



# City Research Online

## City St George's, University of London

**Citation:** Hening, A., Nguyen, D. H., Ta, T. & Ungureanu, S. C. (2025). Long-term behavior of stochastic SIQRS epidemic models. *Journal of Mathematical Biology*, 90(4), 41. doi: 10.1007/s00285-025-02204-1

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

**Permanent repository link:** <https://city-ref-test.eprints-hosting.org/id/eprint/34905/>

**Link to published version:** <https://doi.org/10.1007/s00285-025-02204-1>

**Copyright and Reuse:** Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

# LONG-TERM BEHAVIOR OF STOCHASTIC SIQRS EPIDEMIC MODELS

ALEXANDRU HENING, DANG H. NGUYEN, TRANG TA, AND SERGIU C. UNGUREANU

**ABSTRACT.** In this paper we analyze and classify the dynamics of SIQRS epidemiological models with susceptible, infected, quarantined, and recovered classes, where the recovered individuals can become reinfected. We are able to treat general incidence functional responses. Our models are more realistic than what has been studied in the literature since they include two important types of random fluctuations. The first type is due to small fluctuations of the various model parameters and leads to white noise terms. The second type of noise is due to significant environment regime shifts in the that can happen at random. The environment switches randomly between a finite number of environmental states, each with a possibly different disease dynamic. We prove that the long-term fate of the disease is fully determined by a real-valued threshold  $\lambda$ . When  $\lambda < 0$  the disease goes extinct asymptotically at an exponential rate. On the other hand, if  $\lambda > 0$  the disease will persist indefinitely. We end our analysis by looking at some important examples where  $\lambda$  can be computed explicitly, and by showcasing some simulation results that shed light on real-world situations.

**Keywords.** switching diffusion, epidemic model, ergodicity, invariant measure, quarantine, temporary immunity.

## 1. INTRODUCTION

Severe pandemics of infectious diseases have historically been, and will likely continue to be, a serious worldwide systemic risk. This risk needs to be understood and controlled. The policy debates around the most recent Covid-19 outbreak are a reminder that we need to have better and more realistic models. In this paper we analyze realistic stochastic epidemic models that can help to better understand and predict the evolution of outbreaks. The main goal is to be able to offer a straightforward and easy way to make policy recommendations.

The history of the mathematical modelling of epidemic models is long. Kermack and McKendrick, considered the fathers of mathematical epidemiology, were the first to analyze rigorously deterministic epidemic models - this happened in series of papers called “Contributions to the Mathematical theory of epidemics” [Kermack and McKendrick \[1927, 1932\]](#). In the first paper written in 1927, they assumed that complete immunity is conferred by a single attack, and that an individual is not infective before the moment he gets infected. They supposed that the population is subdivided into three distinct classes: the susceptible class ( $S$ ), infective class ( $I$ ) and the recovered, or removed, class ( $R$ ). Individuals will be transferred consecutively between the classes,  $S \rightarrow I \rightarrow R$ . This kind of model is known in the literature as an SIR model.<sup>1</sup> The dynamics can be modeled by the system of ordinary differential equations

$$(1.1) \quad \begin{cases} dS(t) = (A - I(t)F(S(t), I(t)) - \mu S(t) + \gamma R(t))dt, \\ dI(t) = I(t)(F(S(t), I(t)) - (\mu + \gamma_1 + \gamma_2))dt, \\ dR(t) = (\gamma_2 I(t) - (\mu + \gamma)R(t))dt, \end{cases}$$

where  $A > 0$  is the birth rate of the population,  $\mu$  is the independent death rate of all the groups,  $\gamma_1 > 0$  is the death rate of infected individuals,  $\gamma_2 > 0$  is the recovery rate of infected individuals,  $\gamma > 0$  is the rate of loss of immunity in the recovered group, and  $I(t)F(S(t), I(t))$  indicates the incidence rate, which describes the number of new cases per unit of time. There are many different

---

<sup>1</sup>A person who recovered will have permanent immunity in this type of model.

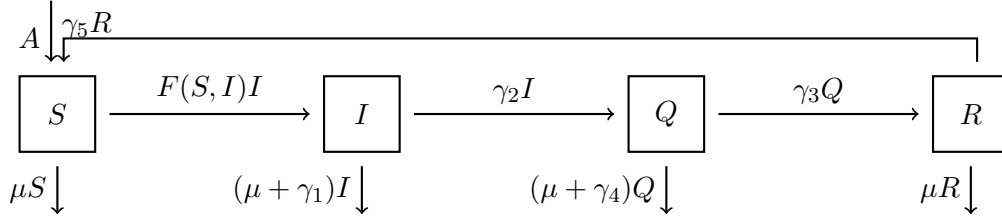


FIGURE 1. Schematic representation of the system (1.2) and of the nonstochastic effects on the population groupings from (1.4).

types of incidence rates in the literature – see [Nguyen et al. \[2020c\]](#), [Phu et al. \[2020\]](#), [Huang et al. \[2009\]](#), [Dieu et al. \[2016\]](#), [Du and Nhu \[2017\]](#), [Du et al. \[2019\]](#), [Jiang et al. \[2016\]](#), [Liu et al. \[2018\]](#), [Nguyen et al. \[2020a\]](#), [Nguyen and Yin \[2019\]](#), [Nie et al. \[2018\]](#), [Ruan and Wang \[2003\]](#), [Zhang et al. \[2013, 2015\]](#), [Tang et al. \[2008\]](#) and the references therein.

It is not always the case that SIR models work well, as some of their assumptions could get violated. For some diseases, like syphilis or Covid-19 [Moein et al. \[2021\]](#), recovered patients can become susceptible again. In order to have more realistic models which can capture such diseases people have started looking at SIRS epidemic models where one can have dynamics of the form  $S \rightarrow I \rightarrow R \rightarrow S$ . During the Covid pandemic governments made use of quarantines, where infected individuals were separated in order to lessen the spread of the pandemic to the susceptible population. Quarantines have been widely used throughout the ages so it is natural to consider models where there is a quarantined class  $Q$  of individuals. Quarantine leads to the isolation of infected individuals in order to try and contain the spread of the disease. We will assume that isolation makes it impossible for the quarantined to get in contact with susceptible individuals. As a result the transmissions of the infection to the susceptible class is reduced. This method has been used for thousands of years to reduce the transmission of diseases among human beings, from leprosy to the plague, TBC, Ebola, smallpox, etc. Models that include quarantine were thought to be appropriate for childhood diseases, where quarantine seems to destabilize the epidemic. However, this is sometimes detrimental as the destabilization can lead to oscillations and recurrent outbreaks of childhood diseases [Feng and Thieme \[1995\]](#). There are various ways to model the individuals who recover from a disease. In certain diseases, the recovered individuals are still susceptible to infection (SIQS models), while in others, recovered individuals are better modeled as permanently immune to the infection (SIQR models). SIQRS models combine features of these models for more generality, letting parametrisation decide the correct evolution of the susceptible and recovered categories.

The SIQRS dynamics can be modelled, in the absence of noise, by the system of ordinary differential equations

$$(1.2) \quad \begin{cases} dS(t) = [A - F(S(t), I(t))I(t) - \mu S(t) + \gamma_5 R(t)]dt, \\ dI(t) = [F(S(t), I(t)) - \mu - \gamma_1 - \gamma_2]I(t)dt, \\ dQ(t) = [\gamma_2 I(t) - (\mu + \gamma_3 + \gamma_4)Q(t)]dt, \\ dR(t) = [\gamma_3 Q(t) - (\mu + \gamma_5)R(t)]dt. \end{cases}$$

Here,  $A$  is again the population growth rate and  $\mu$  is the independent death rate, common to all groups.  $\gamma_1$  is the death rate specific to the infected category, attributed to the disease,  $\gamma_2$  is the rate of transfer of the infected into the quarantined category,  $\gamma_3$  is the rate of recovery of the

quarantined individuals,  $\gamma_4$  is the rate of death in the quarantined group, also attributed to the disease, and finally  $\gamma_5$  is the rate of transfer of the recovered back to the susceptible group. If this last parameter is 0, then the recovered stay immune. The term  $F(S, I)I(t)$  is the incidence rate and it describes the number of new infections per unit of time.  $F(S, I)$  is the incidence rate per infective individual. In the most common model  $F(S, I)I = \beta IS$ , but it turns out that it is very important to consider more general nonlinear functions  $F$ .

There are many plausible mechanisms that can lead to nonlinear incidence rates. Many common pathogens cannot live separated from the host for long periods of time. In addition, susceptible individuals can usually be infected only once the concentration of the pathogen reaches a certain level. At low densities of infected individuals, low  $I$ , the threshold for infectivity may not be reached often. If  $I$  increases, the threshold will be reached more often and infections will rise faster. If the disease is vectored and the vector must attack on average  $q$  infective individuals in order to make its next attack infective, one can show that the incidence rate will be proportional to  $I^q$  when one makes the natural assumption that the attack times in a given time period follow a Poisson distribution. Nonlinear incidence rates also arise naturally in models of soil-transmitted parasitic worm infections – helminth infections – when the probability of pairing is taken into account [Liu et al. \[1986\]](#). As showcased by [Liu et al. \[1986\]](#), SIR and SIRS models exhibit very different behavior if the incidence rate is not  $F(S, I)I = \beta IS$ , i.e., it is nonlinear. The dynamical behavior can include Hopf, saddle-node and homoclinic bifurcations.

Populations are influenced by random environmental fluctuations, and these fluctuations have a significant impact on growth dynamics. In order to have robust dynamic models, it is important to include these fluctuations in the model. Furthermore, environmental fluctuations can have a significant impact on whether a species survives in the long-term. For example, in certain settings, persistence can be reversed into extinction by environmental fluctuations, while in others, extinction can be transformed into persistence [Benaïm and Lobry \[2016\]](#), [Hening and Nguyen \[2020\]](#), [Hening et al. \[2022\]](#). Common ways to model environmental fluctuations in population dynamics are systems of stochastic differential equations and Markov processes [Chesson \[1982\]](#), [Evans et al. \[2013, 2015\]](#), [Lande et al. \[2003\]](#), [Schreiber and Lloyd-Smith \[2009\]](#), [Schreiber et al. \[2011\]](#), [Benaïm and Schreiber \[2009\]](#), [Benaïm et al. \[2008\]](#), [Benaïm \[2018\]](#), [Hening et al. \[2021\]](#). A typical approach is to transform ordinary differential equations (ODE) models into stochastic differential equations (SDE) models. This boils down to saying that the various rates in the disease ecosystem are not constant, but fluctuate around their average values according to white noise. There is a well established general theory of coexistence and extinction for systems that come from ecology, when they are in Kolmogorov form [Schreiber et al. \[2011\]](#), [Hening and Nguyen \[2018\]](#), [Hening et al. \[2021\]](#).

In [Hethcote et al. \[2002\]](#) the authors study the role of quarantine in a series of SIQR models. They show that ‘After examining many parameter sets and the corresponding periodic solutions, we have found that as one aspect becomes more realistic, another becomes less realistic. Thus we have not been able to find a parameter set that matches every feature of the observed data.’ Nevertheless, they follow [Feng and Thieme \[2000\]](#) and say that stochastic models might be significantly more realistic.

It is very important to use stochastic models in epidemiological modelling, when possible. From the start it is natural to define the probability of disease transmission between two individuals instead of stating certainly that transmission will or will not happen. Furthermore, as the authors of [Andersson and Britton \[2012\]](#) explain, when one considers the extinction of endemic diseases, this phenomenon can only be analyzed using stochastic models since ‘extinction occurs when the epidemic process deviates from the expected level’.

In [Nguyen et al. \[2020c\]](#) the authors consider a model which is perturbed by white noise and has the general incidence rate  $S(t)I(t)/F(S(t), I(t))$ , where  $F(S(t), I(t))$  is a locally Lipschitz continuous function in both variables. They show that if the real-valued threshold  $\lambda < 0$ , then the disease will eventually disappear, and if  $\lambda > 0$  the epidemic will become permanent.

In certain systems, it makes more sense to assume that when the environment changes, the dynamics also changes significantly. In a deterministic setting this can be modelled by periodic vector fields which can be interpreted to mimic seasonal fluctuations. In the random setting, these types of fluctuations are captured by piecewise deterministic Markov processes (PDMP) – see [Davis \[1984\]](#) for an introduction to PDMP. In a PDMP, the environment switches between a fixed finite number of states to each of which we associate an ODE. In each state the dynamics is given by the flow of its associated ODE. After a random time, the environment switches to a different state, and the dynamics is governed by the ODE from that state.

Our goal is to capture both white noise and discrete types of fluctuations. As such we will look at SSDE (stochastic differential equations with switching). These processes involve a discrete component that keeps track of the environment and which changes at random times. In a fixed environmental state the system is modelled by stochastic differential equations. This way we can capture the more realistic behaviour of two types of environmental fluctuations: (1) large changes of the environment (daily or seasonal changes) and (2) small fluctuations within each environment. Random switching has been used in epidemiological settings [Lan et al. \[2019\]](#), [Lin and Jiang \[2014\]](#), [Wang and Liu \[2023\]](#). In [Phu et al. \[2020\]](#) the authors considered the system

$$(1.3) \quad \begin{cases} dS(t) = [\Lambda(r(t)) - F(S(t), I(t), r(t))I(t) - \mu(r(t))S(t) + \gamma(r(t))I(t)]dt \\ \quad + \sigma_1(r(t))S(t)dW_1(t), \\ dI(t) = I(t)[F(S(t), I(t), r(t)) - \tilde{\gamma}(r(t))]dt \\ \quad + \sigma_2(r(t))I(t)dW_2(t), \end{cases}$$

where  $\sigma_1, \sigma_2 \neq 0$  parametrise the variance of the white noise,  $\tilde{\gamma} = \mu_1 + \gamma$ , where  $\mu_1$  is a separate death rate for the infected, and  $r(t)$  is a Markov chain taking values in a finite space.

Our goal is to study (1.2) when there are both types of environmental fluctuations present. In its stochastic form, system (1.2) becomes

$$(1.4) \quad \begin{cases} dS(t) = [A(r(t)) - F(S(t), I(t), r(t))I(t) - \mu(r(t))S(t) + \gamma_5(r(t))R(t)]dt \\ \quad + \sigma_1(r(t))S(t)dW_1(t), \\ dI(t) = [F(S(t), I(t), r(t)) - \mu(r(t)) - \gamma_1(r(t)) - \gamma_2(r(t))]I(t)dt \\ \quad + \sigma_2(r(t))I(t)dW_2(t), \\ dQ(t) = [\gamma_2(r(t))I(t) - (\mu(r(t)) + \gamma_3(r(t)) + \gamma_4(r(t)))Q(t)]dt \\ \quad + \sigma_3(r(t))Q(t)dW_3(t), \\ dR(t) = [\gamma_3(r(t))Q(t) - (\mu(r(t)) + \gamma_5(r(t)))R(t)]dt \\ \quad + \sigma_4(r(t))R(t)dW_4(t), \end{cases}$$

where  $\sigma_i \neq 0$  for  $i = 1, \dots, 4$ , and the other parameters are interpreted as for (1.2).

Here  $\{r(t), t \geq 0\}$  is a right-continuous Markov chain on  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ , taking values in the finite state space  $\mathcal{M} := \{1, \dots, m_0\}$ . The Markov chain  $\{r(t), t \geq 0\}$  is supposed to have an irreducible generator  $\mathbf{Q} = (q_{kl})_{m_0 \times m_0}$ . This implies in particular that  $\{r(t), t \geq 0\}$  has a unique invariant probability measure  $\pi = (\pi_1, \dots, \pi_{m_0})$  which can be found by solving  $\pi \mathbf{Q} = 0$  with  $\sum_k \pi_k = 1$ . We will also assume that the Markov chain is independent of the Brownian motions so that:

$$(1.5) \quad \begin{aligned} \mathbb{P}\{r(t + \Delta) = j | r(t) = i, r(s), s \leq t\} &= q_{ij}\Delta + o(\Delta), & \text{if } i \neq j, \text{ and} \\ \mathbb{P}\{r(t + \Delta) = i | r(t) = i, r(s), s \leq t\} &= 1 + q_{ii}\Delta + o(\Delta). \end{aligned}$$

Intuitively, the above equations tells us that if we know that at  $t$  the environment is in state  $i$ , then the probability of jumping in a small time  $\Delta$  to state  $j$  is approximately  $q_{ij}\Delta$ .

For the field of epidemiology, one of the fundamental questions is analyzing when a disease is endemic, i.e., persisting over long periods of time, and when it is epidemic, i.e., when it only lasts a short time. Our main goal is to provide sharp conditions which tell us when the disease is endemic and when it is epidemic. Furthermore, we are able to characterize endemic diseases by proving that the stochastic process converges to a stationary distribution exponentially fast. When the disease is an epidemic, we can prove that the disease disappears exponentially fast and can find the rate of extinction.

The paper is structured as follows. In Section 2, we present the main results. The proofs for the extinction results appear in Section 3 while those for the persistence of the disease appear in Section 4. The discussion of the results and the simulations is in Section 5.

## 2. MAIN RESULTS

The following assumption is supposed to hold throughout the paper.

**Assumption 2.1.**  $F(s, i, k)$  is a locally Lipschitz function in  $(s, i) \in \mathbb{R}_+^2$  and continuous in  $i$  at  $i = 0$  uniformly in  $s$ , i.e.,

$$(2.1) \quad \lim_{i \rightarrow 0} \sup_{s \geq 0, k \in \mathcal{M}} \{|F(s, i, k) - F(s, 0, k)|\} = 0.$$

Suppose further that  $F(s, 0, k)$  is non-decreasing in  $s$  for each fixed  $k \in \mathcal{M}$  and  $F(s, 0, k) \leq C_F(s + 1)$ ,  $\forall (s, k) \in [0, \infty) \times \mathcal{M}$  for some constant  $C_F > 0$ .

**Remark 2.1.** The functional forms  $F(S, I) = \beta S/(1 + m_1 S)$  (Holling type II),  $F(S, I) = \beta S$  (bilinear functional response), and  $F(S, I) = \beta S/(1 + m_1 S + m_2 I)$  (Beddington-DeAngelis functional response) satisfy Assumption 2.1. Moreover, we can also include functional responses of the form  $IS/F(S, I)$ , from [Nguyen et al. \[2020c\]](#), by replacing  $S/F(S, I)$  with  $F(S, I)$ . As a result one can include, for example, functional responses of the form  $IS/(S + I)$  and of the form  $\beta IS/(1 + pI + qI^2)$ .

Consider the dynamics of the population without disease by letting  $I(t) = Q(t) = R(t) = 0$  in the first equation of (1.4). We obtain

$$(2.2) \quad d\bar{S}(t) = [A(r(t)) - \mu(r(t))\bar{S}(t)]dt + \sigma_1(r(t))\bar{S}(t)dW_1(t)$$

Lemma 3.1 will show that the process  $(\bar{S}(t), r(t))$  has a unique invariant probability measure  $\nu_0$  supported on  $(0, \infty) \times \mathcal{M}$ . Moreover, the measure  $\nu_0$  can be regarded as the unique invariant measure on the boundary of  $(S(t), I(t), Q(t), R(t), r(t))$  by embedding  $[0, \infty) \times \mathcal{M}$  to  $[0, \infty) \times \{0\}^3 \times \mathcal{M}$ . Let

$$(2.3) \quad \lambda = \sum_{k \in \mathcal{M}} \int_{(0, \infty)} \left[ F(s, 0, k) - c_2(k) - \frac{\sigma_2^2(k)}{2} \right] \nu_0(ds, \{k\}),$$

where  $c_2(k) = \mu(k) + \gamma_1(k) + \gamma_2(k)$ , be the disease invasion rate.

**Remark 2.2.** In particular, if there is no switching, one has

$$(2.4) \quad \lambda = \int_{(0, \infty)} \left[ F(s, 0) - c_2 - \frac{\sigma_2^2}{2} \right] \nu_0(ds).$$

The SDE

$$(2.5) \quad d\bar{S}(t) = [A - \mu\bar{S}(t)]dt + \sigma_1\bar{S}(t)dW_1(t)$$

has a unique stationary distribution which can be found explicitly (see Lemma 2.2 in [Hening et al. \[2023\]](#)) so that if  $F(s, i) = \beta s$  then

$$\lambda = \int_{(0, \infty)} \left[ \beta s - c_2 - \frac{\sigma_2^2}{2} \right] \nu_0(dx) = \beta \frac{A}{\mu} - c_2 - \frac{\sigma_2^2}{2}.$$

**Remark 2.3.** If  $F(s, i, k) = \beta(k)s$  and we set  $\mathbf{Q} = (q_{kl})_{m_0 \times m_0}$ ,  $\mathbf{U} := \text{diag}(\mu(1), \dots, \mu(m_0))$ ,  $\beta := (\beta(1), \dots, \beta(m_0))^T$  then one can show Wang and Liu [2023] that

$$\lambda = \sum_{k=1}^{m_0} \pi_k \left( \kappa_k A(k) - c_2(k) - \frac{\sigma_2^2(k)}{2} \right)$$

where  $K = (\kappa_1, \dots, \kappa_{m_0})^T$  is the unique solution to

$$(\mathbf{U} - \mathbf{Q})K = \beta.$$

Specifically, if  $m_0 = 2$  we get

$$\lambda = \pi_1 \left( \kappa_1 A(1) - c_2(1) - \frac{\sigma_2^2(1)}{2} \right) + \pi_2 \left( \kappa_2 A(2) - c_2(2) - \frac{\sigma_2^2(2)}{2} \right)$$

where

$$\kappa_1 = \frac{(\mu(2) - q_{22})\beta(1) + q_{12}\beta(2)}{(\mu(1) - q_{11})(\mu(2) - q_{22}) - q_{12}q_{21}},$$

$$\kappa_2 = \frac{(\mu(1) - q_{11})\beta(2) + q_{21}\beta(1)}{(\mu(1) - q_{11})(\mu(2) - q_{22}) - q_{12}q_{21}},$$

and

$$\pi_1 = \frac{q_{21}}{q_{12} + q_{21}}, \pi_2 = \frac{q_{12}}{q_{12} + q_{21}}.$$

**Remark 2.4.** We note that  $\lambda$  from (2.3) would be unchanged if we would ignore the class  $Q$  and would just look at the SIR model

$$(2.6) \quad \begin{cases} dS(t) = \left[ A(r(t)) - F(S(t), I(t), r(t))I(t) - \mu(r(t))S(t) + \gamma_5(r(t))R(t) \right] dt \\ \quad + \sigma_1(r(t))S(t)dW_1(t), \\ dI(t) = \left[ F(S(t), I(t), r(t)) - \mu(r(t)) - \gamma_1(r(t)) - \gamma_2(r(t)) \right] I(t) dt \\ \quad + \sigma_2(r(t))I(t)dW_2(t), \\ dR(t) = \left[ \gamma_2(r(t))I(t) - (\mu(r(t)) + \gamma_5(r(t)))R(t) \right] dt \\ \quad + \sigma_4(r(t))R(t)dW_4(t). \end{cases}$$

This happens because  $\lambda$  measures the ‘invasion rate’ of the infected class  $I$  into the susceptible class  $S$  at stationarity. Note that (2.2) looks at the susceptible class  $S$  in the absence of any of the other classes so that the stationary measure  $\nu_0$  only depends on  $A$ ,  $\mu$ , and  $\sigma_1$ . Similarly, in (2.3) we only get dependence on the coefficients from the SDE governing the infected class  $I$ . Since we will show that the sign of  $\lambda$  determines the persistence or extinction of the diseases, the only influence  $Q$  can have will be on the distribution at stationarity of the infected individuals, if the disease persists.

The following is our general persistence result.

**Theorem 2.1.** If  $\lambda > 0$  then there exists a unique invariant measure  $\nu^*$  of  $(S(t), I(t), Q(t), R(t), r(t))$  on  $\mathbb{R}_+^{4,\circ} \times \mathcal{M}$ . Moreover, the rate of convergence is exponential: there exists  $\rho > 0$  such that for any  $j := (s, i, q, r, k) \in \mathbb{R}_+^{4,\circ} \times \mathcal{M}$  we have

$$\lim_{t \rightarrow \infty} e^{\rho t} \|P(t, j, \cdot) - \nu^*(\cdot)\| = 0$$

where  $\|\cdot\|$  is the total variation norm.

This result tells us that if  $\lambda > 0$  then the process converges exponentially fast to the unique invariant probability measure  $\nu^*$ . Define the randomized occupation measures

$$\tilde{\Pi}^t(\cdot) := \frac{1}{t} \int_0^t \mathbf{1}_{\{(S(u), I(u), Q(u), R(u), r(u)) \in \cdot\}} du, \quad t > 0.$$

For a Borel set  $A$  the quantity  $\tilde{\Pi}^t(A)$  tells us what fraction of the time  $t$  the process has spent in the set  $A$ . We can next state the extinction result.

**Theorem 2.2.** *If  $\lambda < 0$  then for any  $j := (s, i, q, r, k) \in \mathbb{R}_+^{4,0} \times \mathcal{M}$  we have*

$$(2.7) \quad \mathbb{P}_j \left\{ \tilde{\Pi}^t \rightarrow \nu_0 \text{ weakly}, \lim_{t \rightarrow \infty} \frac{\ln I(t)}{t} = \lambda < 0, \lim_{t \rightarrow \infty} \frac{\ln Q(t)}{t} < 0, \lim_{t \rightarrow \infty} \frac{\ln R(t)}{t} < 0 \right\} = 1.$$

### 3. PROOFS FOR EXTINCTION RESULTS

Let  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$  be a complete probability space with a filtration  $\{\mathcal{F}_t\}_{t \geq 0}$ , which satisfies the usual conditions. Define  $c_1 = \mu$ ,  $c_2 = \mu + \gamma_1 + \gamma_2$ ,  $c_3 = \mu + \gamma_3 + \gamma_4$ ,  $c_4 = \mu + \gamma_5$ , which all depend on  $r(t)$ , and set  $j = (s, i, q, r, k)$ ,  $j_0 = (s, 0, q, r, k)$ ,  $J(t) := (S(t), I(t), Q(t), R(t), r(t))$ . Throughout this paper, we use the lower case letters  $s, i, q, r$  to represent the initial values of  $S(t), I(t), Q(t), R(t)$  respectively. Denote  $\mathbb{R}_+^4 := \{(s, i, q, r) \in \mathbb{R}^4 : s \geq 0, i \geq 0, q \geq 0, r \geq 0\}$ ,  $\mathbb{R}_+^{4,0} := \{(s, i, q, r) \in \mathbb{R}^4 : s > 0, i > 0, q > 0, r > 0\}$ , and  $\mathbb{R}_+^{4,*} := \{(s, i, q, r) \in \mathbb{R}^4 : s \geq 0, i > 0, q \geq 0, r \geq 0\}$ . Moreover, to simplify the notation, we let  $\tilde{a} := \max_{k \in \mathcal{M}} \{a(k)\}$ , and  $\hat{a} := \min_{k \in \mathcal{M}} \{a(k)\}$ .

**Theorem 3.1.** *The system (1.4) has the following properties:*

- (i) *For any initial value  $(s, i, q, r, k) \in \mathbb{R}_+^4 \times \mathcal{M}$ , there exists a unique global solution  $(S(t), I(t), Q(t), R(t))$  to (1.4) such that  $\mathbb{P}_j \{(S(t), I(t), Q(t), R(t)) \in \mathbb{R}_+^4, \forall t \geq 0\} = 1$ , where  $j$  denotes the initial condition. Moreover,  $\mathbb{P}_{j_0} \{I(t) = 0, \forall t \geq 0\} = 1$  and*

$$(3.1) \quad \mathbb{P}_j \left\{ (S(t), I(t), Q(t), R(t)) \in \mathbb{R}_+^{4,0}, \forall t > 0 \right\} = 1, \text{ for any } j \in \mathbb{R}_+^{4,*} \times \mathcal{M}.$$

*In addition, the solution process  $(S(t), I(t), Q(t), R(t), r(t))$  is a Markov-Feller process with transition probability denoted by  $P(t, s, i, q, k, \cdot)$ .*

- (ii) *For any  $p > 0$  sufficiently small, there exist  $C_p > 0$  and  $D_p > 0$  such that*

$$(3.2) \quad \mathbb{E}_j (S(t) + I(t) + Q(t) + R(t))^{1+p} \leq \frac{(1 + s + i + q + r)^{1+p}}{e^{D_p t}} + \frac{C_p}{D_p}, \quad \forall t \geq 0.$$

*Moreover, for any  $H, \varepsilon, T > 0$ , there exists an  $M_{H, \varepsilon, T} > 0$  such that*

$$(3.3) \quad \mathbb{P}_j \left\{ \sup_{t \in [0, T]} \{S(t) + I(t) + Q(t) + R(t)\} \leq M_{H, \varepsilon, T} \right\} \geq 1 - \varepsilon, \quad \forall j \in [0, H]^4 \times \mathcal{M}.$$

**3.1. Extinction.** In this subsection we will look at the case  $\lambda < 0$  and will prove Theorem 2.2.

We will make use of the following result from [Phu et al. \[2020\]](#).

**Lemma 3.1.** *(Lemma 3.1 in [Phu et al. \[2020\]](#)) For any  $\epsilon_0 \in [0, 1]$ , there exists a unique invariant probability measure  $\nu_{\epsilon_0}$  to  $(\bar{S}^{\epsilon_0}(t), r(t))$ , where  $(\bar{S}^{\epsilon_0}(t), r(t))$  is the solution of*

$$(3.4) \quad d\bar{S}^{\epsilon_0}(t) = \left[ A(r(t)) - \mu(r(t))\bar{S}^{\epsilon_0}(t) + \gamma_5(r(t))\epsilon_0 \right] dt + \sigma_1(r(t))\bar{S}^{\epsilon_0}(t)dW_1(t),$$

*In addition, any measure function  $f(s, k)$  satisfying  $f(s, k) \leq C_f(s+1)$  is  $\nu_{\epsilon_0}$ -integrable. Moreover, if  $\lambda < 0$ , there exists  $\epsilon_0 > 0$  such that*

$$(3.5) \quad \tilde{\lambda} \leq \frac{3}{4}\lambda,$$

where

$$\tilde{\lambda} := \sum_{k \in \mathcal{M}} \int_{(0, \infty)} \left[ F(s, 0, k) - c_2(k) - \frac{\sigma_2^2(k)}{2} \right] \nu_{\epsilon_0}(ds, k).$$

and that

$$(3.6) \quad |F(s, i, k) - F(s, 0, k)| \leq \frac{|\lambda|}{4}, \quad \text{for all } 0 \leq i \leq \epsilon_0, s \geq 0, k \in \mathcal{M}.$$

Finally, it is noted that when  $\epsilon_0 = 0$ , (3.4) becomes (2.2). As a consequence, the above conclusions hold for the solution of (2.2).

The next result tells us that if  $\lambda < 0$  and we start with small densities of quarantined and recovered individuals then the number of infected, quarantined and recovered individuals tends asymptotically to zero with high probability.

**Lemma 3.2.** *For any  $\varepsilon > 0$  and  $H > 0$ , there exists a  $\delta_1 > 0$  such that for all  $(s, i, q, r, k) \in [0, H] \times (0, \delta_1]^3 \times \mathcal{M}$ , we have*

$$\mathbb{P}_j \left\{ \lim_{t \rightarrow \infty} \frac{\ln I(t)}{t} < 0, \lim_{t \rightarrow \infty} \frac{\ln Q(t)}{t} < 0, \lim_{t \rightarrow \infty} \frac{\ln R(t)}{t} < 0 \right\} \geq 1 - \varepsilon.$$

*Proof.* Denote  $g(s, i, k) := F(s, i, k) - c_2(k) - \frac{\sigma_2^2(k)}{2}$ . By the ergodicity of  $(\bar{S}^{(\epsilon_0)}(t), r(t))$ ,

$$(3.7) \quad \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t g(\bar{S}^{(\epsilon_0)}(u), 0, r(u)) du = \sum_{k \in \mathcal{M}} \int_{(0, \infty)} g(s, 0, k) \nu_{\epsilon_0}(ds, k) = \tilde{\lambda} \text{ a.s.}$$

Therefore, for any  $\varepsilon > 0$ , there exists a  $T_1 = T_1(H, \varepsilon) > 0$  such that  $\mathbb{P}_{H,i}(\Omega_1) \geq 1 - \frac{\varepsilon}{4}$ , where

$$(3.8) \quad \Omega_1 = \left\{ \omega \in \Omega : \frac{1}{t} \int_0^t g(\bar{S}_s^{(\epsilon_0)}(u), 0, r(u)) du \leq \tilde{\lambda} + \frac{|\lambda|}{4} \quad \text{for all } t \geq T_1 \right\}$$

where the subscript in  $\mathbb{P}_{H,i}$  indicates the initial value of  $(\bar{S}^{(\epsilon_0)}(u), r(u))$ . Because of the uniqueness of solutions, we have  $\bar{S}_{s,k}^{(\epsilon_0)}(u) \leq \bar{S}_{H,i}^{(\epsilon_0)}(u), \forall u \geq 0$  almost surely where the subscript of  $\bar{S}_{s,k}^{(\epsilon_0)}$  indicates the initial value  $(\bar{S}^{(\epsilon_0)}(0), r(0))$ , which implies  $\mathbb{P}_{s,k}(\Omega_1) \geq 1 - \frac{\varepsilon}{4}$  for  $(s, k) \in [0, H] \times \mathcal{M}$ .

Let  $\epsilon_1$  be a positive constant such that  $0 < \epsilon_1 < \min \left\{ \frac{-\lambda}{16}, \hat{c}_3 + \frac{\hat{\sigma}_3^2}{2}, \hat{c}_4 + \frac{\hat{\sigma}_4^2}{2} \right\}$ .

From the strong law of large numbers for Brownian motions we get

$$(3.9) \quad \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t \sigma_k(r(u)) dW_k(u) = 0, \text{ a.s. for } k = 1, 2, 3, 4.$$

This means there is a  $T_2(\varepsilon) > 0$  such that  $\mathbb{P}(\Omega_2) \geq 1 - \frac{\varepsilon}{4}$ , where

$$(3.10) \quad \Omega_2 = \left\{ \omega \in \Omega : \frac{1}{t} \left| \int_0^t \sigma_k(r(u)) dW_k(u) \right| \leq \min \left\{ \frac{|\lambda|}{4}, \epsilon_1 \right\} \quad \text{for all } t \geq T_2, k = 1, 2, 3, 4 \right\}.$$

Let  $T = \max\{T_1, T_2\}$ . As a consequence of (3.3) and the continuity of  $F(s, i, k)$ , there exists  $M := M(\epsilon, T, H) > 0$  such that  $\mathbb{P}(\Omega_3) \geq 1 - \frac{\varepsilon}{4}$ , for all  $(s, i, q, r, k) \in (0, H]^4 \times \mathcal{M}$ , where

$$(3.11) \quad \Omega_3 = \left\{ \omega \in \Omega : \int_0^T F(\bar{S}^{(\epsilon_0)}(u), 0, r(u)) du \leq \frac{M}{2} \right\}.$$

Moreover, we can choose  $M$  large enough so that  $\mathbb{P}_{s,i,k}(\Omega_4) \geq 1 - \frac{\varepsilon}{4}$ , where

$$(3.12) \quad \Omega_4 = \left\{ \omega \in \Omega : \left| \int_0^t \sigma_k(r(u)) dW_k(u) \right| \leq \frac{M}{2}, \text{ for all } t \in [0, T] \text{ and } k = 1, 2, 3, 4 \right\}.$$

Let  $\delta_1 > 0$  such that

$$\begin{aligned} & \delta_1 \left( 1 + e^M + e^{M/2} + \check{\gamma}_2 e^{2M} T + e^{M/2} \check{\gamma}_3 (e^{M/2} + \check{\gamma}_2 e^{2M} T) e^{M/2} T \right. \\ & \quad \left. + \frac{\check{\gamma}_2}{\widehat{c}_3 + \frac{\widehat{\sigma}_3^2}{2} + \epsilon_1 + \frac{\lambda}{4}} + \frac{\check{\gamma}_3 \delta_1 \check{\gamma}_2}{\left( \frac{\lambda}{4} + \widehat{c}_3 + \frac{\widehat{\sigma}_4^2}{2} + \epsilon_1 \right) \left( \frac{\lambda}{4} + \widehat{c}_4 + \frac{\widehat{\sigma}_2^2}{2} + 3\epsilon_1 \right)} \right) \\ & < \epsilon_0. \end{aligned}$$

From the second equation of (1.4) we get

$$(3.13) \quad I(t) = I(0) \exp \left\{ \int_0^t \left( F(S(u), I(u), r(u)) - c_2(r(u)) - \frac{\sigma_2^2(r(u))}{2} \right) du + \int_0^t \sigma_2(r(u)) dW_2(u) \right\}.$$

Using (3.11) and (3.12) we obtain

$$(3.14) \quad \begin{aligned} I(t) & \leq I(0) \exp \left\{ \int_0^t F(S(u), I(u), r(u)) du + \int_0^t \sigma_2(r(u)) dW_2(u) \right\} \\ & \leq \delta_1 e^M < \epsilon_0 \text{ for any } t \in [0, T], I(0) \leq \delta_1 \text{ and } \omega \in \Omega_3 \cap \Omega_4. \end{aligned}$$

Define

$$\Phi_t := \exp \left\{ - \int_0^t \left( c_3(r(s)) + \frac{\sigma_3^2(r(s))}{2} \right) ds + \int_0^t \sigma_3(r(s)) dW_3(s) \right\}.$$

The third equation of (1.4) implies that

$$Q(t) = \Phi_t \left( Q(0) + \int_0^t \gamma_2(r(s)) I(s) \Phi_s^{-1} ds \right)$$

For  $t \in [0, T]$  and  $\omega \in \Omega_3 \cap \Omega_4$ ,

$$e^{-\int_0^t \left( c_3(r(s)) + \frac{\sigma_3^2(r(s))}{2} \right) ds - \frac{M}{2}} \leq \Phi_t \leq e^{-\int_0^t \left( c_3(r(s)) + \frac{\sigma_3^2(r(s))}{2} \right) ds + \frac{M}{2}}.$$

Hence, when  $(i, q) \in (0, \delta_1]^2$ ,

$$(3.15) \quad \begin{aligned} Q(t) & \leq e^{M/2} (\delta_1 + \int_0^t \check{\gamma}_2 \delta_1 e^M e^{M/2} ds) \\ & \leq e^{M/2} \delta_1 + e^{M/2} \check{\gamma}_2 \delta_1 e^M e^{M/2} T \\ & < \epsilon_0. \end{aligned}$$

Now, from the forth equation of (1.4) we obtain

$$R(t) = \Psi_t \left( R(0) + \int_0^t \gamma_3(r(s)) Q(s) \Psi_s^{-1} ds \right),$$

where

$$\Psi_t := \exp \left\{ - \int_0^t \left( c_4(r(s)) + \frac{\sigma_4^2(r(s))}{2} \right) ds + \int_0^t \sigma_4(r(s)) dW_4(s) \right\}.$$

Similarly, for  $t \in [0, T]$  and  $\omega \in \Omega_3 \cap \Omega_4$  we obtain the following estimation for  $\Psi_t$

$$e^{-\int_0^t \left( c_4(r(s)) + \frac{\sigma_4^2(r(s))}{2} \right) ds - \frac{M}{2}} \leq \Psi_t \leq e^{-\int_0^t \left( c_4(r(s)) + \frac{\sigma_4^2(r(s))}{2} \right) ds + \frac{M}{2}}.$$

This implies that for  $(i, q, r) \in (0, \delta_1]^3$ ,

$$(3.16) \quad \begin{aligned} R(t) & \leq e^{M/2} \delta_1 + e^{M/2} \check{\gamma}_3 (e^{M/2} \delta_1 + e^{M/2} \check{\gamma}_2 \delta_1 e^M e^{M/2} T) e^{M/2} T \\ & < \epsilon_0. \end{aligned}$$

Now, we define the stopping time

$$(3.17) \quad \tilde{\tau} := \inf \{t \geq 0 : I(t) \vee Q(t) \vee R(t) \geq \epsilon_0\}.$$

As a consequence, for  $\omega \in \Omega_3 \cap \Omega_4$  we have  $\tilde{\tau} > T$ .

Note that, one has  $S(t) \leq \bar{\mathcal{S}}^{(\epsilon_0)}(t)$  for  $t \in [T, \tilde{\tau}]$  given that they have the same initial value. Then for almost every  $\omega \in \bigcap_{j=1}^4 \Omega_j$ , and  $i \in (0, \delta_1]$ ,

$$(3.18) \quad \begin{aligned} I(t) &= I(0) \exp \left\{ \int_0^t \left( F(S(u), 0, r(u)) - c_2(r(u)) - \frac{\sigma_2^2(r(u))}{2} \right) du + \int_0^t \sigma_2(r(u)) dW_2(u) \right\} \\ &\leq I(0) \exp \left\{ \int_0^t g(\bar{\mathcal{S}}^{(\epsilon_0)}(u), 0, r(u)) du + \int_0^t \sigma_2(r(u)) dW_2(u) \right\} \\ &\leq I(0) \exp \left\{ \tilde{\lambda}t + \frac{|\lambda|t}{4} + \frac{|\lambda|t}{4} \right\} \\ &\leq \delta_1 \exp \left\{ t \frac{\lambda}{4} \right\} < \epsilon_0, \text{ for all } t \in [T, \tilde{\tau}). \end{aligned}$$

Furthermore,

$$\begin{aligned} Q(t) &= \Phi_t \left( Q(0) + \int_0^t \gamma_2(r(s)) I(s) \Phi_s^{-1} ds \right) \\ &= \Phi_t \left( Q(0) + \int_0^T \gamma_2(r(s)) I(s) \Phi_s^{-1} ds \right) + \Phi_t \int_T^t \gamma_2(r(s)) I(s) \Phi_s^{-1} ds. \end{aligned}$$

Note that for all  $\omega \in \bigcap_{j=1}^4 \Omega_j$ ,

$$(3.19) \quad \begin{aligned} \Phi_t \int_T^t \gamma_2(r(s)) I(s) \Phi_s^{-1} ds &= \int_T^t \gamma_2(r(s)) I(s) e^{-\int_s^t \left( c_3(r(u)) + \frac{\sigma_3^2(r(u))}{2} \right) du + \int_s^t \sigma_3(r(u)) dW_3(u)} ds \\ &\leq \int_T^t \tilde{\gamma}_2 \delta_1 e^{\frac{s\lambda}{4}} e^{-\int_s^t \left( \hat{c}_3 + \frac{\hat{\sigma}_3^2}{2} \right) du + \epsilon_1(t+s)} ds \\ &= \int_T^t \tilde{\gamma}_2 \delta_1 e^{\frac{s\lambda}{4}} e^{-\left( \hat{c}_3 + \frac{\hat{\sigma}_3^2}{2} \right)(t-s) + \epsilon_1(t+s)} ds \\ &= \tilde{\gamma}_2 \delta_1 e^{(-\hat{c}_3 - \frac{\hat{\sigma}_3^2}{2} + \epsilon_1)t} \int_T^t e^{(\frac{\lambda}{4} + \hat{c}_3 + \frac{\hat{\sigma}_3^2}{2} + \epsilon_1)s} ds \\ &\leq \tilde{\gamma}_2 \delta_1 e^{(-\hat{c}_3 - \frac{\hat{\sigma}_3^2}{2} + \epsilon_1)t} \frac{e^{(\frac{\lambda}{4} + \hat{c}_3 + \frac{\hat{\sigma}_3^2}{2} + \epsilon_1)t}}{\frac{\lambda}{4} + \hat{c}_3 + \frac{\hat{\sigma}_3^2}{2} + \epsilon_1} \\ &= \frac{\tilde{\gamma}_2 \delta_1}{\frac{\lambda}{4} + \hat{c}_3 + \frac{\hat{\sigma}_3^2}{2} + \epsilon_1} e^{(\frac{\lambda}{4} + 2\epsilon_1)t} \\ &\leq \frac{\tilde{\gamma}_2 \delta_1}{\frac{\lambda}{4} + \hat{c}_3 + \frac{\hat{\sigma}_3^2}{2} + \epsilon_1}. \end{aligned}$$

Let  $n_1 > T$  be an positive integer. For  $t \in [0, \tilde{\tau} \wedge n_1]$ , by (3.15) and (3.19) we obtain

$$(3.20) \quad Q(t) \leq \delta_1 \left( e^{M/2} + e^{2M} \tilde{\gamma}_2 T + \frac{\tilde{\gamma}_2}{\frac{\lambda}{4} + \hat{c}_3 + \frac{\hat{\sigma}_3^2}{2} + \epsilon_1} \right) < \epsilon_0.$$

Similarly,

$$R(t) = \Psi_t \left( R(0) + \int_0^T \gamma_3(r(s))Q(s)\Psi_s^{-1}ds \right) + \Psi_t \int_T^t \gamma_3(r(s))Q(s)\Psi_s^{-1}ds.$$

For the last term,

$$\begin{aligned} \Psi_t \int_T^t \gamma_3(r(s))Q(s)\Psi_s^{-1}ds &= \int_T^t \gamma_3(r(s))Q(s)e^{-\int_s^t \left( c_4(r(u)) + \frac{\sigma_4^2(r(u))}{2} \right) du + \int_s^t \sigma_4(r(u))dW_4(u)} ds \\ &\leq \int_T^t \check{\gamma}_3 \delta_1 \frac{\check{\gamma}_2}{\frac{\lambda}{4} + \hat{c}_3 + \frac{\hat{\sigma}_3^2}{2} + \epsilon_1} e^{(\frac{\lambda}{4} + 2\epsilon_1)s} e^{-\int_s^t \left( \hat{c}_4 + \frac{\hat{\sigma}_4^2}{2} \right) du + \epsilon_1(t+s)} ds \\ &= \int_T^t \check{\gamma}_3 \delta_1 \frac{\check{\gamma}_2}{\frac{\lambda}{4} + \hat{c}_3 + \frac{\hat{\sigma}_3^2}{2} + \epsilon_1} e^{(\frac{\lambda}{4} + 2\epsilon_1)s} e^{-\left( \hat{c}_4 + \frac{\hat{\sigma}_4^2}{2} \right)(t-s) + \epsilon_1(t+s)} ds \\ (3.21) \quad &= \check{\gamma}_3 \delta_1 \frac{\check{\gamma}_2}{\frac{\lambda}{4} + \hat{c}_3 + \frac{\hat{\sigma}_3^2}{2} + \epsilon_1} e^{(-\hat{c}_4 - \frac{\hat{\sigma}_4^2}{2} + \epsilon_1)t} \int_T^t e^{(\frac{\lambda}{4} + \hat{c}_4 + \frac{\hat{\sigma}_4^2}{2} + 3\epsilon_1)s} ds \\ &\leq \check{\gamma}_3 \delta_1 \frac{\check{\gamma}_2}{\frac{\lambda}{4} + \hat{c}_3 + \frac{\hat{\sigma}_3^2}{2} + \epsilon_1} e^{(-\hat{c}_4 - \frac{\hat{\sigma}_4^2}{2} + \epsilon_1)t} \frac{e^{(\frac{\lambda}{4} + \hat{c}_4 + \frac{\hat{\sigma}_4^2}{2} + 3\epsilon_1)t}}{\frac{\lambda}{4} + \hat{c}_4 + \frac{\hat{\sigma}_4^2}{2} + 3\epsilon_1} \\ &= \frac{\check{\gamma}_3 \delta_1 \check{\gamma}_2}{\left( \frac{\lambda}{4} + \hat{c}_3 + \frac{\hat{\sigma}_3^2}{2} + \epsilon_1 \right) \left( \frac{\lambda}{4} + \hat{c}_4 + \frac{\hat{\sigma}_4^2}{2} + 3\epsilon_1 \right)} e^{(\frac{\lambda}{4} + 4\epsilon_1)t} \\ &\leq \frac{\check{\gamma}_3 \delta_1 \check{\gamma}_2}{\left( \frac{\lambda}{4} + \hat{c}_3 + \frac{\hat{\sigma}_3^2}{2} + \epsilon_1 \right) \left( \frac{\lambda}{4} + \hat{c}_4 + \frac{\hat{\sigma}_4^2}{2} + 3\epsilon_1 \right)} \end{aligned}$$

Let  $n_2 > T$  be an positive integer, using (3.16) and (3.21) yields,

(3.22)

$$R(t) \leq \delta_1 \left( e^{M/2} + e^{M/2} \check{\gamma}_3 (e^{M/2} + e^{2M} \check{\gamma}_2 T) e^{M/2} T + \frac{\check{\gamma}_3 \delta_1 \check{\gamma}_2}{\left( \frac{\lambda}{4} + \hat{c}_3 + \frac{\hat{\sigma}_3^2}{2} + \epsilon_1 \right) \left( \frac{\lambda}{4} + \hat{c}_4 + \frac{\hat{\sigma}_4^2}{2} + 3\epsilon_1 \right)} \right) < \epsilon_0,$$

for  $t \in [0, \tilde{\tau} \wedge n_2]$ .

By the definition of  $\tilde{\tau}$ , and combining (3.18), (3.20), (3.22) we have  $\tilde{\tau} > \max\{n_1, n_2\}$ . Note that  $n_1, n_2$  are arbitrary, hence  $\tilde{\tau} = \infty$  for all  $\omega \in \cap_{j=1}^4 \Omega_j$ , and  $(I(0), Q(0), R(0)) \in (0, \delta_1]^3$ .

Therefore,  $\lim_{t \rightarrow \infty} \frac{\ln I(t)}{t} \leq \frac{\lambda}{4} < 0$ ,  $\lim_{t \rightarrow \infty} \frac{\ln Q(t)}{t} \leq \frac{\lambda}{4} + 2\epsilon_1 < 0$ , and  $\lim_{t \rightarrow \infty} \frac{\ln R(t)}{t} \leq \frac{\lambda}{4} + 4\epsilon_1 < 0$ .

The proof of the lemma is therefore complete by noting that  $P(\cap_{j=1}^4 \Omega_j) \geq 1 - \epsilon$ .

□

*Proof of Theorem 2.2.* With probability 1, any weak-limit (if it exists) of  $\tilde{\Pi}^t(\cdot)$  as  $t \rightarrow \infty$  is an invariant probability measure of the process  $(S(t), I(t), Q(t), R(t), r(t))$  on  $\mathbb{R}_+^4 \times \mathcal{M}$ . From the Lemma 3.2 we see that the collection of measures  $\{\tilde{\Pi}^t(\cdot; \omega), t > 0, \omega \in \cap_{j=1}^4 \Omega_j\}$  is tight in  $\mathbb{R}_+^4 \times \mathcal{M}$  and any weak limit of  $\tilde{\Pi}^t(\cdot)$  as  $t \rightarrow \infty$  must have support on  $[0, \infty) \times \{0\} \times \{0\} \times \{0\} \times \mathcal{M}$ . Furthermore,  $\nu_0$  – regarded as an invariant measure of  $(S(t), I(t), Q(t), R(t), r(t))$  – is the unique invariant probability measure on  $[0, \infty) \times \{0\} \times \{0\} \times \{0\} \times \mathcal{M}$ . Therefore,  $\tilde{\Pi}^t(\cdot)$  converges weakly

to  $\nu_0$  for almost every  $\omega \in \cap_{j=1}^4 \Omega_j$  as  $t \rightarrow \infty$ . This means that

$$(3.23) \quad \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t g(S(u), I(t), r(u)) du = \sum_{k \in \mathcal{M}} \int_{(0, \infty)} g(s, 0, k) \nu_0(ds, k) = \lambda,$$

for almost every  $\omega \in \cap_{j=1}^4 \Omega_j$ .

In addition, note that

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \int_0^t (S(u))^{1+\tilde{p}} du \leq \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t (\bar{S}^{(\epsilon_0)}(u))^{1+\tilde{p}} du = \sum_{k \in \mathcal{M}} \int_{(0, \infty)} s^{1+\tilde{p}} \nu_{\epsilon_0}(ds, k) < \infty,$$

for some  $\tilde{p} > 0$ . Therefore, in light of [Henning and Nguyen, 2018, Lemma 5.6], the limit (3.23) is valid.

Using the definition of  $g(s, i, k)$  and (3.13), we obtain

$$\frac{\ln I(t)}{t} = \frac{\ln I(0)}{t} + \frac{1}{t} \int_0^t g(S(u), I(t), r(u)) du + \frac{1}{t} \int_0^t \sigma_2(r(u)) dW_2(u).$$

By letting  $t \rightarrow \infty$  and using (3.9) and (3.23), we obtain that for almost every  $\omega \in \cap_{j=1}^4 \Omega_j$ ,  $\lim_{t \rightarrow \infty} \frac{\ln I(t)}{t} = \lambda$ .

In view of Lemma 3.2, the process  $(S(t), I(t), Q(t), R(t))$  is transient in  $\mathbb{R}_+^{4, \circ}$ . As a result the process has no invariant probability measure in  $\mathbb{R}_+^{4, \circ}$  and  $\nu_0$  is the unique invariant probability measure of  $(S(t), I(t), Q(t), R(t), r(t))$  in  $\mathbb{R}_+^4 \times \mathcal{M}$ . Let  $H$  be sufficiently large so that  $\nu_0(\{s \in (0, H)\}) > 1 - \frac{\varepsilon}{2}$ . Due to (3.2), the process  $(S(t), I(t), Q(t), R(t), r(t))$  is tight. This implies that the occupation measure

$$\Pi_j^t(\cdot) := \frac{1}{t} \int_0^t \mathbb{P}_j \{(S(u), I(u), Q(u), R(u), r(u)) \in \cdot\} du$$

is tight in  $\mathbb{R}_+^4 \times \mathcal{M}$ . By using the fact that any weak-limit of  $\Pi_j^t$  as  $t \rightarrow \infty$  must be an invariant probability measure of  $(S(t), I(t), Q(t), R(t), r(t))$ , we have that  $\Pi_j^t$  converges weakly to  $\nu_0$  as  $t \rightarrow \infty$ . Thus, for any  $\delta > 0$ ,

$$\liminf_{t \rightarrow \infty} \Pi_{s,i,k}^t((0, H) \times (0, \delta) \times (0, \delta) \times (0, \delta) \times \mathcal{M}) \geq \nu_0((0, H) \times (0, \delta) \times (0, \delta) \times (0, \delta) \times \mathcal{M}) > 1 - \frac{\varepsilon}{2}.$$

Thus, there exists a positive  $\hat{T}$  such that

$$\Pi_j^{\hat{T}}((0, H) \times (0, \delta) \times (0, \delta) \times (0, \delta) \times \mathcal{M}) > 1 - \varepsilon,$$

i.e.,

$$\frac{1}{\hat{T}} \int_0^{\hat{T}} \mathbb{P}_{s,i,k} \{(S(t), I(t), Q(t), R(t)) \in (0, H) \times (0, \delta)^3\} dt > 1 - \varepsilon.$$

Let  $\hat{\tau} = \inf\{t \geq 0 : (S(t), I(t), Q(t), R(t)) \in (0, H) \times (0, \delta)^3\}$ , then

$$\mathbb{P}_j\{\hat{\tau} \leq \hat{T}\} > 1 - \varepsilon.$$

By the strong Markov property and Lemma 3.2, we have that

$$\mathbb{P}_j \left\{ \lim_{t \rightarrow \infty} \frac{\ln I(t)}{t} = \lambda < 0, \lim_{t \rightarrow \infty} \frac{\ln Q(t)}{t} < 0, \lim_{t \rightarrow \infty} \frac{\ln R(t)}{t} < 0 \right\} \geq 1 - \varepsilon.$$

for any  $(s, i, q, r, k) \in \mathbb{R}_+^{4,*} \times \mathcal{M}$ . Since  $\varepsilon > 0$  is arbitrary, (2.7) follows. This finishes the proof.  $\square$

## 4. PROOF OF PERSISTENCE RESULTS

In this section we look at the setting when  $\lambda > 0$ . The following lemma will be relevant in the proof of the main result.

**Lemma 4.1.** (*Hening and Nguyen [2018], Lemma 3.5*) *Let  $Y$  be a random variable,  $\theta_0 > 0$  a constant and suppose*

$$\mathbb{E} \exp(\theta_0 Y) + \mathbb{E} \exp(-\theta_0 Y) \leq K_1.$$

*Then the log-Laplace transform  $\phi(\theta) = \ln \mathbb{E} \exp(\theta Y)$  is twice differentiable on  $\left[0, \frac{\theta_0}{2}\right)$  and*

$$\frac{d\phi}{d\theta}(0) = \mathbb{E}Y, \quad 0 \leq \frac{d^2\phi}{d\theta^2}(0) \leq K_2, \quad \theta \in \left[0, \frac{\theta_0}{2}\right),$$

*for some  $K_2 > 0$  depending only on  $K_1$ .*

We prove Theorem 2.1.

*Proof.* The existence of the invariant measure  $\nu^*$  follows by arguments similar to those from the proof of Theorem 2.2 from Nguyen et al. [2020b]. As such, we only provide the details of the proof that the rate of convergence is exponential.

Let  $g(s, i, k) := F(s, i, k) - \mu(k) - \gamma_1(k) - \gamma_2(k) - \frac{\sigma_2^2(k)}{2}$ .

Denote by  $\pi_r$  the invariant measure of  $r(t)$ . Since  $g(s, 0, k)$  is increasing and

$$\lambda = \sum_{k \in \mathcal{M}} \int_{(0, \infty)} [g(s, 0, k)] \nu_0(ds, \{k\}) > 0,$$

we obtain

$$\sum_{k \in \mathcal{M}} (g(s, 0, k)) \pi_r(k) > 0.$$

Thus there exists  $s_0 > 0$  such that

$$\sum_{k \in \mathcal{M}} \left( \inf_{s \geq s_0} g(s, 0, k) \right) \pi_r(k) > 0.$$

Moreover, there exists  $\epsilon_1 > 0$  such that

$$\lambda' = \sum_{k \in \mathcal{M}} h_k \pi_r(k) > 0,$$

where  $h_k := \inf_{(s, i) \in [s_0, \infty) \times [0, \epsilon_1]} g(s, i, k)$ .

By Assumption (2.1) and an application of the Fredholm alternative, there exists  $\gamma_k > 0$  such that

$$\sum_{j \in \mathcal{M}} q_{kj} \gamma_j = \lambda' - h_k, \quad \text{for any } k \in \mathcal{M}.$$

Now choose  $\rho_1$  small enough such that

$$\rho_1 \gamma_k (-\lambda' + h_k) < \frac{3\lambda'}{4} (1 - \rho_1 \gamma_k), \quad \rho_1 \sigma_2^2(k) < \lambda' \quad \text{and } \rho_1 \gamma_k < 1,$$

for any  $k \in \mathcal{M}$ .

Define  $V_3(j) = (1 - \rho_1 \gamma_k) i^{-\rho_1}$  and note that

$$\begin{aligned}
 \mathcal{L}V_3(j) &= \rho_1 V_3(j) \left( -F(s, i, k) + c_2(k) + \frac{\rho_1 + 1}{2} \sigma_2^2(k) \right) \\
 &\quad + \rho_1(-\lambda' + h_k) V_3(j) + \rho_1^2 \gamma_k (-\lambda' + h_k) i^{-\rho_1} \\
 (4.1) \quad &= \rho_1 V_3(j) \left( -F(s, i, k) + c_2(k) + \frac{\rho_1 + 1}{2} \sigma_2^2(k) - \lambda' + h_k \right) + \rho_1^2 \gamma_k (-\lambda' + h_k) i^{-\rho_1} \\
 &\leq \rho_1 V_3(j) (-\lambda') + \rho_1 \frac{3\lambda'}{4} (1 - \rho_1 \gamma_k) i^{-\rho_1} \\
 &\leq -\frac{\rho_1 \lambda' V_3(j)}{4},
 \end{aligned}$$

for any  $k \in \mathcal{M}$ ,  $s \geq s_0$  and  $i \leq \epsilon_1$ .

Now, since

$$\left[ \mathcal{L} \frac{1}{i} \right] (j) \leq \frac{c_2(k) + \sigma_2^2(k)}{i} \leq \frac{\check{c}_2 + \check{\sigma}_2^2}{i}$$

we obtain

$$\mathbb{E}_j I^{-1}(t) \leq \exp\{(\check{c}_2 + \check{\sigma}_2^2)t\} i^{-1},$$

for any  $t \geq 0, j \in \mathbb{R}_+^{4,0} \times \mathcal{M}$ .

Since  $\lambda = \sum_{k \in \mathcal{M}} \int_{(0, \infty)} \left[ F(s, 0, k) - c_2(k) - \frac{\sigma_2^2(k)}{2} \right] \nu_0(ds, \{k\})$ , we see that there exists  $T_1 > 1$  satisfying

$$(4.2) \quad -\ln(1 - \rho_1 \gamma_k) < \frac{\rho_1 \lambda T_1}{4}, \quad \forall k \in \mathcal{M},$$

such that

$$\frac{1}{T} \mathbb{E}_{j_0} \int_0^T \left( F(s, i, k) - c_2(k) - \frac{\sigma_2^2(k)}{2} \right) dt > \frac{3\lambda}{4},$$

for any  $k \in \mathcal{M}$ ,  $T \geq T_1$ ,  $s \leq s_0$ . Choose  $n_* \in \mathbb{Z}_+$  such that

$$n_* > \frac{4(\check{c}_2 + \check{\sigma}_2^2)}{\lambda'} + 1.$$

From Ito's formula,

$$\ln I(t) = \ln I(0) + G(t),$$

for some  $t \in [T_1, n_* T_1]$ . Using Lemma 3.1 from [Hening and Nguyen \[2018\]](#) and applying the log-Laplace transform for  $G(t)$  one can see that there exists  $\widehat{K} \geq 0$  such that

$$0 \leq \frac{d^2 \widehat{\phi}_t}{d\rho_2^2}(\rho_2) \leq \widehat{K} \quad \text{for all } \rho_2 \in (0, \rho_1), t \in [T_1, n_* T_1],$$

where  $\widehat{\phi}_t(\rho_2) = \ln \mathbb{E}_j \exp(-\rho_2 G(t))$ .

Similarly to the proof of Proposition 4.1 from [Hening and Nguyen \[2018\]](#) one can show that there exist  $\rho_2 \in (0, \rho_1)$  and  $\epsilon_2 \in (0, \epsilon_1)$  such that

$$(4.3) \quad \mathbb{E}_j I^{-\rho_2}(t) \leq e^{\frac{-\rho_2 \lambda t}{2}} i^{-\rho_2},$$

for any  $k \in \mathcal{M}$ ,  $t \in [T_1, n_* T_1]$ ,  $s \leq s_0$  and  $i < \epsilon_2$ .

Therefore, by (4.2) and (4.3) we have the following estimation

$$\begin{aligned}
\mathbb{E}_j V_3^{\frac{\rho_2}{\rho_1}}(J(t), r(t)) &\leq \mathbb{E}_j I^{-\rho_2}(t) \\
&\leq e^{\frac{-\rho_2 \lambda t}{2}} i^{-\rho_2} \\
&= e^{\frac{-\rho_2 \lambda t}{2}} V_3^{\frac{\rho_2}{\rho_1}}(j) (1 - \rho_1 \gamma_k)^{-\frac{\rho_2}{\rho_1}} \\
(4.4) \quad &= e^{\frac{-\rho_2 \lambda t}{2}} V_3^{\frac{\rho_2}{\rho_1}}(j) \exp\left\{\frac{\rho_2}{\rho_1} (-\ln(1 - \rho_1 \gamma_k))\right\} \\
&\leq e^{\frac{-\rho_2 \lambda t}{2}} V_3^{\frac{\rho_2}{\rho_1}}(j) e^{\frac{\rho_2 \lambda T_1}{4}} \\
&\leq e^{\frac{-\rho_2 \lambda t}{2}} V_3^{\frac{\rho_2}{\rho_1}}(j) e^{\frac{\rho_2 \lambda t}{4}} \\
&\leq e^{\frac{-\rho_2 \lambda t}{4}} V_3^{\frac{\rho_2}{\rho_1}}(j),
\end{aligned}$$

for any  $k \in \mathcal{M}$ ,  $t \in [T_1, n_* T_1]$ ,  $s \leq s_0$  and  $i < \epsilon_2$ .

Since  $\rho_2 < \rho_1$  one can see from (4.1) that

$$(4.5) \quad \mathcal{L} V_3^{\frac{\rho_2}{\rho_1}}(j) \leq -\frac{\rho_2 \lambda' V_3^{\frac{\rho_2}{\rho_1}}(j)}{4}, \quad \text{for any } k \in \mathcal{M}, s \geq s_0, i \leq \epsilon_2.$$

Since  $g$  is bounded below,  $\mathcal{L} V_3(j) \leq C V_3(j)$ , which implies

$$\mathbb{E}_j V_3(J(t)) \leq \bar{K}, t \leq n_* T_1, i \geq \epsilon_2$$

for some  $\bar{K}$  which depends on  $\epsilon_2$  and  $T_1, n_*$ . Next, let  $\tau = \inf\{t > 0 : S(t) \leq s_0\}$  and  $\xi = \inf\{t > 0 : I(t) > \epsilon_2\}$ . Dynkin's formula yields

$$\begin{aligned}
\mathbb{E}_j [e^{\rho_2 \lambda' (\tau \wedge \xi \wedge n_* T_1)} V_3^{\rho_2/\rho_1}(J(\tau \wedge \xi \wedge n_* T_1))] &\leq V_3^{\frac{\rho_2}{\rho_1}}(j) + \mathbb{E}_j \int_0^{\tau \wedge \xi \wedge n_* T_1} e^{\rho_2 \lambda' s} \left( \mathcal{L} V_3^{\frac{\rho_2}{\rho_1}}(j(s)) + \rho_2 \lambda' V_3^{\frac{\rho_2}{\rho_1}}(j(s)) ds \right) \\
&\leq V_3^{\frac{\rho_2}{\rho_1}}(j).
\end{aligned}$$

Hence,

$$\begin{aligned}
V_3^{\frac{\rho_2}{\rho_1}}(j) &\geq \mathbb{E}_j [\exp(\rho_2 \lambda' (\tau \wedge \xi \wedge n_* T_1)) V_3^{\rho_2/\rho_1}(J(\tau \wedge \xi \wedge n_* T_1))] \\
&\geq \mathbb{E}_j \left[ \mathbf{1}_{\{\tau \wedge \xi \wedge (n_* - 1) T_1 = \tau\}} V_3^{\frac{\rho_2}{\rho_1}}(J(\tau)) \right] + \mathbb{E}_j \left[ \mathbf{1}_{\{\tau \wedge \xi \wedge (n_* - 1) T_1 = \xi\}} V_3^{\frac{\rho_2}{\rho_1}}(J(\xi)) \right] \\
&\quad + \exp(\rho_2 \lambda' (n_* - 1) T_1) \mathbb{E}_j \left[ \mathbf{1}_{\{(n_* - 1) T_1 < \tau \wedge \xi < n_* T_1\}} V_3^{\frac{\rho_2}{\rho_1}}(J(\tau \wedge \xi)) \right] \\
&\quad + \exp(\rho_2 \lambda' n_* T_1) \mathbb{E}_j \left[ \mathbf{1}_{\{(n_* - 1) T_1 \geq n_* T_1\}} V_3^{\frac{\rho_2}{\rho_1}}(J(n_* T_1)) \right].
\end{aligned}$$

Using the strong Markov property and (4.4) we obtain the following estimates

$$\begin{aligned}
\mathbb{E}_j \left[ \mathbf{1}_{\{\tau \wedge \xi \wedge (n_* - 1) T_1 = \tau\}} V_3^{\frac{\rho_2}{\rho_1}}(J(n_* T_1)) \right] &\leq \mathbb{E}_j \left[ \mathbf{1}_{\{\tau \wedge \xi \wedge (n_* - 1) T_1 = \tau\}} e^{\frac{-\rho_2 \lambda (n_* T_1 - \tau)}{4}} V_3^{\frac{\rho_2}{\rho_1}}(J(\tau)) \right] \\
(4.6) \quad &\leq \mathbb{E}_j \left[ \mathbf{1}_{\{\tau \wedge \xi \wedge (n_* - 1) T_1 = \tau\}} V_3^{\frac{\rho_2}{\rho_1}}(J(\tau)) \right].
\end{aligned}$$

Furthermore,

$$(4.7) \quad \mathbb{E}_j \left[ \mathbf{1}_{\{(n_* - 1) T_1 < \tau \wedge \xi < n_* T_1\}} V_3^{\frac{\rho_2}{\rho_1}}(J(n_* T_1)) \right] \leq \bar{K} \cdot \mathbb{E}_j \left[ \mathbf{1}_{\{(n_* - 1) T_1 < \tau \wedge \xi < n_* T_1\}} V_3^{\frac{\rho_2}{\rho_1}}(J(\tau \wedge \xi)) \right].$$

Therefore,

$$\begin{aligned}
V_3^{\frac{\rho_2}{\rho_1}}(j) &\geq \mathbb{E}_j \left[ \mathbf{1}_{\{\tau \wedge \xi \wedge (n_* - 1)T_1 = \tau\}} V_3^{\frac{\rho_2}{\rho_1}}(J(n_*T_1)) \right] + \mathbb{E}_j \left[ \mathbf{1}_{\{\tau \wedge \xi \wedge (n_* - 1)T_1 = \xi\}} V_3^{\frac{\rho_2}{\rho_1}}(J(n_*T_1)) \right] \\
&\quad + (\overline{K})^{-1} \exp(\rho_2 \lambda' (n_* - 1)T_1) \mathbb{E}_j \left[ \mathbf{1}_{\{(n_* - 1)T_1 < \tau \wedge \xi < n_*T_1\}} V_3^{\frac{\rho_2}{\rho_1}}(J(n_*T_1)) \right] \\
&\quad + \exp(\rho_2 \lambda' n_*T_1) \mathbb{E}_j \left[ \mathbf{1}_{\{(n_* - 1)T_1 \geq n_*T_1\}} V_3^{\frac{\rho_2}{\rho_1}}(J(n_*T_1)) \right] \\
&\geq M \mathbb{E}_j V_3^{\frac{\rho_2}{\rho_1}}(J(n_*T_1)),
\end{aligned}$$

where  $M \geq 2$ . Hence,  $\mathbb{E}_j V_3^{\rho_2/\rho_1}(n_*T_1) \leq \kappa V_3^{\rho_2/\rho_1}(j)$  for some  $\kappa \leq 1/2$ .

Finally, by Theorem 2.1 from [Tuominen and Tweedie \[1994\]](#) we get

$$\lim_{t \rightarrow \infty} e^{\rho t} \|P(t, j, \cdot) - \nu^*(\cdot)\| = 0, \forall j := (s, i, q, r, k) \in \mathbb{R}_+^{4,*} \times \mathcal{M}, \text{ for some constant } \rho > 0$$

and the proof is complete.  $\square$

## 5. APPLICATIONS AND SIMULATIONS

**5.1. Monotonicity of the disease invasion rate.** Suppose there is no switching. By standard arguments (see [Phu et al. \[2020\]](#), [Nguyen et al. \[2020c\]](#), [Dieu et al. \[2016\]](#)) the SDE

$$(5.1) \quad d\overline{S}(t) = [A - \mu \overline{S}(t)]dt + \sigma_1 \overline{S}(t)dW_1(t)$$

has a unique strong solution on  $(0, \infty)$  and a unique stationary distribution  $\nu_0$  on  $(0, \infty)$  which is of inverse Gamma type. Let  $\mathbb{E}_\nu$  denote the expectation if the initial distribution of the process is  $\nu$ . We get

$$(5.2) \quad \mathbb{E}_{\nu_0}(\overline{S}) := \int s \nu_0(ds) = \frac{A}{\mu}.$$

Set  $\mathbb{E}_{\nu_0}(F(\overline{S}, 0)) := \int_{(0, \infty)} F(s, 0) \nu_0(ds)$ . One has

$$\begin{aligned}
(5.3) \quad \lambda &= \int_{(0, \infty)} \left[ F(s, 0) - c_2 - \frac{\sigma_2^2}{2} \right] \nu_0(ds) = \int_{(0, \infty)} F(s, 0) \nu_0(ds) - c_2 - \frac{\sigma_2^2}{2} \\
&= \mathbb{E}_{\nu_0}(F(\overline{S}, 0)) - c_2 - \frac{\sigma_2^2}{2}
\end{aligned}$$

where one remembers that

$$c_2 = \mu + \gamma_1 + \gamma_2.$$

Note that in the deterministic setting (1.2) if  $I = Q = R = 0$  we get that the stable equilibrium is  $\widehat{S} = \frac{A}{\mu}$  and the ‘invasion rate’ the disease becomes

$$(5.4) \quad \lambda_0 = F(\widehat{S}, 0) - c_2.$$

Define the function

$$g(s) := F(s, 0)$$

and assume  $g$  is not linear. By Jensen’s inequality we can see that if  $g$  is convex then

$$\mathbb{E}_{\nu_0} g(S) > g(\mathbb{E}_{\nu_0} S) = g(\widehat{S})$$

while if  $g$  is concave

$$\mathbb{E}_{\nu_0} g(S) < g(\mathbb{E}_{\nu_0} S) = g(\widehat{S})$$

Using (5.3) and (5.4) we see that when  $g$  is concave

$$\lambda_0 = g(\hat{S}) - c_2 > \mathbb{E}_{\nu_0} g(S) - c_2 - \frac{\sigma_2^2}{2} = \lambda.$$

Using (5.3) and (5.4) we see that when  $g$  is convex then it is possible to have

$$\lambda = \mathbb{E}_{\nu_0}(g(S)) - c_2 - \frac{\sigma_2^2}{2} > 0 > g(\hat{S}) = \lambda_0$$

**Biological interpretation:** The noise has detrimental effects for the disease when  $g(s) = F(s, 0)$  is concave, while in the convex case one can get situations when  $\lambda_0 < 0 < \lambda$  i.e. when in the deterministic setting the disease goes extinct while with fluctuations the disease persists. This shows we need to carefully see what type of functional response best fits a particular disease, if we want to know whether environmental fluctuations help or impede the spread of the disease.

5.1.1. *Dependence on  $\sigma_2$  and  $c_2$ .* From equation (5.3) we see that  $\lambda$  decreases with  $\sigma_2$  and  $c_2$ .

5.1.2. *Dependence on  $A$ .* Suppose that  $g(s)$  is increasing in  $s$ . Intuitively, by looking at (1.4) we expect that  $\lambda$  increases with  $A$ . We can prove this as follows. Let the dependence on  $A$  be denoted by a subscript, and let  $0 < A_1 < A_2$ . Comparing the drifts of (5.1) for birth rates  $A_1$  and  $A_2$  we can see using well-known comparison results Geiß and Mantney [1994] that if  $\bar{S}_{A_1}(0) = \bar{S}_{A_2}(0)$  then with probability one

$$\bar{S}_{A_1}(t) \leq \bar{S}_{A_2}(t), t \geq 0$$

which then implies that

$$\lambda_{A_1} \leq \lambda_{A_2}.$$

5.2. **Comparison with the literature.** There are a plethora of results for stochastic epidemic models with various types of noise. However, most of the available results have artificial assumptions and do not provide sharp results. In our paper we can consider a large class of models and are able to prove under natural assumptions that a single parameter,  $\lambda$ , determines the long term fate of the disease. A paper that is closely related to ours is Zhou et al. [2021] where the authors study a SIQRS model when there is only white noise. The authors showcase conditions under which the pandemic persists but do not provide any results for extinction. Moreover, their results are not sharp because they do not establish necessary and sufficient conditions for persistence. Nevertheless, Zhou et al. [2021] provided the inspiration for the current paper. A reviewer also brought Wang and Wang [2024] to our attention, a paper where a similar SIQRS model is studied. We note that our paper has significantly weaker assumptions and we are also able to prove that when  $\lambda > 0$  the process converges exponentially fast to its unique invariant distribution, something which is not shown in Wang and Wang [2024].

5.3. **Linear incidence rate.** In this section we will assume that  $F(S, I) = \beta S$ .

5.3.1. *The deterministic model.* The behavior of (1.2) is well known. Set

$$R_0 = \frac{A\beta}{\mu(\mu + \gamma_1 + \gamma_2)}.$$

Using the results from Zhou et al. [2021] one can show that if  $R_0 > 1$  then there is a globally stable equilibrium  $E$  in  $(0, \infty)^4$  and the disease persists. Instead, if  $R_0 \leq 1$  the disease goes extinct. This expression shows that if  $\gamma_2$  is large it is harder for the disease to persist. If  $\gamma_1$  is large it is also harder for the disease to persist but this would mean the disease kills people very fast and therefore cannot spread.

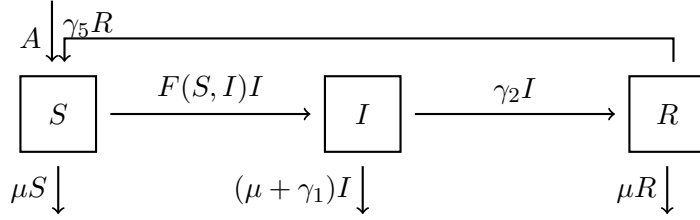


FIGURE 2. Schematic representation of the counterfactual SIRS system used in numerical experiments in Section 5.

5.3.2. *The effects of white noise.* Note that if there is no switching and  $F(S, I) = \beta S$  our criterion boils down to checking whether

$$(5.5) \quad \lambda = \frac{A\beta}{\mu} - (\mu + \gamma_1 + \gamma_2) - \frac{\sigma_2^2}{2}$$

is positive or negative. In this setting it seems that the environmental fluctuations are detrimental to the spread of the disease since when  $\lambda < 0$  the disease dies out. Only when  $\lambda > 0$  can the disease persist long term. We note that in Zhou et al. [2021] the authors prove that the disease persists when

$$R := \frac{A\beta}{\mu + \frac{\sigma_1^2}{2}} - (\mu + \gamma_1 + \gamma_2) - \frac{\sigma_2^2}{2} > 0.$$

However, since  $R < \lambda$ , this result is not sharp. If  $R < 0 < \lambda$  the results from Zhou et al. [2021] are not able to show the persistence of the disease, whereas our results do show that the disease persists.

5.3.3. *The impact of quarantine.* We simulate two simple models with linear incidence and no switching, one with a quarantined group and one without, to determine the impact of introducing a quarantined group. The crucial assumption is that the excess mortality is shifted to the quarantined group in the quarantine model. No other change is introduced. In particular, the quarantined group *does not* have an improved mortality rate.

Consider the system (1.4) in the case of no switching,  $F(S, I) = \beta S$  and  $\lambda > 0$ . It can be show, by an argument similar to the proof of (3.2), that if  $\mu > \sigma_1$  then  $\limsup_{t \rightarrow \infty} \mathbb{E}(S(t) + I(t) + Q(t) + R(t))^2$  is uniformly bounded with respect to the initial value. Because of the ergodicity (Theorem 2.1), the following non-random limits exit almost surely:

$$\begin{aligned} x_1 &:= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T S(t)I(t)dt, \\ x_2 &:= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T S(t)dt, \\ x_3 &:= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I(t)dt, \\ x_4 &:= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T Q(t)dt, \\ x_5 &:= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T R(t)dt. \end{aligned}$$

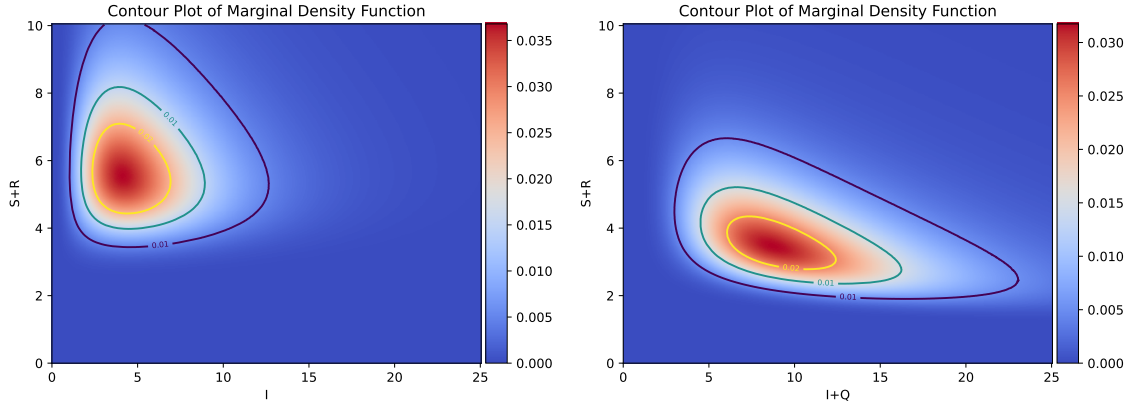


FIGURE 3. Side by side comparison of the estimation of two SSDE models, one without a quarantined group (SIRS) on the left, and one with (SIQRS), on the right. Both models have a linear infection incidence rate, given by  $F(S, I) = \beta S$ . The parameters for the SIRS model are  $A = 20$ ,  $\beta = 1$ ,  $\sigma_1 = 1$ ,  $\mu = \gamma_1 = \gamma_2 = \gamma_5 = 1$ . For the SIQRS model, the other parameters are the same, but now  $\gamma_1 = 0$  and  $\gamma_3 = \gamma_4 = 1$ . Using (5.5) we note that for both models  $\lambda = 16.5 > 0$ . The approximate population expectations at stationarity for the SIRS model are  $\mathbb{E}\bar{S} \approx 3.5$ ,  $\mathbb{E}\bar{I} \approx 6.6$ , and  $\mathbb{E}\bar{R} \approx 3.3$ . The expectation of the total population is  $\approx 13.4$ . In the SIQRS model, the approximate population expectations at stationarity are  $\mathbb{E}\bar{S} \approx 2.5$ ,  $\mathbb{E}\bar{I} \approx 9.5$ ,  $\mathbb{E}\bar{Q} \approx 3.2$ , and  $\mathbb{E}\bar{R} \approx 1.6$ . The expectation of the total population is  $\approx 16.8$ .

From the equation of  $S(t)$  in (1.4) in the integration form and the fact that  $\lim_{T \rightarrow \infty} \frac{S(T)}{T} = 0$  we obtain

$$(5.6) \quad A - \beta x_1 - \mu x_2 + \gamma_5 x_5 = 0$$

From the equation of  $I(t)$  and  $\lim_{T \rightarrow \infty} \frac{I(T)}{T} = 0$  we have

$$(5.7) \quad \beta x_1 - c_2 x_3 = 0.$$

By Itô formula,

$$\lim_{T \rightarrow \infty} \frac{\ln I(T)}{T} = \lim_{T \rightarrow \infty} \left( \frac{1}{T} \int_0^T \beta S(t) dt - \left( c_2 + \frac{\sigma_2^2}{2} \right) - \frac{\sigma_2 W_2(T)}{T} \right) = \beta x_2 - \left( c_2 + \frac{\sigma_2^2}{2} \right).$$

Since the process  $I(t)$  persist, we must have  $\lim_{T \rightarrow \infty} \frac{\ln I(T)}{T} = 0$ , which yields

$$(5.8) \quad \beta x_2 - \left( c_2 + \frac{\sigma_2^2}{2} \right) = 0.$$

Similarly, from the equations of  $Q(t)$ ,  $R(t)$  and  $\lim_{T \rightarrow \infty} \frac{Q(T)}{T} = \lim_{T \rightarrow \infty} \frac{R(T)}{T} = 0$  we get the following equations

$$(5.9) \quad \gamma_2 x_3 - c_3 x_4 = 0.$$

$$(5.10) \quad \gamma_3 x_4 - c_4 x_5 = 0.$$

By solving the system of equations (5.6)-(5.10), we find

$$\begin{aligned}
 (5.11) \quad x_{2,Q} &:= x_2 = \frac{c_2 + \frac{\sigma_2^2}{2}}{\beta}, \\
 x_{3,Q} &:= x_3 = \frac{A - \frac{\mu}{2\beta}(2c_2 + \sigma_2^2)}{c_2 - \frac{\gamma_2\gamma_3\gamma_5}{c_3c_4}}, \\
 x_{5,Q} &:= x_5 = \frac{\gamma_2\gamma_3}{c_3c_4}x_3.
 \end{aligned}$$

Now consider the system (2.6) with linear incidence rate and no switching. From the equations of  $S(t)$ ,  $I(t)$  and  $R(t)$  we obtain the following system of equations:

$$(5.12) \quad \begin{cases} A - \beta x_1 - \mu x_2 + \gamma_5 x_5 = 0, \\ \beta x_1 - c_2 x_3 = 0, \\ \beta x_2 - \left(c_2 + \frac{\sigma_2^2}{2}\right) = 0, \\ \gamma_2 x_3 - c_4 x_5 = 0. \end{cases}$$

Hence, in this case,

$$\begin{aligned}
 (5.13) \quad x_{2,-Q} &:= x_2 = \frac{c_2 + \frac{\sigma_2^2}{2}}{\beta}, \\
 x_{3,-Q} &:= x_3 = \frac{A - \frac{\mu}{2\beta}(2c_2 + \sigma_2^2)}{c_2 - \frac{\gamma_2\gamma_5}{c_4}}, \\
 x_{5,-Q} &:= x_5 = \frac{\gamma_2}{c_4}x_{3,-Q}.
 \end{aligned}$$

Let us first consider a straightforward extension of adding a quarantined group. While keeping all old parameter values the same, we change the model from SIRS to SIQRS. The diagram in Figure 1 summarises the model and the parameters used in the SIQRS model, while Figure 2 summarises the SIRS model. That is, we only have to choose two new values for the new parameters  $\gamma_3$  and  $\gamma_4$ . We observe that the long-term average number of individuals in the susceptible group is the same, with or without quarantine. Moreover, since  $\frac{\gamma_3}{c_3} < 1$ , it follows that  $x_{3,Q} < x_{3,-Q}$  if all other parameters stay the same. Similarly, we have  $x_{5,Q} < x_{5,-Q}$ , which indicates that the number of recovered individuals is also reduced with such an assumption.

Unfortunately, the quarantine extension above likely has unrealistic assumptions. Adding a quarantined group while keeping the old transfer parameters the same means that we have another stage between the infected and the recovered groups with excess mortality. As an example of what can go wrong, our addition of the quarantined group will actually lower the total population for some parameter choices. It is also very easy to find such cases, e.g.,  $\gamma_1 = 1$ ,  $A = 20$ ,  $\mu = 1$ ,  $\beta = 1$ ,  $\sigma_1 = 1$ . Intuitively, if the sick are removed from the infected group to be quarantined, we should be lowering effective mortality in the infected group, therefore reducing  $\gamma_1$ . The correct parametrization is of course a matter of empirical analysis. Another issue is that we would like to know what happens to all non-healthy individuals, not just the infected group, so we would like to compare the sums of infected and quarantined, if any.

For our simulations we will make two strong simplifying assumptions. First, the infected group will have no excess mortality after adding the quarantined group, so  $\gamma_1 = 0$ . That is, we assume all the infected join the quarantined group during disease progression. Second, the quarantined group has no advantage in treatment outcomes compared to the baseline model infected. In other words, the quarantined group is not for treatment, but merely a containment measure for the infected, removing them from the group that affects the susceptible population.

Marginal Density Function

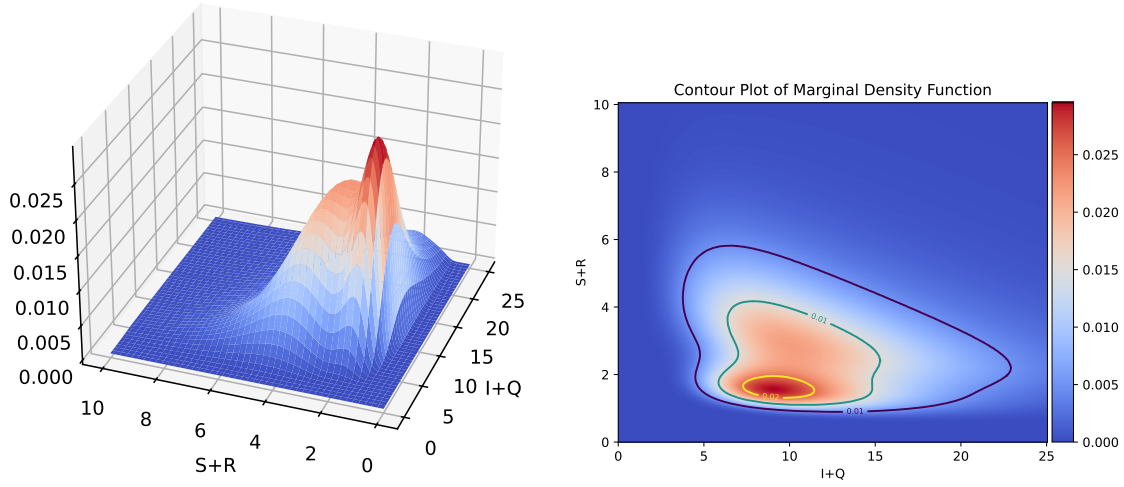


FIGURE 4. The estimation of a SIQRS SSDE model with switching, on the left the 3D plot of marginal density of a cross-section, and on the right the contour equivalent plot. Note the fact that the density is bimodal. The model has a linear infection incidence rate, given by  $F(S, I) = \beta S$ . The parameter values are those summarised in the caption of Figure 3, but with the exception that now  $\beta$  switches between the values 1 and 4 according to the switching rate  $q_{ij} = q = 0.1$ . The approximate population expectations at stationarity for the model are  $\mathbb{E}\bar{S} \approx 1.6$ ,  $\mathbb{E}\bar{I} \approx 10.1$ , and  $\mathbb{E}\bar{Q} \approx 3.4$ , and  $\mathbb{E}\bar{R} \approx 1.7$ . The expectation of the total population is  $\approx 16.6$ .

Figure 3 shows the marginal probability density functions at stationarity. We focus on population composites of interest:  $S + R$ , the healthy group, and  $I$  or  $I + Q$ , the sick group. An approximation for the stationary distribution is obtained using the Monte Carlo method with the Milstein refinement Giles [2006]. We verify that the simulation population expectations at stationarity match the derived formulas for  $x_{\cdot, \pm Q}$  in (5.11), and (5.13). Note that by slightly modifying Theorem 2.1 we can show that the randomized occupation measures (empirical measures) converge exponentially fast to the stationary distribution. This implies that MC methods will work well.

The unusual conclusion that we can draw is that simply isolating individuals in a quarantined group – without providing any improvements in treatment leading to decreased mortality – can increase the population overall, at the expense of having a larger share of infected individuals. This can intuitively be understood by noting that the rate of movement into the  $I$  category is in proportion to the population size already in the  $I$  category. By making individuals non-infective – i.e., by moving them into a category  $Q$  – we are therefore reducing the burden of the disease on total population.

**5.3.4. The effects of switching.** Figure 4 shows a plot where a non-trivial effect of switching is added to the simple linear SIQRS model from Figure 3. Now the parameter  $\beta$  switches between the values  $\{1, 4\}$ , according to the switching rate  $q_{ij} = q = 0.1$  ( $i \neq j$ ). We can observe that the marginal density cross-section shows a bimodal density. Numerical experiments suggest that this is a common occurrence in models with switching, as the density tends to approximate an overlap of the densities of two simpler models without switching – in this case two models with fixed parameters  $\beta = 1$  and  $\beta = 4$ .

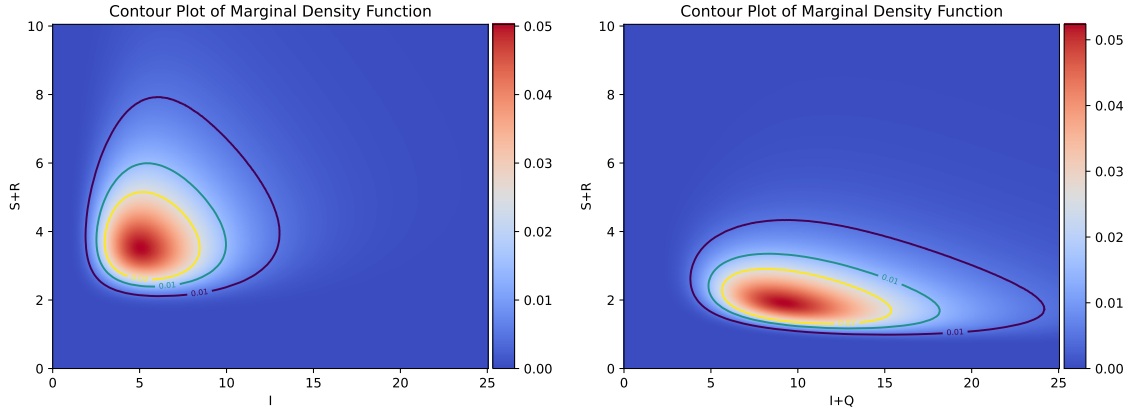


FIGURE 5. Side by side comparison of the estimation of two SSDE models, one without a quarantined group (SIRS) on the left, and one with (SIQRS), on the right. Both models have a Holling type II infection incidence rate, given by  $F(S, I) = \beta S / (1 + m_1 S)$ , where  $m_1 = 0.1$ . The other parameters for the model, summarised in Figure 1, are  $A = 20$ ,  $\beta = 3$ ,  $\sigma = 1$ ,  $\mu = \gamma_1 = \gamma_2 = \gamma_5 = 1$ . For the SIQRS model, the other parameters are the same, but now  $\gamma_1 = 0$  and  $\gamma_3 = \gamma_4 = 1$ . The approximate population expectations at stationarity for the SIRS model are  $\mathbb{E}\bar{S} \approx 1.4$ ,  $\mathbb{E}\bar{I} \approx 7.5$ , and  $\mathbb{E}\bar{R} \approx 3.7$ . The expectation of the total population is  $\approx 12.5$ . In the SIQRS model, the approximate population expectations at stationarity are  $\mathbb{E}\bar{S} \approx 0.9$ ,  $\mathbb{E}\bar{I} \approx 10.4$ ,  $\mathbb{E}\bar{Q} \approx 3.5$ , and  $\mathbb{E}\bar{R} \approx 1.7$ . The expectation of the total population is  $\approx 16.5$ .

**5.4. Nonlinear incidence rates.** We note that most of papers which look at related SIQS/SIQRS models are only able to treat linear incidence rates Wang and Liu [2023], Zhou et al. [2021], and thus do not capture some of the most important models. The seminal paper Liu et al. [1986] has shown that non-linearities appear realistically and that the dynamics changes significantly, even in the absence of noise. In Ruan and Wang [2003] the authors show that stable and unstable limit cycles can appear in SIRS models with nonlinear incidence rates.

**5.4.1. Holling type II functional response.** We explore in our simulations other types of incidence rates used in the wider literature, for example  $F(S, I) = \beta S / (1 + m_1 S)$ , also known as Holling type II. Figure 5 shows that the effect of introducing a quarantined group in such a model with no switching is broadly similar to the linear response model. That is, we see an increase in the total average population, but at the expense of having a larger fraction of sick individuals.

**5.4.2. Beddington-DeAngelis functional response.** Here we consider the Beddington-DeAngelis functional response  $F(S, I) = \beta S / (1 + m_1 S + m_2 I)$ . We note that in this setting one cannot compute  $\lambda$  explicitly. The explicit computations from Dieu [2018] do not hold in our setting because our noise terms in (1.4) are linear whereas in Dieu [2018] the noise terms also have a Beddington-DeAngelis form. We estimate the time evolution of such models in Figure 6. The parameters are such that we have a positive  $\lambda$  for the SIRS model, which changes to a slight negative  $\lambda$  for the SIQRS model. The number of infected and recovered goes to 0 in the SIRS model, but stays positive in the SIQRS model, since  $c_2$  is decreased by the value change of  $\gamma_1$ , from 1 to 0. Observe that, in this case, the introduction of a quarantined group has a negative effect on the long term prevalence of the disease. The long term average of the population is also slightly reduced in the model with quarantine. This shows how quarantine can be beneficial for certain diseases.

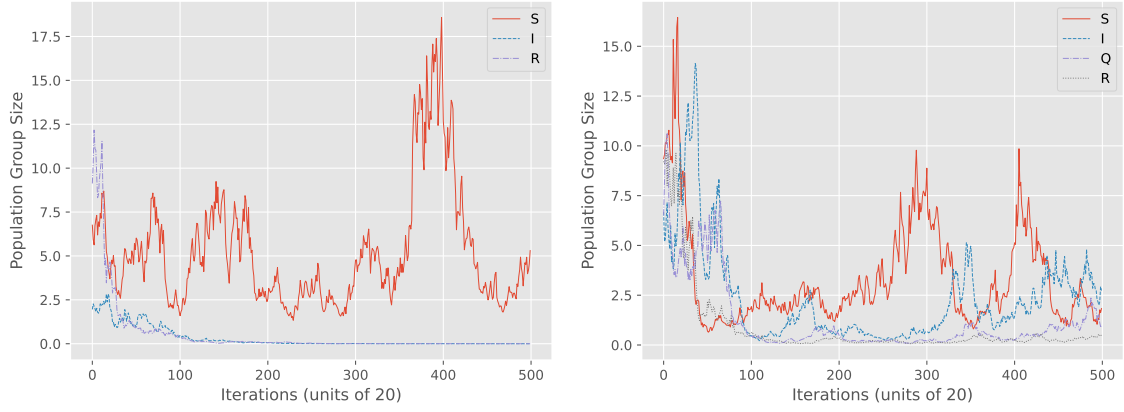


FIGURE 6. Side by side comparison of the estimation of two SSDE models, one without a quarantined group (SIRS) on the left, and one with (SIQRS), on the right. Both models have a Beddington-DeAngelis infection incidence rate, given by  $F(S, I) = \beta S / (1 + m_1 S + m_2 I)$ , where  $m_1 = m_2 = 0.1$ . The common parameters for the two models are  $A = 5$ ,  $\beta = 1$ ,  $\sigma = 1$ ,  $\mu = \gamma_1 = \gamma_2 = \gamma_5 = 1$ . For the SIQRS model, the other parameters are the same, but now  $\gamma_1 = 0$  and  $\gamma_3 = \gamma_4 = 1$ . The horizontal represents periods of simulation, with one plotted point for every 20 periods. One period of simulation corresponds to a discrete increment of time  $\Delta t = 0.001$ . We note that  $\lambda < 0$  for the SIRS model on the left, and  $\lambda > 0$  for the SIQRS model on the right. The long term average of the total population for the SIRS model is  $\approx 5.0$ . For the SIQRS model, it is  $\approx 4.8$ .

**5.5. Alternative formulations.** In the previous experiments, we assumed a particular way in which the introduction of the quarantined group ‘affects’ the baseline SIRS system. Specifically, we have shifted  $\gamma_1$  to 0, to capture the somewhat strong assumption that excess mortality appears only in the quarantined group. In other words, all members of the population are quarantined before they suffer greatly from the disease. There are no other changes in parameters in our contrasting comparisons, although we could have made other parameter changes, say justified on empirical grounds.

Another debatable modeling choice is in the way the quarantined group is a staging group for all infected individuals before they reach the recovered group. This excludes applications in which the quarantine is implemented imperfectly, or in which the disease leads to infected individuals that are able to spread the disease but fall below a certain diagnosis threshold, before fully recovering. E.g., the symptoms of a diseased individual may be too light to lead to detection and quarantine. To include such set-ups, we would need a branching in the way the individuals move between the groups.

**Acknowledgments:** A. Hening and D. Nguyen acknowledge support from the NSF through the grants DMS-2147903 and DMS-2407668 respectively.

**Conflict of interest:** The authors declare no conflict of interest.

## REFERENCES

- Hakan Andersson and Tom Britton. *Stochastic epidemic models and their statistical analysis*, volume 151. Springer Science & Business Media, 2012.
- M. Benaïm. Stochastic persistence. 2018. preprint.

- M. Benaïm and C. Lobry. Lotka Volterra in fluctuating environment or “how switching between beneficial environments can make survival harder”. *The Annals of Applied Probability*, 26(6):3754–3785, 2016.
- M. Benaïm and S. J. Schreiber. Persistence of structured populations in random environments. *Theoretical Population Biology*, 76(1):19–34, 2009.
- M. Benaïm, J. Hofbauer, and W. H. Sandholm. Robust permanence and impermanence for stochastic replicator dynamics. *J. Biol. Dyn.*, 2(2):180–195, 2008.
- Peter L Chesson. The stabilizing effect of a random environment. *Journal of Mathematical Biology*, 15(1):1–36, 1982.
- M. H. A. Davis. Piecewise-deterministic markov processes: A general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society: Series B (Methodological)*, 46(3):353–376, 1984.
- Nguyen Thanh Dieu. Asymptotic properties of a stochastic sir epidemic model with beddington–deangelis incidence rate. *Journal of Dynamics and Differential Equations*, 30(1):93–106, 2018.
- Nguyen Thanh Dieu, Dang Hai Nguyen, Nguyen Huu Du, and G Yin. Classification of asymptotic behavior in a stochastic sir model. *SIAM Journal on Applied Dynamical Systems*, 15(2):1062–1084, 2016.
- Nguyen Huu Du and Nguyen Ngoc Nhu. Permanence and extinction of certain stochastic sir models perturbed by a complex type of noises. *Applied Mathematics Letters*, 64:223–230, 2017.
- Nguyen Huu Du, Nguyen Thanh Dieu, and Nguyen Ngoc Nhu. Conditions for permanence and ergodicity of certain sir epidemic models. *Acta Applicandae Mathematicae*, 160:81–99, 2019.
- S. N. Evans, P. L. Ralph, S. J. Schreiber, and A. Sen. Stochastic population growth in spatially heterogeneous environments. *J. Math. Biol.*, 66(3):423–476, 2013.
- S. N. Evans, A. Hening, and S. J. Schreiber. Protected polymorphisms and evolutionary stability of patch-selection strategies in stochastic environments. *J. Math. Biol.*, 71(2):325–359, 2015. ISSN 0303-6812.
- Zhilan Feng and Horst R Thieme. Recurrent outbreaks of childhood diseases revisited: the impact of isolation. *Mathematical biosciences*, 128(1-2):93–130, 1995.
- Zhilan Feng and Horst R Thieme. Endemic models with arbitrarily distributed periods of infection ii: fast disease dynamics and permanent recovery. *SIAM Journal on Applied Mathematics*, 61(3):983–1012, 2000.
- Christel Geiß and Ralf Manthey. Comparison theorems for stochastic differential equations in finite and infinite dimensions. *Stochastic processes and their applications*, 53(1):23–35, 1994.
- Mike Giles. Improved multilevel monte carlo convergence using the milstein scheme. In *Monte Carlo and Quasi-Monte Carlo Methods 2006*, pages 343–358. Springer, 2006.
- A. Hening and D. Nguyen. Coexistence and extinction for stochastic Kolmogorov systems. *The Annals of applied probability*, 28(3):1893–1942, 2018.
- A. Hening and D. H. Nguyen. The competitive exclusion principle in stochastic environments. *Journal of Mathematical Biology*, 80:1323–1351, 2020.
- A Hening, D. Nguyen, and P Chesson. A general theory of coexistence and extinction for stochastic ecological communities. *Journal of Mathematical Biology*, 82(6):1–76, 2021.
- A. Hening, D. H. Nguyen, and S. J. Schreiber. A classification of the dynamics of three-dimensional stochastic ecological systems. *Annals of Applied Probability*, 32(2), 2022.
- Alexandru Hening, Nguyen Trong Hieu, Dang Hai Nguyen, and Nhu Ngoc Nguyen. Stochastic nutrient-plankton models. *Journal of Differential Equations*, 376:370–405, 2023.
- Herbert Hethcote, Ma Zhien, and Liao Shengbing. Effects of quarantine in six endemic models for infectious diseases. *Mathematical biosciences*, 180(1-2):141–160, 2002.
- Gang Huang, Wanbiao Ma, and Yasuhiro Takeuchi. Global properties for virus dynamics model with beddington–deangelis functional response. *Applied Mathematics Letters*, 22(11):1690–1693, 2009.

- Jiehui Jiang, Sixing Gong, Bing He, et al. Dynamical behavior of a rumor transmission model with holling-type ii functional response in emergency event. *Physica A: Statistical Mechanics and its Applications*, 450:228–240, 2016.
- William Ogilvy Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.
- William Ogilvy Kermack and Anderson G McKendrick. Contributions to the mathematical theory of epidemics. ii.—the problem of endemicity. *Proceedings of the Royal Society of London. Series A, containing papers of a mathematical and physical character*, 138(834):55–83, 1932.
- Guijie Lan, Ziyang Lin, Chunjin Wei, and Shuwen Zhang. A stochastic sirs epidemic model with non-monotone incidence rate under regime-switching. *Journal of the Franklin Institute*, 356(16): 9844–9866, 2019.
- R. Lande, S. Engen, and B.-E. Saether. *Stochastic population dynamics in ecology and conservation*. Oxford University Press on Demand, 2003.
- Yuguo Lin and Daqing Jiang. Threshold behavior in a stochastic sis epidemic model with standard incidence. *Journal of Dynamics and Differential Equations*, 26:1079–1094, 2014.
- Qun Liu, Daqing Jiang, Tasawar Hayat, and Bashir Ahmad. Analysis of a delayed vaccinated sir epidemic model with temporary immunity and lévy jumps. *Nonlinear Analysis: Hybrid Systems*, 27:29–43, 2018.
- Wei-min Liu, Simon A Levin, and Yoh Iwasa. Influence of nonlinear incidence rates upon the behavior of sirs epidemiological models. *Journal of mathematical biology*, 23:187–204, 1986.
- Shiva Moein, Niloofar Nickaeen, Amir Roointan, Niloofar Borhani, Zarifeh Heidary, Shaghayegh Haghjooy Javanmard, Jafar Ghaisari, and Yousof Gheisari. Inefficiency of sir models in forecasting covid-19 epidemic: a case study of isfahan. *Scientific reports*, 11(1):4725, 2021.
- Dang H Nguyen, Nhu N Nguyen, and George Yin. Analysis of a spatially inhomogeneous stochastic partial differential equation epidemic model. *Journal of Applied Probability*, 57(2):613–636, 2020a.
- Dang H Nguyen, Nhu N Nguyen, and George Yin. General nonlinear stochastic systems motivated by chemostat models: Complete characterization of long-time behavior, optimal controls, and applications to wastewater treatment. *Stochastic Processes and their Applications*, 130(8):4608–4642, 2020b.
- Dang Hai Nguyen, George Yin, and Chao Zhu. Long-term analysis of a stochastic sirs model with general incidence rates. *SIAM Journal on Applied Mathematics*, 80(2):814–838, 2020c.
- Nhu N Nguyen and George Yin. Stochastic partial differential equation sis epidemic models: modeling and analysis. *Communications on Stochastic Analysis*, 13(3):8, 2019.
- Lin-Fei Nie, Jing-Yun Shen, and Chen-Xia Yang. Dynamic behavior analysis of sirs epidemic models with state-dependent pulse vaccination. *Nonlinear Analysis: Hybrid Systems*, 27:258–270, 2018.
- Nguyen Dinh Phu, Donal O’Regan, and Tran Dinh Tuong. Longtime characterization for the general stochastic epidemic sis model under regime-switching. *Nonlinear Analysis: Hybrid Systems*, 38: 100951, 2020.
- Shigui Ruan and Wendi Wang. Dynamical behavior of an epidemic model with a nonlinear incidence rate. *Journal of differential equations*, 188(1):135–163, 2003.
- S. J. Schreiber and J. O. Lloyd-Smith. Invasion dynamics in spatially heterogeneous environments. *The American Naturalist*, 174(4):490–505, 2009.
- S. J. Schreiber, M. Benaïm, and K. A. S. Atchadé. Persistence in fluctuating environments. *J. Math. Biol.*, 62(5):655–683, 2011.
- Yilei Tang, Deqing Huang, Shigui Ruan, and Weinian Zhang. Coexistence of limit cycles and homoclinic loops in a sirs model with a nonlinear incidence rate. *SIAM Journal on Applied Mathematics*, 69(2):621–639, 2008.
- Pekka Tuominen and Richard L Tweedie. Subgeometric rates of convergence of f-ergodic markov chains. *Advances in Applied Probability*, 26(3):775–798, 1994.

- Feng Wang and Zaiming Liu. Dynamical behavior of a stochastic siqs model via isolation with regime-switching. *Journal of Applied Mathematics and Computing*, 69(2):2217–2237, 2023.
- Shan Wang and Feng Wang. Asymptotic behavior of a stochastic hybrid siqrs model with vertical transmission and nonlinear incidence. *AIMS Mathematics*, 9(5):12529–12549, 2024.
- Liang Zhang, Zhi-Cheng Wang, and Xiao-Qiang Zhao. Threshold dynamics of a time periodic reaction–diffusion epidemic model with latent period. *Journal of Differential Equations*, 258(9):3011–3036, 2015.
- Xiao-Bing Zhang, Hai-Feng Huo, Hong Xiang, and Xin-You Meng. An sirs epidemic model with pulse vaccination and non-monotonic incidence rate. *Nonlinear Analysis: Hybrid Systems*, 8:13–21, 2013.
- Baoquan Zhou, Daqing Jiang, Yucong Dai, and Tasawar Hayat. Stationary distribution and density function expression for a stochastic siqrs epidemic model with temporary immunity. *Nonlinear Dynamics*, 105(1):931–955, 2021.

DEPARTMENT OF MATHEMATICS, TEXAS A&M UNIVERSITY, MAILSTOP 3368, COLLEGE STATION, TX 77843-3368, UNITED STATES

*Email address:* `ahening@tamu.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ALABAMA, 345 GORDON PALMER HALL, BOX 870350, TUSCALOOSA, AL 35487-0350, UNITED STATES

*Email address:* `dangnh.maths@gmail.com`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ALABAMA, 345 GORDON PALMER HALL, BOX 870350, TUSCALOOSA, AL 35487-0350, UNITED STATES

*Email address:* `ttta@crimson.ua.edu`

DEPARTMENT OF ECONOMICS, CITY, UNIVERSITY OF LONDON, NORTHAMPTON SQUARE, LONDON EC1V 0HB, UNITED KINGDOM

*Email address:* `Sergiu.Ungureanu.1@city.ac.uk`