SIAM J. APPL. MATH.
Vol. 81, No. 5, pp. 2195–2217

EXTINCTION AND QUASI-STATIONARITY FOR DISCRETE-TIME, ENDEMIC SIS AND SIR MODELS*

SEBASTIAN J. SCHREIBER†, SHUO HUANG‡, JIFA JIANG‡, AND HAO WANG§

Abstract. Stochastic discrete-time susceptible–infected–susceptible (SIS) and susceptible–infected–recovered (SIR) models of endemic diseases are introduced and analyzed. For the deterministic, mean-field model, the basic reproductive number R_0 determines their global dynamics. If $R_0 \leq 1$, then the frequency of infected individuals asymptotically converges to zero. If $R_0 > 1$, then the infectious class uniformly persists for all time; conditions for a globally stable, endemic equilibrium are given. In contrast, the infection goes extinct in finite time with a probability of 1 in the stochastic models for all R_0 values. To understand the length of the transient prior to extinction as well as the behavior of the transients, the quasi-stationary distributions and the associated mean time to extinction are analyzed using large deviation methods. When $R_0 > 1$, these mean times to extinction are shown to increase exponentially with the population size N. Moreover, as N approaches ∞, the quasi-stationary distributions are supported by a compact set bounded away from extinction; sufficient conditions for convergence to a Dirac measure at the endemic equilibrium of the deterministic model are also given. In contrast, when $R_0 < 1$, the mean times to extinction are bounded above $1/(1-\alpha)$, where $\alpha < 1$ is the geometric rate of decrease of the infection when rare; as N approaches ∞, the quasi-stationary distributions converge to a Dirac measure at the disease-free equilibrium for the deterministic model. For several special cases, explicit formulas for approximating the quasi-stationary distribution and the associated mean extinction are given. These formulas illustrate how for arbitrarily small R_0 values, the mean time to extinction can be arbitrarily large, and how for arbitrarily large R_0 values, the mean time to extinction can be arbitrarily large.

Key words. infectious diseases, discrete-time SIS model, discrete-time SIR model, quasi-stationary distributions, large deviations, times to extinction

AMS subject classifications. 92D30, 60J10, 37N25

DOI. 10.1137/20M1339015

1. Introduction.

Infectious disease modeling has been one of the most important topics in mathematical biology (Keeling and Rohani, 2011). A recent Google scholar search[1] reveals over a million and a half studies referencing SIS (susceptible-infected-susceptible) and SIR (susceptible-infected-recovered) models. Most of these studies use deterministic, continuous-time equations. However, discrete-time models may be more appropriate when epidemiological events have a characteristic time scale or when the model is calibrated by epidemiological measurements taken at regular time intervals (Anderson and May, 1991; Allen, 1994; Finkenstädt and Grenfell, 2000;

*Received by the editors May 19, 2020; accepted for publication (in revised form) April 26, 2021; published electronically October 5, 2021.
https://doi.org/10.1137/20M1339015
Funding: The research of the first author is partially supported by U.S. National Science Foundation grant DMS-1716803. The research of the third author is supported by the National Natural Science Foundation of China (NSFC) grants 12171321 and 11771295. The research of the fourth author is partially supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Individual Discovery Grant RGPIN-2020-03911 and Discovery Accelerator Supplement Award RGPAS-2020-00090.

†Department of Evolution and Ecology and Center for Population Biology, University of California at Davis, Davis, CA 95616 USA (sschreiber@ucdavis.edu).
‡Mathematics and Science College, Shanghai Normal University, Shanghai 200234, China (1000441256@smail.shnu.edu.cn, jiangjf@shnu.edu.cn).
§Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta T6G 2G1, Canada (hao8@ualberta.ca).

[1]On January 30, 2021, a search on Google scholar with the search term "disease model SIR OR SIS" returned 1.60 million results.

Allen and Burgin, 2000; Salomon et al., 2006; Klepac et al., 2009; Keeling and Rohani, 2011; Baker et al., 2020; Prem et al., 2020; Karatayev, Anand, and Bauch, 2020). For example, Finkenstädt and Grenfell (2000) introduced a discrete-time model, the time series SIR (TSIR) model, to account for the biweekly time scale of infection to recovery and lifelong immunity in measles. Similarly, Baker et al. (2020) used a TSIR model whose weekly time interval corresponds to the generation time of respiratory syncytial virus. During the current coronavirus disease 2019 pandemic, daily counts of cases, recoveries, and deaths are used to parameterize discrete-time models with a daily time step (Prem et al., 2020; Karatayev et al., 2020). Discrete-time models also prove a simple way to carefully prescribe the durations of individuals in different epidemiological states. For example, Salomon et al. (2006) used discrete-time models to evaluate the potential epidemiological effects of reducing the duration of anti-tuberculosis drug treatments. For both continuous-time and discrete-time deterministic models, the basic reproductive number, R_0, often determines the fate of the modeled disease. If $R_0 > 1$, persistence often occurs. While if $R_0 < 1$, the disease-free state (i.e., extinction) often is globally stable, and the infection is lost asymptotically as time marches to infinity.

When considering finite populations without external sources of infection, Markov chain models typically predict that the disease goes extinct in finite time whether $R_0 > 1$ or < 1. To understand this puzzling difference between the asymptotic behaviors of the deterministic and stochastic models (Bartlett, 1966; Keeling and Rohani, 2011; Diekmann et al., 2012), one can use the concept of quasi-stationarity that describes the long-term behavior of the stochastic model conditioned on nonextinction (Darroch and Seneta, 1965, 1967). For finite-state models, the quasi-stationary distribution corresponds to a normalized left eigenvector π of the transition matrix of the Markov chain restricted to the transient states, i.e., the states where the infection persists. In discrete-time models, if the disease dynamics follow the quasi-stationary distribution, then the eigenvalue λ associated with this eigenvector corresponds to the probability of disease persistence over the next time step (respectively, a time interval of length one). Thus, when the stochastic model follows the quasi-stationary distribution, the time to extinction is exponentially distributed with a mean time of extinction $1/(1-\lambda)$. Grimm and Wissel (2004) call $1/(1-\lambda)$ the intrinsic mean time to extinction. To understand the link between the stochastic and deterministic models, it is natural to ask: How does the intrinsic mean time to extinction increase as the population size gets larger? How is the quasi-stationary distribution related to the asymptotic dynamics of the deterministic model as the population size gets larger? More generally, how do these quantities depend on the parameters such as R_0?

For continuous-time, stochastic SIS models, there exists a dichotomy in the mean time to extinction when a fixed, positive fraction of the population is infected (Weiss and Dishon, 1971; Barbour, 1975; Kryscio and Lefèvre, 1989; Foxall, 2020). When $R_0 > 1$, this time increases exponentially with the population size N in the limit of large population sizes. When $R_0 < 1$, these extinction times remain bounded in the limit of large population size. However, to the best of our knowledge, similar statements for the intrinsic mean extinction times have not been rigorously proven for these continuous-time SIR models. However, in a series of papers, Nåsell (1996, 1999, 2001, 2002) provided methods to approximate the intrinsic mean extinction times as well as the quasi-stationary distributions. His approximations support the existence of a similar dichotomy for intrinsic mean times to extinction. Moreover, they highlight a remarkable dichotomy about the qualitative behavior of the quasi-stationary distributions. When $R_0 > 1$, these distributions are well approximated by a normal distribution centered near the endemic equilibrium. When $R_0 < 1$, the quasi-

stationary distribution is best approximated by a discrete, geometric distribution. Despite these advances, mathematically rigorous results for discrete-time, stochastic SIS and SIR models are lacking.

In this paper, we introduce a new class of discrete-time SIS and SIR deterministic and stochastic models that have several desirable properties including (i) they are derived with individual-based rules and, consequently, preserve nonnegativity of all populations, (ii) the deterministic models are the mean-field model of the stochastic models, and (iii) the deterministic models converge to the classical continuous-time models in an appropriate limit. For these models, we analyze the global dynamics of deterministic models and then use this analysis in conjunction with large deviation results from Faure and Schreiber (2014) to rigorously characterize the behavior of the intrinsic mean times to extinction and quasi-stationary distributions in the limit of large population sizes. Moreover, for some special cases, we derive explicit approximations for the quasi-stationary distributions and extinction times that apply for all population sizes.

Our paper is structured as follows. Section 2 introduces and analyzes the discrete-time, deterministic SIS model. Section 3 presents mathematical and numerical findings on quasi-stationary distributions and intrinsic mean extinction times for the stochastic SIS model. Section 4 introduces the discrete-time SIR model and proves results about its global attractors. Section 5 presents mathematical and numerical findings on quasi-stationary distributions and intrinsic mean extinction times for the stochastic SIR model. Section 6 discusses our main findings and future challenges. The mathematical proofs are given in sections 7 through 10.

2. The dynamics of a deterministic SIS model. We begin with a discrete-time version of the classical SIS model where individuals are either susceptible or infected. Let I_n denote the fraction of individuals that are infected at time step n, in which case the fraction of susceptible individuals equals $1 - I_n$. Individuals escape natural mortality with the probability $e^{-\mu}$ while infected individuals escape recovery with the probability $e^{-\gamma}$, where $\mu > 0$ and $\gamma > 0$. Susceptible individuals from the previous time step who have not died escape infection with the probability $e^{-\beta I_n}$, where $\beta > 0$ is the contact and transmission rate. To keep the population size constant, dying individuals are replaced via birth or immigration with new susceptible individuals. The disease dynamics are given by

$$I_{n+1} = F(I_n) := e^{-\mu}(1 - I_n)\left(1 - e^{-\beta I_n}\right) + e^{-\mu-\gamma} I_n,\ 0 \le I_n \le 1. \tag{2.1}$$

This discrete-time formulation of the SIS model has several advantages. First, it is straightforward to verify that the dynamics of I_n remain in the interval $[0, 1]$ provided the initial value I_0 lies in this interval. Second, these dynamics, as described in the next section, correspond to the mean-field dynamics of an individual-based model. Finally, if Δt is the length of a time step, and $\beta = \tilde{\beta}\Delta t$, $\gamma = \tilde{\gamma}\Delta t$, and $\mu = \tilde{\mu}\Delta t$, then

$$I(t+\Delta t) := I_{n+1} = (1-I(t))\tilde{\beta} I(t)\Delta t + I(t) - (\tilde{\mu}+\tilde{\gamma})\Delta t I(t) + O(\Delta t^2),\ \text{where } I(t) := I_n.$$

Hence, in the limit $\Delta t \to 0$, we get the classical SIS ordinary differential equation:

$$\frac{dI}{dt} = \lim_{\Delta t \to 0} \frac{I(t+\Delta t) - I(t)}{\Delta t} = (1 - I)\tilde{\beta} I - (\tilde{\mu} + \tilde{\gamma}) I.$$

To understand the dynamics of (2.1), we can linearize at the origin to obtain the per-capita growth rate of the infection at the disease-free equilibrium

$$\alpha = \alpha(\mu, \beta, \gamma) := \beta e^{-\mu} + e^{-\mu-\gamma}. \tag{2.2}$$

The basic reproduction, alternatively, is given by

$$R_0 = \beta e^{-\mu}/(1 - e^{-\mu-\gamma}). \tag{2.3}$$

As $\alpha > 1$ if and only if $R_0 > 1$, we can use the basic reproductive number to characterize the global dynamics as the following theorem shows.

THEOREM 2.1. (i) *If $R_0 \leq 1$, then the origin is globally asymptotically stable.*

(ii) *If $R_0 > 1$, then there is a unique positive equilibrium in $(0, 1]$ such that it is globally asymptotically stable in $(0, 1]$.*

3. Metastability and extinction in a stochastic SIS model. For the individual-based stochastic model, we require the additional parameter of the total population size N. Given this population size, the state space corresponds to the possible fractions of infected individuals in the population:

$$\mathcal{S} = \left\{0, \frac{1}{N}, \frac{2}{N}, \ldots, \frac{N-1}{N}, 1\right\}.$$

Let $I_n \in \mathcal{S}$ be the fraction of individuals infected at time step n. To determine the fraction infected in the next time step, we assume that each infected individual remains infected with probability $e^{-\mu-\gamma}$ independent of each other, and each susceptible individual lives and becomes infected with the probability $e^{-\mu}(1-e^{-\beta I_n})$ independent of each other. Hence,

$$\begin{aligned} I_{n+1} &= \frac{X_{n+1} + Y_{n+1}}{N}, \text{ where} \\ X_{n+1} &\sim \mathrm{Binom}(\mathrm{NI_n}, \mathrm{e}^{-\mu-\gamma}) \text{ and} \\ Y_{n+1} &\sim \mathrm{Binom}\Big(\mathrm{N}(1-\mathrm{I_n}), \mathrm{e}^{-\mu}\big(1-\mathrm{e}^{-\beta \mathrm{I_n}}\big)\Big). \end{aligned} \tag{3.1}$$

Taking the expectation of I_{n+1} of (3.1) conditioned on its value at time n gives

$$\mathbb{E}[I_{n+1}|I_n = I] = e^{-\mu}(1 - I)(1 - e^{-\beta I}) + e^{-\mu-\gamma} I,$$

which corresponds to the update rule $F(I)$ of the deterministic model (2.1). Despite this connection with the deterministic model, the disease goes extinct in finite time with a probability of 1 for the stochastic model. The following proposition follows from standard results in stochastic processes (see e.g., Harier, 2018, Theorem 3.20).

PROPOSITION 3.1. *Assume that $\mu + \gamma$ and β are positive. With a probability of 1, $I_n = 0$ for some $n \geq 1$.*

Even though extinction is inevitable, it may be preceded by long-term transients. To characterize these transients, we use quasi-stationary distributions introduced by Darroch and Seneta (1965). For all $1 \leq i$ and $j \leq N$, let $Q_{ij} = \mathbb{P}[I_{n+1} = j/N|I_n = i/N]$ be the transition probabilities restricted to the transient states $\mathcal{S} \setminus \{0\}$ and $Q = (Q_{ij})$ be the associated $N \times N$ matrix. As Q is a substochastic, positive matrix, there exists a dominant eigenvalue $\lambda \in (0, 1)$ and associated dominant eigenvector $\pi = (\pi_1, \ldots, \pi_N)$ (depending on N) such that $\sum_i \pi_i = 1$, $\pi_i > 0$ for all i, and $\pi Q = \lambda \pi$. π is the *quasi-stationary distribution* which satisfies (Darroch and Seneta, 1965)

$$\lim_{n\to\infty} \mathbb{P}[I_n = j/N|I_n > 0] = \pi_j;$$

i.e., the probability of having j individuals in the long term given the disease hasn't gone extinct equals π_j. Furthermore,

$$\sum_{i=1}^{N} \mathbb{P}[I_{n+1} > 0|I_n = i/N]\pi_i = \lambda;$$

i.e., λ is the probability the disease persists to the next time step given the process is following the quasi-stationary distribution. Hence, the mean time to extinction when following the quasi-stationary distribution is $\frac{1}{1-\lambda}$, which is what Grimm and Wissel (2004) call the *intrinsic mean time to extinction.*

Our main result for the stochastic SIS model concerns the behavior of the quasi-stationary distribution and the intrinsic mean time to extinction for large population size N.

THEOREM 3.2. *Assume $\mu+\gamma>0, \beta>0$. For each $N\geq 1$, let π^N be the quasi-stationary distribution and λ^N the corresponding eigenvalue. Let $\alpha=\beta e^{-\mu}+e^{-\mu-\gamma}$. Then*

(i) *If $\alpha\leq 1$ (equivalently $R_0\leq 1$), then $\lambda^N\leq\alpha$ for all $N\geq 1$ and $\lim_{N\to\infty}\pi^N=\delta_0$, where δ_0 is a Dirac measure at 0 and convergence is in the weak* topology for the probability measures on $[0,1]$.*

(ii) *If $\alpha>1$ (equivalently $R_0>1$), then $\lim_{N\to\infty}\pi^N=\delta_{I^*}$, where δ_{I^*} is the Dirac measure at the unique positive equilibrium I^* of (2.1) and $\limsup_{N\to\infty}\frac{1}{N}\log(1-\lambda^N)<0$.*

The first assertion of Theorem 3.2 concerns the case ($\alpha<1$) when the disease approaches extinction at an exponential rate for the deterministic model. In this case, when the population size is large for the stochastic model, then any long-term transients mostly involve low frequencies of infected individuals, and the mean time to extinction after these transients is less than $\frac{1}{1-\alpha}$. We conjecture that in the limit of large population size, $N\to\infty$, the intrinsic mean time to extinction equals $\frac{1}{1-\alpha}$. The second assertion of Theorem 3.2 concerns the case ($\alpha>1$) when the deterministic model has a positive stable equilibrium. In this case, the stochastic model exhibits transient fluctuations around this equilibrium and only escapes its basin of attraction via large fluctuations in frequencies. The likelihood of these large fluctuations decreases, exponentially with the population size. Consequently, the mean time to extinction increases exponentially with population size; i.e., there exist $c_1, c_2>0$ such that $\frac{1}{1-\lambda^N}\geq c_1 e^{c_2 N}$ for all $N\geq 1$. Figure 3.1 illustrates these conclusions numerically.

Given that Theorem 3.2 describes the effect on increasing population size on the intrinsic mean time to extinction for a fixed value of α, it is natural to ask what effect increasing α has on these extinction times for a fixed population size. In general, this is a challenging question. However, we can answer this question for two special cases. First, we consider the case of low recovery rates $\gamma=0$ and the very high $\beta\gg 1$ contact rates. In the limit of $\beta\to\infty$, the update rule for I_n for $I_n>0$ is approximately $I_{n+1}\sim\frac{1}{N}\mathrm{Binom(N,e^{-\mu})}$. Namely, provided there is at least one infected individual at time step n, all individuals that have not died get infected. In this case, the quasi-stationary distribution is approximately $\pi_i=\binom{N}{i}e^{-\mu i}(1-e^{-\mu})^{N-i}/\lambda^N$ for $i=1,2,\ldots,N$ with the persistence eigenvalue $\lambda^N=1-(1-e^{-\mu})^N$. In particular, in this case, the mean intrinsic extinction time is bounded despite $\alpha\to\infty$ and $R_0\to\infty$ as $\beta\to\infty$. The accuracy of this approximation is illustrated in Figure 3.2. Second, in the limit of no recovery and no mortality (i.e., $\mu=\gamma=0$), the disease (not surprisingly!) never goes extinct whenever $I_0>0$. Indeed, in this case, $I_n\to 1$ as $n\to\infty$ with the probability of 1 provided $\beta>0$ and $I_0>0$. These two special cases highlight that the magnitude of increasing α on the increase of the intrinsic mean time to extinction depends in a subtle way on how α increases.

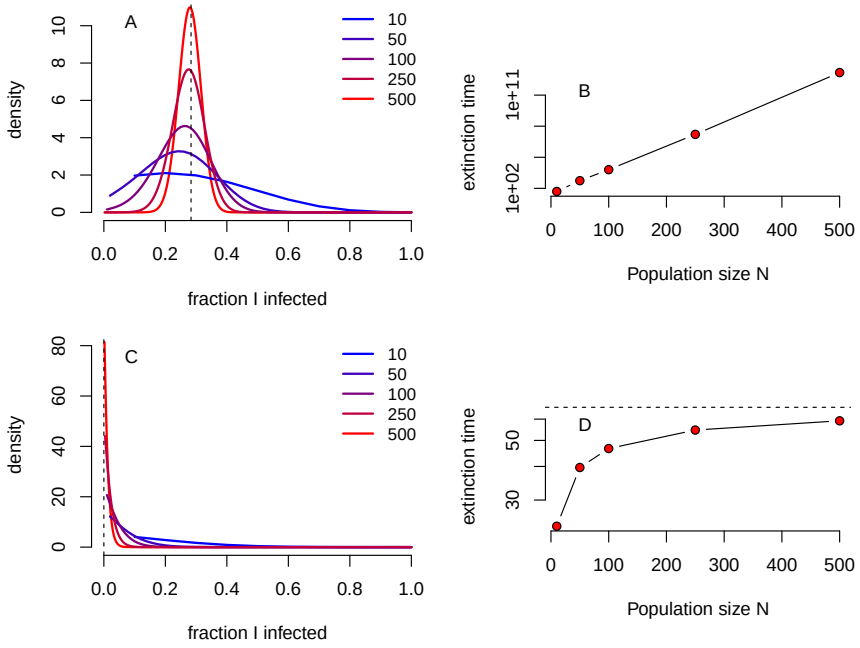

FIG. 3.1. *Quasi-stationary distributions and intrinsic mean extinction times for the stochastic SIS model for* $R_0 > 1$ (A,B) *and* $R_0 < 1$ (C,D). *In* (A) *and* (C), *the quasi-stationary distributions approach a Dirac distribution at the equilibrium density (dashed line). In* (B), *the intrinsic mean time to extinction increases exponentially with a population size for* $R_0 > 1$. *In* (D), *the intrinsic mean time to extinction saturates at* $1/(1-\alpha)$ *(dashed line) as the population size increases. Parameter values:* $\gamma = 0.1$, $\mu = 0.01$, *and* $\beta = 0.15$ *for* (A, B) *and* $\beta = 0.09$ *for* (C, D).

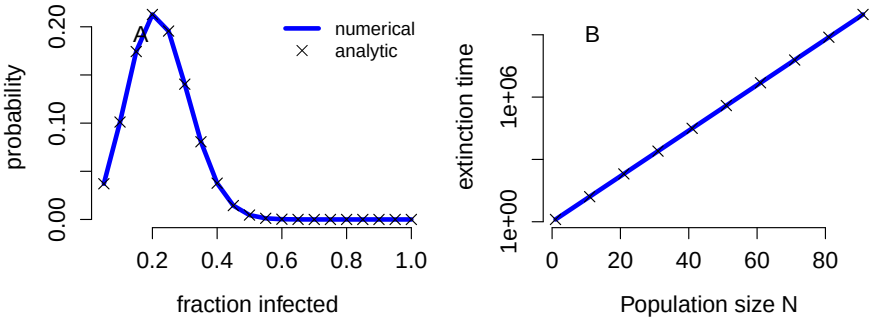

FIG. 3.2. *Quasi-stationary distributions and intrinsic mean extinction times for the stochastic SIS model for high contact rates and low recovery rates. In* (A), *the numerically computed quasi-stationary distribution (solid blue curve) and the analytical approximation (X marks)* $\pi_i = \binom{N}{i} e^{-\mu i}(1-e^{-\mu})^{N-i}/(1-e^{-\mu})^N$. *In* (B), *the numerically computed intrinsic mean time to extinction (solid blue curve) and the analytical approximation (X marks) is* $1/(1-e^{-\mu})^N$. *Parameter values:* $\gamma = 0$, $\mu = 1.5$, $\beta = 100$, *and* $N = 20$ *for* (A).

4. The dynamics of a deterministic SIR model. As the discrete-time model counterpart to the classical SIR model, we assume all individuals escape natural mortality with a probability $e^{-\mu}$, infected individuals escape recovery with a probability $e^{-\gamma}$, and susceptible individuals escape infection with a probability $e^{-\beta I}$, where I is the frequency of infected individuals and $\beta > 0$ is the contact and transmission rate. If the population size is constant, then the discrete-time dynamics are given by

$$\begin{cases} S_{n+1} = 1 - e^{-\mu} + S_n e^{-\mu-\beta I_n}, \\ I_{n+1} = e^{-\mu} S_n \big(1 - e^{-\beta I_n}\big) + e^{-\mu-\gamma} I_n. \end{cases} \tag{4.1}$$

Like the discrete-time SIS model, this discrete-time formulation of the SIR model has several advantages. First, by adding the two equations of (4.1) together, we obtain that the trajectories of (4.1) remain in the domain

$$X := \{(S, I) : S \geq 0,\ I \geq 0,\ S + I \leq 1\},$$

provided the initial value (S_0, I_0) lies in this domain. Furthermore, if we define $\partial X_0 := \{(S, 0) : 0 \leq S \leq 1\}$ and $X_0 := X \setminus \partial X_0$, then X_0 and ∂X_0 are positively invariant. Second, these dynamics, as described in the next section, correspond to the mean-field dynamics of an individual-based model. Finally, if Δt is the length of a time step, and $\beta = \tilde{\beta}\Delta t$, $\gamma = \tilde{\gamma}\Delta t$, and $\mu = \tilde{\mu}\Delta t$, then

$$S(t + \Delta t) := S_{n+1} = \tilde{\mu}\Delta t + S(t) - (\tilde{\mu} + \tilde{\beta} I(t))\Delta t S(t) + O(\Delta t^2), \text{ where } S(t) := S_n,$$
$$I(t + \Delta t) := I_{n+1} = S(t) I(t) \tilde{\beta}\Delta t + I(t) - (\tilde{\mu} + \tilde{\gamma}) I(t)\Delta t + O(\Delta t^2), \text{ where } I(t) := I_n.$$

Hence, in the limit $\Delta t \to 0$, we get the classical SIR system of ordinary differential equations:

$$\frac{dS}{dt} = \lim_{\Delta t \to 0} \frac{S(t + \Delta t) - S(t)}{\Delta t} = \tilde{\mu} - (\tilde{\mu} + \tilde{\beta} I) S,$$
$$\frac{dI}{dt} = \lim_{\Delta t \to 0} \frac{I(t + \Delta t) - I(t)}{\Delta t} = \tilde{\beta} I S - (\tilde{\mu} + \tilde{\gamma}) I.$$

For our discrete-time SIR model, the disease-free equilibrium is $(1, 0)$. At this equilibrium, the per-capita growth rate of the disease still equals $\alpha = \beta e^{-\mu} + e^{-\mu-\gamma}$, and the reproductive number still equals $R_0 = \beta e^{-\mu}/(1 - e^{-\mu-\gamma})$. We will show that if $R_0 > 1$, the disease persists, and if $R_0 < 1$, the disease-free equilibrium is globally stable. Furthermore, we will show that when the recovery rate γ is sufficiently small, there is a globally stable endemic equilibrium. To state these results precisely, we define the parameter space as $P := \{\lambda := (\mu, \beta, \gamma) : \mu > 0, \beta > 0, \gamma \geq 0\}$. Let $C_P^0 := \{\lambda = (\mu, \beta, 0) \in P : \alpha(\lambda) > 1\}$ be the parameters corresponding to no recovery ($\gamma = 0$) and $\alpha > 1$. Finally, define

$$C_P := \{\lambda \in P : \alpha(\lambda) > 1,\ (4.1) \text{ admits a globally stable equilibrium in } X_0\}.$$

Using these definitions, we have the following theorem.

THEOREM 4.1. (i) *If $\alpha < 1$, then the disease-free equilibrium $(1, 0)$ is globally asymptotically stable.*

(ii) *If $\alpha > 1$, then $F : X_0 \to X_0$ admits a global and compact attractor K contained in the interior of X_0.*

(iii) *$C_P \supset C_P^0$ is a nonempty open subset in P.*

In addition we conjecture that there is a globally stable endemic equilibrium when $\mu = \tilde{\mu}\Delta t$, $\beta = \tilde{\beta}\Delta t$, $\gamma = \tilde{\gamma}\Delta t$, $\Delta t > 0$ is sufficiently small, and $\alpha > 1$ (equivalently, $R_0 > 1$).

5. Metastability and extinction in a stochastic SIR model. As with the stochastic SIS model, the stochastic SIR model requires the additional parameter of the total population size N. For a given population size, the state space $\mathcal{S}$ corresponds to the possible fraction of susceptible and infected individuals in the population; i.e.,

$$\mathcal{S} = \{(i/N, j/N) : i, j \in \{0, 1, \ldots, N\}, i + j \leq N\} \subset X.$$

Let $(S_n, I_n) \in \mathcal{S}$ be the fractions of susceptible and infected individuals at the time step n. The fraction of removed individuals at the time step n equals $R_n = 1-S_n-I_n$. Consistent with the deterministic SIR model, we assume (i) each susceptible individual lives and becomes infected with a probability $e^{-\mu}(1-e^{-\beta I_n})$ independent of each other, (ii) each infected individual either remains infected, dies and gets replaced with a susceptible individual, or enters the removed class with probabilities $e^{-\mu-\gamma}$, $1-e^{-\mu}$, or $e^{-\mu}(1-e^{-\gamma})$ independent of each other, and (iii) each removed individual dies and creates a new susceptible individual with a probability $1-e^{-\mu}$. To account for these transitions, let W_{n+1} be a binomial random variable with NS_n trials and a probability of success $e^{-\mu}(1-e^{-\beta I_n})$ (i.e., susceptible individuals that will become infected), X_{n+1} be a binomial random variable with NI_n trails and a probability of success $1-e^{-\mu}$ (i.e., infected individuals that die and get replaced by a susceptible individual), Y_{n+1} be a binomial random variable with $NI_n - X_{n+1}$ trials with a probability of success $e^{-\gamma}$ (i.e., nondying infected individuals that will not enter the removed class), and Z_{n+1} be a binomial random variable with $N(1-I_n-S_n)$ trials with a probability of success $1-e^{-\mu}$ (i.e., removed individuals that die and get replaced with a susceptible individual). Under these assumptions, the stochastic SIR model is

$$
\begin{aligned}
S_{n+1} &= \frac{1}{N}\left(NS_n - W_{n+1} + X_{n+1} + Z_{n+1}\right),\\
I_{n+1} &= \frac{1}{N}\left(W_{n+1} + Y_{n+1}\right), \text{ where}\\
W_{n+1} &\sim \mathrm{Binom}(\mathrm{NS_n}, \mathrm{e}^{-\mu}\left(1-\mathrm{e}^{-\beta \mathrm{I_n}}\right)),\\
X_{n+1} &\sim \mathrm{Binom}(\mathrm{NI_n}, 1-\mathrm{e}^{-\mu}),\\
Y_{n+1} &\sim \mathrm{Binom}(\mathrm{NI_n} - \mathrm{X_{n+1}}, \mathrm{e}^{-\gamma}), \text{ and}\\
Z_{n+1} &\sim \mathrm{Binom}(\mathrm{N}(1-\mathrm{I_n}-\mathrm{S_n}), 1-\mathrm{e}^{-\mu}).
\end{aligned}
\tag{5.1}
$$

Taking the expectations of S_{n+1} and I_{n+1} conditioned on the values of S_n and I_n gives

$$
\begin{aligned}
\mathbb{E}[S_{n+1}|S_n = S, I_n = I] &= 1 - e^{-\mu} + Se^{-\mu-\beta I},\\
\mathbb{E}[I_{n+1}|S_n = S, I_n = I] &= e^{-\mu}S\left(1-e^{-\beta I}\right) + e^{-\mu-\gamma}I,
\end{aligned}
$$

which corresponds to the update rule for the deterministic model (4.1). Despite this relationship, the disease always goes extinct in finite time with a probability of 1 for the stochastic model. The following proposition follows from standard results in stochastic processes (see, e.g., Harier, 2018, Theorem 3.20).

PROPOSITION 5.1. *Assume that $\mu+\gamma$ and β are positive. With probability one, $I_n = 0$ for some $n \geq 1$.*

To characterize metastability and extinction times, define $\mathcal{S}_+ = \{(x_1, x_2) \in \mathcal{S} : x_2 > 0\}$ to be all the states where the disease persists. For all pairs of states $x, y \in \mathcal{S}_+$, let $Q_{xy} = \mathbb{P}[(S_{n+1}, I_{n+1}) = y|(S_n, I_n) = x]$ be the transition probabilities restricted to the transient states and $Q = (Q_{xy})_{x,y\in\mathcal{S}_+}$ be the associated matrix. Let $\pi = (\pi_x)_{x\in\mathcal{S}_+}$ be the quasi-stationary distribution with an associated persistence probability λ, i.e., $\sum_{x\in\mathcal{S}_+}\pi_x = 1$, $\pi_x > 0$ for all $x \in \mathcal{S}_+$ and $\pi Q = \lambda\pi$. Our main result for the stochastic SIR model concerns the behavior of the quasi-stationary distribution and the intrinsic mean time to extinction for large population size N.

THEOREM 5.2. *Assume $\mu+\gamma > 0, \beta > 0$. For each $N \geq 1$, let π^N be the quasi-stationary distribution and λ^N the corresponding eigenvalue for* (5.1). *Let $\alpha = \beta e^{-\mu} + e^{-\mu-\gamma}$. Then*

(i) *If* $\alpha < 1$, *then* $\lambda^N \leq \alpha$ *for all* $N \geq 1$ *and* $\lim_{N\to\infty} \pi^N = \delta_{(1,0)}$, *where* $\delta_{(1,0)}$ *is a Dirac measure at the disease-free equilibrium* $(1, 0)$ *and convergence is in the weak* topology.*

(ii) *If* $\alpha > 1$, *then* $\limsup_{N\to\infty} \frac{1}{N} \log(1 - \lambda^N) < 0$, *and there exists a compact set* $K \subset (0, 1)^2$ *such that* $\pi^*(K) = 1$ *for every weak* limit point* π^* *of* π^N *as* $N \to \infty$ *and where* π^* *is invariant for the deterministic model* (4.1). *Moreover, if* $(\mu, \beta, \gamma) \in C_P$, *then* $\lim_{N\to\infty} \pi^N = \delta_{(S^*,I^*)}$, *where* $\delta_{(S^*,I^*)}$ *is the Dirac measure at the unique positive equilibrium* (S^*, I^*) *of* (4.1).

The first assertion of Theorem 5.2 implies that if $\alpha < 1$ and the population size is large, then any long-term transient mostly involves low frequencies of infected individuals, and the mean time to extinction after these transients is less than $\frac{1}{1-\alpha}$. Furthermore, whenever permanence of the deterministic model corresponds to a globally stable equilibrium (see Theorem 4.1), the quasi-stationary distributions concentrate on a Dirac measure at this equilibrium; i.e., it supports the only invariant measure in K for the deterministic dynamics. The second assertion of Theorem 5.2 implies that if $\alpha > 1$ and the population size is large, then the long-term transients fluctuate away from the disease-free equilibrium and the mean time to extinction following these transients increases exponentially with population size; i.e., there exist $c_1, c_2 > 0$ such that $\frac{1}{1-\lambda^N} \geq c_1 e^{c_2 N}$ for all $N \geq 1$. Figure 5.1 illustrates these conclusions.

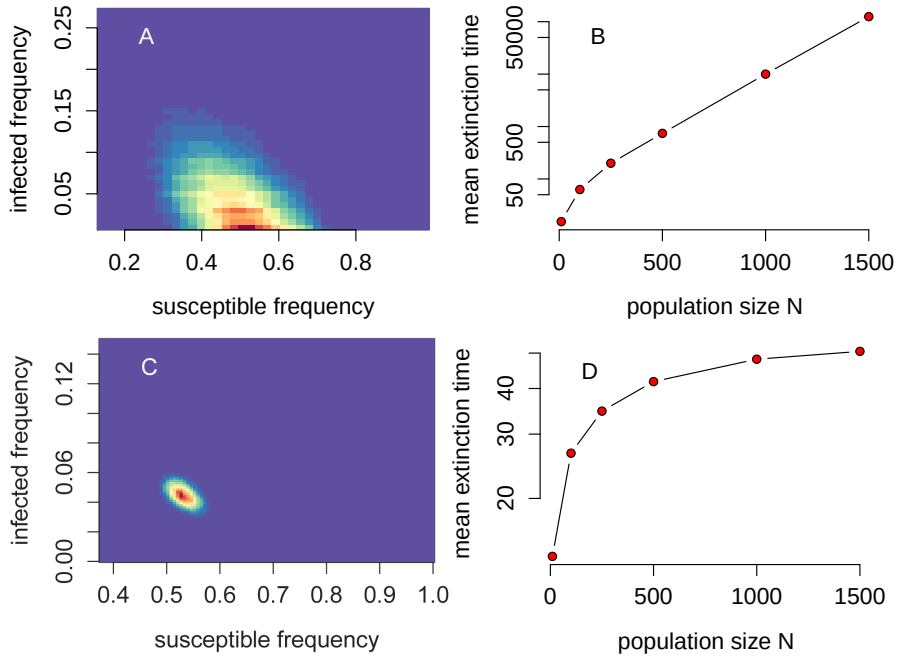

FIG. 5.1. *Quasi-stationary distributions and mean intrinsic extinction times for the stochastic SIR model. For parameters with* $\alpha > 1$, *the quasi-stationary distributions, estimated numerically using the method of* Aldous, Flannery, and Palacios (1988), *concentrate on the stable equilibrium as the population size goes from* $N = 100$ (A) *to* $N = 10,000$ (C). *For this* $\alpha > 1$, *the associated intrinsic mean extinction times increase exponentially with the population size in* (B). *For parameters with* $\alpha < 1$, *the mean extinction times are bounded by* $1/(1 - \alpha)$ *in* (D). *Parameter values:* $\mu = 0.01$, $\gamma = 0.1$, *and* $\beta = 0.2$ *for* (A)–(C) *and* $\beta = 0.09$ *for* (D).

6. Discussion. This paper formulates and provides a mathematically rigorous analysis of deterministic and stochastic, discrete-time SIS and SIR models. The stochastic models are based on probabilistic, individual-based update rules. The conditional expected change in the fraction of infected and susceptible individuals given the current values of these fractions determines the update rule for the deterministic model and, in this sense, the deterministic model is the mean-field model for the stochastic models.

Many earlier discrete-time epidemic models of SIS and SIR dynamics have been derived using numerical approximation schemes for differential equations (Allen 1994; Finkenstädt and Grenfell 2000; Castillo-Chavez and Yakubu 2001, Satsuma et al., 2004; Enatsu, Nakata, and Muroya 2010; Ma, Zhou, and Cao 2013). These models, including the one-dimensional ones, can exhibit oscillatory dynamics. In contrast, our models, which are based on individual-based update rules and use an exponential escape function, are most similar to higher-dimensional, discrete-time epidemiological models that have been used for applications to specific diseases (Emmert and Allen, 2004, 2006; Allen and van den Driessche, 2008). Consistent with Allen and van den Driessche (2008), our analysis and numerical simulations suggests that our models do not exhibit oscillatory behavior. This may be biologically realistic as sustainable oscillations in infectious diseases are normally driven by periodic events such as seasonality, school dates, and holidays (Keeling and Rohani, 2011; Pollicott, Wang, and Weiss, 2012; Kong, Jin, and Wang, 2015), none of which are considered in our models. When $R_0 < 1$, we prove that all trajectories asymptotically approach the disease-free equilibrium for both the SIS and SIR models. When $R_0 > 1$, we prove the disease persists for both models, approaches a globally stable, endemic equilibrium for the SIS model, and provides sufficient conditions for global stability of the endemic equilibrium for the SIR model. Based on extensive numerical simulations, we conjecture that $R_0 > 1$ always implies global stability of the endemic equilibrium for the SIR model.

Unlike the deterministic model for which the disease persists indefinitely when $R_0 > 1$ and only goes asymptotically extinct over an infinite time horizon when $R_0<1$, the fraction of infected in the stochastic model becomes zero in finite time for all values of R_0. To understand this well-know discrepancy (Bartlett, 1966; Keeling and Rohani, 2011; Diekmann et al., 2012) for our model, we characterized the long-term transients using quasi-stationary distributions for finite-state, discrete-time Markov chains (Darroch and Seneta, 1965). For these characterizations, we used the per-capita growth rate $\alpha = \beta e^{-\mu} + e^{-\mu-\gamma}$ of the infection at the disease-free equilibrium. When $\alpha < 1$ (equivalently, $R_0 < 1$), the mean time to extinction, when following the quasi-stationary distribution, is less than or equal to $1/(1-\alpha)$ for all population sizes and for both the SIS and SIR models. Indeed, we conjecture that as $N \to \infty$, this mean time to extinction converges to $1/(1-\alpha)$. While $R_0 < 1$ if and only if $\alpha < 1$, we have $\alpha = (1 - e^{-\mu-\gamma})R_0 + e^{-\mu-\gamma} > R_0$ whenever $R_0 < 1$. Hence, even if R_0 is very small, the mean times to extinction can be arbitrarily long. For example, given any $0 < x < y < 1$, we can make $R_0 = x$ and $\alpha = y$ by choosing $\gamma = 0$, $e^{-\mu} = \frac{y-x}{1-x}$, and $\beta = x(e^{\mu} - 1)$.

When $R_0 > 1$ (equivalently $\alpha > 1$), we show that the mean extinction times increase exponentially with the population size N, and the quasi-stationary distributions concentrate on positive invariant sets for the deterministic model for large N. In particular, coupled with our analysis of the deterministic dynamics, our results imply that the quasi-stationary behavior for large N always concentrates near the globally stable, endemic equilibrium of the SIS model. We provide sufficient conditions for

the same conclusion for the SIR model and conjecture that this always occurs for the SIR model. These conclusions are consistent with earlier studies of continuous-time Markov models where the analysis was done using diffusion approximations (Barbour, 1975; Kryscio and Lefèvre, 1989; Nåsell, 1996, 1999; Andersson and Britton, 2000; Nåsell, 2002; Lindholm and Britton, 2007; Andersson and Lindenstrand, 2011; Clancy and Tjia, 2018) or using large deviation estimates of extinction times for fixed initial frequencies (Kratz and Pardoux, 2018). In contrast, our results apply large deviation methods from Faure and Schreiber (2014) to understand the quasi-stationary distributions and the mean time to extinction from these distributions. An open question for the stochastic model with $R_0 > 1$ concerns the asymptotic rate at which the extinction times increase exponentially with N. Specifically, what is the value of $c > 0$ such that the mean time to extinction grows like $\exp(cN)$ for large N? The diffusion approximations provide one approach to find potential candidates for c.

When $R_0 > 1$, we found that the mean time to extinction can be arbitrarily large even for a fixed population size. For example, this occurs when recovery and mortality rates approach zero in which case R_0 also increases without bound but α remains bounded above by $\beta+1$. In contrast, increasing contact rates (which increase α and R_0 without bound) leads to extinction times that are constrained by population size, recovery rates, and mortality rates.

In addition to the open mathematical questions that we have already raised, future challenges include analyzing extensions of our models. These extensions could include additional compartments such as SEIR models, multi-age group epidemic models, and SIR type models with vaccination (Anderson and May 1991; Keeling and Rohani, 2011; Kong, Jin, and Wang, 2015). More generally, when the discrete-time system is autonomous, the mathematical approaches used here should be applicable to study quasi-stationary distributions and the intrinsic extinction times. However, when population sizes or transmission rate change stochastically over time (Anderson and May 1979; Pollicott, Wang, and Weiss, 2012), new mathematical approaches are required for studying the impact of these environmentally driven random fluctuations on intrinsic extinction times.

7. Proof of Theorem 2.1. (i) First, we shall prove that

$$F(I) := e^{-\mu}(1-I)\big(1-e^{-\beta I}\big) + e^{-\mu-\gamma} I < \big(\beta e^{-\mu} + e^{-\mu-\gamma}\big) I =: L(I),\ I \in (0,1] \tag{7.1}$$

is equivalent to

$$(1-I)\big(1-e^{-\beta I}\big) < \beta I,\ I \in (0,1]. \tag{7.2}$$

In fact, $1 - e^{-x} < x$ for all $x > 0$. This implies that

$$(1-I)\big(1-e^{-\beta I}\big) \le (1-I)\beta I < \beta I \text{ for all } I \in (0,1].$$

This shows that (7.2) holds.

Fix $I_0 \in (0,1]$, and set $I_n := F^n(I_0),\ n = 1, 2, \ldots$. Suppose that $\alpha \le 1$. Then we claim that

$$I_{n+1} < I_n,\ n = 0, 1, 2, \ldots. \tag{7.3}$$

For $n = 0$, $I_1 = F(I_0) < L(I_0) = \alpha I_0 \le I_0$; that is, (7.3) is true for $n = 0$. Suppose that (7.3) holds for n. Then by (7.1), we have that

$$I_{n+2} = F(I_{n+1}) < L(I_{n+1}) = \alpha I_{n+1} \le I_{n+1}.$$

By mathematical induction, (7.3) is valid for all positive integers. Combining $\alpha \leq 1$ with (7.1), we get that F has no positive equilibrium in $(0, 1]$. Therefore, (7.3) implies that $\lim_{n\to\infty} I_n = 0$; that is, 0 is globally asymptotically stable.

(ii) Suppose that $\alpha > 1$. Then we claim that $F(I)$ has a unique positive equilibrium denoted by I^*. In fact, define $G(I) := F(I) - I$. Then $G(0) = 0, G(1) = e^{-\mu-\gamma} - 1 < 0, G'(0) = \alpha - 1 > 0, G'(1) = -e^{-\mu}(1 - e^{-\beta}) + e^{-\mu-\gamma} - 1 < 0$, and $G''(I) = -2\beta e^{-\mu}e^{-\beta I} - \beta^2 e^{-\mu}(1 - I)e^{-\beta I} < 0$. This implies the claim. If F has no critical point in the interval $(0, 1)$, then $F'(I) > 0$ on $[0, 1]$, which implies that F is a strongly monotone map on $(0, 1]$ with every orbit closure being compact. By Theorem 3 of Jiang and Yu (1996), I^* is globally asymptotically stable on $(0, 1]$. In the following, we assume that $F(I)$ has a positive critical point I_c^*. The similar proof to the uniqueness of I^* shows that I_c^* must be unique if it exists. Therefore, F possesses the following properties:

$$\begin{cases} F(I) > I & \text{if } 0 < I < I^*, \\ F(I) < I & \text{if } I^* < I \leq 1, \text{ and} \end{cases} \tag{7.4}$$

$$\begin{cases} F(I) \text{ is strictly increasing on } [0, I_c^*], \\ F(I) \text{ is strictly decreasing on } [I_c^*, 1]. \end{cases} \tag{7.5}$$

Now we will divide (ii) into two cases:

$$\text{(iia) } I^* \leq I_c^*, \textit{ and } \text{(iib) } I^* > I_c^*.$$

Let $I_0 \in (0, 1]$. Then we denote $F^n(I_0)$ by I_n.

(iia) Take $I_0 \in (0, I^*)$. Then we claim that

$$I_0 < I_1 < I_2 < \cdots < I_n < I^*, \text{ for } n = 1, 2, \ldots. \tag{7.6}$$

By (7.4), the assumption of (iia), and (7.5), $I_0 < F(I_0) = I_1 < F(I^*) = I^*$; this means that (7.6) is true for $n = 1$. Suppose that (7.6) holds for n. Then again using (7.4), the assumption of (iia), and (7.5), we get that

$$I_0 < F(I_0) < F(I_1) < F(I_2) < \cdots < F(I_n) < F(I^*),$$

which implies that (7.6) holds for $n + 1$. By induction, the claim holds. Thus as $n \to \infty$, I_n increasingly tends to the unique equilibrium I^*.

If $I_0 \in (I^*, I_c^*]$ with $I^* < I_c^*$, then we can inductively prove

$$I^* < I_n < I_{n-1} < \cdots < I_1 < I_0 \text{ for } n = 1, 2, \ldots$$

by the assumption of (iia), (7.5), and (7.4). This deduces that I_n decreasingly tends to the unique equilibrium I^* as $n \to \infty$.

Assume that $I_0 \in (I_c^*, 1]$. Then by (7.4), the assumption of (iia), and (7.5), $I_c^* \geq F(I_c^*) > F(I_0) = I_1$; that is, $I_1 \in (0, I_c^*]$. By the two cases discussed above, we conclude that $\lim_{n\to\infty} I_n = I^*$.

The above analytic proof can be achieved by Feigenbaum's graphical analysis (see page 106 of Feigenbaum (1983)) performed in the following steps: To iterate an initial I_0 successively,

(1) move vertically to the graph of $S = F(I)$;
(2) move horizontally to the graph of $S = I$; and
(3) repeat steps 1 and 2, etc.
Figure 7.1(a)–(c) depicts these processes corresponding to the three cases of (iia) discussed above.

(iib) By (7.5), $F(I)$ is strictly decreasing when $I > I_c^*$; that is, $F'(I) < 0,\ I \in (I_c^*, 1]$. The computation shows that

$$F'(I) \geq e^{-\mu}\big(-1 + e^{-\beta I} + e^{-\gamma}\big) > -1.$$

Therefore,

$$-1 < F'(I) < 0,\ I \in (I_c^*, 1],\ \kappa = \sup_{I \in [I_c^*, 1]} |F'(I)| \in (0, 1), \tag{7.7}$$

whose graph is shown in Figure 7.2(a). This means that the mapping is contractive from the equilibrium on the interval $[I_c^*, 1]$, but its orientation is reversing. Using (7.7), we can prove that

$$I_1 > I^* \text{ and } I_0 < I_2 < I^* \text{ if } I_c^* \leq I_0 < I^*. \tag{7.8}$$

In fact, for any $I_0 \in [I_c^*, I^*)$, we have

$$I_1 - I^* = F(I_0) - F(I^*) = F'(\xi)(I_0 - I^*) \text{ with } \xi \in (I_0, I^*).$$

Equation (7.7) implies that $I_1 > I^*$ and $I_1 - I^* < I^* - I_0$. Similarly,

$$I_2 - I^* = F(I_1) - F(I^*) = F'(\eta)(I_1 - I^*) \text{ with } \eta \in (I^*, I_1).$$

Again, by (7.7), $I_2 < I^*$ and $I^* - I_2 < I_1 - I^*$. This proves (7.8) as shown in Figure 7.2(b).

As a summary, we conclude that if $I_0 \in [I_c^*, I^*]$, then

$$|F^{2n}(I_0) - I^*| \leq \kappa^{2n} |I_0 - I^*|.$$

This shows $\lim_{n\to\infty} F^{2n}(I_0) = I^*$ and $\lim_{n\to\infty} F^{2n+1}(I_0) = F(I^*) = I^*$ for any $I_0 \in [I_c^*, I^*]$. We conclude that I_n oscillates around I^* and converges to I^* in this case.

Let F_l denote the restriction of F on $[0, I_c^*]$. Then $F_l : [0, I_c^*] \to [0, F(I_c^*)]$ is a homeomorphism. Denote $F_l^{-n}(I_c^*)$ by I_c^{-n} with $I_c^0 = I_c^*$. Then $\{I_c^{-n} : n \geq 0\} \subset [0, I_c^*]$ is strictly decreasing with respect to n. Therefore, $\lim_{n\to\infty} I_c^{-n} = \tilde{I}$ exists, and $\tilde{I} \in [0, I_c^*]$ is an equilibrium of F. This proves $\{I_c^{-n} : n \geq 0\}$ decreasingly tends to the origin. Let J_n denote the interval $[I_c^{-n}, I_c^{-(n-1)})$, $n = 1, 2, \ldots$. Then $(0, I_c^*) = \bigcup_{n=1}^{\infty} J_n$. We shall prove that for each positive integer n and $I_0 \in J_n$, I_n will eventually oscillate around the equilibrium I^* and tend to I^*, as shown in Figure 7.2(c).

We first prove the result holds for J_1. Let $I_0^* \in (0, I_c^*)$ such that $F(I_0^*) = I^*$. Then $J_1 = [I_0^*, I_c^*) \cup [I_c^{-1}, I_0^*]$. If $I_0 \in [I_0^*, I_c^*)$, then $I_1 = F(I_0) \in [I^*, F(I_c^*))$, and there exists $I_0^+ \in (I_c^*, I^*]$ such that $I_1 = F(I_0^+)$. Thus, $I_n (n \geq 0)$ will oscillate around the equilibrium I^* and tend to it as $n \to \infty$ as proved above. By the definition of I_c^{-1} and I_0^*, $F([I_c^{-1}, I_0^*]) = [I_c^*, I^*]$. This implies that $I_1 \in [I_c^*, I^*]$ if $I_0 \in [I_c^{-1}, I_0^*]$. Therefore, $I_n (n \geq 1)$ will oscillate around the equilibrium I^* and converge to it as $n \to \infty$ as proved above.

Now we prove the result holds for $J_m(m > 1)$. Since $F^{m-1}(J_m) = J_1$, using the result proved in the previous paragraph, we conclude that $I_n(n \geq m)$ will oscillate around the equilibrium I^* and approach it. This proves the result for $I_0 \in (0, I_c^*)$.

It remains to consider the case $I_0 \in (I^*, 1]$. We can see that $I_1 = F(I_0) \in (0, I^*)$ in this case. The above arguments have shown that as $n \to \infty$, I_n tends to I^* through such an I_1. This completes the proof.

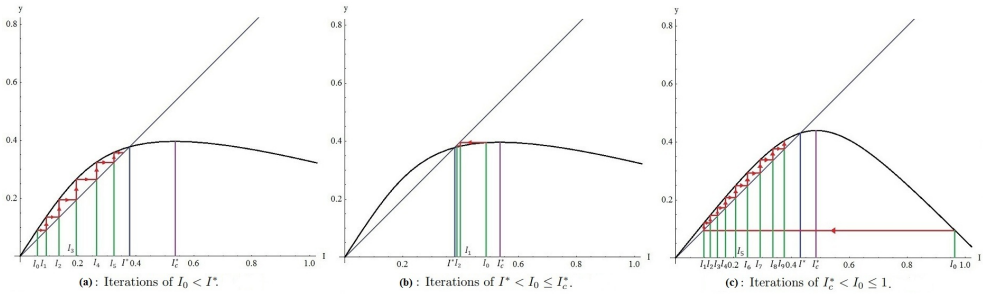

FIG. 7.1. $I^* \leq I_c^*$.

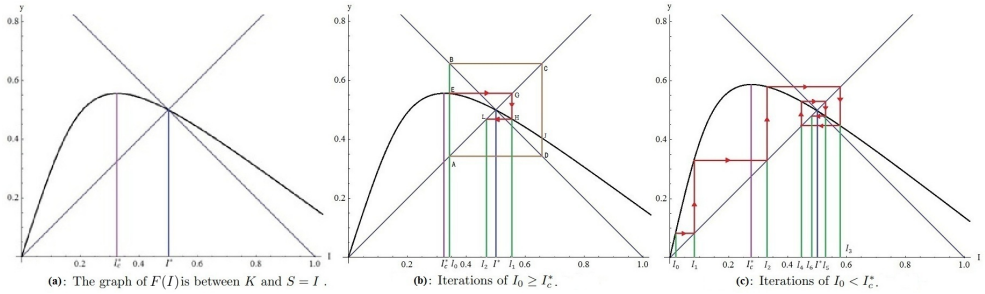

FIG. 7.2. $I^* > I_c^*$.

8. Proof of Theorem 3.2. We use results from Faure and Schreiber (2014) to prove Theorem 3.2. We begin by verifying that Standing Hypothesis 1.1 of Faure and Schreiber (2014) holds. In their notation, "ε" corresponds to $\frac{1}{N}$ in our models; i.e., small demographic noise corresponds to large population size N. For all $\delta > 0$ and $N \geq 1$, define

$$\beta_\delta(N) = \sup_{x \in [0,1]} \mathbb{P}\left[|I_{n+1} - F(x)| \geq \delta | I_n = x\right],$$

where $F(x) = e^{-\gamma-\mu}x + e^{-\mu}(1-e^{-\beta x})(1-x)$ corresponds to the right-hand side of the deterministic model in (2.1). Standing Hypothesis 1.1 of Faure and Schreiber (2014) requires that $\lim_{N\to\infty} \beta_\delta(N) = 0$ for all $\delta > 0$. The following proposition proves something stronger using large deviation estimates.

PROPOSITION 8.1. *There exists a function* $\rho : (0, \infty) \to (0, \infty)$ *such that*

$$\beta_\delta(N) \leq \exp(-N\rho(\delta))$$

for all $N \geq 1$ *and* $\delta > 0$.

Proof. While we use standard large deviation estimates, we go through the details to ensure that the estimates can be taken uniformly in $x \in [0,1]$. Define $a = e^{-\mu-\gamma}$ and $b(x) = e^{-\mu}(1-e^{-\beta x})$. By the exponential Markov inequality, we have for all t

$$\begin{aligned}\mathbb{P}[N(I_{n+1}-F(x))\geq N\delta|I_n=x] &\leq e^{-tN\delta}\mathbb{E}[e^{t(N(I_{n+1}-F(x)))}|I_n=x]\\ &= e^{-tF(x)N-t\delta N}\left(1-a+ae^t\right)^{Nx}\left(1-b(x)+b(x)e^t\right)^{N(1-x)}.\end{aligned} \tag{8.1}$$

Define

$$\psi(t,x) = -tF(x) + x\log(1-a+ae^t) + (1-x)\log(1-b(x)+b(x)e^t).$$

As $a, b(x) \in [0,1]$ for all x, $\psi(t,x)$ is a smooth function of (t,x); i.e., partial derivatives of all orders exist and are continuous in $(-\infty,\infty)\times[0,1]$. Taking log of (8.1) and dividing by N yields

$$\frac{1}{N}\log\mathbb{P}[N(I_{n+1}-F(x)) \geq N\delta|I_n=x] \leq -\delta t + \psi(t,x) \tag{8.2}$$

for all $t \neq 0$, $x \in [0,1]$, and $\delta > 0$. Similarly, one can show that

$$\frac{1}{N}\log\mathbb{P}[N(F(x)-I_{n+1}) \geq N\delta|I_n=x] \leq -\delta t + \psi(-t,x) \tag{8.3}$$

for all $t \neq 0$, $x \in [0,1]$, and $\delta > 0$. Taking the first- and second-order derivatives of ψ with respect to t, we get

$$\frac{\partial\psi}{\partial t}(t,x) = -F(x) + \frac{xae^t}{1-a+ae^t} + \frac{(1-x)b(x)e^t}{1-b(x)+b(x)e^t}$$

and

$$\frac{\partial^2\psi}{\partial t^2}(t,x) = ax\frac{e^t(1-a)}{(1-a+ae^t)^2} + (1-x)b(x)\frac{e^t(1-b(x))}{(1-b(x)+b(x)e^t)^2} \geq 0.$$

As $\psi(0,x) = \frac{\partial\psi}{\partial t}(0,x) = 0$ and $\frac{\partial^2\psi}{\partial t^2}(0,x) \geq 0$ for $x \in [0,1]$, continuity of the partial derivatives of ψ and the compactness of $\{0\}\times[0,1]$ implies that for any $\delta > 0$ there is $t^*(\delta) > 0$ such that $|\frac{\partial\psi}{\partial t}(t,x)| \leq \delta/2$ and $t\frac{\partial\psi(t,x)}{\partial t} \geq 0$ for all $(t,x) \in [-t^*(\delta), t^*(\delta)]\times[0,1]$. Thus, for any $(t,x) \in [0,t^*(\delta)]\times[0,1]$, the fundamental theorem of calculus implies

$$\psi(t,x) = \int_0^t \frac{\partial\psi}{\partial t}(s,x)ds \leq t\delta/2$$

and

$$\psi(-t,x) = \int_0^t -\frac{\partial\psi}{\partial t}(-s,x)ds \leq t\delta/2.$$

Define

$$\rho(\delta) = \delta t^*(\delta) - \max_{x\in[0,1]}\psi(t^*(\delta),x).$$

Our estimates imply that $\rho(\delta) \geq \delta t^*(\delta)/2 > 0$. Furthermore, (8.2) and (8.3) imply that

$$\mathbb{P}[|I_{n+1}-F(x)| \geq \delta|I_n=x] \leq \exp(-N\rho(\delta))$$

for all $x \in [0,1]$ and $\delta > 0$. ∎

To prove the first result of Theorem 3.2, assume that $\alpha \leq 1$. Theorem 2.1 implies that 0 is globally stable for the deterministic model $I \mapsto F(I)$. Theorem 3.12 of Faure and Schreiber (2014), which only requires Standing Hypothesis 1.1, implies that $\lim_{N\to\infty} \pi^N = \delta_0$. Define $R(x) = F(x)/x$ for $x \in (0,1]$. Equation (7.1) in the proof of Theorem 2.1 implies that $R(x) \leq \alpha$ for $x \in (0,1]$. For $N \geq 1$, the quasi-stationarity of π^N implies

$$\begin{aligned}\lambda^N \sum_{i=1}^N \frac{i}{N}\pi_i^N &= \sum_{i=1}^N \mathbb{E}\left[I_{n+1}\Big| I_n = \frac{i}{N}\right]\pi_i^N \\ &= \sum_{i=1}^N F\left(\frac{i}{N}\right)\pi_i^N \\ &= \sum_{i=1}^N \frac{i}{N} R\left(\frac{i}{N}\right)\pi_i^N \\ &\leq \alpha \sum_{i=1}^N \frac{i}{N}\pi_i^N.\end{aligned}$$

Since $\sum_{i=1}^N \frac{i}{N}\pi_i^N > 0$, $\lambda^N \leq \alpha$ for all $N \geq 1$ as claimed.

To prove the second result of Theorem 3.2, assume $\alpha > 1$, in which case Theorem 2.1 implies that there exists a unique positive globally stable equilibrium I^* for the map $I \mapsto F(I)$. Assertion (b) of Lemma 3.9 of Faure and Schreiber (2014) implies that there exists $\delta > 0$ such that $1 - \lambda^N \leq \beta_\delta(N)$ for all $N \geq 1$. Proposition 8.1 implies that

$$\limsup_{N\to\infty} \frac{1}{N}\log(1-\lambda^N) \leq -\rho(\delta).$$

To complete the proof of the second assertion, we need to verify the assumption in assertion (b′) of Lemma 3.9 of Faure and Schreiber (2014). Choose $\eta > 0$ sufficiently small so that

$$\min_{x\in[0,\eta]}(1-e^{-\mu-\gamma})^x \geq \exp(-\rho(\delta)/3) \text{ and } \min_{x\in[0,\eta]}(1-e^{-\mu}(1-e^{-\beta x}))^{1-x} \geq \exp(-\rho(\delta)/3).$$

Then

$$\min_{x\in[0,\eta]} \mathbb{P}[I_{n+1} = 0 | I_n = x] \geq \exp(-2N\rho(\delta)/3)$$

and

$$\lim_{N\to\infty} \frac{\beta_\delta(N)}{\min_{x\in[0,\eta]}\mathbb{P}[I_{n+1}=0|I_n=x]} \leq \lim_{N\to\infty} \frac{\exp(-N\rho(\delta))}{\exp(-2N\rho(\delta)/3)} = 0,$$

which verifies the assumption of (b′) of Lemma 3.9 of Faure and Schreiber (2014) and implies that

$$\lim_{N\to\infty} \sum_{i/N\leq\eta} \pi_i^N = 0;$$

i.e., for any weak* limit point π^* of π^N, $\pi^*([0,\eta]) = 0$. As $\lambda^N \to 1$, Proposition 3.11 of Faure and Schreiber (2014) implies that any weak* limit point of π^N is invariant for the dynamics $x \mapsto F(x)$. As these weak* limit points are supported on $[\eta, 1]$ and the only invariant measure for $x \mapsto F(x)$ on this interval is the Dirac measure δ_{I^*} and the unique positive equilibrium, it follows that π^N converges in the weak* topology to δ_{I^*} as claimed.

9. Proof of Theorem 4.1. Let $E_f = (1, 0)$. Denote by

$$F(S, I) := \begin{pmatrix} 1 - e^{-\mu} + Se^{-\mu-\beta I} \\ e^{-\mu} S\big(1 - e^{-\beta I}\big) + e^{-\mu-\gamma} I \end{pmatrix}.$$

Then

$$DF(E_f) := \begin{pmatrix} e^{-\mu} & -\beta e^{-\mu} \\ 0 & \beta e^{-\mu} + e^{-\mu-\gamma} \end{pmatrix}.$$

It follows that the per-capita growth rate of the disease is still given by (2.2), and the basic reproduction number of (4.1) is still the expression in (2.3). Define the competitive cone:

$$K := \{(u, v) : u \le 0,\ v \ge 0\}.$$

Then the competitive order is defined by

$$(u_1, v_1) \le_K (u_2, v_2) \Longleftrightarrow (u_2 - u_1, v_2 - v_1) \in K \Longleftrightarrow u_2 \le u_1,\ v_1 \le v_2.$$

It is easy to check that $DF(E_f)$ keeps K invariant (see Wang and Jiang (2001)). The competitive models describe systems between two competitors in which an increase of one competitor's density has a negative effect on the other. Their solutions will preserve the order $\le_K$.

Define

$$L(S, I) := \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} e^{-\mu} & -\beta e^{-\mu} \\ 0 & \beta e^{-\mu} + e^{-\mu-\gamma} \end{pmatrix} \begin{pmatrix} S - 1 \\ I \end{pmatrix}.$$

We shall verify that

$$F(S, I) \le_K L(S, I),\ (S, I) \in X. \tag{9.1}$$

By definition, (9.1) is equivalent to

$$\begin{cases} 1 - e^{-\mu} + Se^{-\mu-\beta I} \ge 1 + e^{-\mu}(S - 1) - \beta I e^{-\mu}, \\ e^{-\mu} S\big(1 - e^{-\beta I}\big) + e^{-\mu-\gamma} I \le \big(\beta e^{-\mu} + e^{-\mu-\gamma}\big) I \end{cases} \tag{9.2}$$

on X. By simple computation, (9.2) is equivalent to $S(1 - e^{-\beta I}) \le \beta I,\ (S, I) \in X$, which is obviously true. This shows the competitive ordering relation (9.1) holds.

Let $P_0 := (S_0, I_0) \in X,\ P_n := F^n(P_0) = (S_n(P_0), I_n(P_0))$, and $Q_n := L^n(P_0)$. We shall show that

$$(1, 0) \le_K P_n \le_K Q_n,\ n = 1, 2, \ldots. \tag{9.3}$$

The left inequality is obvious by the definition of competitive order and X. So we will prove the right one. Equation (9.1) deduces that (9.3) holds for $n = 1$. Suppose that (9.3) holds for n. Then using (9.1) and the order preserving for $DF(E_f)$, we obtain

$$P_{n+1} = F(P_n) \le_K L(P_n) \le_K L(Q_n) = Q_{n+1}.$$

By mathematical induction, (9.3) holds.

Let $\alpha < 1$. Then $Q_n \to (1, 0)$ as $n \to \infty$. Therefore, (i) is proved by (9.3).

Suppose $\alpha > 1$. Then we shall prove that the system (4.1) is uniformly persistent with respect to $(X_0, \partial X_0)$; that is, there exists $\eta > 0$ such that

$$\liminf_{n\to\infty} I_n(P_0) \ge \eta \text{ for all } P_0 = (S_0, I_0) \in X_0. \tag{9.4}$$

We can see that E_f is the maximal compact invariant set in ∂X_0, which is positively invariant with respect to F and lies on the stable manifold of E_f. Recalling the Hofbauer and So uniform persistence theorem (see Theorem 4.1 of Hofbauer and So (1989)), the system (4.1) is uniformly persistent if and only if

(a) E_f is isolated in X in the sense that there exists a closed neighborhood U of E_f such that E_f is the largest invariant set in U, and

(b) $W^s(E_f) \subset \partial X_0$, where

$$W^s(E_f) = \{P_0 := (S_0, I_0) \in X : \lim_{n\to\infty} (S_n(P_0), I_n(P_0)) = (1, 0)\}.$$

Since $DF(E_f)$ has the eigenvalues $e^{-\mu} < 1$ and $\alpha > 1$, the stable (unstable) manifold of E_f is one-dimensional; see Theorem 1.4.2 of Guckenheimer and Holmes (1983). This means that E_f is a hyperbolic equilibrium and a saddle. Thus, (a) follows immediately from the Hartman and Grobman theorem (see Theorem 1.4.1 of Guckenheimer and Holmes (1983)) because F is conjugate to the hyperbolic linear map $DF(E_f)$ on a neighborhood of E_f. Besides, we have that $\partial X_0 := \{(S, 0) : 0 \le S \le 1\}$ is positively invariant and $W^s(E_f) = \{(S, 0) : 0 < S \le 1\} \subset \partial X_0$; that is, (b) holds. This verifies the uniform persistence, and hence the system (4.1) admits an attractor in X_0.

It follows from (9.4) that (4.1) contains a compact attractor $K \subset \{(S, I) \in X : I \ge \eta\}$. Besides, by the first equality of (4.1), we have $S_n(P_0) \ge 1 - e^{-\mu}$ for $n \ge 1$ and $P_0 \in X$. This implies that $K \subset \{(S, I) \in X : S \ge 1 - e^{-\mu}\}$. As a result,

$$K \subset \{(S, I) \in X : S \ge 1 - e^{-\mu},\ I \ge \eta,\ S + I \le 1\}.$$

This proves (ii).

It remains to prove (iii). First, we consider the system (4.1) with $\lambda_0 = (\mu_0, \beta_0, 0)$ and $\alpha(\mu_0, \beta_0, 0) > 1$. We shall prove that it admits a globally stable equilibrium $(1 - I^*, I^*)$ in X, where $F(I^*) = I^*$ with $0 < I^* < 1$; that is, $K = \{(1 - I^*, I^*)\}$.

Let $\Delta_n := S_n + I_n$. Then from (4.1) it follows that

$$\Delta_{n+1} = 1 - e^{-\mu_0} + e^{-\mu_0}\Delta_n. \tag{9.5}$$

It is easy to see that (9.5) has the positive equilibrium 1 and all positive orbits tend to 1 as $n \to +\infty$. Therefore, the system (4.1) is reduced to the system (2.1) with $\mu = \mu_0, \beta = \beta_0, \gamma = 0$. Applying Theorem 2.1(ii), we get that the system (2.1) has a globally stable equilibrium I^* in $(0, 1)$. Thus the system (4.1) admits a globally stable equilibrium $(1 - I^*, I^*)$ in X, where $F(I^*) = I^*$ with $0 < I^* < 1$. Recalling the proof of Theorem 2.1(ii) (see Figure 7.1 and Figure 7.2(a)), we have

$$|F'(I^*)| = \left|e^{-\mu_0-\beta_0 I^*}\left(1 + \beta_0(1 - I^*)\right)\right| < 1. \tag{9.6}$$

Next, we will prove that the spectral radius of the Jacobian matrix for $F(S, I)$ at the positive equilibrium $E^*(S^*, I^*) := (1 - I^*, I^*)$ is strictly less than 1.

An easy calculation yields that

$$DF(E^*) := \begin{pmatrix} e^{-\mu_0-\beta_0 I^*} & -\beta_0 S^* e^{-\mu_0-\beta_0 I^*} \\ e^{-\mu_0}(1 - e^{-\beta_0 I^*}) & \beta_0 S^* e^{-\mu_0-\beta_0 I^*} + e^{-\mu_0} \end{pmatrix},$$

$$\begin{pmatrix} e^{-\mu_0-\beta_0 I^*} & -\beta_0 S^* e^{-\mu_0-\beta_0 I^*} \\ e^{-\mu_0}(1 - e^{-\beta_0 I^*}) & \beta_0 S^* e^{-\mu_0-\beta_0 I^*} + e^{-\mu_0} \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = F'(I^*) \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

and $\det DF(E^*) = e^{-\mu_0}F'(I^*)$. This proves that $F'(I^*)$ and $e^{-\mu_0}$ are two eigenvalues of $DF(E^*)$; the spectral radius of $DF(E^*)$ is strictly less than 1.

In what follows, we shall use Theorem 2.1 of Smith and Waltman (1999) to prove that C_P is open in the parameter space P.

For this purpose, denote by $\|\cdot\|$ the Euclidean norm of $\mathbb{R}^3$ and $B_C(\lambda_0, s)$ the open ball in P of radius s about the point λ_0. We fix a $\lambda_0 \in C_P$ and an $s_0 \in (0, \mu_0)$ such that $\alpha(\lambda) > 1$ for any $\lambda \in B_C(\lambda_0, s_0)$. In order to consider the perturbed systems for parameters, we set $F_{\lambda_0}(S, I)$ the given mapping and $F_\lambda(S, I) = (S_\lambda(S, I), I_\lambda(S, I))$ the mappings for $\lambda \in B_C(\lambda_0, s_0)$. Define

$$R_\lambda(S, I) = \begin{cases} \frac{I_\lambda(S,I)}{I} & \text{if } I > 0, \\ \beta e^{-\mu} S + e^{-\mu-\gamma} & \text{if } I = 0. \end{cases}$$

Then $R_\lambda(S, I)$ is continuous on X. By induction, it is not difficult to prove that

$$\begin{aligned} F_\lambda^n(S, 0) =& \big(1 - e^{-n\mu} + e^{-n\mu}S, 0\big), \\ R_\lambda(F_\lambda^n(S, 0)) =& \beta e^{-\mu}\big(1 - e^{-n\mu} + e^{-n\mu}S\big) + e^{-\mu-\gamma} \end{aligned} \tag{9.7}$$

for $n = 1, 2, \ldots$. We claim that there exist an $s_1 \in (0, s_0]$, an integer $N > 0$, $\delta > 0$, and $\rho > 1$, all only depending on λ_0, such that

$$I_\lambda^N(S, I) \geq \rho I \text{ for all } \lambda \in \overline{B_C(\lambda_0, s_1)} \text{ and } I \in [0, \delta], \tag{9.8}$$

where $F_\lambda^n(S, I) = (S_\lambda^n(S, I), I_\lambda^n(S, I))$.

From (9.7) it follows that

$$R_\lambda(F_\lambda^n(S, 0)) \geq \alpha(\lambda) - \beta e^{-(n+1)\mu}.$$

By the continuity of $\alpha(\lambda)$, there exists an $s_1 \in (0, s_0]$ such that $\alpha(\lambda) > \frac{\alpha(\lambda_0)+1}{2}$ for all $\lambda \in \overline{B_C(\lambda_0, s_1)}$, and hence

$$R_\lambda(F_\lambda^n(S, 0)) \geq \frac{\alpha(\lambda_0) + 1}{2} - (\beta_0 + s_0)e^{-(n+1)(\mu_0 - s_0)} \text{ for all } \lambda \in \overline{B_C(\lambda_0, s_1)}.$$

This implies that there is an integer $N > 0$, only depending on λ_0, such that

$$R_\lambda(F_\lambda^N(S, 0)) > \frac{\alpha(\lambda_0) + 3}{4} \text{ for all } \lambda \in \overline{B_C(\lambda_0, s_1)}. \tag{9.9}$$

Using (9.9) and the uniform continuity of $R_\lambda(F_\lambda^N(S, I))$ on $\overline{B_C(\lambda_0, s_1)} \times X$, we obtain that there is a $\delta > 0$, only depending on λ_0, such that

$$R_\lambda(F_\lambda^N(S, I)) > \frac{\alpha(\lambda_0) + 7}{8} := \rho \text{ for all } \lambda \in \overline{B_C(\lambda_0, s_1)} \text{ and } 0 \leq I \leq \delta,$$

which implies that (9.8) holds; thus the claim is proved.

By (9.8), we get that for each $\lambda \in \overline{B_C(\lambda_0, s_1)}$,

$$I_\lambda^{mN}(S, I) \geq \rho^m I \text{ if } F_\lambda^{kN}(S, I) \in [0, 1] \times (0, \delta] \text{ for } k = 0, 1, \ldots, m-1.$$

This shows that there exists at least a positive integer m with the property

$$F_\lambda^{mN}(S, I) \in [0, 1] \times (\delta, 1] \text{ if } (S, I) \in [0, 1] \times (0, \delta].$$

Let $U = X_0,\ \Lambda = \overline{B_C(\lambda_0, s_1)}$, and $B_\lambda \equiv B = X_0 \cap ([0,1] \times [\delta, 1])$, and take

$$N((S,I),\lambda) = \begin{cases} mN & \text{if } (S,I) \in B, \\ 0 & \text{if } X_0 \backslash B. \end{cases}$$

Then the assumption (H1) of Theorem 2.1 of Smith and Waltman (1999) holds. Set $C = \overline{\bigcup_{\lambda\in\Lambda} F_\lambda(B)}$. It follows from the compactness of $\Lambda \times B \subset P \times X_0$ and the continuity of $F_\lambda(S,I)$ that $\bigcup_{\lambda\in\Lambda} F_\lambda(B)$ is compact in X_0. This implies that the assumption (H2) of Theorem 2.1 of Smith and Waltman (1999) still holds. Applying Theorem 2.1 of Smith and Waltman (1999), we conclude that $\overline{B_C(\lambda_0, s_1)} \subset C_P$. The proof is complete.

10. Proof of Theorem 5.2. As in the proof of Theorem 3.2, we use results from Faure and Schreiber (2014) to prove Theorem 5.2. We begin by verifying that Standing Hypothesis 1.1 of Faure and Schreiber (2014) holds. For all $\delta > 0$ and $N \geq 1$, define

$$\beta_\delta(N) = \sup_{x,y\geq 0, x+y\leq 1} \mathbb{P}\left[\|(S_{n+1}, I_{n+1}) - F(x,y)\|_\infty \geq \delta | (S_n, I_n) = (x,y)\right],$$

where $F(x,y) = (1 - e^{-\mu} + e^{-\mu-\beta y}x, xe^{-\mu}(1-e^{-\beta y}) + ye^{-\mu-\gamma})$ corresponds to the right-hand side of the deterministic model in (4.1) and $\|(x,y)\|_\infty = \max\{|x|, |y|\}$ corresponds to the sup norm. Standing Hypothesis 1.1 of (Faure and Schreiber, 2014) requires that $\lim_{N\to\infty} \beta_\delta(N) = 0$ for all $\delta > 0$. Like Proposition 8.1 for the stochastic SIS model, the following proposition proves something stronger using large deviation estimates.

PROPOSITION 10.1. *There exists a function $\rho : (0,\infty) \to (0,\infty)$ such that*

$$\beta_\delta(N) \leq \exp(-N\rho(\delta))$$

for all $N \geq 1$ and $\delta > 0$.

Proof. We begin by observing that $NS_n - W_{n+1}$ and $X_{n+1} + Z_{n+1}$ in (5.1) conditioned on $(S_n, I_n) = (x,y)$ are independent binomials where $NS_n - W_{n+1}$ has Nx trials with a probability of success $a_1(y) = 1 - e^{-\mu}(1 - e^{-\beta y})$ and $X_{n+1} + Z_{n+1}$ has $N(1-x)$ trials with a probability of success $b_1 = 1 - e^{-\mu}$. Let F_1 and F_2 denote the first and second coordinates of the function F. Using the exponential Markov inequality as in the proof of Proposition 8.1, we get

$$\begin{aligned} &\frac{1}{N}\log \mathbb{P}[N(S_{n+1} - F_1(x,y)) \geq N\delta | (S_n, I_n) = (x,y)] \leq -\delta t + \psi_1(t,x,y), \\ &\frac{1}{N}\log \mathbb{P}[N(F_1(x,y) - S_{n+1}) \geq N\delta | (S_n, I_n) = (x,y)] \leq -\delta t + \psi_1(-t,x,y) \end{aligned} \tag{10.1}$$

for all t, δ and where

$$\psi_1(t,x,y) = -tF_1(x,y) + x\log(1 - a_1(y) + a_1(y)e^t) + (1-x)\log(1 - b_1 + b_1 e^t).$$

As $a_1(y), b_1 \in [0,1]$ for all $y \geq 0$, $\psi_1(t,x,y)$ is a smooth function, i.e., has continuous derivatives of all orders. We have $\psi_1(0,x,y) = \frac{\partial\psi_1}{\partial t}(0,x,y) = 0$, and $\frac{\partial^2\psi_1}{\partial t^2}(t,x,y) \geq 0$ for all $t,\ x, y \in X$ where one recalls that $X = \{(x,y) | x \geq 0, y \geq 0, x + y \leq 1\}$. Given $\delta > 0$, the compactness of $\{0\} \times X$ implies there exists $t^*(\delta) > 0$ such that

$|\frac{\partial \psi_1}{\partial t}(t,x,y)| \leq \delta/2$ and $t\frac{\partial \psi_1}{\partial t}(t,x,y) \geq 0$ for all $(t,x,y) \in [-t^*(\delta), t^*(\delta)] \times X$. Thus, for any $(t,x,y) \in [0,t^*(\delta)] \times X$, the fundamental theorem of calculus implies

$$0 \leq \psi_1(t,x,y) = \int_0^t \frac{\partial \psi_1}{\partial t}(s,x,y)ds \leq \frac{t\delta}{2}$$

and

$$0 \leq \psi_1(-t,x,y) = -\int_0^t \frac{\partial \psi_1}{\partial t}(-s,x,y)ds \leq \frac{t\delta}{2}.$$

Define

$$\rho_1(\delta) = \delta t^*(\delta) - \max_{x,y\in[0,1],x+y\leq 1} \psi_1(t^*(\delta),x,y) > 0.$$

Our estimates about ψ_1 imply that $\rho_1(\delta) \geq \delta t^*(\delta)/2 > 0$. Furthermore, (10.1) implies that

$$\mathbb{P}[|S_{n+1} - F_1(x,y)| \geq \delta | (S_n, I_n) = (x,y)] \leq \exp(-N\rho_1(\delta))$$

for all $x, y \in [0,1]$ with $x + y \leq 1$ and $\delta > 0$.

W_{n+1} and Y_{n+1} conditioned on $(S_n, I_n) = (x,y)$ in (5.1) are also independent binomial random variables where W_{n+1} has Nx trials with a probability of success $e^{-\mu}(1 - e^{-\beta y})$ and Y_{n+1} has Ny trials with a probability of success $e^{-\gamma}$. Therefore using the exponential Markov inequality and similar estimates used for $\psi_1(t,x,y)$, one can show there exists a function $\rho_2 : (0,\infty) \to (0,\infty)$ such that

$$\mathbb{P}[|I_{n+1} - F_2(x,y)| \geq \delta | (S_n, I_n) = (x,y)] \leq \exp(-N\rho_2(\delta))$$

for all $x, y \in [0,1]$ with $x + y \leq 1$ and $\delta > 0$. Setting $\rho(\delta) = \min\{\rho_1(\delta), \rho_2(\delta)\}$ completes the proof of the proposition. ∎

To prove the first result of Theorem 5.2, assume that $\alpha \leq 1$. Theorem 4.1 implies that $(1,0)$ is globally stable for the deterministic model $(S,I) \mapsto F(S,I)$. Theorem 3.12 of Faure and Schreiber (2014), which only requires Standing Hypothesis 1.1, implies that $\lim_{N\to\infty} \pi^N = \delta_{(1,0)}$ in the weak* topology. Define $R(x,y) = F_2(x,y)/y$ for $y \in (0,1], x \in [0,1]$, and $x + y \leq 1$. Equation (9.2) in the proof of Theorem 4.1 implies that $R(x,y) \leq \alpha$. For $N \geq 1$, the quasi-stationarity of π^N implies

$$\begin{aligned}
\lambda^N \sum_{x,y\in\mathcal{S}_+} y\pi^N_{x,y} &= \sum_{x,y\in\mathcal{S}_+} \mathbb{E}\left[I_{n+1} \middle| (S_n, I_n) = (x,y)\right] \pi^N_{x,y} \\
&= \sum_{x,y\in\mathcal{S}_+} F_2(x,y)\, \pi^N_{x,y} \\
&= \sum_{x,y\in\mathcal{S}_+} yR(x,y)\, \pi^N_{x,y} \\
&\leq \alpha \sum_{x,y\in\mathcal{S}_+} y\pi^N_{x,y}.
\end{aligned}$$

Since $\sum_{x,y\in\mathcal{S}_+} y\pi^N_{x,y} > 0$, $\lambda^N \leq \alpha$ for all $N \geq 1$ as claimed.

To prove the second result of Theorem 5.2, assume $\alpha > 1$, in which case Theorem 4.1 implies that there exists a global, compact attractor $K \subset (0,1) \times (0,1)$ for $(S,I) \mapsto F(S,I)$. Assertion (b) of Lemma 3.9 of Faure and Schreiber (2014) implies that there exists $\delta > 0$ such that $1 - \lambda^N \leq \beta_\delta(N)$ for all $N \geq 1$. Proposition 10.1 implies that

$$\limsup_{N\to\infty} \frac{1}{N} \log(1 - \lambda^N) \leq -\rho(\delta).$$

The assumption in assertion (b′) of Lemma 3.9 of Faure and Schreiber (2014) holds for an argument similar to the proof of Theorem 3.2, and consequently any weak* limit point π^* of π^N satisfies $\pi^*(K) = 1$. As $\lambda^N \to 1$, Proposition 3.11 of Faure and Schreiber (2014) implies that any weak* limit point of π^N is invariant for the dynamics $(x, y) \mapsto F(x, y)$.

REFERENCES

D. Aldous, B. Flannery, and J. L. Palacios, *Two applications of urn processes the fringe analysis of search trees and the simulation of quasi-stationary distributions of Markov chains*, Probab. Eng. Inf. Sci., 2 (1988), pp. 293–307.

L. J. S. Allen, *Some discrete-time SI, SIR, and SIS epidemic models*, Math. Biosci., 124 (1994), pp. 83–105, https://doi.org/10.1016/0025-5564(94)90025-6.

L. J. S. Allen and A. M. Burgin, *Comparison of deterministic and stochastic SIS and SIR models in discrete time*, Math. Biosci., 163 (2000), pp. 1–33, https://doi.org/10.1016/s0025-5564(99)00047-4.

L. J. S. Allen and P. van den Driessche, *The basic reproduction number in some discrete-time epidemic models*, J. Differ. Equ. Appl., 14 (2008), pp. 1127–1147.

R. M. Anderson and R. M. May, *Population biology of infectious diseases: Part* I, Nature, 280 (1979), pp. 361–367.

R. M. Anderson and R. M. May, *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press, New York, 1991.

H. Andersson and T. Britton, *Stochastic epidemics in dynamic populations: Quasi-stationarity and extinction*, J. Math. Biol., 41 (2000), pp. 559–580.

P. Andersson and D. Lindenstrand, *A stochastic sis epidemic with demography: Initial stages and time to extinction*, J. Math. Biol., 62 (2011), pp. 333–348.

R. E. Baker, S. W. Park, W. Yang, G. A. Vecchi, C. J. E. Metcalf, and B. T. Grenfell, *The impact of COVID-19 nonpharmaceutical interventions on the future dynamics of endemic infections*, Proc. Natl. Acad. Sci. USA, 117 (2020), pp. 30547–30553.

A. D. Barbour, *The duration of the closed stochastic epidemic*, Biometrika, 62 (1975), pp. 477–482.

M. S. Bartlett, *An Introduction to Stochastic Processes With Special Reference to Methods and Applications*, Cambridge University Press, London-New York, 1966.

C. Castillo-Chavez and A.-A. Yakubu, *Discrete-time sis models with complex dynamics*, Nonlinear Anal., Theory Methods Appl., 47 (2001), pp. 4753–4762.

D. Clancy and E. Tjia, *Approximating time to extinction for endemic infection models*, Method. Comput. Appl. Probab., 20 (2018), pp. 1043–1067, https://doi.org/10.1007s11009-018-9621-8.

J. N. Darroch and E. Seneta, *On quasi-stationary distributions in absorbing discrete-time finite Markov chains*, J. Appl. Prob., 2 (1965), pp. 88–100, https://doi.org/10.2307/3211876.

J. N. Darroch and E. Seneta, *On quasi-stationary distributions in absorbing continuous-time finite Markov chains*, J. Appl. Probab., 4 (1967), pp. 192–196.

O. Diekmann, H. Heesterbeek, and T. Britton, *Mathematical Tools for Understanding Infectious Disease Dynamics*, in Princetion Ser. Theor. Comput. Biol. Vol. 7, Princeton University Press, Princeton, NJ, 2012.

K. E. Emmert and L. J. S. Allen, *Population persistence and extinction in a discrete-time, stage-structured epidemic model*, J. Differ. Equ. Appl., 10 (2004), pp. 1177–1199.

K. E. Emmert and L. J. S. Allen, *Population extinction in deterministic and stochastic discrete-time epidemic models with periodic coefficients with applications to amphibian populations*, Nat. Res. Model., 19 (2006), pp. 117–164.

Y. Enatsu, Y. Nakata, and Y. Muroya, *Global stability for a class of discrete SIR epidemic models*, Math. Biosci. Eng., 7 (2010), pp. 347–361.

M. Faure and S. J. Schreiber, *Quasi-stationary distributions for randomly perturbed dynamical systems*, Ann. Appl. Probab., 24 (2014), pp. 553–598.

M. J. Feigenbaum, *Universal Behaviour in Nonlinear Systems*, Phys. D, 7(1983), pp. 16–39.

B. F. Finkenstädt and B. T. Grenfell, *Time series modelling of childhood diseases: A dynamical systems approach*, J. R. Stat. Soc. Ser. C Appl. Stat., 49 (2000), pp. 187–205.

E. Foxall, *Extinction Time of the Logistic Process*, preprint, arXiv:1805.08339, 2020, http://arxiv.org/abs/1805.08339.

V. Grimm and C. Wissel, *The intrinsic mean time to extinction: A unifying approach to analysing persistence and viability of populations*, Oikos, 105 (2004), pp. 501–511.

J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer-Verlag, New York, 1983.

M. Hairer, *Ergodic Properties of Markov Processes*, Lecture notes, 2018, http://www.hairer.org/notes/Markov.pdf.

J. Hofbauer and J. W.-H. So, *Uniform persistence and repellors for maps*, Proc. Amer. Math. Soc., 107 (1989), pp. 1137–1142.

J. Jiang and S. Yu, *Stable cycles for attractors of strongly monotone discrete-time dynamical systems*, J. Math. Anal. Appl., 202 (1996), pp. 349–362.

V. A. Karatayev, M. Anand, and C. T. Bauch, *Local lockdowns outperform global lockdown on the far side of the COVID-19 epidemic curve*, Proc. Natl. Acad. Sci. USA, 117 (2020), pp. 24575–24580.

M. J. Keeling and P. Rohani, *Modeling Infectious Diseases in Humans and Animals*, Princeton University Press, Princeton, NJ, 2011.

P. Klepac, L. W. Pomeroy, O. N. Bjørnstad, T. Kuiken, A. D. M. E. Osterhaus, and J. M. Rijks, *Stage-structured transmission of phocine distemper virus in the Dutch 2002 outbreak*, Proc. Royal Soc. B, 276 (2009), pp. 2469–2476, https://doi.org/10.1098/rspb.2009.0175.

J. D. Kong, C. Jin, and H. Wang, *The inverse method for a childhood infectious disease model with its application to pre-vaccination and post-vaccination measles data*, Bull. Math. Biol., 77 (2015), pp. 2231–2263.

P. Kratz and E. Pardoux, *Large Deviations for Infectious Diseases Models*, in Séminaire de Probabilités XLIX, C. Danati-Mortin, A. Lejay, and A. Rouault, eds., Springer, Basel, Switzerland, 2018, pp. 221–327.

R. J. Kryscio and C. Lefèvre, *On the extinction of the SIS stochastic logistic epidemic*, J. Appl. Probab., 26 (1989), pp. 685–694 https://doi.org/10.2307/3214374.

M. Lindholm and T. Britton, *Endemic persistence or disease extinction: The effect of separation into sub-communities*, Theor. Popul. Biol., 72 (2007), pp. 253–263, https://doi.org/10.1016%2Fj.tpb.2007.05.001.

X. Ma, Yi. Zhou, and H. Cao, *Global stability of the endemic equilibrium of a discrete SIR epidemic model*, Adv. Differ. Equ. 42, (2013), pp. 1–19.

I. Nåsell, *The quasi-stationary distribution of the closed endemic SIS model*, Adv. Appl. Probab., 28 (1996), pp. 895–932, https://doi.org/10.2307%2F1428186.

I. Nåsell, *On the quasi-stationary distribution of the stochastic logistic epidemic*, Math. Biosci., 156 (1999), pp. 21–40.

I. Nåsell, *Extinction and quasi-stationarity in the Verhulst logistic model*, J. Theoret. Biol., 211 (2001), pp. 11–27.

I. Nåsell, *Stochastic models of some endemic infections*, Math. Biosci., 179 (2002), pp. 1–19.

M. Pollicott, H. Wang, and H. Weiss, *Extracting the time-dependent transmission rate from infection data via solution of an inverse ODE problem*, J. Biol. Dyn., 6 (2012), pp. 509–523.

K. Prem, Y. Liu, T. W. Russell, A. J. Kucharski, R. M. Eggo, and N. Davies, *The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: A modelling study*, The Lancet Public Health, 5 (2020), pp. e261–e270.

J. A. Salomon, J. O. Lloyd-Smith, W. M. Getz, S. Resch, M. S. Sánchez, T. C. Porco, and M. W. Borgdorff, *Prospects for advancing tuberculosis control efforts through novel therapies*, PLoS Med., 3 (2006), p. e273.

J. Satsuma, R Willox, A Ramani, B Grammaticos, and A. S. Carstea, *Extending the SIR epidemic model*, Phys. A, 336 (2004), pp. 369–375.

H. L. Smith and P. Waltman, *Perturbation of a globally stable steady state*, Proc. Amer. Math. Soc., 127 (1999), pp. 447–453.

Y. Wang and J. Jiang, *The general properties of discrete-time competitive dynamical systems*, J. Differ. Equ., 24 (2001), pp. 470–493.

G. H. Weiss and M. Dishon, *On the asymptotic behavior of the stochastic and deterministic models of an epidemic*, Math. Biosci., 11 (1971), pp. 261–265, https://doi.org/10.1016/0025-5564(71)90087-3.