$ of the total number of infections, n, returning again to the case of a single initial infected individual, $n_o = 1$. For the subcritical and supercritical cases, these are fairly simple. In the subcritical case, almost surely, the infection goes extinct before the percentage of infected individuals, I/N, is significant. In this case, the up transition probabilities are essentially constant: $p_k^+ = R_0(N - I)/(R_0(N - I) + N) \approx R_0$ and $p_k^- = 1 - p_k^+ \approx 1 - R_0$. Therefore, the random walk reduces to that of a constant leftward bias, with $P(n)$ falling exponentially with n:

$$P(n) = \frac{R_0^{n-1}}{(1 + R_0)^{2n-1}} \left(\binom{2n - 2}{n - 1} - \binom{2n - 2}{n} \right)$$

$$\approx \frac{1}{\sqrt{\pi n^3}} \frac{(4R_0)^{n-1}}{(1 + R_0)^{2n-1}}, \qquad n \gg 1. \tag{8}$$

Above the threshold, the system spends an exponentially long time in the metastable state. Thus, $P(n)$, for macroscopic ns, is an exponential distribution, the waiting time distribution for the decay of the metastable state. For small epidemics, where the space-dependent drift is not yet relevant, the system can again be approximated by a random walk with constant bias, this time to the right. Thus, for $n \ll N$, $P(n)$ is given by (8). For larger N, $P(n)$ crosses over to a pure-exponential decay, normalized to $1 - 1/R_0$, the probability of the infection surviving to macroscopic size. This behavior is exhibited in Figure 9. It is important to note the difference
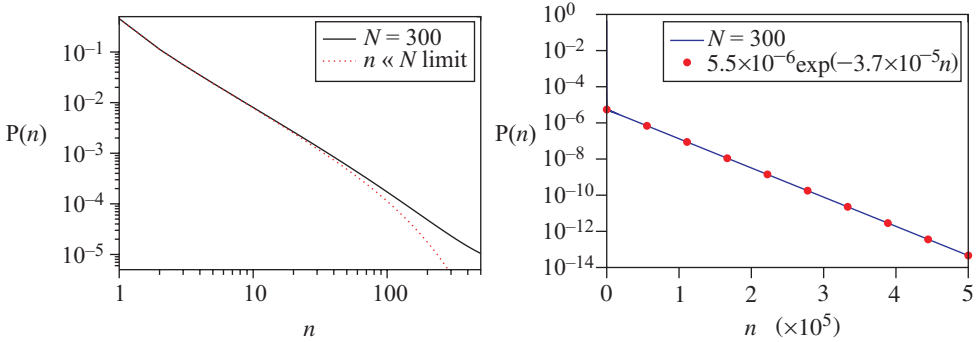
FIGURE 9: Probability distribution P(n) for the total size of the epidemic in the supercritical case, $R_0 = 1.2$ and $N = 300$. *Left*: behavior for $n < N$, together with the constant bias approximation, (8). The beginning of the crossover to exponential behavior is visible for the largest ns. *Right*: behavior for n on the scale of the average epidemic size, together with an exponential fit. The normalization is seen to be $(5.5 \times 10^{-6})/(3.7 \times 10^{-5}) = 0.2$, consistent with our expectation of $1 - 1/1.2 = 0.17$.

between this behavior and that exhibited above the threshold in the SIR model [13], [16], [19]. There the distribution has a second peak (in addition to the one at the origin) at the number of infections predicted by the deterministic dynamics. In the SIS model, the total number of infections above the threshold predicted by the deterministic dynamics is infinite. Rather, the behavior for major epidemics is the pure-exponential waiting-time distribution, with its peak at the origin.

We now turn to investigate the behavior in the crossover regime, $R_0 = 1 + \delta/\sqrt{N}$. The essential simplification here is in the transition probabilities, which in this regime can be approximated by

$$p_\pm = \frac{1}{2} \mp \frac{I}{4N} \pm \frac{\delta}{4\sqrt{N}}.$$

The critical regime is characterized by the scaling $I \sim \mathcal{O}(N)$, so that the bias is small, of order $1/\sqrt{N}$. For relatively small Times ($n \ll \sqrt{N}$), the bias is irrelevant, and the problem reduces to the unbiased random walk, given by substituting $R_0 = 1$ into (8):

$$P(n) = \frac{1}{2^{2n-1}} \left(\binom{2n-2}{n-1} - \binom{2n-2}{n} \right) \approx \frac{1}{\sqrt{4\pi n^3}}, \qquad n \gg 1. \tag{9}$$

We now study how, for larger Times, the bias, resulting from the reduction of the susceptible pool with increasing I and the small deviation from criticality, modifies this answer.

As the bias is very weak, however, and only effective at large Times, we are justified in passing to a Fokker–Planck description for the probability distribution $K(I, T; I_o)$ for I, given that there were I_o infected individuals at $T = 0$. For a typical major epidemic, T at extinction is of order N, this is despite the fact that $\bar{n}$ is of order $\sqrt{N}$, since the probability of a major epidemic is of order $1/\sqrt{N}$. We thus define $t \equiv T/2N$ and $x \equiv I/\sqrt{N}$, and consider $K(x, t; x_o)$:

$$\frac{\partial}{\partial t} K(x, t; x_o) = \frac{\partial^2}{\partial x^2} K + \frac{\partial}{\partial x} x K - \delta \frac{\partial K}{\partial x} \tag{10}$$

with the initial condition $K(x, t; x_o) = \delta(x - x_o)$ and the absorbing boundary condition $K(0, t; x_o) = 0$. In terms of K, the probability distribution for the epidemic size, P(n), is given
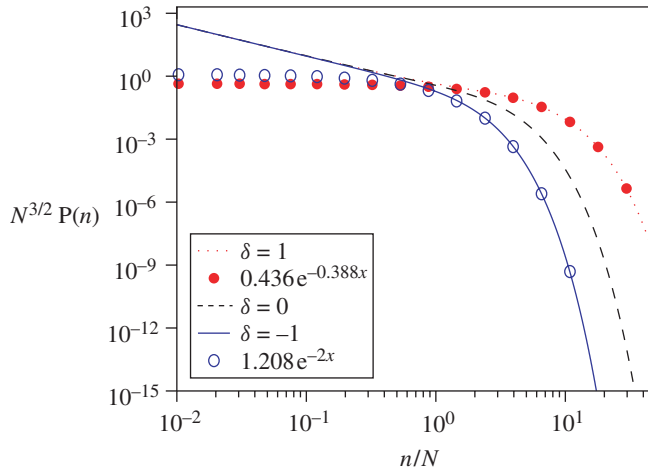
FIGURE 10: The scaled probability density of epidemics of size n in the critical regime, $N^{3/2} P(n)$, for the cases $\delta = -1$, 0, and 1, as a function of the scaled epidemic size n/N, along with the asymptotic large-n/N behavior for the $\delta = \pm 1$ cases.

by

$$P(n) = N^{-3/2} \frac{\partial^2}{\partial x \partial x_o} K\left(x, \frac{n}{N}; x_o\right)\bigg|_{x=x_o=0}, \tag{11}$$

since $x_o = I_o/\sqrt{N} = 1/\sqrt{N}$. This critical regime distribution function, obtained from a numerical solution of (10) and (11), is presented in Figure 10 for $\delta = -1$, 0, and 1. We see that, for small n, all three curves collapse into a universal power law. For large n, the distributions fall off rapidly, with the speed of falloff decreasing with increasing δ.

To make more progress, we recognize (10) as the imaginary-time Schrödinger equation for a harmonic oscillator potential [18, Chapter 7], after factoring out of K the square root of the Wronskian, in this case a Gaussian. Defining

$$K(x, t; x_o) = \exp\left(\frac{(x - x_o)\delta}{2} - \frac{x^2 - x_o^2}{4} + \frac{t}{2}\right) G(x, t; x_0),$$

we have

$$\dot{G} = G'' - \tfrac{1}{4}(x - \delta)^2 G$$

with $G(x, 0; x_o) = \delta(x - x_o)$ and $G(0, t; x_0) = 0$, in terms of which

$$P(n) = N^{-3/2} e^{n/2N} \frac{\partial^2}{\partial x \partial x_o} G\left(x, \frac{n}{N}; x_o\right)\bigg|_{x=x_o=0}.$$

The only complication is the presence of the absorbing boundary condition at $x = 0$, which breaks the reflection symmetry of the potential around $x = \delta$. Exactly at the threshold, $\delta = 0$, however, this is not a problem, as the boundary condition can be enforced by the method of images (see, for example, [10, Chapter 2]). The solution is

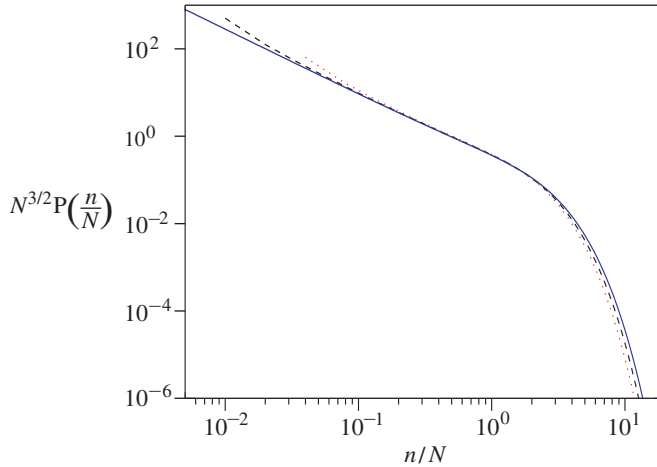$$G(x, x_o, t) = A(t) \sinh(2c(t)x x_m(t)) \exp(-c(t)(x^2 - x_m(t)^2)),$$

FIGURE 11: The analytic scaling solution for the threshold $P(n)$, (12) (*solid line*), together with data for $N = 25$ (*dotted line*) and 100 (*dashed line*).

where

$$c(t) = \tfrac{1}{4}\coth t,$$

$$x_m(t) = \frac{x_o}{\cosh t},$$

$$A(t) = \frac{1}{\sqrt{\pi}}\sinh^{-1/2} t \exp\left(-\frac{x_o^2}{4}\tanh t\right).$$

This gives the probability distribution

$$P(n) = \frac{1}{\sqrt{4\pi N^3}} e^{n/2N} \sinh^{-3/2}\left(\frac{n}{N}\right). \tag{12}$$

This clearly reproduces the expected behavior, (9), for $1 \ll n \ll N$, and then decays exponentially for $n \sim N$. This result is shown in Figure 11, together with data for $N = 25$ and 100. Even for these small Ns, the agreement is excellent, except at the smallest ns, where the discreteness of n factors in.

Knowing $P(n)$ gives us another way to calculate the mean epidemic size, $\bar{n}_1$:

$$\begin{aligned}
\bar{n}_1 &= \int_0^\infty n\, P(n) \\
&= \sqrt{\frac{2N}{\pi}} \int_0^\infty dx\, \frac{x e^{-x}}{(1 - e^{-2x})^{3/2}} \\
&= \sqrt{\frac{2N}{\pi}} \int_0^\infty dx\, \frac{e^{-x}}{\sqrt{1 - e^{-2x}}} \\
&= \sqrt{\frac{2N}{\pi}} \sin^{-1}(e^{-x})|_0^\infty \\
&= \sqrt{\frac{\pi N}{2}},
\end{aligned}$$

which of course agrees with our previous result, (2), specialized to $\delta = 0$.

We now return to the distribution of n for general δ. Formally, we can decompose $G(x, t; x_o)$ into a sum over eigenfunctions of the Schrödinger operator,

$$H = -\frac{d^2}{dx^2} + \frac{1}{4}(x - \delta)^2,$$

as follows:

$$G(x, x_o, t) = \sum_n \phi_n(x)\phi_n(x_o) \exp(E_n t).$$

If we define Green's function, $G(x, x', E)$, as usual by

$$G(x, x', E) = \sum_n \frac{\phi_n(x)\phi_n(x')}{E_n - E},$$

then $G(x, x', -E)$ is the Laplace transform of $G(x, x', t)$ with respect to time:

$$G(x, x', E) = \int_0^\infty dt e^{Et} G(x, x', t),$$

and $G(x, x', E)$ satisfies

$$HG - EG = \delta(x - x').$$

We can recover $G(t)$ from $G(E)$ by an inverse Laplace transform:

$$G(x, x', t) = \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} dE e^{-Et} G(x, x', E),$$

where γ lies to the left of all the poles of G. Defining $P(E)$ as

$$P(E) \equiv \frac{\partial^2}{\partial x \partial x'} G(x, x', E)\Big|_{x=x'=0},$$

we obtain the following expression for our desired probability distribution $P(n)$:

$$P(n) = \frac{1}{2\pi i N^{3/2}} \int_{\gamma-i\infty}^{\gamma+i\infty} e^{-(E-1/2)n/N} P(E).$$

Thus, all we need to do is calculate Green's function, $G(x, x', E)$. There is a nice formula relating $G(x, x', E)$ to $G^0(x, x', E)$, Green's function of the problem without the wall. Green's function for the system with a wall on the left-hand side of the system at $x = a$ is (see [9])

$$G(x, x', E) = G^0(x, x', E) - \frac{G^0(x, a, E)G^0(x', a, E)}{G^0(a, a, E)}.$$

Denoting $f(x, E)$ as the solution of $Hf = Ef$, which decays as $x \to +\infty$, and g as the solution which decays as $x \to -\infty$, then, in general,

$$G^0(x, x', E) = \frac{f(x_>)g(x_<)}{\text{Wr}[f, g]},$$

where $\mathrm{Wr}[f, g]$ is the Wronskian, and $x_>$ and $x_<$ denote the larger, respectively smaller, of x and x'. Then, using the fact that $\mathrm{Wr}[f, g]$ is constant,

$$
\begin{aligned}
\mathrm{P}(E) &\equiv \left.\frac{\partial^2}{\partial x \partial x'} G(x, x', E)\right|_{x=x'=0} \\
&= \frac{1}{\mathrm{Wr}[f, g]} \frac{\partial^2}{\partial x \partial x'}\left(f(x_>, E)\left(g(x_<, E) - \frac{f(x_<, E)g(0, E)}{f(0, E)}\right)\right)_{x=x'=0} \\
&= \frac{1}{\mathrm{Wr}[f, g]} f'(0, E)\left(g'(0, E) - \frac{f'(0, E)g(0, E)}{f(0, E)}\right) \\
&= \frac{f'(0, E)}{f(0, E)}.
\end{aligned}
$$

In our case, $f(x, E) = U(-E, x-\delta)$, where U is the parabolic cylinder function [1, Chapter 19] that decays for positive argument. Thus,

$$
\mathrm{P}(E) = \frac{U'(-E, -\delta)}{U(-E, -\delta)}.
$$

Green's function is also of course the moment generating function. Thus, already at this stage, we can use Green's function to recover the mean epidemic size. The answer is

$$
\begin{aligned}
\bar{n}_1 &= \int_0^\infty n \, \mathrm{P}(n) \, \mathrm{d}n \\
&= N^{-3/2} \frac{\partial^2}{\partial x \partial x'}\left(\int_0^\infty \mathrm{d}n e^{n/2N} n K\left(x, x', \frac{n}{N}\right)\right)_{x=x'=0} \\
&= N^{1/2} \frac{\partial^2}{\partial x \partial x'} \frac{\partial}{\partial E} G(x, x', E)\bigg|_{x=x'=0; E=1/2} \\
&= N^{1/2} \frac{\partial}{\partial E} \mathrm{P}(E)\bigg|_{E=1/2} \\
&= -N^{1/2} \frac{\partial^2}{\partial a \partial x} \ln(U(a, x))|_{a=-1/2, \, x=-\delta}.
\end{aligned}
$$

We thus need an expression for $U(a, x)$ for a near $-\frac{1}{2}$. We can obtain this by perturbing about the Gaussian solution at $a = -\frac{1}{2}$, writing

$$
\psi = A \exp\left(-\frac{x^2}{4}\right) + \psi_1,
$$

where ψ_1 satisfies the inhomogeneous equation

$$
-\psi_1'' + \frac{1}{4}\psi_1 - \frac{1}{2}\psi_1 = \delta_a A \exp\left(-\frac{x^2}{4}\right)
$$

and δ_a is the shift in a. Then, since the two modes of the homogeneous equation are $f_1 = \exp(-x^2/4)$ and $f_2 = \exp(-x^2/4)\int_0^x \mathrm{d}t \exp(t^2/2)$, the solution for ψ_1 that decays as $x \to +\infty$ is

$$
\psi_1 = -\delta_a f_1(x) \int_0^x \mathrm{d}x' A \exp\left(-\frac{x'^2}{4}\right) f_2(x') - \delta_a f_2(x) \int_x^\infty \mathrm{d}x' A \exp\left(-\frac{x'^2}{2}\right).
$$

The derivative with respect to a is then

$$
\frac{\partial}{\partial a} \ln(U(a, x))\Big|_{a=-1/2}
$$

$$
= \frac{\psi_1}{\delta_a \psi_0}
$$

$$
= -\int_0^x dx' \exp\left(-\frac{x'^2}{4}\right) f_2(x') - f_2(x) \exp\left(\frac{x^2}{4}\right) \int_x^\infty dx' \exp\left(-\frac{x'^2}{2}\right).
$$

We can now differentiate with respect to x and obtain

$$
\bar{n}_1 = -N^{1/2}\left(-\exp\left(-\frac{x^2}{4}\right) f_2(x) - \exp\left(\frac{x^2}{2}\right) \int_x^\infty dx' \exp\left(-\frac{x'^2}{2}\right)\right.
$$

$$
\left. + f_2(x) \exp\left(-\frac{x'^2}{4}\right)\right)\Big|_{x=-\delta}
$$

$$
= \sqrt{\frac{\pi N}{2}} \exp\left(\frac{\delta^2}{2}\right)\left(1 + \mathrm{erf}\left(\frac{\delta\sqrt{2}}{2}\right)\right),
$$

reproducing of course our previous result, (2).

The last order of business is to take the inverse Laplace transform. We have

$$
\mathrm{P}(n) = \frac{1}{2\pi i N^{3/2}} e^{n/2N} \int_{\gamma-i\infty}^{\gamma+i\infty} e^{an/N} \frac{U'(a, -\delta)}{U(a, -\delta)}.
$$

For the threshold case, $\delta = 0$, we have

$$
\frac{U'(a, 0)}{U(a, 0)} = -\sqrt{2}\frac{\Gamma(3/4 + a/2)}{\Gamma(1/4 + a/2)}.
$$

The integral can then be evaluated by residues, giving

$$
\mathrm{P}(n) = \sqrt{\frac{2}{\pi N^3}} \sum_{k=0}^\infty e^{-(2k+1)n} \frac{(2k+1)!!}{k!\, 2^k} = \sqrt{\frac{2}{\pi}} \frac{e^{-n}}{(1 - e^{-2n})^{3/2}},
$$

in agreement with our previous result. For general δ, however, we have to evaluate the integral numerically. The most important information however, the asymptotic behavior for small and large n, can be gleaned analytically.

We can obtain the small-n expansion of the distribution function by using the large-a expansion of the integrand. Using the Hankel contour integral,

$$
\frac{1}{2\pi i} \int_{\mathcal{C}} t^{-z} e^t \, dt = \frac{1}{\Gamma(z)},
$$

we have

$$
\begin{aligned}
P(n) &\approx e^{n/2N}\frac{1}{2\pi i N^{3/2}} \\
&\times \int_{\gamma-i\infty}^{\gamma+i\infty} da\, e^{an}\left(-\sqrt{a}-\frac{\delta^2}{8}a^{-1/2}+\frac{\delta}{8}a^{-1}+\left(-\frac{1}{16}+\frac{\delta^4}{128}\right)a^{-3/2}+\cdots\right) \\
&\approx N^{-3/2}e^{n/2N}\left(\frac{-1}{\Gamma(-1/2)}\left(\frac{n}{N}\right)^{-3/2}-\frac{\delta^2}{8\Gamma(1/2)}\left(\frac{n}{N}\right)^{-1/2}+\frac{\delta}{8\Gamma(1)}\left(\frac{n}{N}\right)^{0}\right. \\
&\left.\quad+\left(-\frac{1}{16}+\frac{\delta^4}{128}\right)\frac{1}{\Gamma(3/2)}\left(\frac{n}{N}\right)^{1/2}+\cdots\right) \\
&\approx \frac{e^{n/2N}}{\sqrt{4\pi n^3}}\left(1-\frac{\delta^2}{4}\frac{n}{N}+\frac{\delta\sqrt{\pi}}{4}\left(\frac{n}{N}\right)^{3/2}-\frac{1}{4}\left(1-\frac{\delta^4}{8}\right)\left(\frac{n}{N}\right)^{2}+\cdots\right).
\end{aligned}
$$

Thus, we find that the leading-order behavior at small n is $P(n)\approx(4\pi n^3)^{-1/2}$, independent of δ. We can also verify that this small-n series reproduces the full $P(n)$ for the $\delta=0$ case.

The large-n asymptotics is clearly given by the ground state of the wall problem. In the limit $\delta\to\infty$, the wall, relative to the bottom of the potential, moves to $-\infty$ and the ground state energy goes to $\frac{1}{2}$ in our units. This translates to a decay rate of 0, once the $\exp(n/2N)$ factor is taken into account. As δ decreases, the wall moves closer to the potential minimum, and the energy (and the decay rate) rises monotonically. The energy is $\frac{3}{2}$ in our units when the wall hits the potential minimum at $\delta=0$. This leads to the decay behavior $\exp(-n/N)$ at the threshold, in accord with the full solution in this case. The energy continues to rise as δ is decreased below 0, leading to a faster decay in n, with the energy diverging in the limit $\delta\to-\infty$.

For example, the second excited state of the harmonic oscillator has 0s located a distance $1/\sqrt{2\omega}=1$ to the right and left of the energy minimum. Thus, for $\delta=-1$, we expect the decay $\exp(-2n/N)$. Since the normalized eigenvector in our units is

$$
\phi(x)=\mathcal{N}(x^2-1)\exp\left(-\frac{x^2}{4}\right),
$$

where $\mathcal{N}^{-2}=2e^{-1/2}+\sqrt{2\pi}\,\mathrm{erfc}(\sqrt{2}/2)$ so that $(\phi'(1))^2=4\mathcal{N}^2e^{-1/2}\approx1.208$, then, for large n,

$$
P(n)\sim 1.208e^{-2n/N}.
$$

For $\delta=+1$, on the other hand, the solution of the Schrödinger problem is given by the first 0 of $U(a,-1)$ at $a_1=-0.888\,24=-E_1$, leading to a decay exponent of $0.388\,824$. The coefficient of the exponent is given by $U'(a_1,-1)/(\partial U(a_1,-1)/\partial a)=0.4365$. This can be seen in Figure 10, where the correct exponential falloff in both cases is seen for large n. For $n/N<2$, the effect of the higher eigenvectors leads to deviation from a pure-exponential behavior. In general, the large-n approximation is accurate as long as the next higher eigenvector has decayed. This 'energy gap' is approximately 1 for large positive δ, rises to 2 at $\delta=0$, and continues to rise as δ becomes more negative. However, as the ratio of the energy gap to the ground state energy falls as δ decreases, the role of the higher excited states becomes more pronounced as δ decreases.

Examining the limit of large positive δ in more detail, the energy is very slightly above $\frac{1}{2}$, so there is an extremely small decay rate. The actual ground state energy can be calculated as follows. We first shift x by δ, so that the quadratic potential is centered at the origin. Then,

since the ground state energy in the presence of the wall is close to the wall-free value of $\frac{1}{2}$, we can write

$$\psi = A \exp\left(-\frac{x^2}{4}\right) + \psi_1,$$

where ψ_1 satisfies the inhomogeneous equation

$$-\psi_1'' + \frac{1}{4}\psi_1 - \frac{1}{2}\psi_1 = \varepsilon A \exp\left(-\frac{x^2}{4}\right)$$

and ε is the shift in the energy. As before, the two modes of the homogeneous equation are $f_1 = \exp(-x^2/4)$ and $f_2 = \exp(-x^2/4)\int_0^x dt \exp(t^2/2)$. The solution for ψ_1 that decays as $x \to +\infty$ is

$$\psi_1 = -\varepsilon f_1(x)\int_0^x dx' A \exp\left(-\frac{x'^2}{4}\right) f_2(x') - \varepsilon f_2(x)\int_x^\infty dx' A \exp\left(-\frac{x'^2}{2}\right).$$

The second term dominates for large negative x, so that

$$\psi(x) \approx A \exp\left(-\frac{x^2}{4}\right) + \varepsilon A \frac{\sqrt{2\pi}}{x} \exp\left(\frac{x^2}{4}\right),$$

and then the boundary condition $\psi(-\delta) = 0$ gives

$$\varepsilon \approx \frac{\delta}{\sqrt{2\pi}} \exp\left(-\frac{\delta^2}{2}\right).$$

Then

$$\frac{d\psi}{da} = -\frac{d\psi_1}{d\varepsilon} = \frac{\sqrt{2\pi}}{\delta} A \exp\left(\frac{\delta^2}{4}\right)$$

and

$$\psi'(-\delta) = A\left(\frac{\delta}{2}\exp\left(-\frac{\delta^2}{4}\right) + \varepsilon\sqrt{\frac{\pi}{2}}\exp\left(\frac{\delta^2}{4}\right)\right) = A\delta\exp\left(-\frac{\delta^2}{4}\right).$$

For the leading-order asymptotics of P(n), this gives, for large n and large δ,

$$P(n) \approx \frac{\delta^2}{\sqrt{2\pi N^3}}\exp\left(-\frac{\delta^2}{2}\right)\exp\left(\frac{n}{N\sqrt{2\pi}}\delta\exp\left(-\frac{\delta^2}{2}\right)\right),$$

so that the total probability of a major epidemic is the integral of P(n), which is $\delta/\sqrt{N} \approx 1 - 1/R_0$, as expected, and $\bar{n} \approx \sqrt{2\pi}\exp(\delta^2/2)$, in accord with our previous result.

In the limit of large n and large negative δ, we find that the ground state energy is large, approximately $E \approx \delta^2/4$, so that $a \approx -\delta^2/4$. However, there are many states with approximately this energy. It is best to work directly from our integral formulation. For large negative δ, we can use a WKB ansatz (see, for example, [5, Chapter 10]) to write $U(a, x) \approx \exp(\sqrt{-a}\, S(x/\sqrt{-a}))$, which yields

$$\frac{U'(a, x)}{U(a, x)} = S' = -\sqrt{a + \frac{x^2}{4}}.$$

Expanding this in a power series for large a, we obtain

$$S'(-\delta) = -\sqrt{a} \sum_{k=0}^{\infty} \frac{\Gamma(3/2)}{k!\,\Gamma(3/2-k)} \left(\frac{\delta^2}{4a}\right)^k.$$

Performing the integral over a gives

$$
\begin{aligned}
P(n) &= -N^{-3/2} e^{n/2N} \sum_{k=0}^{\infty} \frac{\Gamma(3/2)}{k!\,\Gamma(3/2-k)\Gamma(k-1/2)} \left(\frac{\delta^2}{4}\right)^k \left(\frac{n}{N}\right)^{k-3/2} \\
&= -e^{n/2N} n^{-3/2} \sum_{k=0}^{\infty} \frac{\Gamma(3/2)\sin(\pi(k-1/2))}{k!\,\pi} \left(\frac{\delta^2 n}{4N}\right)^k \\
&= \frac{e^{n/2N}}{\sqrt{4\pi n^3}} \exp\left(-\frac{n\delta^2}{4N}\right).
\end{aligned}
$$

We see that we have successfully summed all the leading-order contributions. For large negative δ, this is cut off at ns of order $1/\delta^2$, so we may drop the $e^{n/2N}$ term. Then we have the result for a pure-constant drift, and so matches on to the subcritical distribution (which is the same as the SIR case). It, of course, reproduces the correct mean as well, since in the integral over n, small ns predominate, and

$$\bar{n} \approx \int_0^{\infty} dn\, n \exp\left(-\frac{n\delta^2}{4N}\right) \frac{1}{2\sqrt{\pi}n^{3/2}} = \frac{\sqrt{N}}{(-\delta)}.$$

This in turn matches on to the subcritical result, $\bar{n} \approx 1/(1-R_0)$, as R_0 approaches 1 from below.

An interesting subtlety arises if we consider the zeroth moment of the distribution. In normal circumstances this would be unity, but the scaling behavior of $P(n)$ dictates that the normalization integral is formally of order $N^{-1/2}$ and, furthermore, diverges, owing to the $n^{-3/2}$ behavior of P for small n. Nevertheless, if we blindly forge ahead, we find that

$$
\begin{aligned}
P_{\text{tot}} &= \int_0^{\infty} dn\, P(n) \\
&= N^{-1/2} P\left(\tfrac{1}{2}\right) \\
&= N^{-1/2} \frac{U'(-1/2,-\delta)}{U(-1/2,-\delta)} \\
&= \frac{\delta}{2\sqrt{N}}.
\end{aligned}
$$

Clearly, this finite answer is the result of an analytic continuation. To understand its significance, let us consider the difference between $P(n)$ and the distribution for the constant bias random walk with the same δ. For small n, these distributions as we have seen are identical, so the difference is integrable. Integrating the constant bias random walk, we obtain

$$P_{\text{tot}}^{\text{CB}} = \int_0^{\infty} dn\, \frac{1}{\sqrt{4\pi n^3}} \exp\left(-\frac{n\delta^2}{4N}\right) = -\frac{|\delta|}{2\sqrt{N}}.$$

Thus, the difference is $(\delta+|\delta|)/2\sqrt{N}$. This is 0 for $\delta \leq 0$, which is correct, since even without the space-dependent drift toward the origin, every walker will eventually hit the origin. On

the other hand, for $\delta > 0$, the difference is $\delta/\sqrt{N}$, which reflects the fact that, with the added space-dependent drift, all walkers are guaranteed to return to the origin, while without the added space-dependent drift, only a fraction $1/R_0 \approx 1 - \delta/\sqrt{N}$ of walkers return to the origin. Thus, looking at the difference between distributions provides an excellent way to make rigorous the concept of 'major epidemics', even slightly above the threshold in the critical regime. Even below the threshold, it highlights the added role of the space-dependent drift in reducing the probability of larger epidemics in favor of smaller epidemics.

6. Concluding discussion

We have exhibited an exact expression for the mean epidemic size in the SIS model of endemic infection. We have evaluated this in the limit of large population size, and shown the crossover behavior that occurs in the vicinity of the critical infection number, $R_0 = 1$. We have also calculated the distribution function for the epidemic size, again focusing on the crossover regime.

It is important to note that the crossover behavior is universal, independent of the details of the model. What is important is the existence of two fixed points of the rate equation dynamics and a critical parameter where the two fixed points interchange stability. In the case of the functional form of the mean epidemic size as a function of δ, the scaled distance to the critical point will be the same, along with the scaling behavior with N. This is also true in general for any first passage time problem where the transition rates are constant (independent of N and location) at the transition. The first passage time, i.e. the mean (physical) time to extinction, in the SIS model also exhibits a crossover behavior at the transition, albeit different than that of the mean epidemic size. This is due to the fact that the transition rates in time are location dependent, $p_k^+ = \alpha k(N - k)/N$ and $p_k^- = \beta k$. As Doering *et al.* [6] did not investigate the transition behavior of the mean extinction time, we for completeness present it here. The mean extinction time (starting with one infected individual) in the crossover regime is given by

$$\bar{t}(\delta) \approx \sum_{k=1}^{\infty} \frac{1}{\beta k} \exp\left(-\frac{k^2}{2N} + \frac{\delta k}{\sqrt{N}}\right)$$

$$\approx \gamma + \frac{1}{2} \ln N + \int_0^{\infty} \ln x \exp\left(-\frac{x^2}{2} + \delta x\right)(x - \delta)\,\mathrm{d}x.$$

In particular, at the threshold, the integral can be performed analytically and we obtain

$$\bar{t}(\delta = 0) \approx \tfrac{1}{2}(\ln 2N + \gamma).$$

The SIR model, as has been demonstrated in [4], [13], and [16], exhibits a different scaling in the threshold regime, since its fixed point structure is different. The SIR model has a line of fixed points at $I = 0$, but no fixed point at positive I. In fact, any tendency to immunity (or death for that matter) will cause an otherwise SIS model to exhibit SIR behavior in the threshold regime for large enough N. This is due to the fact that the Time-dependent bias in the SIR model, no matter how small in strength, overwhelms the space-dependent bias [13] for large enough N.

Acknowledgements

References

[1] ABROWOWITZ, M. AND STEGUN, I. A. (eds) (1972). *Handbook of Mathematical Functions*. United States Government Printing Office, Washington, DC.

[2] ANDERSON, H. AND DJEHICHE, B. (1998). A threshold limit theorem for the stochastic logistic epidemic. *J. Appl. Prob.* **35,** 662–670.

[3] ANTIA, R., REGOES, R. R., KOELLA, J. C. AND BERGSTROM, C. T. (2003). The role of evolution in the emergence of infectious diseases. *Nature (London)* **426,** 658–661.

[4] BEN-NAIM, E. AND KRAPIVSKY, P. L. (2004). Size of outbreaks near the epidemic threshold. *Phys. Rev. E* **69,** 050901.

[5] BENDER, C. M. AND ORSZAG, S. A. (1999). *Advanced Mathematical Methods for Scientists and Engineers*. Springer, New York.

[6] DOERING, C. R., SARGSYAN, K. V. AND SANDER, L. M. (2005). Extinction times for birth–death processes: Exact results, continuum asymptotics and the failure of the Fokker–Planck approximation. *Multiscale Model. Simul.* **3,** 283–299.

[7] DOLGOARSHINNYKH, R. G. AND LALLEY, S. P. (2006). Critical scaling for the SIS stochastic epidemic. *J. Appl. Prob.* **43,** 892–898.

[8] ELGART, V. AND KAMENEV, A. (2004). Rare event statistics in reaction-diffusion systems. *Phys. Rev. E* **70,** 041106.

[9] GROSCHE, C. (1993). Path integration via summation of perturbation expansions and applications to totally reflecting boundaries, and potential steps. *Phys. Rev. Lett.* **71,** 1–4.

[10] JACKSON, J. D. (1998). *Classical Electrodynamics*, 3rd edn. John Wiley, New York.

[11] KERMACK, W. O. AND MCKENDRICK, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proc. R. Soc. A* **115,** 700–721.

[12] KESSLER, D. A. AND SHNERB, M. N. (2007). Extinction rates for fluctuation-induced metastabilities: a real-space WKB approach. *J. Statist. Phys.* **127,** 861–886.

[13] KESSLER, D. A. AND SHNERB, M. N. (2007). Solution of an infection model near threshold. *Phys. Rev. E* **76,** 010901.

[14] KRYSCIO, R. J. AND LEFÉVRE, C. (1989). On the extinctions of the S-I-S stochastic logistic epidemic. *J. Appl. Prob.* **27,** 685–694.

[15] LEIGH, E. J. (1981). The average lifetime of a population in a varying environment. *J. Theoret. Biol.* **90,** 213–239.

[16] MARTIN-LÖF, A. (1998). The final size of a nearly critical epidemic, and the first passage time of a Wiener process to a parabolic barrier. *J. Appl. Prob.* **35,** 671–682.

[17] NÅSELL, I. (1996). The quasi-stationary distribution of the closed endemic SIS model. *Adv. Appl. Prob.* **28,** 895–932.

[18] SHANKAR, R. (1994). *Principles of Quantum Mechanics*. Springer, New York.

[19] WATSON, R. (1980). On the size distribution for some epidemic models. *J. Appl. Prob.* **17,** 912–921.

[20] WEISS, G. H. AND DISHON, M. (1971). On the asymptotic behavior of the stochastic and deterministic models of an epidemic. *Math. Biosci.* **11,** 261–265.