Journal of Mathematical Biology (2023) 87:85
https://doi.org/10.1007/s00285-023-02018-z

Mathematical Biology

Stochastic dynamics of human papillomavirus delineates cervical cancer progression

Tuan Anh Phan[1] · Farhana Sarower[2] · Jinqiao Duan[3] · Jianjun Paul Tian[2]

Received: 1 June 2022 / Revised: 5 July 2023 / Accepted: 17 October 2023

Abstract
Starting from a deterministic model, we propose and study a stochastic model for human papillomavirus infection and cervical cancer progression. Our analysis shows that the chronic infection state as random variables which have the ergodic invariant probability measure is necessary for progression from infected cell population to cervical cancer cells. It is shown that small progression rate from infected cells to precancerous cells and small microenvironmental noises associated with the progression rate and viral infection help to establish such chronic infection states. It implicates that large environmental noises associated with viral infection and the progression rate in vivo can reduce chronic infection. We further show that there will be a cervical cancer if the noise associated with precancerous cell growth is large enough. In addition, comparable numerical studies for the deterministic model and stochastic model, together with Hopf bifurcations in both deterministic and stochastic systems, highlight our analytical results.

✉ Jianjun Paul Tian
jtian@nmsu.edu

Tuan Anh Phan
tphan@uidaho.edu

Farhana Sarower
farhana8@nmsu.edu

Jinqiao Duan
duan@gbu.edu.cn

[1] Institute for Modeling Collaboration and Innovation, University of Idaho, Moscow, ID 83844, USA

[2] Department of Mathematical Sciences, New Mexico State University, Las Cruces, NM 88001, USA

[3] Departments of Mathematics, School of Sciences, Great Bay University, Dongguan 523000, Guangdong, China

Published online: 12 November 2023

Keywords Chronic infection state · Ergodic invariant probability measure · Cervical cancer · Stochastic dynamical bifurcation

Mathematics Subject Classification 60H10 · 34F10

1 Introduction

About a quarter of all human cancers worldwide are caused by infectious agents (Zapatka et al. 2020). The World Health Organization estimates that 15.4% of all cancers are attributable to infections and 9.9% are linked to viruses (Parkin 2006; Plummer 2016). There are eleven pathogens which have been identified as carcinogenic agents in humans (Bouvard 2009). The first two are Helicobacter pylori (associated with 770,000 cases worldwide) and human papilloma virus (HPV) (associated with 640,000 cases worldwide) (Munoz et al. 2006). Various characteristics have been proposed to define human viruses that cause cancer while understanding the progression from viral infection to cancer development is still in need (White et al. 2014; Butel 2000). As an important example, we will study the malignant progression of human papilloma virus infection in this research.

HPVs infect epithelial cells, causing hyperproliferative lesions such as warts and condylomas. More than 100 HPV types have been identified (de Villiers 1994). According to their tissue tropism, they are categorized into two major groups, the cutaneous and mucosal HPVs. The mucosal HPVs are further grouped into high-risk and low-risk types. Lesions caused by high-risk HPVs have a propensity to progress to malignant tumors, most prominently cervical carcinomas. In contrast, lesions caused by low-risk HPVs have a much lower risk for malignant progression (Shah and Howley 1996). HPV can be sexually transmitted. Actually, it is the most common sexually transmitted disease in the world (Frazer et al. 2006). Genital HPV, which is transmitted sexually, is the major etiologic factor in cervical cancer worldwide (Bosch et al. 1995). Cervical cancer is the second most common form of cancer worldwide and HPV types 16, 18, 31 account for approximately 85% of all cervical cancer cases (Clifford et al. 2003). But, most women infected with HPV, even those infected with the types that are most closely associated with cervical dysplasia (types 16 and 18), will not develop invasive cervical cancer (Reingold 2000). Low-grade cervical cell abnormalities usually are cleared spontaneously and rarely progress to cancer, while high-grade cervical cell abnormalities have a lower probability of spontaneous clearance and a higher probability of progression to cancer (greater than 12%) (Ostor 1993). On the other hand, at the molecular level, to initiate infection, HPVs bind basal epithelial cells in the cervix. Following binding and entry of the cell, viral materials migrate to the nucleus and replicate its HPV genome to 20–100 copies as multiple-copy extrachromosomal plasmids. The viral proteins E6 and E7 are expressed by infected cells, which promote their proliferation and deactivate cancer suppressor proteins p53 and pRb. Upon cell division, the viral genome duplicates and divides into two daughter cells. One daughter cell migrates away from the basal layer and starts processes of differentiation. Unlike normal cells, HPV-infected cells undergo differentiation but remain active in the cell cycle (White et al. 2014; Lee and Laimins 2007). At the cellular level, the presence

and persistence of viral DNA in tumor site are necessary conditions to establish an epidemiologically evident tumor (Hausen 2001).

Mathematical modeling is a suitable tool to study malignant progression, the dynamics of virus infection and cancer development. Actually, most mathematical models related to cervical cancer have been focused on epidemiology with emphasis on the transmission between individuals and effectiveness of HPV vaccine, for example, (Barnabas et al. 2006; Brown and White 2011; Elbasha 2008; Goldhaber-Fiebert et al. 2008; Kim et al. 2008; Lee and Tameru 2012; Ziyadi 2017; Bumrungthai 2023). There are two review articles about mathematical modeling of cervical cancer epidemiology (Ryser et al. 2017; Iskandar et al. 2022). Recently, there are three published mathematical models related to cervical cancer at the cellular level, Asih et al. (2016), Murtono et al. (2019) and Sierra-Rojas et al. (2022), as the authors know. The model in Murtono et al. (2019) was based on that in Asih et al. (2016) with a chemotherapy treatment. The model in Sierra-Rojas et al. (2022) is only about the dynamics of HPV infection without cervical cancer development. We will focus on the model in Asih et al. (2016) which was proposed for the dynamics of HPV infected cells and progression to cancer at the cellular level. This model is a system of five ordinary differential equations. There are three parameters, the progression rate from infected cells to precancerous cells, the net death rate of infected cells, the net growth rate of the precancerous cells, which are important for cancer establishment. A deep insight is needed from these parameters. In the current research, we use Ito stochastic differential equations to gain a thorough understanding about how infected epithelial cells progress to cervical cancer cells through these parameters.

Asih et al. (2016) proposed a deterministic mathematical model that describes HPV infection and cancer development in the cervix. This model consists of five compartments, susceptible (normal) cells (S), infected cells (I), free virus (V), precancerous cells (P) and cancer cells (C). It is as follows.

$$\begin{aligned}
\frac{dS}{dt} &= rS\left(1-\frac{S+I}{N}\right)-\alpha SV,\\
\frac{dI}{dt} &= \alpha SV + a_1 I - d_1 I - \delta I,\\
\frac{dV}{dt} &= n_1 d_1 I - d_4 V,\\
\frac{dP}{dt} &= \delta I + a_2 P - d_2 P - \frac{\theta P^2}{K^2+P^2},\\
\frac{dC}{dt} &= \frac{\theta P^2}{K^2+P^2} + a_3 C - d_3 C.
\end{aligned} \tag{1}$$

For the completion, the model is briefly explained here. For the detailed explanation of the model establishment, the reader is referred to Asih et al. (2016). In the cervix, there are several types of cells in different layers of the epithelium which facilitate HPV infections, for example, basal epithelial cells and squamous epithelial cells, granular layer cells, spinous layer cells, and basal layer cells. And, these cells are distributed in different locations of the cervix, ectocervix, squamous columnar junction, and endocervix (Graham 2017). HPV can bind receptors on the basement membrane of the

cervix and go on to infect basal layer cells of the epithelium. Division of an infected basal epithelial cells can give rise to a transit amplifying cell that is capable of differentiation. Viral genomes are segregated into daughter cells upon basal cell division and can be carried into deep epithelial layers. The keratinocyte differentiation process allows an orchestrated pattern of viral gene expression deep into epithelial layers. On the other hand, the infection spread along basement membrane into squamous columnar junction. Most cervical cancers are thought to arise from this zone (Graham 2017; Herfs 2012). To mathematically model, the authors of Asih et al. (2016) made several simplifications and assumptions as follows.

It is assumed that the normal cell population $S(t)$ including basal and squamous epithelial cells follows logistic growth with the intrinsic growth rate r and the carrying capacity N. These cells are infected via a mass action law $\alpha S(t)V(t)$, and the infected cells grow or die linearly at a net per capita rate of $a_1 - d_1$. Infected cells may transit to precancerous cells at per capita rate δ. New viruses are produced at a rate proportional to the death rate of infected cells and decay linearly at rate d_4. Once infected cells become precancerous cells, they stop producing viral particles (Graham 2017). The precancerous cell population grows or decays at the net per capita rate $a_2 - d_2$. These precancerous cells can transit to cancerous cells. This transition is a rare event. There are several ways to model rare events. The work Asih et al. (2016) chose a type III Holling response functional $\frac{\theta P^2}{K^2 + P^2}$. In this term, θ is the maximal progression rate from precancerous cells to cancerous cells, and K is the half-saturation concentration for the progression rate from precancerous to cancerous cells. When the population of precancerous cells is low, there is a small risk of developing cancer cells. However, when the precancerous cells are above half-saturation concentration (K), it is highly likely for cancer cells to be developed quickly and to approach a maximum rate of θ. Once formed, cancer cells have a net per capita growth rate of $a_3 - d_3$.

The assumption of logistic growth of normal cells is a reasonable simplification because the total number of basal epithelial cells remains approximately constant and the squamous cells are completely replaced every 4–5 days during the infection course (Wright and Ferenczy 2002). All other linear growth and death terms are first approximations. The incidence of cervical cancer cells is a rare event. The type III Holling response functional catches some character of rare events, which is acceptable. Because precancerous cells neither produce viruses nor transit back to infected cells, and only transit to cancer cells, these two populations are decoupled from the first three populations (S, I, and V). And, this decoupling is manifested in their different physical locations (Graham 2017; Herfs 2012). We note that spatial interactions or competitions among cells play some roles during the infection process. However, the model only considers temporal dynamics without spital dynamics. This may be because the model if a system of ordinary differential equations (ODEs). Overall, this model is of biological significance, which can serve as a starting point for further study.

Theoretical and numerical results of the system (1) in Asih et al. (2016) indicated that the parameter $a = d_1 - a_1$, the net death rate of infected cells; the parameter δ, the progression rate from infected cells to precancerous cells; and the parameter $b = a_2 - d_2$, the net growth of precancerous cells, are three key parameters for long-term behaviors of the deterministic system (1) as well as gaining insights into possible

strategies for the prevention of cervical cancer development. However, as all deterministic ODE systems, this model only represents mean behaviors of cells and viruses, and it is known that the system is subject to micro-environmental fluctuations in vivo. The system perturbed by such fluctuations may exhibit some long-term behaviors that are different from those of the unperturbed system (Duan 2015). Obviously, three parameters a, δ, and b are subject to micro-environmental fluctuations. It may lead to the fact that effective treatment methods by altering these parameters to prevent cancer development as proposed in Asih et al. (2016) does not work anymore. Hence, in this paper, in order to understand how microenvironmental noises or uncertainties influence the progression dynamics, we incorporate micro-environmental noises into the ODE system (1). In addition, we may think of some noises come from spatial variations, which may provide some complement to the ODE system.

There are several ways to incorporate environmental noise or stochastic effects into deterministic mathematical models. Suppose X is a population, its growth or change is modeled $\frac{dX}{dt} = f(t, X)$ in the deterministic situation. To count for environmental noise and stochastic effects, we may consider that each individual in the population makes almost same contribution to stochastic effects and receives the same environmental noise. Then, we may assume the environmental noise and stochastic effects of the population is proportional to the population X. In other words, the environmental noise and stochastic effects can be represented by $\tau X\xi$, where ξ is the unit noise and τ can be thought as a way to measuring an average variation of each individual. In general, we take the noise to be white noise $\xi = \frac{dW}{dt}$, where $W = W(t)$ is the standard Wiener Process. So, we obtain an Ito stochastic differential equation $dX = f(t, x)dt+\tau X dW$. We may call the noise added this way the linear noise (Phan and Tian 2020; Phan et al. 2021). For the model (1), we will incorporate linear noise into the free virus population $V(t)$, namely $\tau_2 V\frac{dW_2}{dt}$. The second way to incorporate environmental noises is perturbing parameters of interest (Phan and Tian 2022; Phan et al. 2021). As the similar principle established based on law of large numbers for perturbing parameters, we incorporate environmental noises into the net death rate of infected cells a and the progression rate of precancerous cells δ by replacing the sum $a+\delta$ with $a + \delta + \tau_1\frac{dW_1}{dt}$, into the net proliferation rate of precancerous cells b by replacing b with $b + \tau_3\frac{dW_3}{dt}$. We consider $W_1(t)$, $W_2(t)$, and $W_3(t)$ are mutually independent Wiener processes. Then we obtain five dimensional Ito stochastic differential equation system as follows.

$$\begin{aligned}
dS &= \left[rS\left(1-\frac{S+I}{N}\right)-\alpha SV\right]dt,\\
dI &= (\alpha SV + a_1 I - d_1 I - \delta I)dt - \tau_1 I dW_1,\\
dV &= (n_1 d_1 I - d_4 V)dt + \tau_2 V dW_2,\\
dP &= \left(\delta I + a_2 P - d_2 P - \frac{\theta P^2}{K^2+P^2}\right)dt + \tau_3 P dW_3,\\
dC &= \left(\frac{\theta P^2}{K^2+P^2} + a_3 C - d_3 C\right)dt. \qquad (2)
\end{aligned}$$

In this article, we conduct a detailed analysis about this stochastic model with numerical demonstrations, and provide biological interpretations of our analytical results. In a sense, ergodic invariant probability measures in Ito stochastic systems play similar roles as equilibrium solutions in deterministic systems. However, analyzing stochastic systems requires more and deeper knowledge from probability theory. To study the proposed stochastic system above, we utilize the stochastic version of Lyapunov exponent theory (Arnold et al. 1984, 1990) and boundary analysis (Dieu et al. 2016; Du et al. 2016; Hening and Nguyen 2018; Phan and Tian 2022). Since our stochastic system is noise degenerated, we apply Hörmander's theorems (Bellet 2006; Nualart 2006; Hörmander 1967) to check hypoellipticity. To study ergodicity, for example, supports of invariant measures, we use geometric control theory (Bellet 2006; Ikeda and Watanabe 1989; Jurdjevic 1996). In addition, we show Hopf bifurcation occurs in the deterministic system (1). To compare with the deterministic counterpart, we also study the occurrence of stochastic Hopf bifurcations in the stochastic system. There are two types of stochastic bifurcations described in the book (Arnold 1998). The first type is the phenomenological bifurcation (or P-bifurcation), which is concerned with the change in the shape of density functions of a family of invariant probability measures in a stochastic system as one of its parameters changes. The second one is the dynamical bifurcation (or D-bifurcation), which is characterized by sign changes of Lyapunov exponents of a family of invariant probability measures in a stochastic system as one of its parameters changes. We numerically illustrate dynamical bifurcations occur in our systems (2).

Because of the way we incorporate environmental noises, many equilibria of the deterministic system (1) do not have their stochastic counterparts in the stochastic system (2). However, the biological significant equilibria correspond to ergodic invariant probability measures with different thresholds, which encode environmental uncertainties and provide more insights to tumor progression. The chronic infection state or equilibrium is a necessary condition to establish cervical cancer. In our stochastic system, it is an ergodic invariant measure. When the noises associated with viral infection and progression from infected cells to precancerous cells are small, the chronic infection state can be achieved. When these two noises are large, there is a great probability to reduce chronic infection state. When the chronic infection state is established, there is a possibility to establish a cervical cancer. The cancer is an ergodic invariant measure in our stochastic system which corresponds to an equilibrium in the deterministic system. When the noise associated with precancerous cell growth is large enough, there will be a cervical cancer.

The rest of the paper is organized as follows. In Sect. 2, we list our main results and give medical interpretations or implications. In Sect. 3, we prove theorems listed in Sect. 2. In Sect. 4, we conduct numerical studies to demonstrate our analytical results and bifurcations for both deterministic and stochastic models. In Sect. 5, we discuss several aspects of our research. The article ends with Appendix for the relation between thresholds in deterministic and stochastic models.

2 Results and interpretations

We non-dimensionalize the system (2) by setting $S_1 = \frac{S}{N}$, $I_1 = \frac{I}{N}$, $P_1 = \frac{P}{K}$, $C_1 = \frac{C}{K}$, $a = d_1 - a_1$, $b = a_2 - d_2$, and $k = d_3 - a_3$. To reduce the number of parameters, we set $n = n_1 d_1 N$, $e = d_4$, $q = \frac{N}{K}$, and $\overline{\theta} = \frac{\theta}{K}$. (It should be noted that we do not rescale time which avoids changes of Wiener processes.) Then dropping all the indices of the variables and all the bars over parameters, we obtain our non-dimensionalized SDE system

$$\begin{aligned}
dS &= [rS(1 - S - I) - \alpha SV]\,dt,\\
dI &= (\alpha SV - aI - \delta I)dt - \tau_1 I dW_1,\\
dV &= (nI - eV)dt + \tau_2 V dW_2,\\
dP &= \left(\delta q I + bP - \theta \frac{P^2}{1 + P^2}\right) dt + \tau_3 P dW_3,\\
dC &= \left(\theta \frac{P^2}{1 + P^2} - kC\right) dt,
\end{aligned} \tag{3}$$

and the corresponding deterministic system of (3) is

$$\begin{aligned}
\frac{dS}{dt} &= rS(1 - S - I) - \alpha SV,\\
\frac{dI}{dt} &= \alpha SV - aI - \delta I,\\
\frac{dV}{dt} &= nI - eV,\\
\frac{dP}{dt} &= \delta q I + bP - \theta \frac{P^2}{1 + P^2},\\
\frac{dC}{dt} &= \theta \frac{P^2}{1 + P^2} - kC.
\end{aligned} \tag{4}$$

Note that all the parameters are positive except that b might be positive or negative. First, we present some results of the deterministic system (4). In Asih et al. (2016), it was found that, under some certain conditions of parameter b and the basic reproduction number $R_0 := \frac{\alpha n}{e(a + \delta)}$, the system (4) has two positive equilibria, which are $(S^*, I^*, V^*, P_{2i}, C_{2i})$ $(i = 1, 2)$. The one with the lower P and C values, namely $(S^*, I^*, V^*, P_{21}, C_{21})$, is locally asymptotically stable if $R_0 \leq 3$ and possibly for larger values of R_0. The one with higher P and C values is always unstable. However, we found a threshold b_0 for parameter b and a threshold R^* for the reproduction number R_0 so that a Hopf bifurcation arising from the equilibrium point $(S^*, I^*, V^*, P_{21}, C_{21})$ occurs when b is below b_0 and R_0 passes through R^*. We summarize these results in the following theorem in order to compare with results obtained from our stochastic

system, where parts (i) and (iii) are our new results and (ii) is from Asih et al. (2016).

Theorem 2.1 (i) *Assume that* $R_0 > 1$ *and* $\alpha n > \frac{e^2+\sqrt{e^4+4e^3r}}{2}$. *Then, the sub-system* (S, I, V) *of the system* (4) *has a unique positive equilibrium point* $E^* := (S^*, I^*, V^*)$. *Let* $R^* := \frac{2}{\sqrt{A^2-4B}-A}$, *where* $A := \frac{2\alpha ne^2+2\alpha nre+re^3}{\alpha^2 n^2} + \frac{e^2}{er+\alpha n}$ *and* $B := \frac{e^4}{\alpha^2 n^2} - \frac{e^2}{er+\alpha n}$. *When* $R_0 < R^*$, E^* *is locally asymptotically stable; when* $R_0 > R^*$, E^* *is unstable. A Hopf bifurcation occurs at* $R_0 = R^*$ *and this bifurcation gives rise to one family of periodic solutions around* E^*.

(ii) *Assume that the sub-system has a unique positive equilibrium point* $E^* = (S^*, I^*, V^*)$. *Furthermore, we assume that* $0 < b < \frac{\theta}{2}$ *and* $\theta > \delta q I^*$. *Let* $b_0 := \frac{2\theta P_0}{(1+P_0^2)^2}$, *where* $P_0 := \left(\frac{\theta+2\delta q I^*}{2\theta-2\delta q I^*} + \left(\left(\frac{\theta+2\delta q I^*}{2\theta-2\delta q I^*}\right)^2 + \frac{\delta q I^*}{\theta-\delta q I^*}\right)^{1/2}\right)^{1/2}$. *When* $0 < b < b_0$, *the sub-system* (P, C) *has two positive equilibria* (P_{2i}, C_{2i}) $(i = 1, 2)$ *where* (P_{21}, C_{21}) *is locally asymptotically stable and* (P_{22}, C_{22}) *is unstable. When* $b_0 < b < \frac{\theta}{2}$, *there is no positive equilibrium point for the sub-system* (P, C). *When* $b = b_0$, *the sub-system has a unique positive equilibrium point* (P_0, C_0).

(iii) *Assume that* $R_0 > 1$, $\alpha n > \frac{e^2+\sqrt{e^4+4e^3r}}{2}$, $\theta > \delta q I^*$, *and* $0 < b < b_0$. *The whole system* (4) *has the positive equilibrium point* $(S^*, I^*, V^*, P_{21}, C_{21})$ *in which* (P_{21}, C_{21}) *is locally asymptotically stable. There occurs a Hopf bifurcation around this equilibrium point when* R_0 *passes through the threshold* R^*.

Next, we present our main results about the stochastic system (3). Similar to the stochastic model in Phan and Tian (2020), to make analysis of the system (3) easier, we make variable transformation, $S = S$, $I = I$, $V = MI$ or $M = \frac{V}{I}$, $P = P$, $C = C$, using Ito's formula, then the system (3) is equivalent to the following system,

$$\begin{aligned}
dS &= [rS(1 - S - I) - \alpha SIM]\,dt,\\
dI &= (\alpha SM - a - \delta)I dt - \tau_1 I dW_1,\\
dM &= (n - \alpha SM^2 + (a + \delta + \tau_1^2 - e)M)dt + \tau_1 M dW_1 + \tau_2 M dW_2,\\
dP &= \left(\delta q I + bP - \theta\frac{P^2}{1+P^2}\right)dt + \tau_3 P dW_3,\\
dC &= \left(\theta\frac{P^2}{1+P^2} - kC\right)dt.
\end{aligned} \tag{5}$$

Suppose there is a complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual condition for our system. The process given by the solution to (5) is denoted by U or $U(t) = (S(t), I(t), M(t), P(t), C(t))^T$, $t \geq 0$. We denote the drift term and the diffusion term of the system (5), respectively, by

$$f(U) = \begin{bmatrix} rS(1-S-I) - \alpha SIM \\ (\alpha SM - a - \delta)I \\ n - \alpha SM^2 + (a + \delta + \tau_1^2 - e)M \\ \delta qI + bP - \theta \dfrac{P^2}{1+P^2} \\ \theta \dfrac{P^2}{1+P^2} - kC \end{bmatrix}, \quad \text{and} \quad g(U) = \begin{bmatrix} 0 & 0 & 0 \\ -\tau_1 I & 0 & 0 \\ \tau_1 M & \tau_2 M & 0 \\ 0 & 0 & \tau_3 P \\ 0 & 0 & 0 \end{bmatrix}.$$

Let $\mathcal{L}$ be the infinitesimal generator of the process U and, for any smooth enough function $F : \mathbb{R}^5_+ \to \mathbb{R}$, the generator $\mathcal{L}$ acts as

$$\mathcal{L}F(U) := F_U \cdot f(U) + \frac{1}{2}\text{trace}\left(g(U)g(U)^T F_{UU}\right)$$

where F_U is the gradient of F and F_{UU} is the Hessian matrix of F. We denote by $\mathbb{P}_u$ the probability law on Ω when the solution starts at $u = (s, i, m, p, c)^T$ and $\mathbb{E}_u$ the expectation corresponding to $\mathbb{P}_u$. The following theorem guarantees that the non-compact region

$$D := \left\{(S, I, M, P, C)^T : S \geq 0,\ I \geq 0,\ M \geq 0,\ P \geq 0,\ C \geq 0,\ S + I \leq 1\right\}$$

is the a.s. (almost sure) non-negative invariant domain of the system (5) and we refer it to be a global domain.

Theorem 2.2 *For any initial value $u = (s, i, m, p, c)^T \in \mathbb{R}^5_+$ where*

$$\mathbb{R}^5_+ := \{(S, I, M, P, C)^T : S \geq 0,\ I \geq 0,\ M \geq 0,\ P \geq 0,\ C \geq 0\},$$

there exists a unique a.s. continuous solution $U(t)$ of the system (5) that remains in $\mathbb{R}^5_+$ for all times $t \geq 0$ (i.e. the explosion time $\tau_e = \infty$ a.s.) and $U(t)$ is a strong Markov process that satisfies the Feller property. Furthermore, if $u \in D^\circ$ in which

$$D^\circ := \left\{(S, I, M, P, C)^T \in \mathbb{R}^{5,\circ}_+ :\ S + I < 1\right\}$$

then $U(t) \in D^\circ$ for all $t \geq 0$ a.s.

By boundary analysis in Sect. 3, when the solution $U(t)$ of the system (5) starts in

$$\{S = 0\} := \{(S, I, M, P, C)^T \in D :\ S = 0\},$$

there is a unique ergodic invariant probability measure $\mu_0 := \delta_0^* \times \delta_0^* \times \pi_1 \times \delta_0^* \times \delta_0^*$ for the system (5) provided $2(e-a-\delta) > \tau_1^2 - \tau_2^2$ and $b < \frac{\tau_3^2}{2}$ in which

$$\pi_1 \sim \text{IG}\left(\frac{2(e-a-\delta-\tau_1^2)}{\tau_1^2+\tau_2^2}+1, \frac{2n}{\tau_1^2+\tau_2^2}\right)$$

is an inverse gamma distribution. When the solution $U(t)$ of the system (5) starts in

$$\{S>0, I=0\} := \{(S,I,M,P,C)^T \in D : S>0,\ I=0\},$$

there exists a unique ergodic invariant probability measure $\mu_1 := \delta_1^* \times \delta_0^* \times \pi_2 \times \delta_0^* \times \delta_0^*$ for the system (5) provided $b < \frac{\tau_3^2}{2}$ where $\pi_2 \sim \text{GIG}(\Theta, \chi, \psi)$ is a generalized inverse Gaussian distribution with parameters $\Theta = \frac{2(a+\delta+\tau_1^2-e)}{\tau_1^2+\tau_2^2} - 1$, $\chi = \frac{4n}{\tau_1^2+\tau_2^2}$, and $\psi = \frac{4\alpha}{\tau_1^2+\tau_2^2}$. Now suppose the solution $U(t)$ of the system (5) starts in

$$\{S>0, I>0\} := \{(S,I,M,P,C)^T \in D : S>0,\ I>0\},$$

by Theorem 2.2, $U(t) \in D^\circ$ for all $t \geq 0$ a.s. Define the threshold

$$\lambda := \int_{\partial D} (\alpha SM - a - \delta) d\mu_1 = \sqrt{\alpha n}\, R_\Theta(w) - a - \delta - \frac{\tau_1^2}{2}$$

where $w := \sqrt{\chi\psi} = \frac{4\sqrt{\alpha n}}{\tau_1^2+\tau_2^2}$ and $R_\Theta(w) = \frac{K_{\Theta+1}(w)}{K_\Theta(w)}$ with $K_\Theta(\cdot)$ is the modified Bessel function of the third kind given by

$$K_\Theta(\phi) := \frac{1}{2}\int_0^\infty x^{\Theta-1} \exp\left\{-\frac{1}{2}\phi\left(x+\frac{1}{x}\right)\right\} dx, \quad \phi > 0.$$

Since the dynamics of the first three equations of the system (5) is independent of the last two equations, and the dynamics of the last two equations can be derived once we know the dynamics of the first three equations, we can separate the system (5) into two subsystems

$$\begin{aligned} dS &= [rS(1-S-I) - \alpha SIM]\,dt, \\ dI &= (\alpha SM - a - \delta) I dt - \tau_1 I dW_1, \\ dM &= (n - \alpha SM^2 + (a+\delta+\tau_1^2-e)M)dt + \tau_1 M dW_1 + \tau_2 M dW_2, \end{aligned} \tag{6}$$

and

$$dP = \left(\delta q I + bP - \theta \frac{P^2}{1+P^2}\right) dt + \tau_3 P dW_3,$$
$$dC = \left(\theta \frac{P^2}{1+P^2} - kC\right) dt. \tag{7}$$

Then the long-term behavior of the system (6) is determined by the value of λ while the long-term behavior of the system (7) is governed by the value of b and the dynamics of the variable $I(t)$. We summarize the complete picture of the system (5) based on the dynamics of these two subsystems in the following theorems.

Theorem 2.3 *Let $U(t) = (U_1(t), U_2(t))^T$ be the solution of the system (5) with initial value $u = (u_1, u_2)^T \in D^\circ$ in which $U_1(t)$ solves the system (6) with initial value $u_1 \in D_1^\circ := \{(s, i, m)^T : s > 0,\ i > 0,\ m > 0,\ s + i < 1\}$ and $U_2(t)$ solves the system (7) with initial value $u_2 \in D_2^\circ := \{(p, c)^T : p > 0,\ c > 0\}$. Assume that $\lambda > 0$ and $0 < b < \frac{\tau_3^2}{2}$. Then*

(i) *the system (6) has a unique invariant probability measure Π_1^* in D_1° whose support is*

$$supp(\Pi_1^*) = \left\{\left(1 - i - \frac{\alpha}{r} im,\ i,\ m\right)^T : im < \frac{r}{\alpha},\ i \in (0, 1),\ m > 0\right\}$$

and the solution $U_1(t)$ is exponentially ergodic with respect to Π_1^*;

(ii) *the system (7) has a unique invariant probability measure Π_2^* in D_2° whose support is*

$$supp(\Pi_2^*) = \left\{\left(p,\ \frac{\theta}{k} \frac{p^2}{1+p^2}\right)^T : p > 0\right\}$$

and the solution $U_2(t)$ is ergodic with respect to Π_2^*;

(iii) *$\mu^* = \Pi_1^* \times \Pi_2^*$ is the unique ergodic invariant probability measure of the system (5) with the support*

$$supp(\mu^*) = \left\{\left(1 - i - \frac{\alpha}{r} im,\ i,\ m,\ p,\ \frac{\theta}{k} \frac{p^2}{1+p^2}\right)^T : im < \frac{r}{\alpha},\ i \in (0, 1),\ m > 0,\ p > 0\right\}$$

and for any μ^*-integrable function h and $u \in D^\circ$

$$\mathbb{P}_u \left\{\lim_{t\to\infty} \frac{1}{t} \int_0^t h(U(s))ds = \int_D h(U)\mu^*(dU)\right\} = 1.$$

Theorem 2.4 *Let $U(t)$ be the solution of the system (5) with initial value $u \in D^\circ$. Assume that $\lambda < 0$ and $b < \frac{\tau_3^2}{2}$. Then $I(t)$ converges to 0 a.s., $S(t)$ converges to 1 a.s., $M(t)$ converges weakly to π_2, $P(t)$ converges to 0 a.s., and $C(t)$ converges to 0 a.s. Moreover,*

$$\lim_{t\to\infty} \frac{\ln I(t)}{t} = \lambda \ a.s.$$

As a consequence, we can derive the complete dynamic picture of the stochastic system (3) from Theorems 2.3 and 2.4 in the following theorem.

Theorem 2.5 *Suppose the initial values of the solution $(S(t), I(t), V(t), P(t), C(t))^T$ to the stochastic system (3) is in $\left\{(S, I, V, P, C)^T \in \mathbb{R}_+^{5,\circ} : S + I < 1\right\}$. Then, according to the threshold λ and parameter b, the solution will evolve over time as follows.*

(i) *If $\lambda < 0$ and $b < \frac{\tau_3^2}{2}$, then the solution $(S(t), I(t), V(t), P(t), C(t))^T$ converges to $\overline{\mu}_1 = \delta_1^* \times \delta_0^* \times \delta_0^* \times \delta_0^* \times \delta_0^*$ a.s.*

(ii) *If $\lambda > 0$ and $0 < b < \frac{\tau_3^2}{2}$, then the system (3) is stochastically persistent in the sense that the solution $(S(t), I(t), V(t), P(t), C(t))^T$ converges weakly to a unique invariant probability measure $\overline{\mu}^*$ supported by*

$$supp(\overline{\mu}^*) = \left\{ \left(1 - i - \frac{\alpha}{r} v,\ i,\ v,\ p,\ \frac{\theta}{k} \frac{p^2}{1+p^2}\right)^T : v < \frac{r}{\alpha},\ i \in (0,1),\ v > 0,\ p > 0 \right\}.$$

(iii) *If $b > \frac{\tau_3^2}{2}$, then $P(t)$ converges to ∞ and $C(t)$ converges to $\frac{\theta}{k}$ a.s.*

In order to interpret our results from the stochastic model and to understand how environmental noises and randomness affect the dynamical behaviors of the deterministic system (4), we need to find the relation between the basic reproduction number R_0 and the Lyapunov exponent λ. The following propositions give such relations, which are proved in Appendix.

Proposition 2.1 *The threshold $\lambda = \sqrt{\alpha n}\, R_\Theta(w) - a - \delta - \frac{\tau_1^2}{2}$ is a decreasing function of the progression rate δ and hence λ is an increasing function of the basic reproduction number R_0. The threshold λ as a function of noise intensities τ_1 and τ_2, the limit $\lim_{(\tau_1,\tau_2)\to(0,0)} \lambda$ exists, and denote $\overline{\lambda} := \lim_{(\tau_1,\tau_2)\to(0,0)} \lambda$. Then, $\overline{\lambda} = 0$ iff $R_0 = 1$, $\overline{\lambda} < 0$ iff $R_0 < 1$, and $\overline{\lambda} > 0$ iff $R_0 > 1$; in other words, $\overline{\lambda}$ is an increasing function of R_0.*

Proposition 2.2 *As the basic reproduction number $R_0 = \frac{\alpha n}{e(a+\delta)}$, the threshold λ is a decreasing function of the progression rate δ, death rate of infected cells a, and virus decay rate e, and λ is an increasing function of infection rate α and viral burst size n. When $a + \delta - e < \frac{\tau_2^2}{4} - \frac{3\tau_1^2}{4}$, λ is a decreasing function of noise intensities τ_1 and τ_2.*

Interpretation 2.1 (1) *From Proposition 2.1, when the intensities of two noises τ_1 and τ_2 approach zero, the threshold λ approaches $\overline{\lambda}$ which corresponds to R_0 while the sub-system of the first three Eq. (3) reduces to the sub-system of the first three Eq. (4).*

(2) *The locally asymptotically stable equilibrium point $(1, 0, 0, 0, 0)$ of the deterministic system (4) under condition $b < 0$ and $R_0 < 1$ corresponds to the invariant probability measure $\overline{\mu}_1 = \delta_1^* \times \delta_0^* \times \delta_0^* \times \delta_0^* \times \delta_0^*$ of the stochastic system (3). However, the condition under which all solutions approach the invariant ergodic distribution are $\lambda < 0$ and $b < \frac{\tau_3}{2}$. These conditions are more realistic since the net growth rate of precancerous cells is not necessarily zero.*

(3) *For the deterministic system (4), the positive equilibrium point $(S^*, I^*, V^*, P_{21}, C_{21})$ is locally asymptotically stable when $R_0 > 1$, $\alpha n > \frac{e^2+\sqrt{e^4+4e^3r}}{2}$, $\theta > \delta q I^*$, $R_0 < R^*$, and $0 < b < b_0$. This equilibrium point corresponds to the invariant ergodic distribution $\overline{\mu}^*$ where all solutions of the stochastic system (3) weakly approach when $\lambda < 0$ and $b < \frac{\tau_3^2}{2}$. There is no explicit restrict for virus infection αn in stochastic case. In addition, the point $(S^*, I^*, V^*, P_{21}, C_{21})$ belongs to the support of this invariant measure.*

(4) *There are other equilibrium points, particularly, one with higher P and C value which is unstable for the deterministic system (4),* (Asih et al. 2016)*. However, for the stochastic system (4), no matter which value λ takes, as long as $b > \frac{\tau_3^2}{2}$, the precancerous cell population will indefinitely grow, and cancer cell population will approach a fixed value a.s.*

(5) *As in the deterministic case, we numerically verify there is stochastic Hopf bifurcation in the dynamical sense for the stochastic system (3) when the progression rate δ passes some value.*

Interpretation 2.2 (1) *To establish a cervical cancer, there must be states of chronic infection first. In the deterministic model, the chronic infection state is represented by (S^*, I^*, V^*). Three conditions $R_0 > 1$, $\alpha n > \frac{e^2+\sqrt{e^4+4e^3r}}{2}$, and $R_0 < R^*$ guarantee that the chronic infection state exists and is stable. In the stochastic model, the chronic infection state is represented by random variables whose distribution are an ergodic invariant distribution with support $\left\{\left(1 - i - \frac{\alpha}{r}v,\ i,\ v\right)^T : 0 < v < \frac{r}{\alpha},\ i \in (0, 1)\right\}$, and the conditions $\lambda > 0$ guarantees its existence and stability. The noise intensities τ_1, τ_2, and τ_3 have great ranges to allow existence and stability of this ergodic distribution. More importantly, the threshold λ is a decreasing function of τ_1 and τ_2 under some condition in Proposition 2.2, which means that even small noises associated with the progression rate and viral infection can help to establish an ergodic invariant distribution of the chronic infection state. In other words, large noises of these two types may help to reduce chronic infection state.*

(2) *Since the threshold λ is a decreasing function of the progression rate δ, the small δ will help to establish the chronic infection state, and small noise also help chronic infection state formation. Medically, this may explain why it take decades to develop cervical cancer from chronic infection of HPV. On the other hand, the*

noise intensity τ_1 and τ_2 may not be easy to increase in vivo or reality. Therefore, the noise in the progression from infected cells to precancerous cells and the noise received or made by free viruses in vivo are difficult factors to treat in order to reduce the probability of cervical cancer development.

(3) *When the chronic infection state is established, there is a possibility to establish a cervical cancer. In the deterministic model, it is represented by an equilibrium point (P_{21}, C_{21}) under the conditions $\theta > \delta q I^*$ and $0 < b < b_0$. In the stochastic model, the cancer is represented by an ergodic invariant distribution with support $\left\{(p, \frac{\theta}{k}\frac{p^2}{1+p^2}),\ p > 0\right\}$ under condition $0 < b < \frac{\tau_3^2}{2}$. So, the noise associated with precancerous cell growth rate suppresses other conditions in deterministic model. If this noise is large enough, there is always cervical cancer; as it is even larger, precancerous cell population will grow indefinitely.*

3 Analysis of the model

This section is devoted to proving Theorems 2.1, 2.2, 2.3 and 2.4 in Sect. 2. We organize the section into four subsections. In the first subsection, we give a brief proof of Theorem 2.1 by using Routh–Hurwitz's criterion and a theorem of Hopf bifurcation in Hassard et al. (1981). In the second subsection, we give a detailed proof of Theorem 2.2. In the third subsection, we conduct boundary analysis for the system (5). The purpose of this analysis is to investigate the set of ergodic invariant probability measures of the system (5) when its solutions start in ∂D. Finally, the last subsection provides detailed proofs of Theorems 2.3 and 2.4 by using control theory, support theorem of diffusion processes, and some results of ergodicity in the theory of homogeneous Markov processes.

3.1 Proof of Theorem 2.1

(i) Consider the sub-system (S, I, V) of the system (4) and its positive equilibrium point $E^* = (S^*, I^*, V^*)$. The Jacobian matrix of the system (S, I, V) at E^* has the form

$$J_1 = \begin{bmatrix} r - 2rS^* - rI^* - \alpha V^* & -rS^* & -\alpha S^* \\ \alpha V^* & -a - \delta & \alpha S^* \\ 0 & n & -e \end{bmatrix},$$

and its corresponding characteristic polynomial is

$$p(\lambda) = \lambda^3 + a_1\lambda^2 + a_2\lambda + a_3$$

where $a_1 = rS^* + e + a + \delta$, $a_2 = rS^*(e + a + \delta + \alpha V^*)$, and $a_3 = \alpha S^* V^*(\alpha n + er)$. By Routh–Hurwitz's criterion, $p(\lambda)$ has three roots with negative real parts iff $a_1 a_2 - a_3 >$

0. This inequality is equivalent to

$$(e+a+\delta)^2+(rS^*+\alpha V^*)(e+a+\delta)+\alpha rS^*V^*-\frac{\alpha V^*}{r}(\alpha n+er)>0. \quad (8)$$

Since $S^*=\frac{1}{R_0}$, $V^*=\frac{rn}{re+\alpha n}\left(1-\frac{1}{R_0}\right)$, and $I^*=\frac{re}{re+\alpha n}\left(1-\frac{1}{R_0}\right)$, we obtain

$$\begin{aligned} rS^*+\alpha V^* &= r\frac{1}{R_0}+\frac{r\alpha n}{re+\alpha n}\left(1-\frac{1}{R_0}\right), \quad \alpha rS^*V^* \\ &= \frac{r^2\alpha n}{re+\alpha n}\frac{1}{R_0}\left(1-\frac{1}{R_0}\right), \\ -\frac{\alpha V^*}{r}(\alpha n+er) &= -\alpha n\left(1-\frac{1}{R_0}\right), \quad \text{and} \quad e+a+\delta=e+\frac{\alpha n}{e}\frac{1}{R_0}. \end{aligned}$$

Then (8) is equivalent to

$$\frac{1}{R_0^2}+\left(\frac{2\alpha ne^2+2\alpha nre+re^3}{\alpha^2n^2}+\frac{e^2}{er+\alpha n}\right)\frac{1}{R_0}+\frac{e^4}{\alpha^2n^2}-\frac{e^2}{er+\alpha n}>0. \quad (9)$$

Let $A:=\frac{2\alpha ne^2+2\alpha nre+re^3}{\alpha^2n^2}+\frac{e^2}{er+\alpha n}$ and $B:=\frac{e^4}{\alpha^2n^2}-\frac{e^2}{er+\alpha n}$. Since $\alpha n>\frac{e^2+\sqrt{e^4+4e^3r}}{2}$, using the quadratic formula, we have $B<0$. This implies that $R^*:=\frac{2}{\sqrt{A^2-B}-A}>0$. Then (9) is equivalent to $R_0<R^*$. In other words, all the eigenvalues of J_1 have negative real parts iff $R_0<R^*$.

To study the Hopf bifurcation that may occur from the equilibrium point E^* as R_0 changes and passes through the threshold R^*, we look into the roots of $p(\lambda)=0$ when its coefficients a_1, a_2, and a_3 are considered as functions of R_0. And

$$p(\lambda)=\lambda^3+a_1\lambda^2+a_2\lambda+a_3$$

where

$$\begin{aligned} a_1 &= a_1(R_0)=r\frac{1}{R_0}+e+\frac{\alpha n}{e}\frac{1}{R_0}, \\ a_2 &= a_2(R_0)=r\frac{1}{R_0}\left[e+\frac{\alpha n}{e}\frac{1}{R_0}+\frac{r\alpha n}{re+\alpha n}\left(1-\frac{1}{R_0}\right)\right], \\ a_3 &= a_3(R_0)=r\alpha n\frac{1}{R_0}\left(1-\frac{1}{R_0}\right). \end{aligned}$$

Now suppose $p(\lambda) = 0$ has a pair of complex roots $\lambda = u \pm iv$ and one real root $\lambda = \lambda_0$. Then

$$(\lambda^2 - 2u\lambda + u^2 + v^2)(\lambda - \lambda_0) = \lambda^3 + a_1\lambda^2 + a_2\lambda + a_3.$$

Expanding the left-hand side and then equating both sides, we get $a_1 = -(\lambda_0 + 2u)$, $a_2 = u^2+v^2+2u\lambda_0$, and $a_3 = -(u^2+v^2)\lambda_0$. It implies that $a_1a_2-a_3 = -2u(\lambda_0^2+a_2)$. So $p(\lambda) = 0$ has a pair of pure imaginary roots iff $u = 0$ and $v \neq 0$ iff $a_1a_2 - a_3 = 0$ and $a_2 > 0$. When $p(\lambda) = 0$ has pure imaginary roots, these imaginary roots are given by $\lambda = \pm i\sqrt{a_2}$ while the real root is given by $\lambda = -a_1$ with $a_1a_3 > 0$.

Next, suppose $\lambda(R_0) = \alpha(R_0) \pm i\beta(R_0)$ are two complex roots of the polynomial $p(\lambda)$. When $R_0 = R^*$, $a_1a_2 - a_3 = 0$ and $a_2 > 0$. This follows that $\alpha(R^*) = 0$ and $\beta(R^*) \neq 0$. Now we claim that $\frac{d\alpha}{dR_0}\big|_{R_0=R^*} \neq 0$. To show this, we use proof by contradiction. Suppose that $\alpha'(R^*) := \frac{d\alpha}{dR_0}\big|_{R_0=R^*} = 0$. Differentiating both sides of $p(\lambda) = 0$ with respect to R_0, evaluating at R^*, and applying $\alpha(R^*) = \alpha'(R^*) = 0$, and finally equating the real part and imaginary part to zero, we obtain

$$\begin{aligned} -a_1'\beta^2 - 2a_1\beta\beta' + a_3' &= 0, \\ -3\beta^2\beta' + a_2'\beta + a_2\beta' &= 0, \end{aligned}$$

in which $a_i' = \frac{da_i}{dR_0}\big|_{R_0=R^*}$ $(i = 1, 2, 3)$, $\beta' = \frac{d\beta}{dR_0}\big|_{R_0=R^*}$, and $\beta = \beta(R^*)$. Solving these equations for β', we get

$$\frac{a_2'\beta}{3\beta^2 - a_2} = \frac{a_3' - a_1'\beta^2}{2a_1\beta},$$

which infers that, since $\beta^2 = a_2 = \frac{a_3}{a_1}$,

$$a_2'a_3 = a_2(a_3' - a_1'a_2). \tag{10}$$

Let $H(R_0) = a_1(R_0)a_2(R_0) - a_3(R_0)$. Then, by above computation,

$$H(R_0) = \frac{1}{R_0^3} + A\frac{1}{R_0^2} + B\frac{1}{R_0},$$

which implies that

$$H'(R^*) = -\frac{1}{(R^*)^2}\left(\frac{2}{(R^*)^2} + \frac{A}{R^*}\right) < 0.$$

However, by (10), we also have

$$\begin{aligned} H'(R^*) &= a_1'a_2 + a_1a_2' - a_3' = a_1a_2' - (a_3' - a_1'a_2) = a_1a_2' - a_2'\frac{a_3}{a_2} \\ &= a_1a_2' - a_2'a_1 = 0 \end{aligned}$$

which is a contradiction. Therefore $\frac{d\alpha}{dR_0}\big|_{R_0=R^*} \neq 0$.

In summary, at E^* the characteristic polynomial $p(\lambda)$ of the Jacobian matrix J_1 has one negative real root $\lambda = \lambda_0$ and a pair of complex roots $\lambda = \alpha(R_0) \pm i\beta(R_0)$. We assume that $R_0 > 1$ and $\alpha n > \frac{e^2+\sqrt{e^4+4e^3 r}}{2}$. With the threshold R^*, we have shown that

- when $R_0 < R^*$, $\alpha(R_0) < 0$;
- when $R_0 = R^*$, $\alpha(R^*) = 0$, $\beta(R^*) \neq 0$, and $\alpha'(R^*) \neq 0$;
- when $R_0 > R^*$, $\alpha(R_0) > 0$.

Hence the real part $\alpha(R_0)$ of the complex roots of $p(\lambda)$ changes sign when R_0 passes through R^*. Therefore the system (S, I, V) of the system (4) undergoes Hopf bifurcation at $R_0 = R^*$. This completes the proof of part (i).

(ii) When the system (S, I, V) has a unique positive equilibrium $E^* = (S^*, I^*, V^*)$, positive equilibria $(\overline{P}, \overline{C} := \frac{\theta}{k}\frac{\overline{P}^2}{1+\overline{P}^2})$ of the system (P, C) depends on the solutions to the equation $f(\overline{P}) = g(\overline{P})$ where $f(\overline{P}) = \delta q I^* + b\overline{P}$ and $g(\overline{P}) = \theta\frac{\overline{P}^2}{1+\overline{P}^2}$. Since $0 < b < \frac{\theta}{2}$ and $\theta > \delta q I^*$, there are 3 cases:

- The equation $f(\overline{P}) = g(\overline{P})$ has a unique positive solution P_0. Then P_0 can be found by solving the system $f(\overline{P}) = g(\overline{P})$ and $f'(\overline{P}) = g'(\overline{P})$. This system leads to the following equation

$$(\theta - \delta q I^*)\overline{P}^4 - (\theta + 2\delta q I^*)\overline{P}^2 - \delta q I^* = 0.$$

  It is easy to see that this equation has unique positive solution

$$\overline{P} = P_0 := \left(\frac{\theta + 2\delta q I^*}{2\theta - 2\delta q I^*} + \left(\left(\frac{\theta + 2\delta q I^*}{2\theta - 2\delta q I^*}\right)^2 + \frac{\delta q I^*}{\theta - \delta q I^*}\right)^{1/2}\right)^{1/2}.$$

  So, when $b = b_0 := \frac{2\theta P_0}{(1+P_0^2)^2}$, the equation $f(\overline{P}) = g(\overline{P})$ has unique positive solution P_0. Clearly $P_0 > 1$, which implies that $b_0 < \frac{\theta}{2}$.
- When $b_0 < b < \frac{\theta}{2}$, the straight line $f(\overline{P})$ lies above the curve $g(\overline{P})$ and so there is no positive solution.
- When $0 < b < b_0$, there are exactly 2 intersections between $f(\overline{P})$ and $g(\overline{P})$. Hence $f(\overline{P}) = g(\overline{P})$ has exactly 2 positive solutions $P_{21} < P_0 < P_{22}$.

At $(\overline{P}, \overline{C})$ the Jacobian matrix of the (P, C) system is of the form

$$J_4 = \begin{bmatrix} b - \frac{2\theta\overline{P}}{(1+\overline{P}^2)^2} & 0 \\ \frac{2\theta\overline{P}}{(1+\overline{P}^2)^2} & -k \end{bmatrix}$$

and its eigenvalues are real and of the form $\lambda_1 = b - \frac{2\theta\overline{P}}{(1+\overline{P}^2)^2}$ and $\lambda_2 = -k$. When $0 < b < b_0$ and $\overline{P} = P_{21}, \lambda_1 = b - \frac{2\theta P_{21}}{(1+P_{21}^2)^2} = f'(P_{21}) - g'(P_{21}) < 0$ which follows

that (P_{21}, C_{21}) is locally asymptotically stable. When $0 < b < b_0$ and $\overline{P} = P_{22}$, $\lambda_1 = b - \frac{2\theta P_{22}}{(1+P_{22}^2)^2} = f'(P_{22}) - g'(P_{22}) > 0$ and so (P_{22}, C_{22}) is unstable. This completes the proof of part (ii).
(iii) This part can be easily derived from part (i) and part (ii).

3.2 Proof of Theorem 2.2

Since $f(U)$ and $g(U)$ are locally Lipschitz continuous on $\mathbb{R}^5$, there is a unique local a.s. continuous solution $U(t)$ of the system (5) up to the explosion time

$$\tau_e := \inf\{t > 0 : \min\{S(t), I(t), M(t), P(t), C(t)\} = -\infty \\ \max\{S(t), I(t), M(t), P(t), C(t)\} = \infty\}.$$

and, furthermore, $U(t)$ is a strong Markov process that possesses the Feller property. By the first equation and the second equation of the system (5), for all $t \in (0, \tau_e)$

$$S(t) = s \exp\left\{\int_0^t [r(1 - S(s) - I(s)) - \alpha I(s)M(s)]ds\right\} \ a.s. \text{ and}$$
$$I(t) = i \exp\left\{\int_0^t \left[\alpha S(s)M(s) - a - \delta - \frac{\tau_1^2}{2}\right]ds - \tau_1 W_1(t)\right\} \ a.s.$$

If $s = 0$ then $S(t) = 0$ for all $t \in (0, \tau_e)$ a.s. and if $s > 0$ then $S(t) > 0$ for all $t \in (0, \tau_e)$ a.s. If $i = 0$ then $I(t) = 0$ for all $t \in (0, \tau_e)$ a.s. and if $i > 0$ then $I(t) > 0$ for all $t \in (0, \tau_e)$ a.s. The third equation of the system (5) implies for all $t \in (0, \tau_e)$

$$M(t) = \phi(t)\left[m + \int_0^t n\phi^{-1}(s)ds\right] \ a.s.$$

where

$$\phi(t) = \exp\left\{\int_0^t \left[-\alpha S(s)M(s) + a + \delta + \frac{\tau_1^2}{2} - \frac{\tau_2^2}{2} - e\right]ds + \tau_1 W_1(t) + \tau_2 W_2(t)\right\}.$$

If $m \geq 0$ then $M(t) > 0$ for all $t \in (0, \tau_e)$ a.s. From the fourth equation of the system (5), for all $t \in (0, \tau_e)$

$$P(t) = \psi(t)\left[p + \int_0^t \delta q I(s)\psi^{-1}(s)ds\right] \ a.s.$$

where

$$\psi(t) = \exp\left\{\int_0^t \left[b - \theta\frac{P(s)}{1 + P(s)^2} - \frac{\tau_3^2}{2}\right]ds + \tau_3 W_3(t)\right\}.$$

If $i = 0$ and $p = 0$ then $I(t) = 0$ for all $t \in (0, \tau_e)$ a.s. which follows that $P(t) = 0$ for all $t \in (0, \tau_e)$ a.s. If $(i \geq 0$ and $p > 0) \vee (i > 0$ and $p = 0)$ then $P(t) > 0$ for all $t \in (0, \tau_e)$ a.s. The last equation of the system (5) follows, for all $t \in (0, \tau_e)$,

$$C(t) = \exp\{-kt\}\left[c + \int_0^t \frac{\theta P(s)}{1 + P(s)^2} \exp\{ks\}ds\right] \ a.s.$$

If $i = 0$, $p = 0$, and $c = 0$ then $P(t) \equiv 0$ a.s. which implies that $C(t) = 0$ for all $t \in (0, \tau_e)$ a.s. If $(i \geq 0,\ p > 0,\ c \geq 0) \vee (i > 0,\ p = 0,\ c \geq 0)$ then $C(t) > 0$ for all $t \in (0, \tau_e)$ a.s. Thus, we have shown that if $u = (s, i, m, p, c)^T \in \mathbb{R}^5_+$ then $U(t) \in \mathbb{R}^5_+$ a.s.

Next, we show that if $s > 0$, $i > 0$, $m \geq 0$, $p \geq 0$, $c \geq 0$, and $s + i < 1$ then $S(t) + I(t) < 1$ for all $t \in (0, \tau_e)$ a.s. (Note that if $s + i = 1$ then by the first equation of the system (5) $\frac{dS}{dt}(0) = -\alpha sim \leq 0$ which means that, after some time $t > 0$, S(t) decreases a.s. Hence without loss of generality we can assume that $s + i < 1$ in the first place.) We consider the function

$$V_1(s, i, m, p, c) = 2 - s - i - \ln(1 - s - i).$$

By Ito's formula,

$$\begin{aligned}
\mathcal{L}V_1 &= -rs(1 - s - i) + \alpha sim + \frac{rs(1 - s - i)}{1 - s - i} - \frac{\alpha sim}{1 - s - i} - \alpha sim \\
&\quad + (a + \delta)i + \frac{\alpha sim}{1 - s - i} - (a + \delta)\frac{i}{1 - s - i} + \frac{\tau_1^2}{2}\left(\frac{i}{1 - s - i}\right)^2 \\
&= rs(s + i) + (a + \delta)i - (a + \delta)\frac{i}{1 - s - i} + \frac{\tau_1^2}{2}\left(\frac{i}{1 - s - i}\right)^2 \\
&= rs(s + i) - (a + \delta)\frac{(s + i)i}{1 - s - i} + \frac{\tau_1^2}{2}\left(\frac{i}{1 - s - i}\right)^2.
\end{aligned}$$

Since $s + i < 1$, there is a $h > 0$ large enough so that $s + \left(1 + \frac{1}{h}\right)i < 1$ which implies that $\frac{i}{1 - s - i} < h$. Hence for $s > 0$, $i > 0$, and $s + i < 1$

$$\mathcal{L}V_1(s, i, m, p, c) \leq r(s + i) + \frac{h^2\tau_1^2}{2} \leq K_1 V_1(s, i, m, p, c)$$

for some suitable positive constant K_1. Let

$$\xi_k = \inf\{t \in [0, \tau_e) : \ V_1(U(t)) \geq k\}, \ k \in \mathbb{N}.$$

Fix $t \in (0, \tau_e)$, applying Ito's formula for $V_1(U(\xi_k \wedge t))$ gives

$$\begin{aligned}\mathbb{E}_u V_1(U(\xi_k \wedge t)) &= V_1(u) + \mathbb{E}_u \int_0^{\xi_k \wedge t} \mathcal{L} V_1(U(s))ds \\ &= V_1(u) + \int_0^t \mathbb{E}_u \mathcal{L} V_1(U(\xi_k \wedge s))ds.\end{aligned}$$

Since $\xi_k \wedge s < \xi_k$ for all $s \in (0, \xi_k \wedge t)$, $V_1(U(\xi_k \wedge s)) < k < \infty$ which implies that $\ln(1 - S(\xi_k \wedge s) - I(\xi_k \wedge s)) < \infty$ and so $S(\xi_k \wedge s) + I(\xi_k \wedge s) < 1$ for all $s \in (0, \xi_k \wedge t)$. By the argument above, for all $s \in (0, \xi_k \wedge t)$

$$\mathcal{L} V_1(U(\xi_k \wedge s)) \leq K_1 V_1(U(\xi_k \wedge s)).$$

So

$$\mathbb{E}_u V_1(U(\xi_k \wedge t)) \leq V_1(u) + K_1 \int_0^t \mathbb{E}_u V_1(U(\xi_k \wedge s))ds.$$

By Gronwall's inequality (see Theorem 8.1, p. 45 in Mao 1997)

$$\mathbb{E}_u V_1(U(\xi_k \wedge t)) \leq V_1(u) \exp\{K_1 t\}.$$

But

$$\mathbb{E}_u V_1(U(\xi_k \wedge t)) = \int_\Omega V_1(U(\xi_k \wedge t))d\mathbb{P}_u \geq \int_{\{\xi_k \leq t\}} V_1(U(\xi_k))d\mathbb{P}_u \geq k\mathbb{P}_u\{\xi_k \leq t\},$$

which follows that for all $k \geq 1$

$$\mathbb{P}_u\{\xi_k > t\} = 1 - \mathbb{P}_u\{\xi_k \leq t\} \geq 1 - \frac{V_1(u)e^{K_1 t}}{k}.$$

On the other hand, $\xi_k > t$ implies $V_1(U(s)) < k$ for all $s \in [0, t]$. Thus, for all $k \geq 1$

$$\mathbb{P}_u\{V_1(U(s)) < k \ \forall s \in [0, t]\} \geq 1 - \frac{V_1(u)e^{K_1 t}}{k}.$$

Letting $k \to \infty$ yields

$$\mathbb{P}_u\{V_1(U(s)) < \infty \ \forall s \in [0, t]\} = 1.$$

As $V_1(U(s)) < \infty$ implies $S(s) + I(s) < 1$ for all $s \in [0, t]$, so

$$\mathbb{P}_u\{S(s) + I(s) < 1 \ \forall s \in [0, t]\} = 1.$$

Since $t \in (0, \tau_e)$ is arbitrary, $S(t) + I(t) < 1$ for all $t \in (0, \tau_e)$ a.s.

Finally, we prove that $\tau_e = \infty$ a.s. Consider the function

$$V_2(S, I, M, P, C) = S + I + \ln(1 + M) + \ln(1 + P) + C.$$

Then, by Ito's formula, for all $t \in (0, \tau_e)$

$$\begin{aligned}\mathcal{L}V_2(U(t)) &= rS(t)[1 - S(t) - I(t)] - \alpha S(t)I(t)M(t) + \alpha S(t)I(t)M(t)\\ &\quad - (a + \delta)I(t) + \frac{n}{1 + M(t)} - \alpha\frac{S(t)M(t)^2}{1 + M(t)} + (a + \delta + \tau_1^2 - e)\frac{M(t)}{1 + M(t)}\\ &\quad - \left(\frac{\tau_1^2}{2} + \frac{\tau_2^2}{2}\right)\left(\frac{M(t)}{1 + M(t)}\right)^2 + \delta q\frac{I(t)}{1 + P(t)} + b\frac{P(t)}{1 + P(t)}\\ &\quad - \frac{\theta P(t)^2}{(1 + P(t))(1 + P(t)^2)} - \frac{\tau_3^2}{2}\left(\frac{P(t)}{1 + P(t)}\right)^2 + \theta\frac{P(t)^2}{1 + P(t)^2}\\ &\quad - kC(t) \le K_2\end{aligned}$$

for some suitable positive constant K_2. Let

$$\tau_k = \inf\{t \in (0, \tau_e) : M(t) > k \text{ or } P(t) > k \text{ or } C(t) > k\}.$$

Then τ_k increases to τ_∞ as $k \to \infty$ where

$$\tau_\infty = \inf\{t \in (0, \tau_e) : M(t) = \infty \text{ or } P(t) = \infty \text{ or } C(t) = \infty\}.$$

Since $\tau_\infty \le \tau_e$ a.s., it suffices to prove that $\tau_\infty = \infty$ a.s. Fix $t > 0$, applying Ito's formula for $V_2(U(\tau_k \wedge t))$ gives

$$\begin{aligned}\mathbb{E}_u V_2(U(\tau_k \wedge t)) &= V_2(u) + \mathbb{E}_u \int_0^{\tau_k \wedge t} \mathcal{L}V_2(U(s))ds\\ &\le K_3 + K_2\,\mathbb{E}_u(\tau_k \wedge t) \le K_3 + K_2 t,\end{aligned}$$

where $K_3 = V_2(u)$. On the other hand,

$$\mathbb{E}_u V_2(U(\tau_k \wedge t)) \ge \int_{\{\tau_k \le t\}} V_2(U(\tau_k))d\mathbb{P}_u \ge (k \wedge \ln(1 + k))\mathbb{P}_u\{\tau_k < t\}.$$

Thus

$$\mathbb{P}_u\{\tau_k < t\} \le \frac{K_3 + K_2 t}{k \wedge \ln(1 + k)} \to 0 \text{ as } k \to \infty.$$

Since $t > 0$ is arbitrary, $\mathbb{P}_u\{\tau_\infty < \infty\} = 0$. Therefore $\tau_\infty = \infty$ a.s.

3.3 Boundary analysis

A. Suppose that the solution $U(t)$ of the system (5) starts in $\{S = 0\}$, that is, $S(0) = 0$. Then, by the first equation of the system (5), $S(t) \equiv 0$ a.s. So the second equation of the system (5) becomes

$$dI = (-a - \delta)I dt - \tau_1 I dW_1 \tag{11}$$

which implies that

$$I(t) = I(0)\ \exp\left\{\left(-a - \delta - \frac{\tau_1^2}{2}\right)t - \tau_1 W_1(t)\right\}.$$

Hence $I(t)$ converges to 0 a.s. for any initial value $I(0) \in [0, 1]$. Furthermore, $S(t) \equiv 0$ implies, by the third equation of the system (5), that

$$dM = [n + (a + \delta + \tau_1^2 - e)M]dt + \tau_1 M dW_1 + \tau_2 M dW_2 \tag{12}$$

To study the long-term behavior of the solution of the Eq. (12), we consider for $\alpha_1 > 0$ fixed

$$\begin{aligned} s(M) &= \int_{\alpha_1}^{M} \exp\left\{-\int_{\alpha_1}^{y} \frac{2n + 2(a + \delta + \tau_1^2 - e)z}{(\tau_1^2 + \tau_2^2)z^2} dz\right\} dy \\ &= C_1 \int_{\alpha_1}^{M} y^{\frac{2(e-a-\delta-\tau_1^2)}{\tau_1^2+\tau_2^2}} \exp\left\{\frac{2n}{\tau_1^2 + \tau_2^2}\frac{1}{y}\right\} dy. \end{aligned}$$

Since the integrand can be written as

$$y^{\frac{2(e-a-\delta-\tau_1^2)}{\tau_1^2+\tau_2^2}}\left[1 + \frac{2n}{\tau_1^2 + \tau_2^2}\frac{1}{y} + \frac{1}{2!}\frac{4n^2}{(\tau_1^2 + \tau_2^2)^2}\frac{1}{y^2} + \cdots\right],$$

there exists a positive natural number k so that $\dfrac{2(e - a - \delta - \tau_1^2)}{\tau_1^2 + \tau_2^2} - k < -1$. It implies that $s(0+) := \lim\limits_{M \to 0^+} s(M) = -\infty$. On the other hand, the behavior of $s(\infty) = \lim\limits_{M \to \infty} s(M)$ depends upon the sign of $\dfrac{2(e - a - \delta - \tau_1^2)}{\tau_1^2 + \tau_2^2} + 1$. If $\dfrac{2(e - a - \delta - \tau_1^2)}{\tau_1^2 + \tau_2^2} + 1 < 0$ then $s(\infty) < \infty$. By item 2 of Theorem 3.1, p. 447 in Ikeda and Watanabe (1989), $M(t)$ converges to ∞ a.s. If $\dfrac{2(e - a - \delta - \tau_1^2)}{\tau_1^2 + \tau_2^2} + 1 < 0$, which is equivalent to $2(e - a - \delta) > \tau_1^2 - \tau_2^2$, then $s(\infty) = \infty$. By item 1 of Theorem 3.1, p. 447 in Ikeda and Watanabe (1989), there exists a unique ergodic invariant probability measure π_1

for the Eq. (12) whose density $p_1(m)$ is the solution to the associated Fokker–Planck equation of (12)

$$-\frac{d}{dm}[(n+(a+\delta+\tau_1^2-e)m)p_1(m)]+\frac{d^2}{dm^2}\left[\frac{1}{2}(\tau_1^2+\tau_2^2)m^2p_1(m)\right]=0.$$

Solving this equation (see Sect. 5.2 in Phan et al. 2021) yields $p_1(m) = \frac{\beta_*^{\alpha_*}}{\Gamma(\alpha_*)}m^{-\alpha_*-1}e^{-\beta_*/m}$ in which $\alpha_* := \frac{2(e-a-\delta-\tau_1^2)}{\tau_1^2+\tau_2^2}+1$, $\beta_* := \frac{2n}{\tau_1^2+\tau_2^2}$, and $\Gamma(\cdot)$ is the Gamma function. Since any measure whose density is of the form of $p_1(m)$ is the inverse gamma distribution with parameters α_* and β_*, $\pi_1 \sim \mathrm{IG}(\alpha_*, \beta_*)$. Hence $M(t)$ converges weakly to π_1 for any initial value $M(0) \geq 0$. Next, as $I(t)$ converges to 0 a.s., the long-term behavior of the fourth equation of the system (5) is the same as that of the equation

$$d\widetilde{P}=\left(b\widetilde{P}-\theta\frac{\widetilde{P}^2}{1+\widetilde{P}^2}\right)dt+\tau_3\widetilde{P}dW_3 \tag{13}$$

with $\widetilde{P}(0) = P(0) \geq 0$. The Eq. (13) can rewritten as

$$\widetilde{P}(t)=P(0)\exp\left\{\int_0^t\left[b-\theta\frac{\widetilde{P}(s)}{1+\widetilde{P}(s)^2}-\frac{\tau_3^2}{2}\right]ds+\tau_3W_3(t)\right\}. \tag{14}$$

If $P(0) = 0$ then $\widetilde{P}(t) \equiv 0$ a.s. By the last equation of the system (5), $C(t)$ converges to 0 a.s. Now assume that $P(0) > 0$. Then, by (14), $\widetilde{P}(t) > 0$ for all $t > 0$ a.s. and the long-term behavior of $\widetilde{P}(t)$ depends upon the value of b. Consider the following cases.

Case 1. If $b \leq 0$ then, by (14), $\widetilde{P}(t)$ converges to 0 a.s. Hence the long-term behavior of the last equation of the system (5) is the same as that of the equation $\frac{d\widetilde{C}}{dt} = -k\widetilde{C}$, which follows that $\widetilde{C}(t)$ converges to 0 a.s.

Case 2. If $b > \frac{\theta}{2}+\frac{\tau_3^2}{2}$ then, since $\frac{\widetilde{P}(s)}{1+\widetilde{P}(s)^2} \leq \frac{1}{2}$ for all $s \geq 0$,

$$b-\theta\frac{\widetilde{P}(s)}{1+\widetilde{P}(s)^2}-\frac{\tau_3^2}{2}\geq b-\frac{\theta}{2}-\frac{\tau_3^2}{2}>0.$$

This shows that $\widetilde{P}(t)$ converges to ∞ a.s. Hence the last equation of the system (5) implies $C(t)$ converges to θ a.s.

Case 3. Assume that $0 < b \leq \frac{\theta}{2}+\frac{\tau_3^2}{2}$. It is clear that δ_0^* is an ergodic invariant probability measure for the Eq. (13). Suppose that $\widetilde{P}(t)$ is close to a neighborhood of 0 for a long time. Then, by strong law of large numbers, the Lyapunov exponent of δ_0^*

can be computed as

$$\lambda(\delta_0^*) := \lim_{t\to\infty}\frac{1}{t}\ln\widetilde{P}(t) = \lim_{t\to\infty}\frac{1}{t}\int_0^t\left[b-\theta\frac{\widetilde{P}(s)}{1+\widetilde{P}(s)^2}-\frac{\tau_3^2}{2}\right]ds$$
$$= \int_{\{0\}}\left[b-\theta\frac{\widetilde{P}}{1+\widetilde{P}^2}-\frac{\tau_3^2}{2}\right]\delta_0^*(d\widetilde{P}) = b-\frac{\tau_3^2}{2}.$$

We consider two cases:

Case 3.1. If $0 < b < \dfrac{\tau_3^2}{2}$ then, since

$$b-\theta\frac{\widetilde{P}(s)}{1+\widetilde{P}(s)^2}-\frac{\tau_3^2}{2} \le b-\frac{\tau_3^2}{2} < 0,$$

it follows from the Eq. (14) that $\widetilde{P}(t)$ converges to 0 a.s. and hence $C(t)$ converges to 0 a.s.

Case 3.2. If $b > \dfrac{\tau_3^2}{2}$ then $\lambda(\delta_0^*) > 0$ which means that $\widetilde{P}(t)$ will be repelled away from 0 if it gets close to 0. To study the long-term behavior of the solution $\widetilde{P}(t)$ of the Eq. (13), we consider for $\alpha_2 > 0$ fixed

$$s(\widetilde{P}) = \int_{\alpha_2}^{\widetilde{P}}\exp\left\{-\int_{\alpha_2}^{y}\frac{2bz-\frac{2\theta z^2}{1+z^2}}{\tau_3^2z^2}\right\}dy = C_2\int_{\alpha_2}^{\widetilde{P}} y^{-\frac{2b}{\tau_3^2}}\exp\left\{\frac{2\theta}{\tau_3^2}\tan^{-1}y\right\}dy.$$

Since $\lim\limits_{y\to0^+}\exp\left\{\dfrac{2\theta}{\tau_3^2}\tan^{-1}y\right\} = 1$, for any $\epsilon > 0$ there is a $\eta > 0$ so that $0 < y < \eta$, we have $1-\epsilon < \exp\left\{\dfrac{2\theta}{\tau_3^2}\tan^{-1}y\right\} < 1+\epsilon$. Thus, for $0 < \widetilde{P} < \eta$,

$$(1-\epsilon)C_2\int_{\alpha_2}^{\widetilde{P}} y^{-\frac{2b}{\tau_3^2}}dy \le s(\widetilde{P}) \le (1+\epsilon)C_2\int_{\alpha_2}^{\widetilde{P}} y^{-\frac{2b}{\tau_3^2}}dy.$$

Because $-\dfrac{2b}{\tau_3^2}+1 < 0$, we have

$$\int_{\alpha_2}^{\widetilde{P}} y^{-\frac{2b}{\tau_3^2}}dy = \frac{\widetilde{P}^{-\frac{2b}{\tau_3^2}+1}-\alpha_2^{-\frac{2b}{\tau_3^2}+1}}{-\frac{2b}{\tau_3^2}+1} \to -\infty \text{ as } \widetilde{P}\to0^+$$

which implies that $s(0+) = -\infty$. As $\lim_{y\to\infty} \exp\left\{\frac{2\theta}{\tau_3^2}\tan^{-1} y\right\} = \exp\left\{\frac{\theta\pi}{\tau_3^2}\right\}$, there exists a $M_1 > \alpha_2$ so that $y > M_1$ implies $\exp\left\{\frac{2\theta}{\tau_3^2}\tan^{-1} y\right\} < \exp\left\{\frac{\theta\pi}{\tau_3^2}\right\} + \epsilon$. Then, for $\widetilde{P} > M_1$, we get

$$\int_{M_1}^{\widetilde{P}} y^{-\frac{2b}{\tau_3^2}} \exp\left\{\frac{2\theta}{\tau_3^2}\tan^{-1} y\right\} dy \le \left(\exp\left\{\frac{\theta\pi}{\tau_3^2}\right\} + \epsilon\right) \int_{M_1}^{\widetilde{P}} y^{-\frac{2b}{\tau_3^2}} dy$$
$$= \left(\exp\left\{\frac{\theta\pi}{\tau_3^2}\right\} + \epsilon\right) \frac{\widetilde{P}^{-\frac{2b}{\tau_3^2}+1} - M_1^{-\frac{2b}{\tau_3^2}+1}}{-\frac{2b}{\tau_3^2}+1}.$$

It follows that for $\widetilde{P} > M_1$

$$s(\widetilde{P}) = C_2 \int_{\alpha_2}^{M_1} y^{-\frac{2b}{\tau_3^2}} \exp\left\{\frac{2\theta}{\tau_3^2}\tan^{-1} y\right\} dy + C_2 \int_{M_1}^{\widetilde{P}} y^{-\frac{2b}{\tau_3^2}} \exp\left\{\frac{2\theta}{\tau_3^2}\tan^{-1} y\right\} dy$$
$$\le M_2 + C_2 \left(\exp\left\{\frac{\theta\pi}{\tau_3^2}\right\} + \epsilon\right) \frac{\widetilde{P}^{-\frac{2b}{\tau_3^2}+1} - M_1^{-\frac{2b}{\tau_3^2}+1}}{-\frac{2b}{\tau_3^2}+1}$$

where $M_2 := C_2 \int_{\alpha_2}^{M_1} y^{-\frac{2b}{\tau_3^2}} \exp\left\{\frac{2\theta}{\tau_3^2}\tan^{-1} y\right\} dy$. Thus

$$s(\infty) \le M_2 + \frac{C_2(\exp\left\{\frac{\theta\pi}{\tau_3^2}\right\} + \epsilon) M_1^{-\frac{2b}{\tau_3^2}+1}}{\frac{2b}{\tau_3^2} - 1} < \infty.$$

By item 2 of Theorem 3.1, p. 447 in Ikeda and Watanabe (1989), $\widetilde{P}(t)$ converges to ∞ a.s. and so $C(t)$ converges to θ/k a.s.

Therefore we have proved that if the solution $U(t)$ of the system (5) starts in $\{S = 0\}$ and if $2(e - a - \delta) > \tau_1^2 - \tau_2^2$ and $b < \frac{\tau_3^2}{2}$ then $S(t) \equiv 0$ a.s., $I(t)$ converges to 0 a.s., $M(t)$ converges weakly to π_1, $P(t)$ converges to 0 a.s., and $C(t)$ converges to 0 a.s. In other words, $\mu_0 = \delta_0^* \times \delta_0^* \times \pi_1 \times \delta_0^* \times \delta_0^*$ is the unique ergodic invariant probability measure for the system (5) on the boundary $\{S = 0\} \subseteq \partial D$ provided that $2(e - a - \delta) > \tau_1^2 - \tau_2^2$ and $b < \frac{\tau_3^2}{2}$.

B. Suppose that the solution $U(t)$ of the system (5) starts in $\{S > 0,\ I = 0\}$, that is, $S(0) > 0$ and $I(0) = 0$. Since $I(0) = 0$, the second equation of the system (5) implies

$I(t) \equiv 0$ a.s. But then the first equation of the system (5) becomes $dS = rS(1-S)dt$, which follows that $S(t) = \dfrac{Ce^{rt}}{1+Ce^{rt}}$ where $C = \dfrac{S(0)}{1-S(0)} > 0$. Hence $S(t)$ converges to 1 a.s. Then the long-term behavior of the third equation of the system (5) is the same as that of the equation

$$d\widetilde{M} = [n - \alpha\widetilde{M}^2 + (a + \delta + \tau_1^2 - e)\widetilde{M}]dt + \tau_1\widetilde{M}dW_1 + \tau_2\widetilde{M}dW_2 \tag{15}$$

To study the long-term behavior of the solution $\widetilde{M}$ of the Eq. (15), we consider, for a fixed $\alpha_3 > 0$,

$$\begin{aligned} s(\widetilde{M}) &= \int_{\alpha_3}^{\widetilde{M}} \exp\left\{-\int_{\alpha_3}^{y} \frac{2n - 2\alpha z^2 + 2(a+\delta+\tau_1^2-e)z}{(\tau_1^2+\tau_2^2)z^2}dz\right\}dy \\ &= C_3 \int_{\alpha_3}^{\widetilde{M}} y^{-\frac{2(a+\delta+\tau_1^2-e)}{\tau_1^2+\tau_2^2}} \exp\left\{\frac{2n}{\tau_1^2+\tau_2^2}\frac{1}{y} + \frac{2\alpha}{\tau_1^2+\tau_2^2}y\right\}. \end{aligned}$$

Since integrand can be rewritten as

$$\begin{aligned} y^{-\frac{2(a+\delta+\tau_1^2-e)}{\tau_1^2+\tau_2^2}} &\left[1 + \left(\frac{2n}{\tau_1^2+\tau_2^2}\frac{1}{y} + \frac{2\alpha}{\tau_1^2+\tau_2^2}y\right)\right. \\ &\left. + \frac{1}{2!}\left(\frac{2n}{\tau_1^2+\tau_2^2}\frac{1}{y} + \frac{2\alpha}{\tau_1^2+\tau_2^2}y\right)^2 + \cdots\right], \end{aligned}$$

there exists a positive natural number k_1 such that $-\dfrac{2(a+\delta+\tau_1^2-e)}{\tau_1^2+\tau_2^2} - k_1 < -1$ and so $s(0+) = -\infty$. As there is a positive natural number k_2 such that $-\dfrac{2(a+\delta+\tau_1^2-e)}{\tau_1^2+\tau_2^2} + k_2 > -1$, so $s(\infty) = \infty$. By item 1 of Theorem 3.1, p. 447 in Ikeda and Watanabe (1989), there is a unique ergodic invariant probability measure π_2 for the system (5) whose density $p_2(m)$ is the solution to the associated Fokker–Planck equation of (15)

$$-\frac{d}{dm}[(n - \alpha m^2 + (a+\delta+\tau_1^2-e)m)p_2(m)] + \frac{d^2}{dm^2}\left[\frac{1}{2}(\tau_1^2+\tau_2^2)m^2p_2(m)\right] = 0.$$

Solving this equation (see Sect. 3.2 in Phan and Tian 2020) yields

$$p_2(m) = \frac{(\alpha/n)^{\Theta/2}}{2K_\Theta\left(\frac{4\sqrt{\alpha n}}{\tau_1^2+\tau_2^2}\right)} m^{\Theta-1}\exp\left\{-\frac{1}{2}(\chi m^{-1} + \psi m)\right\}, \quad m \in (0,\infty),$$

where $\Theta := \dfrac{2(a+\delta+\tau_1^2-e)}{\tau_1^2+\tau_2^2} - 1$, $\chi := \dfrac{4n}{\tau_1^2+\tau_2^2}$, $\psi := \dfrac{4\alpha}{\tau_1^2+\tau_2^2}$, and K_Θ is the modified Bessel function of the third kind with index Θ. Any measure whose density is of the form of $p_2(m)$ is the generalized inverse Gaussian distribution with parameters Θ, χ, and ψ. Hence $\pi_2 \sim \text{GIG}(\Theta, \chi, \psi)$ and thus $M(t)$ converges weakly to π_2. Furthermore, the first moment of π_2 can be computed as

$$\int_0^\infty m\pi_2(dm) = R_\Theta(w)\sqrt{\frac{n}{\alpha}} =: m^* \tag{16}$$

where $R_\Theta(w) = \dfrac{K_{\Theta+1}(w)}{K_\Theta(w)}$ and $w = \dfrac{4\sqrt{\alpha n}}{\tau_1^2+\tau_2^2}$. Finally, since $I(t) \equiv 0$ a.s., the same argument as in part **A** implies that $P(t)$ approaches 0 a.s. and so $C(t)$ converges to 0 a.s. provided $b < \dfrac{\tau_3^2}{2}$. Therefore we have proved that if the solution $U(t)$ of the system (5) starts in $\{S > 0,\ I = 0\}$ and if $b < \dfrac{\tau_3^2}{2}$ then $S(t)$ converges to 1 a.s., $I(t) \equiv 0$ a.s., $M(t)$ converges weakly to π_2, $P(t)$ converges to 0 a.s., and $C(t)$ converges to 0 a.s. In other words, $\mu_1 = \delta_1^* \times \delta_0^* \times \pi_2 \times \delta_0^* \times \delta_0^*$ is the unique ergodic invariant probability measure for the system (5) on the boundary $\{S > 0,\ I = 0\} \subseteq \partial D$ provided that $b < \dfrac{\tau_3^2}{2}$.

C. Now assume that the solution $U(t)$ of the system (5) starts in $\{S > 0,\ I > 0\}$, that is, $S(0) > 0$ and $I(0) > 0$. By Theorem 2.2, $U(t) \in D^\circ$ for all $t > 0$ a.s. From the third equation of the system (5), $P(t) \geq \widetilde{P}(t)$ a.s. where $\widetilde{P}(t)$ is the solution to the Eq. (13) with $P(0) = \widetilde{P}(0) \geq 0$. By the arguments in part **A**, if $b > \dfrac{\tau_3^2}{2}$ then $\widetilde{P}(t) \to \infty$ a.s. and so $P(t) \to \infty$ a.s. So from now on we assume that $b < \dfrac{\tau_3^2}{2}$. Then the dynamics of the system (5) depends on the threshold

$$\lambda := \sqrt{\alpha n}\, R_\Theta(w) - a - \delta - \frac{\tau_1^2}{2}.$$

To understand why the combined parameter λ determines the long-term behavior of the system (5), we investigate the set of ergodic invariant probability measures of the system (5) on the boundary ∂D, denoted by $\mathcal{M}$, and then compute their Lyapunov exponents. By arguments in part **A** and part **B**, since $b < \dfrac{\tau_3^2}{2}$, $\mathcal{M} = \{\mu_0, \mu_1\}$ if $2(e-a-\delta) > \tau_1^2 - \tau_2^2$ and $\mathcal{M} = \{\mu_1\}$ if $2(e-a-\delta) \geq \tau_1^2 - \tau_2^2$.

For $\mu_0 = \delta_0^* \times \delta_0^* \times \pi_1 \times \delta_0^* \times \delta_0^*$, the Lyapunov exponents of μ_0 along the I, M, P, and C component are equal to zero, that is, $\lambda_2(\mu_0) = \lambda_3(\mu_0) = \lambda_4(\mu_0) = \lambda_5(\mu_0) = 0$. Intuitively, this is because the components I, M, P, and C are inside the support of μ_0 which is

$$\text{supp}(\mu_0) = \{S = 0\} = \{(S, I, M, P, C)^T \in D :\ S = 0\} \subseteq \partial D$$

and so they are at an "equilibrium state". Hence the solution does not decay or grow along these components. Since the Lyapunov exponent of μ_0 along the S component is

$$\lambda_1(\mu_0) = \lim_{t\to\infty} \frac{\ln S(t)}{t} = \lim_{t\to\infty} \frac{1}{t}\int_0^t [r(1 - S(s) - I(s)) - \alpha I(s)M(s)]ds$$
$$= \int_{\{S=0\}} [r(1 - S - I) - \alpha IM]d\mu_0 = r > 0,$$

μ_0 is always a repeller.

For $\mu_1 = \delta_1^* \times \delta_0^* \times \pi_2 \times \delta_0^* \times \delta_0^*$, we have $\lambda_1(\mu_1) = \lambda_3(\mu_1) = \lambda_4(\mu_1) = \lambda_5(\mu_1) = 0$ since the components S, M, P, and C are inside the support of μ_1 which is $\{S > 0,\ I = 0\}$. The Lyapunov exponent of μ_1 along the I component is

$$\lambda_2(\mu_1) = \lim_{t\to\infty} \frac{\ln I(t)}{t} = \lim_{t\to\infty} \frac{1}{t}\int_0^t \left[\alpha S(s)M(s) - a - \delta - \frac{\tau_1^2}{2}\right] ds$$
$$= \int_{\{S>0,\ I=0\}} \left[\alpha SM - a - \delta - \frac{\tau_1^2}{2}\right] d\mu_1 = \lambda.$$

The combined parameter λ provides us with the rate of convergence of the I component when $I(t)$ is getting close to 0 for a long time. Theorem 2.4 guarantees that if $\lambda < 0$ then μ_1 is a global attractor in the sense that, for any initial value in D°, $I(t)$ converges to 0 a.s., $S(t)$ converges to 1 a.s., $M(t)$ converges weakly to π_2, $P(t)$ converges to 0 a.s., and $C(t)$ converges to 0 a.s.

If $\lambda > 0$ then μ_1 becomes a repeller. When the I component is getting close to 0, it will be repelled away from 0 and then wanders around between 0 and 1. This makes sure that the solution will be positive recurrent and finally end up at an equilibrium state, which is characterized by an ergodic invariant probability measure μ^* in D°. This fact is ensured by Theorem 2.3.

Before ending this subsection, we state and prove the following proposition that will be utilized in proof of Theorems 2.3 and 2.4.

Proposition 3.1 *Let $U(t) = (S(t),\ I(t),\ M(t),\ P(t),\ C(t))^T$ be the solution to the system (5) with initial value $u \in D^\circ$. If $I(t)$ converges to 0 a.s. then $S(t)$ converges to 1 a.s., $M(t)$ converges weakly to π_2, $P(t)$ converges to 0 a.s., and $C(t)$ converges to 0 a.s.*

Proof of Proposition 3.1 Since $I(t)$ converges to 0 a.s., for any $\epsilon > 0$ there exists a non-random time $T_1 > 0$ so that $t \geq T_1$ implies $0 < I(t) < \epsilon$ a.s. Since $n - \alpha SM^2 + (a + \delta + \tau_1^2 - e)M \to -\infty$ as $M \to \infty$, there is an $A > 0$ large enough so that $M > A$ implies $\mathcal{L}(M) = n - \alpha SM^2 + (a + \delta + \tau_1^2 - e)M < -M$. Then, by Theorem 3.1 in Meyn and Tweedie (1993b), $\mathbb{P}_u\{M(t) \to \infty\} = 0$. In other words, $\limsup_{t\to\infty} M(t) < \infty$ a.s. Then there is an $\bar{A} > 0$ so that $M(t) \leq \bar{A}$ for all $t \geq 0$ a.s. By the first equation of the system (5), it is easy to see that $\lim_{t\to\infty} S(t)$ exists

and is finite a.s. Furthermore, for $t \geq T_1$ we have

$$rS(1-S) \geq \frac{dS}{dt} \geq rS\left[1 - S - \epsilon(1+\alpha\bar{A})\right] \ a.s. \tag{17}$$

Let $S_\epsilon(t)$ satisfy

$$\frac{dS_\epsilon}{dt} = rS_\epsilon[1 - \epsilon(1+\alpha\bar{A}) - S_\epsilon]$$

with initial value $S_\epsilon(0) = S(0) > 0$. Clearly, $S_\epsilon(t)$ converges to $1 - \epsilon(1+\alpha\bar{A})$ a.s. Notice that ϵ is chosen so that $\epsilon < \frac{1}{1+\alpha\bar{A}}$. By comparison theorem for ODEs, (17) implies, for all $t \geq T_1$, that $S_\epsilon(t) \leq S(t) \leq 1$ a.s. Letting $t \to \infty$ yields $1 - \epsilon(1+\alpha\bar{A}) \leq \lim_{t\to\infty} S(t) \leq 1$ a.s. So letting $\epsilon \downarrow 0$ gives $\lim_{t\to\infty} S(t) = 1$ a.s. Then, for any $\epsilon > 0$, there is a non-random time $T_2 > 0$ so that $t \geq T_2$ implies $1 - \epsilon < S(t) < 1$ a.s. By the third equation of the system (5), we obtain for $t \geq T_2$

$$\widetilde{M}(t) \leq M(t) \leq M_\epsilon(t) \ a.s. \tag{18}$$

where $\widetilde{M}(t)$ solves the Eq. (15) with $\widetilde{M}(0) = M(0) > 0$ and $M_\epsilon(t)$ solves the equation

$$\begin{aligned} dM_\epsilon = &[n - \alpha(1-\epsilon)M_\epsilon^2 + (a + \delta + \tau_1^2 - e)M_\epsilon]dt \\ &+\tau_1 M_\epsilon dW_1 + \tau_2 M_\epsilon dW_2 \end{aligned} \tag{19}$$

with $M_\epsilon(0) = M(0) > 0$. By the same argument as in **B**, we know that $M_\epsilon(t)$ converges weakly to $\pi_{2\epsilon} \sim \text{GIG}(\Theta, \chi, \psi_\epsilon)$ where Θ and χ are defined as in **B** and $\psi_\epsilon = \frac{4\alpha(1-\epsilon)}{\tau_1^2 + \tau_2^2}$. Let $\widetilde{P}(t, M(0), \cdot)$, $P(t, M(0), \cdot)$, and $P_\epsilon(t, M(0), \cdot)$ be the transition probability functions of $\widetilde{M}(t)$, $M(t)$, and $M_\epsilon(t)$, respectively. Due to (18), for any $\bar{m} > 0$,

$$P_\epsilon(t, M(0), (0, \bar{m})) \leq P(t, M(0), (0, \bar{m})) \leq \widetilde{P}(t, M(0), (0, \bar{m})).$$

Since an open set in $\mathbb{R}_+^\circ = (0, \infty)$ is the union of at most countably many disjoint open intervals in R_+° and $\mathcal{B}(\mathbb{R}_+^\circ)$ is generated by the collection of open sets in $\mathbb{R}_+^\circ$, it implies that for all $B \in \mathcal{B}(\mathbb{R}_+^\circ)$

$$P_\epsilon(t, M(0), B) \leq P(t, M(0), B) \leq \widetilde{P}(t, M(0), B). \tag{20}$$

If we denote by $p_\epsilon(m)$ the density of $\pi_{2\epsilon}$ then it is straightforward that $\lim_{\epsilon\downarrow 0} p_\epsilon(m) = p(\epsilon)$ due to continuity of exponential functions. It follows that for any $B \in \mathcal{B}(\mathbb{R}_+^\circ)$

$$\lim_{\epsilon\downarrow 0} \pi_{2\epsilon}(B) = \lim_{\epsilon\downarrow 0} \int_B p_\epsilon(m)dm = \int_B p(m)dm = \pi_2(B)$$

and hence as $\epsilon \downarrow 0$

$$\|\pi_{2\epsilon}(\cdot) - \pi_2(\cdot)\|_{TV} = 2 \sup_{B \in \mathcal{B}(\mathbb{R}_+^\circ)} |\pi_{2\epsilon}(B) - \pi_2(B)| \to 0.$$

Since $\|P_\epsilon(t, M(0), \cdot) - \pi_{2\epsilon}(\cdot)\|_{TV} \to 0$ and $\|\widetilde{P}(t, M(0), \cdot) - \pi_2(\cdot)\|_{TV} \to 0$ as $t \to \infty$,

$$\lim_{\epsilon \downarrow 0} \lim_{t \to \infty} \|P_\epsilon(t, M(0), \cdot) - \widetilde{P}(t, M(0), \cdot)\|_{TV} = 0.$$

Thus, by (20), $\|P(t, M(0), \cdot) - \pi_2(\cdot)\|_{TV} \to 0$ as $t \to \infty$. In other words, $M(t)$ converges weakly to π_2. Next, the fourth equation of the system (5) implies for $t \geq T_1$

$$\tilde{P}(t) \leq P(t) \leq P_\epsilon(t) \ a.s. \tag{21}$$

where $\tilde{P}(t)$ solves the Eq. (13) with $\tilde{P}(0) = P(0) > 0$ and $P_\epsilon(t)$ is the solution to the equation

$$dP_\epsilon = (\delta q \epsilon + bP_\epsilon)dt + \tau_3 P_\epsilon dW_3 \tag{22}$$

with $P_\epsilon(0) = P(0) > 0$. Fix $\alpha_4 > 0$ and consider

$$s(P_\epsilon) = \int_{\alpha_4}^{P_\epsilon} \exp\left\{-\int_{\alpha_4}^{y} \frac{2\delta q\epsilon + 2bz}{\tau_3^2 z^2} dz\right\} dy = C_4 \int_{\alpha_4}^{P_\epsilon} y^{-\frac{2b}{\tau_3^2}} \exp\left\{\frac{2\delta q\epsilon}{\tau_3^2 y}\right\} dy.$$

The integrand can be rewritten as

$$y^{-\frac{2b}{\tau_3^2}} \left[1 + \frac{2\delta q\epsilon}{\tau_3^2}\frac{1}{y} + \frac{1}{2!}\left(\frac{2\delta q\epsilon}{\tau_3^2}\right)^2 \frac{1}{y^2} + \cdots\right].$$

As there is a positive natural number k so that $-\frac{2b}{\tau_3^2} - k < -1$, so $s(0+) = -\infty$. Since $b < \frac{\tau_3^2}{2}, -\frac{2b}{\tau_3^2} + 1 > 0$ which implies that $s(\infty) = \infty$. Thus, by item 1 of Theorem 3.1, p. 447 in Ikeda and Watanabe (1989), there exists a unique ergodic invariant probability measure π_3^ϵ for the Eq. (22) which is the inverse gamma distribution $\pi_3^\epsilon \sim \text{IG}(\beta_1, \beta_2^\epsilon)$ where $\beta_1 = -\frac{2b}{\tau_3^2} + 1$ and $\beta_2^\epsilon = \frac{2\delta q\epsilon}{\tau_3^2}$. Then $P_\epsilon(t)$ converges weakly to π_3^ϵ. But, since $\beta_2^\epsilon \to 0$ as $\epsilon \to 0$, the density of π_3^ϵ

$$f_\epsilon(p) = \frac{(\beta_2^\epsilon)^{\beta_1}}{\Gamma(\beta_1)} p^{-\beta_1-1} \exp\left\{-\frac{\beta_2^\epsilon}{p}\right\} \to 0 \text{ as } \epsilon \to 0.$$

It implies that π_3^ϵ converges weakly to 0 as $\epsilon \to 0$. Hence $\lim_{\epsilon\downarrow 0} \lim_{t\to\infty} P_\epsilon(t) = 0$ a.s. Thus, by (21), $P(t)$ converges to 0 a.s. Finally, by the last equation of the system (5),

$$C(t) = C(0)e^{-kt} + e^{-kt} \int_0^t \theta \frac{P(s)^2}{1+P(s)^2} e^{ks} ds.$$

It follows from L'Hospital's Rule that

$$\lim_{t\to\infty} C(t) = \lim_{t\to\infty} \frac{\theta}{k} \frac{P(t)^2}{1+P(t)^2} = 0.$$

□

3.4 Proofs of Theorems 2.3 and 2.4

In order to prove Theorems 2.3 and 2.4, first of all we need to check whether the system (5) has sufficient noises that can locally push its dynamics in all directions or not. This can be achieved by proving the solutions of the system (5) starting in D° satisfies Hörmander's condition (Bellet 2006; Nualart 2006; Hörmander 1967). Indeed, we rewrite the system (5) in the Stratonovich form

$$\begin{aligned}
dS &= [rS(1-S-I) - \alpha SIM]\,dt,\\
dI &= \left(\alpha SM - a - \delta - \frac{\tau_1^2}{2}\right) I dt - \tau_1 I \circ dW_1,\\
dM &= \left(n - \alpha SM^2 + \left(a + \delta + \frac{\tau_1^2}{2} - \frac{\tau_2^2}{2} - e\right) M\right) dt + \tau_1 M \circ dW_1\\
&\quad + \tau_2 M \circ dW_2,\\
dP &= \left(\delta q I + \left(b - \frac{\tau_3^2}{2}\right) P - \theta \frac{P^2}{1+P^2}\right) dt + \tau_3 P \circ dW_3,\\
dC &= \left(\theta \frac{P^2}{1+P^2} - kC\right) dt.
\end{aligned} \tag{23}$$

Let

$$\bar{f}(U) = \begin{bmatrix} rS(1-S-I) - \alpha SIM \\ \left(\alpha SM - a - \delta - \dfrac{\tau_1^2}{2}\right) I \\ n - \alpha SM^2 + \left(a + \delta + \dfrac{\tau_1^2}{2} - \dfrac{\tau_2^2}{2} - e\right) M \\ \delta q I + \left(b - \dfrac{\tau_3^2}{2}\right) P - \theta \dfrac{P^2}{1+P^2} \\ \theta \dfrac{P^2}{1+P^2} - kC \end{bmatrix},$$

$g_1(U) = [0, -\tau_1 I, \tau_1 M, 0, 0]^T$, $g_2(U) = [0, 0, \tau_2 M, 0, 0]^T$, and $g_3(U) = [0, 0, 0, \tau_3 P, 0]^T$. By definition, the solution $U(t)$ of the system (5) is said to satisfy Hörmander's condition if the set of vector fields

$$g_1, g_2, g_3, [\bar{f}, g_1], [\bar{f}, g_2], [\bar{f}, g_3], [\bar{f}, [\bar{f}, g_1]], \ldots$$

spans $\mathbb{R}^5$ at every point $u = (s, i, m, p, c)^T \in D^\circ$ where $[A, B]$ is the Lie Bracket of two vector fields $A = (A_1, A_2, A_3, A_4, A_5)^T$ and $B = (B_1, B_2, B_3, B_4, B_5)^T$ defined by

$$[A, B] = ([A, B]_1 [A, B]_2, [A, B]_3, [A, B]_4, [A, B]_5)^T \text{ and for } j = 1, 2, 3, 4, 5$$
$$[A, B]_j := \left(A_1 \frac{\partial B_j}{\partial s} - B_1 \frac{\partial A_j}{\partial s}\right) + \left(A_2 \frac{\partial B_j}{\partial i} - B_2 \frac{\partial A_j}{\partial i}\right) + \left(A_3 \frac{\partial B_j}{\partial m} - B_3 \frac{\partial A_j}{\partial m}\right) + \left(A_4 \frac{\partial B_j}{\partial p} - B_4 \frac{\partial A_j}{\partial p}\right) + \left(A_5 \frac{\partial B_j}{\partial s} - B_5 \frac{\partial A_j}{\partial c}\right).$$

By computation,

$$[\bar{f}, g_1] = \begin{bmatrix} -\tau_1 rsi - 2\tau_1 \alpha ims \\ \tau_1 \alpha ims \\ -\tau_1 n - \tau_1 \alpha m^2 s \\ -\tau_1 i \delta q \\ 0 \end{bmatrix}, \quad [\bar{f}, g_2] = \begin{bmatrix} \tau_2 \alpha ims \\ -\tau_2 \alpha ims \\ \tau_2 n + \tau_2 \alpha m^2 s \\ 0 \\ 0 \end{bmatrix}, \quad \text{and}$$

$$[\bar{f}, g_3] = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \tau_3 \delta qi + \tau_3 \theta \frac{p^2(1-p^2)}{(1+p^2)^2} \\ -2\tau_3 \theta \frac{p^2}{(1+p^2)^2} \end{bmatrix}.$$

Clearly, the vectors g_1, g_2, g_3, $[\bar{f}, g_2]$, and $[\bar{f}, g_3]$ span $\mathbb{R}^5$ for any $u \in D^\circ$. So Hörmander's condition holds for the solutions of the system (5) in D°. As a consequence, the transition probability function $P(t, u, \cdot)$ of the solutions $U(t)$ has density $p(t, u, \bar{u})$ which is smooth in $(u, \bar{u}) \in D^\circ \times D^\circ$.

Next, we consider the control system corresponding to the system (23)

$$\begin{aligned}
\dot{S}_\phi &= rS_\phi\left(1 - S_\phi - I_\phi\right) - \alpha S_\phi I_\phi M_\phi,\\
\dot{I}_\phi &= \left(\alpha S_\phi M_\phi - a - \delta - \frac{\tau_1^2}{2}\right) I_\phi - \tau_1 I_\phi \phi_1,\\
\dot{M}_\phi &= n - \alpha S_\phi M_\phi^2 + \left(a + \delta + \frac{\tau_1^2}{2} - \frac{\tau_2^2}{2} - e\right) M_\phi + \tau_1 M_\phi \phi_1 + \tau_2 M_\phi \phi_2,\\
\dot{P}_\phi &= \delta q I_\phi + \left(b - \frac{\tau_3^2}{2}\right) P_\phi - \theta \frac{P_\phi^2}{1 + P_\phi^2} + \tau_3 P_\phi \phi_3,\\
\dot{C}_\phi &= \theta \frac{P_\phi^2}{1 + P_\phi^2} - kC_\phi.
\end{aligned} \tag{24}$$

where $\phi = (\phi_1, \phi_2, \phi_3)^T$ is from the set of piecewise continuous vector functions taking values in $\mathbb{R}^3$ and defined on $\mathbb{R}_+$. Let $(S_\phi(t, u), I_\phi(t, u), M_\phi(t, u), P_\phi(t, u), C_\phi(t, u))^T$ be the solution to the system (24) with control ϕ and initial value $u = (s, i, m, p, c)^T \in D^\circ$. We define a reachable set of $u \in D^\circ$ as the collection of

$$(S_\phi(t, u), I_\phi(t, u), M_\phi(t, u), P_\phi(t, u), C_\phi(t, u))^T$$

under all piecewise continuous controls $\phi(\cdot)$ (time t is fixed). By investigating the reachable sets of different initial values in D° and using support theorem, we can obtain the desired properties of invariant probability measures of the system (5). For convenience, let

$$\begin{aligned}
f_1(U_\phi) &:= rS_\phi\left(1 - S_\phi - I_\phi\right) - \alpha S_\phi I_\phi M_\phi,\\
f_2(U_\phi) &:= \left(\alpha S_\phi M_\phi - a - \delta - \frac{\tau_1^2}{2}\right) I_\phi,\\
f_3(U_\phi) &:= n - \alpha S_\phi M_\phi^2 + \left(a + \delta + \frac{\tau_1^2}{2} - \frac{\tau_2^2}{2} - e\right) M_\phi,\\
f_4(U_\phi) &:= \delta q I_\phi + \left(b - \frac{\tau_3^2}{2}\right) P_\phi - \theta \frac{P_\phi^2}{1 + P_\phi^2},\\
f_5(U_\phi) &:= \theta \frac{P_\phi^2}{1 + P_\phi^2} - kC_\phi,
\end{aligned}$$

where we denote $U_\phi = (S_\phi,\ I_\phi,\ M_\phi,\ P_\phi,\ C_\phi)^T$. Then the system (24) is equivalent to

$$\begin{aligned}
\dot{S}_\phi &= f_1(U_\phi),\\
\dot{I}_\phi &= f_2(U_\phi) - \tau_1 I_\phi \phi_1,\\
\dot{M}_\phi &= f_3(U_\phi) + \tau_1 M_\phi \phi_1 + \tau_2 M_\phi \phi_2,\\
\dot{P}_\phi &= f_4(U_\phi) + \tau_3 P_\phi \phi_3,\\
\dot{C}_\phi &= f_5(U_\phi).
\end{aligned}$$

The dynamics of the system (24) is presented in the following claims.

Claim 3.1 *Let* $u_0 := (s_0,\ i_0,\ m_0,\ p_0,\ c_0)^T \in D^\circ$ *and* $(i_1,\ m_1,\ p_1)^T \in (0,1) \times (0,\infty) \times (0,\infty)$. *Then for any* $\epsilon > 0$, *there are a control* $\phi(\cdot)$ *and a time* $T > 0$ *such that*

$$\begin{aligned}
&\left|S_\phi(T,u_0) - s_0\right| < \epsilon,\\
&I_\phi(T,u_0) = i_1,\\
&M_\phi(T,u_0) = m_1,\\
&P_\phi(T,u_0) = p_1,\ \textit{and}\\
&\left|C_\phi(T,u_0) - c_0\right| < \epsilon.
\end{aligned}$$

Remark This claim indicates that we can control the solution of the system (24) to move back and forth along the I-direction, M-direction, and P-direction; while the other directions still remain within a small neighborhood of their initial values.

Proof of Claim 3.1 Suppose that $i_0 < i_1, m_0 < m_1$, and $p_0 < p_1$. (The other cases are treated similarly.) Let

$$\begin{aligned}
\rho := \sup_{j\in\{1,2,3,4,5\}} \Big\{ &|f_j(u)| :\ |s - s_0|\\
&\le \epsilon,\ |c - c_0| \le \epsilon,\ (i,m,p)^T \in [i_0,i_1] \times [m_0,m_1] \times [p_0,p_1] \Big\}
\end{aligned}$$

where $u = (s,\ i,\ m,\ p,\ c)^T$. We choose $\phi_1(t) \equiv -\rho_1$ with $\rho_1 > 0$ such that

$$0 < i_1 - i_0 < \epsilon \left(\frac{\tau_1 \rho_1 i_0}{\rho} - 1 \right).$$

Then $\tau_1\rho_1 i_0 - \rho > 0$ and so $\dot{I}_\phi(0,u_0) = f_2(u_0) + \tau_1\rho_1 i_0 \ge -\rho + \tau_1\rho_1 i_0 > 0$. By continuity of $\dot{I}_\phi$, there exists a $t_1 > 0$ so that $\dot{I}_\phi(t,u_0) > 0$ for all $t \in (0,t_1)$. In other words, I_ϕ is increasing on $(0,t_1)$. Next, we choose $\phi_2(t) \equiv \rho_2 > 0$ such that

$$0 < m_1 - m_0 < \epsilon \left(\frac{\tau_2\rho_2 m_0 - \tau_1\rho_1 m_0}{\rho} - 1 \right).$$

Then $\dot{M}_\phi(0,u_0) = f_3(u) + (\tau_2\rho_2 - \tau_1\rho_1)m_0 \geq \tau_2\rho_2 m_0 - \tau_1\rho_1 m_0 - \rho > 0$. It implies that there is a $t_2 > 0$ such that M_ϕ is increasing on $(0, t_2)$. Finally, we choose $\phi_3(t) \equiv \rho_3 > 0$ so that

$$0 < p_1 - p_0 < \epsilon\left(\frac{\tau_3\rho_3 p_0}{\rho} - 1\right).$$

Then $\dot{P}_\phi(0,u_0) = f_4(u_0) + \tau_3\rho_3 p_0 \geq \tau_3\rho_3 p_0 - \rho > 0$ which follows that there exists a $t_3 > 0$ such that P_ϕ is increasing on $(0, t_3)$. Now let $t_0 = \min\{t_1,\ t_2,\ t_3\}$ and choose $\epsilon < \rho t_0$. Suppose that there were the first time $t \in (0, \frac{\epsilon}{\rho})$ so that $|S_\phi(t,u_0) - s_0| > \epsilon$. Then, by Mean Value Theorem, we would have

$$\begin{aligned}\epsilon < \left|S_\phi(t,u_0) - s_0\right| &= \left|\dot{S}_\phi(\eta,u_0)\right| t, \quad \text{for some} \quad \eta \in (0,t)\\ &= \left|f_1(U_\phi(\eta,u_0))\right| t \leq \rho\,\frac{\epsilon}{\rho} = \epsilon,\end{aligned}$$

which is a contradiction. Hence for all $t \in (0, \frac{\epsilon}{\rho})$ we get $\left|S_\phi(t,u_0) - s_0\right| \leq \epsilon$. By the same argument, $\left|C_\phi(t,u_0) - s_0\right| \leq \epsilon$ for all $t \in (0, \frac{\epsilon}{\rho})$. Next, if for all $t \in (0, \frac{\epsilon}{\rho})$ we had $I_\phi(t,u_0) < i_1$ then it would imply that $I_\phi(\frac{\epsilon}{\rho}, u_0) = \lim_{t\to\frac{\epsilon}{\rho}} I(t,u_0) \leq i_1$. But then, by Mean Value Theorem,

$$\begin{aligned}\frac{\epsilon}{\rho}(\tau_1\rho_1 i_0 - \rho) > i_1 - i_0 &\geq I_\phi\left(\frac{\epsilon}{\rho}, u_0\right) - I_\phi(0,u_0)\\ &= \dot{I}_\phi(\bar{\eta},u_0)\frac{\epsilon}{\rho} \quad \text{for some} \quad \bar{\eta} \in (0, \frac{\epsilon}{\rho})\\ &\geq (\tau_1\rho_1 i_0 - \rho)\frac{\epsilon}{\rho},\end{aligned}$$

which is a contradiction. Thus, there is a time $T_1 \in (0, \frac{\epsilon}{\rho})$ such that $I_\phi(T_1,u_0) = i_1$. By the same reasonings, there are a time $T_2 \in (0, \frac{\epsilon}{\rho})$ such that $M_\phi(T_2,u_0) = m_1$ and a time $T_3 > 0 \in (0, \frac{\epsilon}{\rho})$ such that $P_\phi(T_3,u_0) = p_1$. If $T_1 = T_2 = T_3$ then take $\bar{\phi} \equiv (-\rho_1, \rho_2, \rho_3)$ and we're done. If T_1, T_2, and T_3 are different, for example, $T_1 < T_2 < T_3$ then we can choose

$$\phi_1(t) = \begin{cases} -\rho_1 & \text{if} \quad t \in [0,T_1],\\ \frac{1}{\tau_1}\left(\alpha\bar{S}(t)\bar{M}(t) - a - \delta - \frac{\tau_1^2}{2}\right) & \text{if} \quad t \in (T_1,T_2],\\ \frac{1}{\tau_1}\left(\alpha\underline{S}(t)\underline{M}(t) - a - \delta - \frac{\tau_1^2}{2}\right) & \text{if} \quad t > T_2,\end{cases}$$

$$\phi_2(t) = \begin{cases} \rho_2 & \text{if} \quad t \in [0,T_1],\\ -\frac{1}{\tau_2}\left(\alpha\bar{S}(t)\bar{M}(t) - a - \delta - \frac{\tau_1^2}{2}\right) + \rho_2 - \frac{\tau_1}{\tau_2}\rho_1 & \text{if} \quad t \in (T_1,T_2],\\ -\frac{1}{\tau_2}\underline{M}(t)^{-1} f_3(\underline{U}(t)) - \frac{\tau_1}{\tau_2}\phi_1(t) & \text{if} \quad t > T_2,\end{cases}$$

and $\phi_3(t) \equiv \rho_3$ for all $t \geq 0$, where $\bar{U}(t) = (\bar{S}(t),\ \bar{I}(t),\ \bar{M}(t),\ \bar{P}(t),\ \bar{C}(t))^T$ solves the system

$$\begin{aligned}
\dot{S} &= f_1(U),\\
\dot{I} &= 0,\\
\dot{M} &= f_3(U) + (\tau_2\rho_2 - \tau_1\rho_1)M,\\
\dot{P} &= f_4(U) + \tau_3\rho_3 P,\\
\dot{C} &= f_5(U),
\end{aligned}$$

with initial value $\left(S_{\bar{\phi}}(T_1),\ i_1,\ M_{\bar{\phi}}(T_1),\ P_{\bar{\phi}}(T_1),\ C_{\bar{\phi}}(T_1)\right)^T$ and $\underline{U}(t) = (\underline{S}(t),\ \underline{I}(t),\ \underline{M}(t),\ \underline{P}(t),\ \underline{C}(t))^T$ solves the system

$$\begin{aligned}
\dot{S} &= f_1(U),\\
\dot{I} &= 0,\\
\dot{M} &= 0,\\
\dot{P} &= f_4(U) + \tau_3\rho_3 P,\\
\dot{C} &= f_5(U),
\end{aligned}$$

with initial value $\left(\bar{S}(T_2),\ i_1,\ m_1,\ \bar{P}(T_2),\ \bar{C}(T_2)\right)^T$. □

Claim 3.2 *Suppose* $u_0 = (s_0,\ i_0,\ m_0,\ p_0,\ c_0)^T \in D^\circ$ *such that* $i_0 m_0 < \dfrac{r}{\alpha}$. *Let*

$$S_* := 1 - i_0 - \frac{\alpha}{r} i_0 m_0 \quad \text{and} \quad C_* := \frac{\theta}{k}\frac{p_0^2}{1+p_0^2}.$$

Then for any $\epsilon > 0$ *there are a control* $\phi(\cdot)$ *and a time* $T > 0$ *so that*

$$\begin{aligned}
&\left|S_\phi(T, u_0) - S_*\right| < \epsilon,\\
&I_\phi(t, u_0) = i_0,\ \forall\, t \in [0, T],\\
&M_\phi(t, u_0) = m_0,\ \forall\, t \in [0, T],\\
&P_\phi(t, u_0) = p_0,\ \forall\, t \in [0, T],\ \textit{and}\\
&\left|C_\phi(T, u_0) - C_*\right| < \epsilon.
\end{aligned}$$

Remark Claim 3.2 states that if we hold the I-direction, M-direction, and P-direction of the solution to the system (24) then the S-direction and C-direction will end up within a small neighborhood of a fixed point in finite time. This helps us to describe exactly the support of the invariant probability measure μ^* of the system (5) in D°.

Proof of Claim 3.2 Consider the ODE system

$$\dot{S} = rS(1 - S - I) - \alpha SIM,$$

$$\begin{aligned}\dot{I} &= 0,\\ \dot{M} &= 0,\\ \dot{P} &= 0,\\ \dot{C} &= \theta\frac{P^2}{1+P^2} - kC,\end{aligned} \tag{25}$$

with initial value $u_0 \in D^\circ$ where $i_0 m_0 < \frac{r}{\alpha}$. The second, the third, and the fourth equations imply $I(t) \equiv i_0$, $M(t) \equiv m_0$, and $P(t) \equiv p_0$. So the system (25) is reduced to 2-dim ODE system

$$\begin{aligned}\dot{S} &= rS(1-S-i_0) - \alpha S i_0 m_0,\\ \dot{C} &= \theta\frac{p_0^2}{1+p_0^2} - kC,\end{aligned}$$

with initial condition $(s_0, c_0)^T$. It is straightforward that $S(t)$ converges to S_* and $C(t)$ converges to C_* as $t \to \infty$. Let $\bar{U}(t) = (\bar{S}(t),\ \bar{I}(t),\ \bar{M}(t),\ \bar{P}(t),\ \bar{C}(t))^T$ be the solution to the system (25) with initial value u_0. With the feedback control $\phi = (\phi_1,\ \phi_2,\ \phi_3)^T$ satisfying

$$\begin{aligned}\phi_1(t) &\equiv \frac{1}{\tau_1}\left(\alpha\bar{S}(t)\bar{M}(t) - a - \delta - \frac{\tau_1^2}{2}\right),\\ \phi_2(t) &\equiv -\frac{1}{\tau_2}\bar{M}(t)^{-1} f_3(\bar{U}(t)) - \frac{\tau_1}{\tau_2}\phi_1(t), \text{ and}\\ \phi_3(t) &\equiv \frac{\theta}{\tau_3}\frac{\bar{P}(t)}{1+\bar{P}(t)^2} - \frac{\delta q}{\tau_3}\frac{\bar{I}(t)}{\bar{P}(t)} - \frac{b}{\tau_3} + \frac{\tau_3}{2},\end{aligned}$$

we have $U_\phi(t) = \bar{U}(t)$ for all $t \geq 0$ where $U_\phi(t)$ is the solution of the system (24) with control $\phi(\cdot)$ above and initial value $u_0 \in D^\circ$ with $i_0 m_0 < \frac{r}{\alpha}$. This completes the proof. □

By Claims 3.1 and 3.2, we can easily derive the following claims that will be used in proofs of Theorems 2.3 and 2.4.

Claim 3.3 *For any* $u = (s, i, m, p, c)^T \in D^\circ$, *we can find a point* $(S_{**}, I_*, M_*, P_*, C_{**})^T$ *in* D° *with the following properties: if* $0 < \zeta < \min\left\{S_{**}, I_*, \frac{1}{\sqrt{2}}(S_{**} + I_*), M_*, P_*, C_{**}\right\}$ *and let*

$$\begin{aligned}V_\zeta := &(S_{**} - \zeta, S_{**} + \zeta) \times (I_* - \zeta, I_* + \zeta) \times (M_* - \zeta, M_* + \zeta)\\ &\times (P_* - \zeta, P_* + \zeta) \times (C_{**} - \zeta, C_{**} + \zeta),\end{aligned}$$

then

(i) there are a control $\phi(\cdot)$ *and a time* $T > 0$ *so that* $U_\phi(T, u) \in V_\zeta$;

(ii) there are a neighborhood $\bar{V}_\zeta \subseteq V_\zeta$ *and a control* $\phi(\cdot)$ *so that* $\bar{V}_\zeta$ *is invariant under the system (24), that is, for all* $t \geq 0$ *and* $u \in \bar{V}_\zeta$ *we have* $U_\phi(t, u) \in \bar{V}_\zeta$.

Claim 3.4 *For any* $u = (s, i, m, p, c)^T \in D^\circ$ *and for any* $0 < \zeta < \min\{m^*, 1\}$, *there are a control* $\phi(\cdot)$ *and a time* $T > 0$ *such that*

$$U_\phi(T, u) \in W_\zeta := (0, 1) \times (0, \zeta) \times (m^* - \zeta, m^* + \zeta) \times (0, \infty) \times (0, \infty),$$

where m^* *is defined by (16).*

Now we recall some technical concepts and results in the theory of homogeneous Markov processes in Meyn and Tweedie (1993a, b). Let X be a locally compact and separable metric space and $\mathcal{B}(X)$ the Borel σ-algebra on X. Let $\Phi = \{\phi_t : t \geq 0\}$ be a homogeneous Markov process with state space $(X, \mathcal{B}(X))$ and transition probability function $P(t, x, \cdot)$. Suppose that Φ is defined on a probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \{\mathbb{P}_x\}_{x \in X})$ where

$$\mathbb{P}_x\{\phi_t \in A\} = P(t, x, A) \text{ for all } x \in X, \ t \geq 0, \ A \in \mathcal{B}(X).$$

Furthermore, assume that Φ is a Feller process. For a probability measure a on $\mathbb{R}_+$, we define a sampled Markov transition function K_a of Φ by

$$K_a(x, B) = \int_0^\infty P(t, x, B) a(dt).$$

K_a is said to possess a nowhere-trivial continuous component if there is a kernel $\mathcal{T} : (X, \mathcal{B}(X)) \to \mathbb{R}_+$ satisfying

- for all $B \in \mathcal{B}(X)$, $\mathcal{T}(\cdot, B)$ is lower semicontinuous, that is, for each $x \in X$ $\liminf_{y \to x} \mathcal{T}(y, B) \geq \mathcal{T}(x, B)$ for all $B \in \mathcal{B}(X)$;
- for each $x \in X$, $\mathcal{T}(x, \cdot)$ is a nontrivial measure satisfying $K_a(x, B) \geq \mathcal{T}(x, B)$ for all $B \in \mathcal{B}(X)$.

Φ is called a T-process if, for some probability measure a, the corresponding transition function K_a admits a nowhere-trivial continuous component. A subset $A \in \mathcal{B}(X)$ is called petite for the δ-skeleton chain $\{\phi_{n\delta}, \ n \in \mathbb{N}\}$ of Φ if there is a probability measure a on $\mathbb{N}$ and a nontrivial measure $\psi(\cdot)$ on X such that for all $x \in A$ and $B \in \mathcal{B}(X)$

$$K_a(x, B) = \sum_{n=1}^\infty P(n\delta, x, B) a(n) \geq \psi(B).$$

We extract the following theorem from Meyn and Tweedie (1993a, b), which is the combination of Theorem 8.1 in Meyn and Tweedie (1993a) and Theorem 6.1 in Meyn and Tweedie (1993b).

Theorem 3.1 *Suppose* $\Phi = \{\phi_t : t \geq 0\}$ *is a T-process with generator* $\mathcal{L}$.

(i) *If Φ is bounded in probability on average, that is, for each $x \in X$ and $\epsilon > 0$, there is a compact set $C_{\epsilon,x}$ satisfying*

$$\liminf_{t\to\infty} \frac{1}{t}\int_0^t P(s, x, C_{\epsilon,x})ds > 1-\epsilon,$$

then for any $x \in X$ and $f \in L^1(X, \mathcal{B}(X), \pi^*)$ we get

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t f(\phi_s)ds = \int_X f d\pi^*$$

where π^* is an invariant probability measure of Φ.

(ii) *If all compact sets are petite for some skeleton chain and there exists a positive norm-like function $V(\cdot) : X \to \mathbb{R}_+$ such that $\pounds V(x) \le -cV(x)+d$ for all $x \in X$ and for some constants $c > 0$, $d \in \mathbb{R}$, then there is a unique invariant probability measure π^* and Φ is exponentially ergodic with respect to π^*, that is, for some constants $b_1>0$ and $b_2 > 0$*

$$\|P(t, x, \cdot) - \pi^*(\cdot)\|_{TV} \le b_1(V(x)+1)e^{-b_2 t} \text{ for all } x \in X,\ t \ge 0,$$

where $\|\cdot\|_{TV}$ is the total variation norm.

We will use Theorem 3.1 to prove our main Theorem 2.3. Before doing so, we need the following lemma.

Lemma 3.1 *The solution $U(t) = (S(t), I(t), M(t), P(t), C(t))^T$ to the system (5) is a T-process. Moreover, every compact set $K \subseteq D^\circ$ is petite for the Markov chain $U(n)$, $n \in \mathbb{N}$.*

Proof of Lemma 3.1 From the argument above, we know that the transition probability function $P(t, u, \cdot)$ of the solution $U(t)$ to the system (5) has a smooth density function $p(t, \cdot, \cdot)$ on $D^\circ \times D^\circ$. By standard argument, we can show that the resolvent kernel

$$R(u, A) = \int_0^\infty e^{-t} P(t, u, A)dt$$

is continuous function in u for each $A \in \mathcal{B}(D^\circ)$. With the probability measure $a(dt) = e^{-t}dt$ on $\mathbb{R}_+$, $R(u, A)$ is its own nowhere-trivial continuous component. Hence $U(t)$ is a T-process.

Next, consider the point $U^* := (S_{**}, I_*, M_*, P_*, C_{**})^T$ as in Claim 3.3. Since D° is invariant under the system (5), $P(1, U^*, D^\circ) = 1$. So for some $u_2 := (s_2, i_2, m_2, p_2, c_2) \in D^\circ$ we get $p(1, U^*, u_2) > 0$. By Claim 3.3(ii) and the smoothness of the density $p(1, \cdot, \cdot)$ on $D^\circ \times D^\circ$, there are a neighborhood $\bar{V}_\zeta$ of U^* in D°, which is invariant under the system (24) with some control $\phi(\cdot)$, and an open set $G \ni u_2$ in D° such that $p(1, u, u') \ge m' > 0$ for all $u \in \bar{V}_\zeta$ and $u' \in G$. Now suppose K is a compact set in D°. Then, for any $u \in K$, it follows from Claim 3.3(i) that there exists a control $\phi(\cdot)$ and a time $T > 0$ such that $U_\phi(T, u) \in \bar{V}_\zeta$. Let

$n_u \in \mathbb{Z}_+$, where $\mathbb{Z}_+$ is the set of positive natural numbers, such that $n_u > T$. Due to Claim 3.3(ii), we can extend the control $\phi(\cdot)$ after time T so that $U_\phi(n_u, u) \in \bar{V}_\zeta$. Then the support theorem implies $P(n_u, u, \bar{V}_\zeta) = 2\rho_u > 0$. As $U(t)$ is Feller, so there is an open set $V_u \ni u$ such that $P(n_u, u', \bar{V}_\zeta) \geq \rho_u$ for all $u' \in V_u$. By compactness of K, there exists a finite number of such open sets V_{u_i} $(i = 1, \ldots, l)$ that satisfies $K \subseteq \bigcup_{i=1}^{l} V_{u_i}$. Let $\rho_K = \min_{1\leq i\leq l} \rho_{u_i}$, then for each $u \in K$ there exists a $n_{u_i} \in \mathbb{Z}_+$ so that $P(n_{u_i}, u, \bar{V}_\zeta) \geq \rho_K$. So for $u \in K$ and $u' \in G$ we get

$$\begin{aligned} p(n_{u_i}, u, u') &= \int_{D^\circ} p(n_{u_i}, u, u'')\, p(1, u'', u')du'' \\ &\geq m' \int_{\bar{V}_\zeta} p(n_{u_i}, u, u'')du'' = m' P(n_{u_i}, u, \bar{V}_\zeta) \end{aligned}$$

which follows that $p(n_{u_i+1}, u, u') \geq m'\rho_K$. Define the probability measure a on $\mathbb{N}$ by

$$a(n) = \begin{cases} \frac{1}{l} & \text{if} \quad n = n_{u_i} + 1 \ (i = 1, \ldots, l), \\ 0 & \text{otherwise,} \end{cases}$$

and define the kernel, for $u \in K$ and $Q \in \mathcal{B}(D^\circ)$,

$$K_a(u, Q) := \sum_{n=0}^{\infty} P(n, u, Q)a(n) = \frac{1}{l} \sum_{i=1}^{l} P(n_{u_i+1}, u, Q).$$

Then for $Q \in \mathcal{B}(D^\circ)$

$$\begin{aligned} K_a(u, Q) &= \frac{1}{l} \sum_{i=1}^{l} \int_Q p(n_{u_i+1}, u, u')du' \geq \frac{1}{l} \sum_{i=1}^{l} \int_{G\cap Q} p(n_{u_i+1}, u, u')du' \\ &\geq \rho_K m' \mu(G \cap Q), \end{aligned}$$

where μ is the Lebesgue measure on $\mathcal{B}(D^\circ)$. Let $\psi(Q) := \rho_K m' \mu(G \cap Q)$ for $Q \in \mathcal{B}(D^\circ)$ then it is clear that ψ is a nontrivial measure on D° and $K_a(u, Q) \geq \psi(Q)$ for all $u \in K$ and $Q \in \mathcal{B}(D^\circ)$. By definition, K is petite for the 1-skeleton chain $U(n)$, $n \in \mathbb{N}$. □

Proof of Theorem 2.3 First, we claim that $\liminf_{t\to\infty} I(t) > 0$ a.s. Indeed, suppose $\Omega_1 = \{\omega \in \Omega : \liminf_{t\to\infty} I(t, \omega) = 0\}$ has positive probability. Let $\omega \in \Omega_1$, then $\liminf_{t\to\infty} I(t, \omega) = 0$ which means that there is an increasing sequence of real numbers $t_k \uparrow \infty$ as $k \to \infty$ such that $I(t_k, \omega) \to 0$ as $k \to \infty$. By proof of Proposition 3.1, $\lim_{k\to\infty} S(t_k, \omega) = 1$. So by letting $\Omega_2 = \{\omega \in \Omega : \lim_{k\to\infty} S(t_k, \omega) = 1\}$ we have $\Omega_1 \subseteq \Omega_2$ and hence $\mathbb{P}(\Omega_2) > 0$. Again from proof of Proposition 3.1 we can show that $M(t_k)$ converges weakly to π_2, $P(t_k)$ converges to 0, and $C(t_k)$ converges

to 0 on Ω_2. In other words, the family of occupation measures

$$\left\{\Pi_{t_k}^{U(t_k)}(\cdot) := \frac{1}{t_k}\int_0^{t_k} \mathbb{P}_{U(t_k)}\{U(s)\in\cdot\}ds\right\}$$

is tight on ∂D and converges weakly to μ_1 on Ω_2. But then, by Lemma 3.4 in Hening and Nguyen (2018), for any $\omega \in \Omega_2$

$$\begin{aligned}\lim_{k\to\infty}\frac{\ln I(t_k,\omega)}{t_k} &= \lim_{k\to\infty}\frac{1}{t_k}\int_0^{t_k}\left[\alpha S(s,\omega)M(s,\omega) - a - \delta - \frac{\tau_1^2}{2}\right]ds\\ &= \int_{\partial D}\left[\alpha SM - a - \delta - \frac{\tau_1^2}{2}\right]\mu_1(dU) = \lambda > 0,\end{aligned}$$

which implies that $\lim_{k\to\infty} I(t_k,\omega) > 0$. So $\liminf_{t\to\infty} I(t) > 0$ on Ω_1, which is a contradiction. Therefore there is a constant $\eta_1^* > 0$ so that $\liminf_{t\to\infty} I(t) \geq \eta_1^*$ a.s. Second, if $\liminf_{t\to\infty} S(t) = 0$ on some positive set Ω_3 then it would imply that there is an increasing sequence of real numbers $t_k \uparrow \infty$ such that $\lim_{k\to\infty} S(t_k) = 0$ on Ω_3. But then, by the second equation of the system (5), $\lim_{k\to\infty} I(t_k) = 0$ on Ω_3 which contradicts the fact that $\liminf_{t\to\infty} I(t) > 0$ a.s. Hence there exists a constant $\eta_2^* > 0$ such that $\liminf_{t\to\infty} S(t) \geq \eta_2^*$ a.s.

Now we will focus on the system (6) and prove the existence and uniqueness of Π_1^*. Indeed, as $\lambda = \alpha m^* - a - \delta - \frac{\tau_1^2}{2} > 0$, so there is a $\gamma \in (0,1)$ such that $\alpha m^* - a - \delta - \frac{\tau_1^2}{2}(\gamma+1) > 0$. Consider the system (6) on the invariant domain

$$\mathcal{M}_1 := \{(s,i,m)^T \in D_1^\circ : s \geq \eta_2^*,\ i \geq \eta_1^*\}.$$

For $(s,i,m)^T \in \mathcal{M}_1$, let $V_3(s,i,m) = s + i^{-\gamma} + i + 1 + m$. Since $V_3 \to \infty$ as $m \to \infty$, V_3 is a positive norm-like function on $\mathcal{M}_1$. Furthermore,

$$\begin{aligned}\mathcal{L}V_3 &= rs(1-s-i) - \alpha ims - \gamma i^{-\gamma}(\alpha ms - a - \delta) + \frac{1}{2}\tau_1^2\gamma(\gamma+1)i^{-\gamma}\\ &\quad + \alpha sim - (a+\delta)i + n - \alpha sm^2 + (a+\delta+\tau_1^2 - e)m\\ &= -\gamma\left[\alpha m^* - a - \delta - \frac{\tau_1^2}{2}(\gamma+1)\right]i^{-\gamma} - (a+\delta)i - s - em + (r+1)s\\ &\quad - rs(s+i) + n - \alpha sm^2 + (a+\delta+\tau_1^2)m + \gamma\alpha m^* i^{-\gamma} - \gamma\alpha s i^{-\gamma} m\\ &\leq -\gamma\left[\alpha m^* - a - \delta - \frac{\tau_1^2}{2}(\gamma+1)\right]i^{-\gamma} - (a+\delta)i - s - em\\ &\quad + r + 1 + n + \gamma\alpha m^*(\eta_1^*)^{-\gamma} - \alpha\eta_2^*\left[m^2 - \frac{a+\delta+\tau_1^2}{\alpha\eta_2^*}m\right]\end{aligned}$$

$$\leq -\gamma\left[\alpha m^* - a - \delta - \frac{\tau_1^2}{2}(\gamma+1)\right] i^{-\gamma} - (a+\delta)i - s - em$$
$$+ r + 1 + n + \gamma\alpha m^*(\eta_1^*)^{-\gamma} + \frac{(a+\delta+\tau_1^2)^2}{\alpha\eta_2^*}$$

which implies that for any $(s, i, m)^T \in \mathcal{M}_1$

$$\mathcal{L}V_3(s, i, m) \leq -\theta_1 V_3(s, i, m) + \theta_2 \tag{26}$$

in which

$$\theta_1 := \min\left\{\gamma\left[\alpha m^* - a - \delta - \frac{\tau_1^2}{2}(\gamma+1)\right], a+\delta, 1, e\right\} > 0$$

and

$$\theta_2 := \theta_1 + s + 1 + n + \gamma\alpha m^*(\eta_1^*)^{-\gamma} + \frac{(a+\delta+\tau_1^2)^2}{\alpha\eta_2^*}.$$

By Theorem 3.1(ii), it follows from Lemma 3.1 and (26) that the solution $U_1(t)$ of the system (6) has a unique invariant probability measure Π_1^* in $\mathcal{M}_1$ such that for some $b_1 > 0$ and $b_2 > 0$ we get

$$\|P_1(t, u_1, \cdot) - \Pi_1^*(\cdot)\|_{TV} \leq b_1(V_3(u_1) + 1)e^{-b_2 t}$$

for all $t \geq 0$ and for all $u_1 \in \mathcal{M}_1$. (Note that $P_1(t, u_1, \cdot)$ denotes the transition probability function of the solution $U_1(t)$.) Moreover, due to Claims 3.1 and 3.2, using the support theorem we obtain the support of Π_1^* as

$$\text{supp}(\Pi_1^*) = \left\{\left(1 - i - \frac{\alpha}{r}im, i, m\right)^T : im < \frac{r}{\alpha},\ i \in (0,1),\ m > 0\right\}.$$

Next, we can easily have the estimate

$$\mathcal{L}V_3(s, i, m) \leq \theta_3 V_3(s, i, m)$$

for any $(s, i, m) \in D_1^\circ$ and some $\theta_3 > 0$. By standard argument, we can prove that there are a $H_1 > 0$ and $b_3 > 0$ such that for all $t > 0$ and $u_1 = (s, i, m)^T \in D_1^\circ$

$$\mathbb{E}V_3(U_1(t, u_1)) \leq H_1\, V_3(u_1)e^{b_3 t}.$$

As $\liminf_{t\to\infty} I(t) \geq \eta_1^*$ a.s. and $\liminf_{t\to\infty} S(t) \geq \eta_2^*$ a.s., so for $u_{10} \in D_1^\circ$ there is a non-random time $t_0 = t_0(u_{10}) > 0$ such that $U_1(t, u_{10}) \in \mathcal{M}_1$ for all $t \geq t_0$ a.s.

Therefore we get the following estimate

$$\begin{aligned}\|P_1(t+t_0,u_{10},\cdot)-\Pi_1^*\|_{TV} &= \left\|\int_{\mathcal{M}_1} P_1(t_0,u_{10},du_1)\,[P_1(t,u_1,\cdot)-\Pi_1^*(\cdot)]\right\|_{TV}\\ &\le \int_{\mathcal{M}_1} p_1(t_0,u_{10},u_1)\|P_1(t,u_1,\cdot)-\Pi_1^*(\cdot)\|_{TV}du_1\\ &\le \int_{\mathcal{M}_1} p_1(t_0,u_{10},u_1)b_1[V_3(u_1)+1]e^{-b_2t}du_1\\ &= b_1[\mathbb{E}V_3(U_1(t_0,u_{10}))+1]e^{-b_2t}\\ &\le b_1[H_1V_3(u_{10})e^{b_3t_0}+1]e^{-b_2t} \text{ for all } t\ge 0.\end{aligned}$$

Thus $U_1(t)$ is exponentially ergodic with respect to Π_1^*. Hence Theorem 2.3(i) is proved.

Second, we turn to the system (7). We claim that $\liminf_{t\to\infty} P(t) > 0$ a.s. Indeed, assume that $\Omega_4 = \{\omega\in\Omega : \liminf_{t\to\infty} P(t) = 0\}$ has positive probability. Then there is an increasing sequence of real numbers $t_k \uparrow \infty$ such that $\lim_{k\to\infty} P(t_k) = 0$ on Ω_4. Since $\liminf_{t\to\infty} I(t) \ge \eta_1^*$ a.s., there is a non-random time $T_2 > 0$ so that $t \ge T_2$ implies $I(t) \ge \eta_1^*$ a.s. Hence, by the first equation of the system (7), $t \ge T_2$ implies $P(t) \ge \underset{\sim}{P}(t)$ a.s. where $\underset{\sim}{P}(t)$ solves the equation

$$d\underset{\sim}{P} = \left[\delta q\eta_1^* + b\underset{\sim}{P} - \theta\frac{\underset{\sim}{P}^2}{1+\underset{\sim}{P}^2}\right]dt + \tau_3\underset{\sim}{P}dW_3.$$

By the same argument as in proof of Proposition 3.1, we can show that $\underset{\sim}{P}(t)$ converges weakly to an ergodic invariant probability measure on $(0,\infty)$. However, since $P(t_k) \ge \underset{\sim}{P}(t_k) \ge 0$ for $t_k > T_2$ on Ω_4, $\lim_{k\to\infty}\underset{\sim}{P}(t_k) = 0$ on Ω_4 with $\mathbb{P}(\Omega_4) > 0$, which is a contradiction. Thus there is a constant $\eta_3^* > 0$ such that $\liminf_{t\to\infty} P(t) \ge \eta_3^*$ a.s.

Now we consider the first equation of the system (7) on the invariant domain $\mathcal{M}_2 = \{p : p \ge \eta_3^*\}$. For $p \in \mathcal{M}_2$, let $V_4(p) = \ln p - \ln\eta_3^*$. As $V_4 \to \infty$ as $p\to\infty$, so V_4 is a positive norm-like on $\mathcal{M}_2$. Moreover,

$$\mathcal{L}V_4 = \frac{\delta qi}{p} + b - \frac{\tau_3^2}{2} - \theta\frac{p}{1+p^2} \le \frac{\delta q}{p} + b - \frac{\tau_3^2}{2}.$$

Since $b - \dfrac{\tau_3^2}{2} < 0$, we can choose $p^* > 0$ large enough so that $\dfrac{\delta q}{p^*} + b - \dfrac{\tau_3^2}{2} < 0$. Then $\mathcal{L}V_4 \le \dfrac{\delta q}{p^*} + b - \dfrac{\tau_3^2}{2}$ for any $p \ge p^*$. By Theorem 4.1, p. 108, Theorem 4.2, p. 110, and Corollary p. 112 in Khasminskii (2012), there is a unique ergodic invariant probability measure π_p^* for the first equation of the system (7) on $(0,\infty)$ and $P(t)$ is ergodic with respect to π_p^*. By the second equation of the system (7),

$$C(t) = C(0)e^{-kt} + e^{-kt}\int_0^t \theta\frac{P(s)^2}{1+P(s)^2}e^{ks}ds.$$

It follows from L'Hospital's Rule that

$$\lim_{t\to\infty} C(t) = \lim_{t\to\infty} \frac{\theta}{k}\frac{P(t)^2}{1+P(t)^2}.$$

Thus $C(t)$ converges weakly to a unique ergodic invariant probability measure π_c^* on $(0,\infty)$. Therefore, the system (7) has a unique invariant probability measure $\Pi_2^* = \pi_p^* \times \pi_c^*$ in D_2° and $U_2(t)$ is ergodic with respect to Π_2^*. By Claims 3.1 and 3.2, using the support theorem gives

$$\text{supp}(\Pi_2^*) = \left\{ \left(p, \frac{\theta}{k}\frac{p^2}{1+p^2} \right)^T : p > 0 \right\}.$$

Hence Theorem 2.3(ii) is shown. Finally, Theorem 2.3(iii) can be derived from Theorem 3.1, Lemma 3.1, and the fact that convergence in total variation norm implies the boundedness in probability on average. □

Proof of Theorem 2.4 First, we consider the system (6) on the invariant domain D_1°. We will show that if $\lambda < 0$ and $b < \frac{\tau_3^2}{2}$ then $I(t) \to 0$ a.s. Indeed, since $\lambda = \alpha m^* - a - \delta - \frac{\tau_1^2}{2}$, there are a $\zeta > 0$ and $\gamma \in (0,1)$ such that $\theta_4 := \alpha(m^*+\zeta) - a - \delta - \frac{\tau_1^2}{2}(1-\gamma) < 0$. Consider the positive definite decrescent function $V_5(u_1) = i^\gamma$, which is twice differentiable on $\overline{W}_\zeta := (0,1)\times(0,\zeta)\times(m^*-\zeta, m^*+\zeta)$. For any $u_1 = (s,i,m)^T \in \overline{W}_\zeta$, since $m - m^* \le |m - m^*| < \zeta$, we obtain $\pounds V_5 \le \theta_4 V_5$. By Theorem 2.3, p. 112 in Mao (1997), for any $\epsilon > 0$ and for any $u_1 \in \overline{W}_\zeta$

$$\mathbb{P}_{u_1}\left\{ \lim_{t\to\infty} I(t) = 0 \right\} \ge 1 - \epsilon. \tag{27}$$

Next, we construct a compact set $\widetilde{K}$ so that the solution $U_1(t)$ is recurrent relative to $\widetilde{K}$, i.e., as the solution $U_1(t)$ starts in D_1° it will visit $\widetilde{K}$ infinitely many times in finite times with probability 1. By using Theorem 3.9, p. 89 in Khasminskii (2012), we construct a non-negative twice differentiable function $V_6(s,i,m)$ and a compact set $\widetilde{K}$ in D_1° such that $\pounds V_6 < 0$ for any $(s,i,m)^T \in \widetilde{K}^c$. Indeed, consider $V_6(s,i,m) = s + i + m$. Then

$$\pounds V_6 = rs(1-s-i) - (a+\delta)i + n - \alpha s m^2 + (a+\delta+\tau_1^2 - e)m.$$

Since there exists an $A > r$ large enough so that $n - \alpha s m^2 + (a+\delta+\tau_1^2 - e)m \le -m$ for $m > A$, letting $\widetilde{K} := \{(s,i,m)^T \in D_1^\circ : m \le A\}$ we have $\pounds V_6 \le r - m < r - A < 0$ for any $(s,i,m)^T \in \widetilde{K}^c$.

Now consider the control system corresponding to the system (6), which is exactly the system of the first three equations of (24). Denote by $U_{1\phi}(t,u_1)$ the solution of this control system with control $(\phi_1,\phi_2)^T$ and initial value u_1. For clarity, we denote

by $U_1(t, u_1)$ the solution to the system (6) with initial value u_1. We use $\mathbb{P}$ to denote the probability law on Ω for the solution $U_1(t, u_1)$ and sometimes we use $\mathbb{P}_{u_1}(A)$ to denote the probability of an event A that involves the solution U_1 when the solution U_1 starts at u_1.

By Claim 3.4, for any $u_1 \in \widetilde{K}$ we can choose a control $(\phi_1, \phi_2)^T$ and a time $T_{u_1} > 0$ such that $U_{1\phi}(T_{u_1}, u_1) \in \overline{W}_\zeta$. Due to the support theorem (Theorem 8.1 in Ikeda and Watanabe 1989), for any $u_1 \in \widetilde{K}$, there is $T_{u_1} > 0$ so that

$$\mathbb{P}\left\{U_1(T_{u_1}, u_1) \in \overline{W}_\zeta\right\} = 2\rho^{u_1} > 0.$$

Using the Markov–Feller property of $U_1(t)$, there is a neighborhood $V_{u_1} \ni u_1$ so that for any $u_1' \in V_{u_1}$ we get

$$\mathbb{P}\left\{U_1(T_{u_1}, u_1') \in \overline{W}_\zeta\right\} > \rho^{u_1}.$$

Since $\widetilde{K}$ is compact, there exists a finite number of such neighborhoods $V_{u_1^i}$ $(i = 1, \ldots, l)$ so that $\widetilde{K} \subseteq \bigcup_{i=1}^{l} V_{u_1^i}$. Put $T^* = \max_{i=1,\ldots,l} T_{u_1^i}$ and $\rho^* = \min_{i=1,\ldots,l} \rho^{u_1^i}$. For $u_1 \in D_1^\circ$, set

$$\tau_\zeta^{u_1} = \inf\left\{t > 0 : U_1(t, u_1) \in \overline{W}_\zeta\right\}.$$

Then, for any $u_1 \in \widetilde{K}$, the event $\tau_\zeta^{u_1} < T^*$ is followed from the fact that there exists an $i \in \{1, \ldots, l\}$ such that $U_1(T_{u_1^i}, u_1) \in \overline{W}_\zeta$. Hence

$$\mathbb{P}\{\tau_\zeta^{u_1} < T^*\} \geq \rho^* \quad \text{for all} \quad u_1 \in \widetilde{K}. \tag{28}$$

As $U_1(t, u_1)$ is recurrent relative to $\widetilde{K}$, we define a sequence of finite stopping times

$$\begin{aligned}
&\zeta_0 = 0, \ \ \zeta_1 = \inf\{t > T^* : U_1(t, u_1) \in \widetilde{K}\}, \ldots \\
&\zeta_k = \inf\{t > \zeta_{k-1} + T^* : U_1(t, u_1) \in \widetilde{K}\}, \quad k = 2, 3, \ldots
\end{aligned}$$

Consider the event

$$A_k = \left\{U_1(t, u_1) \notin \overline{W}_\zeta \ \forall\, t \in [\zeta_k, \zeta_k + T^*]\right\}, \quad k \in \mathbb{N}.$$

It follows from (28) that, for all $k \in \mathbb{N}$, $\mathbb{P}_{u_1}(A_k^c) = \mathbb{P}\{\tau_\zeta^{\bar{u}_1} < T^*\} \geq \rho^*$ where $\bar{u}_1 = U_1(\zeta_k, u_1) \in \widetilde{K}$. So $\mathbb{P}_{u_1}(A_k) \leq 1 - \rho^*$ for all $k \in \mathbb{N}$. By the strong Markov property of $U_1(t, u_1)$, we get

$$\mathbb{P}_{u_1}(A_1 \cap A_2) = \mathbb{P}_{u_1}(A_1)\, \mathbb{P}_{U_1(\zeta_2, u_1)}(A_2) \leq (1 - \rho^*)^2.$$

Then, by induction,

$$\mathbb{P}_{u_1}\left(\bigcap_{k=1}^{n} A_k\right) \leq (1-\rho^*)^n \to 0 \quad \text{as} \quad n \to \infty.$$

As a result, $\mathbb{P}_{u_1}\left(\bigcap_{k=1}^{\infty} A_k\right) = 0$. It implies that

$$\mathbb{P}\{\tau_\zeta^{u_1} < \infty\} = 1 \quad \text{for any} \quad u_1 \in D_1^\circ. \tag{29}$$

Again, using the strong Markov property of $U_1(t, u_1)$, it follows from (27) and (29) that, for any $u_1 \in D_1^\circ$,

$$\mathbb{P}_{u_1}\left\{\lim_{t\to\infty} I(t) = 0\right\} \geq 1 - \epsilon.$$

Since $\epsilon > 0$ is arbitrary, $I(t)$ converges to 0 a.s. for any initial value $u_1 \in D_1^\circ$. By Proposition 3.1, $S(t)$ converges to 1 a.s., $M(t)$ converges weakly to π_2, $P(t)$ converges to 0 a.s., and $C(t)$ converges to 0 a.s. Furthermore, by the second equation of the system (6), using strong law of large numbers yields

$$\begin{aligned}\lim_{t\to\infty} \frac{\ln I(t)}{t} &= \lim_{t\to\infty} \frac{1}{t}\int_0^t \left[\alpha S(s)M(s) - a - \delta - \frac{\tau_1^2}{2}\right] ds \\ &= \int_{\partial D}\left[\alpha SM - a - \delta - \frac{\tau_1^2}{2}\right](\delta_1^* \times \delta_0^* \times \pi_2)(dSdIdM) = \lambda \ \ a.s.\end{aligned}$$

□

4 Numerical Simulation

In this section, we will perform several numerical simulations to demonstrate Hopf bifurcation occurrence of the deterministic system (4) in Theorem 2.1 as well as the long-term dynamics of the stochastic system (3) in Theorem 2.5 including extinction, persistence, and stochastic Hopf bifurcations. We fix the following baseline parameter values: $r = 0.02$, $\alpha = 0.0003$, $a = 0.1$, $n = 10^5$, $e = 5$, $q = 13.44$, $\theta = 2.5$, $k = 1.01$, $\tau_1 = 0.01$, $\tau_2 = 0.2$, and $\tau_3 = 0.15$. Parameters δ and b are varied to obtain different dynamical behaviors of the deterministic system (4) and the stochastic system (3). In these two systems, because of non-dimensionalization, the units of normal cells, infected cells, precancerous cells, and cancerous cells are not absolute number of cells but relative number of cells. So the quantities S and I represent the percentage of the cell number of normal cells and infected cells, respectively. While the quantities P and C are the portions of the cell number of precancerous cells and cancerous cells over the half-saturation concentration for rate of progression from precancerous to cancerous, and so we indicate them as relative precancerous and cancerous cells. The unit of the time is measured in days. In all the figures below, the solution paths of

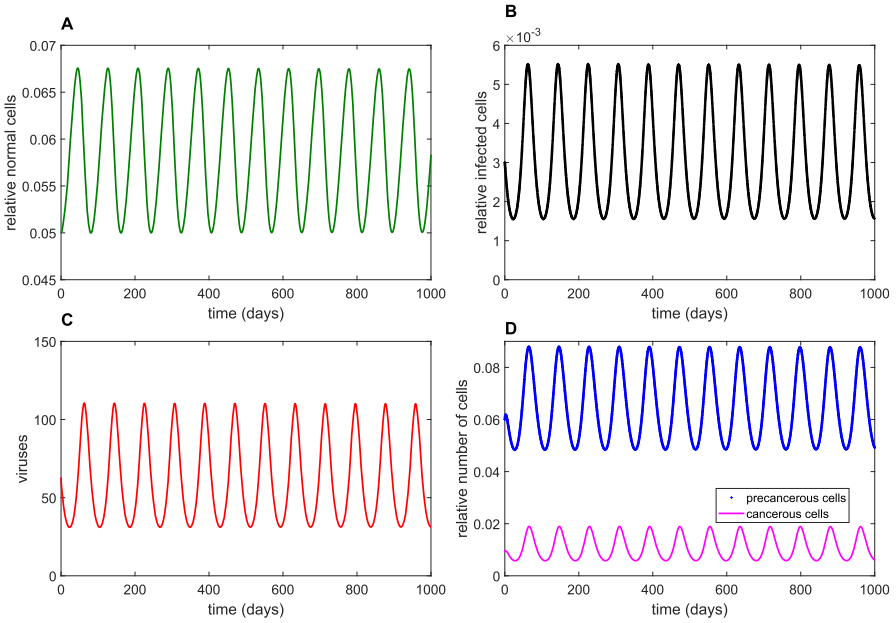

Fig. 1 Dynamics of the deterministic system (4) with the parameters $r = 0.02$, $\alpha = 0.0003$, $a = 0.1$, $\delta = 0.25$, $n = 10^5$, $e = 5$, $q = 13.44$, $\theta = 2.5$, $b = 0.01$, $k = 1.01$, and the initial value $(0.05, 0.003, 63, 0.06, 0.01)$. The parameter set is taken to satisfy the assumptions of Theorem 2.1(i) and (iii). A, B, and C represent the periodic solutions of relative normal cells, relative infected cells, and viruses, respectively, arising around the equilibrium point $E^* = (0.0583, 0.0031, 62.5692)$. D represents the periodic solutions of relative precancerous cells and relative cancerous cells arising from the stable equilibrium point $(P_{21}, C_{21}) = (0.0618, 0.0094)$

the deterministic system (4) and the stochastic system (3) are simulated with different initial values and different values of δ and b.

For the deterministic system (4), we used the algorithm of the fourth order Runge–Kutta method, which can be easily implemented by ode45 in Matlab, to produce Figs. 1 and 2. For the stochastic system (3), we utilized the algorithm of stochastic Runge–Kutta method of strong order 1 to produce Figs. 3, 4 and 5. The details of this algorithm can be found in the supplementary material of our previous work (Phan and Tian 2022). Based on the algorithm, we developed an algorithm of simulating the Lyapunov exponent along each solution component to produce Fig. 6 and check if a stochastic Hopf bifurcation occurs in the stochastic system (3).

To illustrate the occurrence of Hopf bifurcation of the deterministic system (4) as the basic reproduction number R_0 changes and passes through the threshold R^*, we first take $\delta = 0.25$ and $b = 0.01$. From computation, we obtain $R_0 = 17.1429$, $R^* = 18.7818$, one positive stable equilibrium $(0.0583, 0.0031, 62.5692, 0.0670, 0.0111)$, and one positive unstable equilibrium $(0.0583, 0.0031, 62.5692, 248.9448, 2.4752)$. In this case, the behavior of the system (4) depends on where it starts. Figure 1 shows the periodic solutions arising from Hopf bifurcation when the solution starts close to the positive stable equilibrium. In Fig. 2, when the solution starts close to

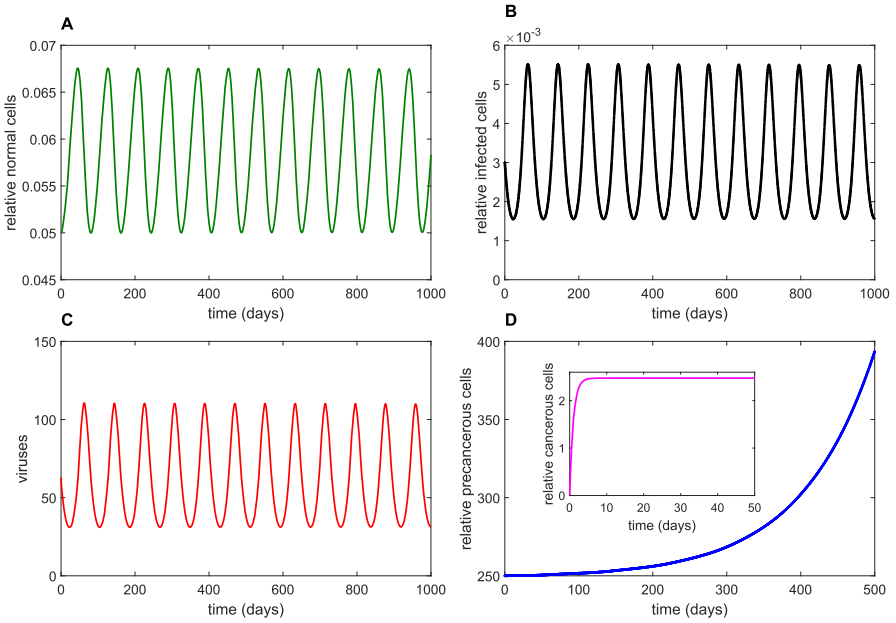

Fig. 2 Dynamics of the deterministic system (4) with the parameters $r = 0.02$, $\alpha = 0.0003$, $a = 0.1$, $\delta = 0.25$, $n = 10^5$, $e = 5$, $q = 13.44$, $\theta = 2.5$, $b = 0.01$, $k = 1.01$, and the initial value $(0.05, 0.003, 63, 250, 0.01)$. A, B, and C represent the periodic solutions of relative normal cells, relative infected cells, and viruses, respectively, arising from the equilibrium point $E^* = (0.0583, 0.0031, 62.5692)$. D shows the relative precancerous cell population (the blue curve) grows without bound when it starts near $P_{22} = 249.1067$ while the relative cancerous cell population (the magenta curve) attains its peak (color figure online)

the positive unstable equilibrium, the precancerous cells blow up and the cancerous cells reach its maximum value even though we obtain the periodic solutions of normal cells, infected cells, and viruses. If we take $\delta = 0.21$ and $b = 0.01$ then, by computation, we get $R_0 = 19.3548$, $R^* = 18.7818$, one positive stable equilibrium $(0.0517, 0.0032, 63.0122, 0.0618, 0.0094)$, and one positive unstable equilibrium $(0.0517, 0.0032, 63.0122, 249.1067, 2.4752)$. For this case, the system (4) behaves similarly to Fig. 1 when it starts near the positive stable equilibrium and similarly to Fig. 2 when it starts close to the positive unstable one.

Next, we demonstrate numerically the extinction and persistence of the stochastic system (3) in Theorem 2.5. First, we take $\delta = 0.25$ and $b = 0.01$. By computation, we have $\lambda = 3.2734 > 0$ and, clearly, $b < \frac{\tau_3^2}{2}$. So the conditions of Theorem 2.5(ii) is fulfilled. Figure 3 shows one solution path of the stochastic system (3) when it starts close to the positive stable equilibrium of the corresponding deterministic system (4) and this path finally ends up at a equilibrium state which is characterized by the ergodic invariant probability measure $\overline{\mu}^*$. Second, if we take $\delta = 6$ and $b = 0.01$, then computation gives us $\lambda = -0.0513 < 0$ and hence the conditions of Theorem 2.5(i) is satisfied. Figure 4 shows one typical solution path of the stochastic system (3) that converges a.s. to $(1, 0, 0, 0, 0)$. Third, if we increase the value of b to pass through

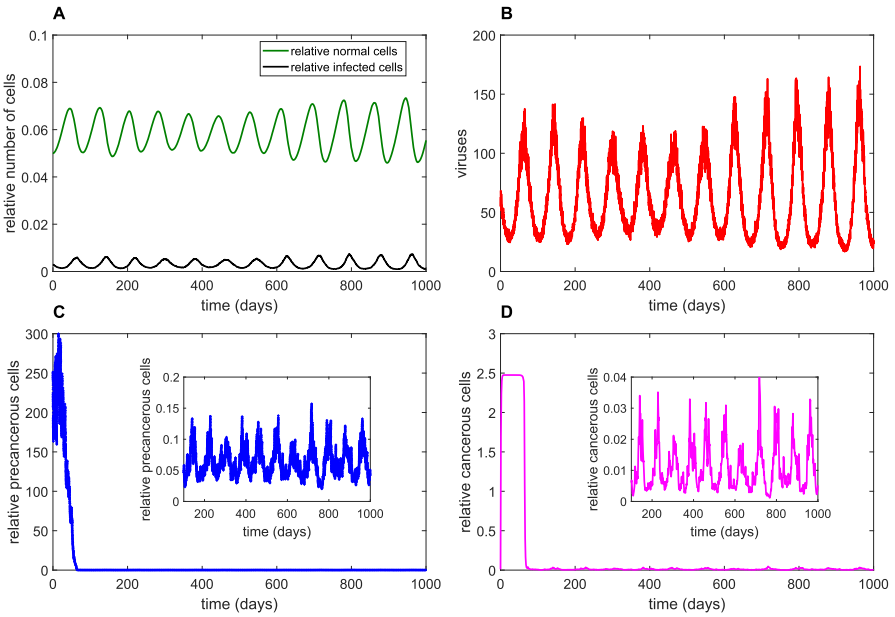

Fig. 3 Dynamics of one solution path of the stochastic system (3) with the parameters $r = 0.02, \alpha = 0.0003$, $a = 0.1, \delta = 0.25, n = 10^5, e = 5, q = 13.44, \theta = 2.5, b = 0.01, k = 1.01$, noise intensities $\tau_1 = 0.01$, $\tau_2 = 0.2, \tau_3 = 0.15$, and the initial value $(0.05, 0.003, 63, 250, 0.01)$, which satisfy the assumptions of Theorem 2.5(ii). The solution path persists for a long time and ends up at an equilibrium state $\overline{\mu}^\star$

$\frac{\tau_3^2}{2}$, then precancerous cells will blow up and cancerous cells approach its carrying capacity no matter how large the value of λ is. Figure 5 shows the situation when $\lambda > 0$ but $b > \frac{\tau_3^2}{2}$. In this figure, even though the normal cells, infected cells, and viruses eventually wind up at an equilibrium state, precancerous cells finally reach a very large value and cancerous cells approach its maximum value.

Finally, it is natural to expect that the stochastic system (3) would also undergo a stochastic Hopf bifurcation as a parameter of the system changes since the deterministic system (4) undergoes Hopf bifurcation when the basic reproduction number R_0 passes through the threshold R^*. Before studying stochastic Hopf bifurcation for the stochastic system (3), we recall the difference between Hopf bifurcation phenomena in the deterministic and stochastic setting. For a deterministic system, Hopf bifurcation occurs when a positive equilibrium point of the system has a pair of complex eigenvalues and its real part crosses zero as a parameter of the system changes. While, in the stochastic setting, we look at the sign changes of Lyapunov exponents of invariant probability measures, whose supports are the interior of a stochastic system's domain, as one of its parameters changes. Notice that an invariant probability measure of a stochastic system plays the same role as an equilibrium point of the corresponding deterministic system and the Lypapunov exponents of an invariant probability measure have the same meaning as eigenvalues of an equilibrium point, which measures the rates of convergence along solution components. So, to verify stochastic Hopf bifurcation for the system (3), we compute the system's Lyapunov exponents along solution

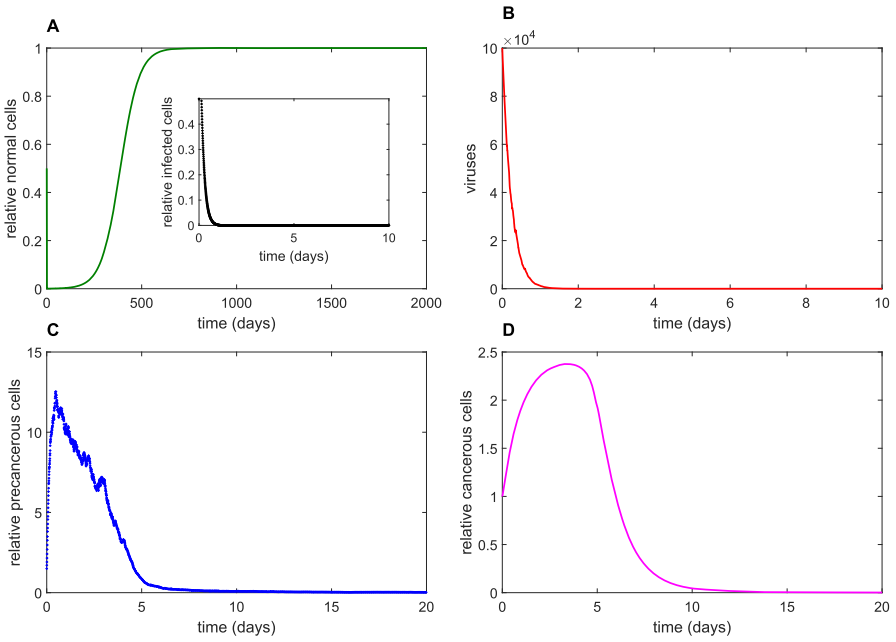

Fig. 4 Dynamics of one solution path of the stochastic system (3) with the parameters $r = 0.02$, $\alpha = 0.0003$, $a = 0.1$, $\delta = 6$, $n = 10^5$, $e = 5$, $q = 13.44$, $\theta = 2.5$, $b = 0.01$, $k = 1.01$, noise intensities $\tau_1 = 0.01$, $\tau_2 = 0.2$, $\tau_3 = 0.15$, and the initial value $(0.5, 0.5, 10^5, 1.5, 1)$, which satisfy the assumptions of Theorem 2.5(i). A relative normal cells reach its carrying capacity 1 while relative infected cells decrease quickly to 0. The other solution paths in B, C, and D also decrease to 0 in a short period of time

components as the parameter δ changes. If there is a stochastic Hopf bifurcation in the dynamical sense, then it is necessary that one Lyapunov exponent has to cross zero (Phan and Tian 2022; Keller 1996). Fig. 6 shows the behavior of Lyapunov exponents λ_i $(i = 1, 2, 3, 4, 5)$ of the five solution components (S, I, V, P, C) when the parameter δ runs through between 0.25 and 5. (The details of how to simulate Lyapunov exponents can be found in Phan and Tian 2022.) Here we take the other parameter as in Fig. 3. As the parameter δ lies between 0.25 and 5, the system (3) is always strongly persistent due to Theorem 2.5(ii), which means that it always has a unique positive invariant probability measure $\overline{\mu}^*$. Actually, we computed the Lyapunov exponents of this invariant probability measure when δ varies from 2.5 to 5 to see if any of them crosses 0. We see that all four Lyapunov exponents λ_1, λ_2, λ_4, and λ_5 are always below zero; while the exponent λ_3 crosses zero twice at some values of δ between 0.5 and 0.7. This means that the invariant probability measure $\overline{\mu}^*$ is unstable at first and it becomes stable as δ crosses 0 for the first time. Then it loses its stability and becomes unstable when δ increases and crosses 0 for the second time. It shows that a stochastic Hopf bifurcation occurs in the stochastic system (3). One interesting point here is that, with the same parameter set where the deterministic system (4) undergoes deterministic Hopf bifurcation as δ is within a certain range, the stochastic system (3) still undergoes stochastic Hopf bifurcation when δ lies within almost the same

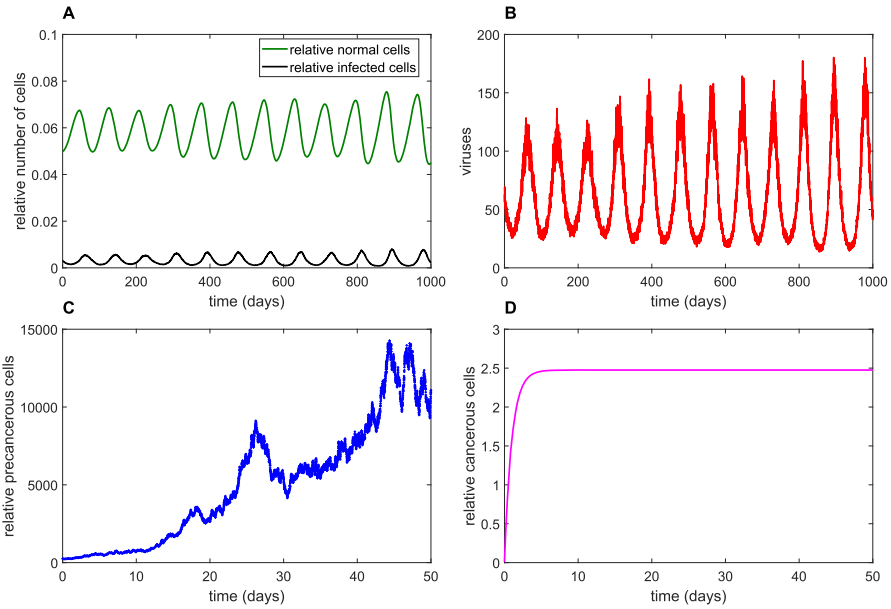

Fig. 5 Dynamics of one solution path of the stochastic system (3) with the parameters $r = 0.02, \alpha = 0.0003$, $a = 0.1, \delta = 0.25, n = 10^5, e = 5, q = 13.44, \theta = 2.5, b = 0.04, k = 1.01$, noise intensities $\tau_1 = 0.01$, $\tau_2 = 0.2$, $\tau_3 = 0.15$, and the initial value $(0.05, 0.003, 63, 250, 0.01)$, which satisfy the assumptions of Theorem 2.5(iii). A and B show the persistence of relative normal cells, relative infected cells, and viruses. C represents the blowup of relative precancerous cells in a short period of time while D shows the peak reaching of relative cancerous cells

range with appropriate chosen noise intensities. From our Theorem 2.5 and Proposition 2.2, decreasing the noise intensities τ_1 and τ_2 while increasing the noise intensity τ_3 appropriately would not destroy Hopf bifurcation of the deterministic system.

5 Discussion

In this research, we examine a deterministic system in terms of five ordinary differential equations through stochastic viewpoint. The system models how human papillomaviruses infect basal epithelia cells in the cervix and how the infection progresses to a cervical cancer. The system captures basic biological characters of both HPV infection dynamics and cancer progress process. Particularly, three parameters, the progression rate from infected cells to precancerous cells, the net death rate of infected cells, and the net growth of precancerous cells, are important for long-term behaviors of the model. However, as all deterministic systems, this system only represents mean behaviors of cells and viruses, and it is known that the system is subject to micro-environmental fluctuations in vivo. The system perturbed by such fluctuations may exhibit some long-term behaviors that are different from those of the unperturbed system. We incorporate three noises associated to the progression rate from infected cells to precancerous cells, precancerous cell growth rate, and free virus population

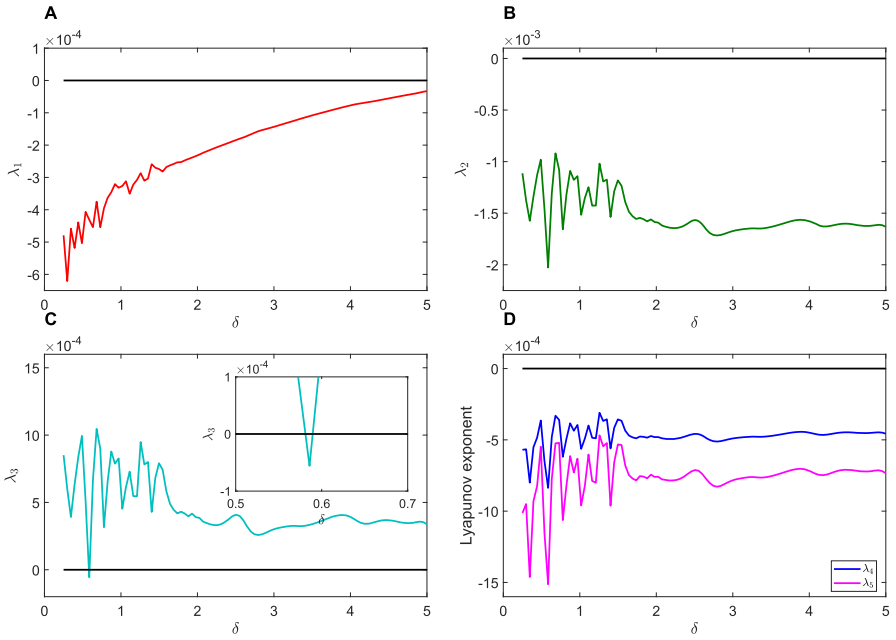

Fig. 6 Lyapunov exponents of μ^* along solution components of the stochastic system (3) as δ changes from 0.25 to 5. The parameters are taken as $r = 0.02, \alpha = 0.0003, a = 0.1, n = 10^5, e = 5, q = 13.44, \theta = 2.5, b = 0.01, k = 1.01$, and the initial value is $(0.05, 0.003, 63, 250, 0.01)$, which satisfy the assumptions of Theorem 2.5(iii) satisfy when δ is between 0.25 and 5

into the deterministic system to study cervical cancer progression from HPV infection. The stochastic system yields some insights.

To have a cervical cancer, the chronic infection state or equilibrium must be established first. In our stochastic system, the chronic infection state is a random variable whose distribution is an ergodic invariant distribution. The small progression rate and noise associated with it will be helpful to establish such chronic infection state, and hence to develop into a cervical cancer. The small noise received or made by free viruses is also in favor of chronic infection establishment. This may explain why it takes decades to develop a cervical cancer from HPV infection. On the other hand, the large of these two types of noises will help to reduce chronic infections. This may provide some hints for medical treatments that increasing instability of cervical microenvironment can help to reduce the probability of chronic infection establishment. When the chronic infection state is established, there is a positive probability to establish a cervical cancer. The cancer is an ergodic invariant measure or random variable with an ergodic invariant distribution in our stochastic system, which corresponds to an equilibrium in the deterministic system. When the noise associated with precancerous cell growth is large enough, there will be a cervical cancer. It is clear that these three noises have different effects on the progression from HPV infection to cervical cancers.

In the deterministic model, there is Hopf bifurcation when the basic reproduction number passes some value, which may represent normal cells, infected cells, and viruses interact indefinitely without settling down. In our stochastic model, we also observe stochastic Hopf bifurcation in the dynamical sense. It is difficult to give a criterion what deterministic Hopf bifurcation will become dynamical or phenomenological bifurcation.

Since the stochastic system (3) is difficult to analyze, we transform it to a different stochastic system (5). The transformation is valid as long as the denominator variable is not zero. For the situation where the denominator variable is zero, we apply limit procedure to give the value of the transformed variable as usual variable transformations. This transformed system has two invariant probability measures on the boundary in that one has an inverse gamma distribution as one of its components, and another one has a generalized inverse Gaussian distribution as one of its components. That means, on the different parts of the boundary, the convergent rate follows different distributions. Because one transformed variable is the ratio of two original variables and these two variables approach delta distributions on the boundary, the two non-delta distribution components disappear after the system is transformed back. This is very interesting mathematical phenomenon. It reveals that there are different time scales in the stochastic system. It may be worth for a deep study in the future.

There may be other ways to incorporate environmental noises into a deterministic model. Our method is based on law of large numbers. It is obvious that there are other parameters in the deterministic model which are worth to investigate. For example, the maximum progression rate from precancerous cells to cancer cells θ is another important parameter. We may consider this parameter and other uncertainties in the future.

As we pointed out in the Sect. 1, spatial interactions among cells play some roles in both HPV infection dynamics and cancer progression process. It is a limitation for ordinary differential equation systems to incorporate spatial variations. Although some environmental noises can be attributed to spatial variations, it is not clear how exactly to connect them. This calls for the need to establish mathematical models in terms of partial differential equations for cervical cancer research.

Acknowledgements TAP would like to acknowledge support by the National Institute Of General Medical Sciences of the National Institutes of Health under Award Number P20GM104420. JPT would like to acknowledge grant National Institutes of Health (U54CA132383) which supported FS for a summer during the grant periods. We thank two anonymous reviewers for their constructive suggestions which help to improve the presentation of this work.

Declarations

Conflict of interest All authors declare that they have no conflicts of interest.

Appendix

To comprehend how environmental noises and randomness affect the dynamical behaviors of the deterministic system (4), we need to study how the threshold λ is

related to the reproduction number R_0. Since $R_0 = \frac{\alpha n}{e(a+\delta)}$ is a decreasing function of the parameter δ, which is the progression rate from infected cells to precancerous cells, we can consider the threshold λ as a function of δ as well as a function of noise intensities. Propositions 2.1 and 2.2 give the behavior of the threshold λ with respect to system parameters and noise intensities. Because Proposition 2.2 is a direct consequence of Proposition 2.1, so we only present the proof of Proposition 2.1

Proof of Proposition 2.1 Since $\frac{\Theta}{w} = \frac{a+\delta-e}{2\sqrt{\alpha n}} + \frac{\tau_1^2-\tau_2^2}{4\sqrt{\alpha n}}$, the threshold λ can be rewritten as

$$\lambda = \sqrt{\alpha n}\left[R_\Theta(w) - 2\frac{\Theta}{w}\right] - e - \frac{\tau_2^2}{2}.$$

Let $D_\Theta(w) := \dfrac{K_{\Theta+1}(w)K_{\Theta-1}(w)}{K_\Theta^2(w)}$ where $K_\Theta(\cdot)$ is the modified Bessel function of third kind with index Θ. From Jorgensen (1982) in p. 172 and p. 175, we have $D_\Theta(w) = R_\Theta(w)R_{-\Theta}(w)$ and $D_\Theta(w) = R_\Theta(w)\left[R_\Theta(w) - 2\frac{\Theta}{w}\right]$. Hence

$$\lambda = \sqrt{\alpha n}R_{-\Theta}(w) - e - \frac{\tau_2^2}{2}.$$

Since $w > 0$, a result in p. 173 in Jorgensen (1982) implies that $R_\Theta(w)$ is an increasing function of Θ and thus λ is a decreasing function of Θ. But, as Θ is an increasing function of δ, so λ is a decreasing function of δ. Since R_0 is a decreasing function of δ, λ is an increasing function of R_0.

Next, since $\overline{\lambda} = \lim\limits_{(\tau_1,\tau_2)\to(0,0)} \lambda$ and $R_\Theta(w) = \frac{\Theta}{w} + \sqrt{\left(\frac{\Theta}{w}\right)^2 + D_\Theta(w)}$,

$$\overline{\lambda} = \lim_{(\tau_1,\tau_2)\to(0,0)} \left\{ \left[\frac{\Theta}{w} + \sqrt{\left(\frac{\Theta}{w}\right)^2 + D_\Theta(w)}\right]\sqrt{\alpha n} - a - \delta - \frac{\tau_1^2}{2} \right\}.$$

As $\frac{\Theta}{w} \to \frac{a+\delta-e}{2\sqrt{\alpha n}}$ and $D_\Theta(w) \to 1$ as $(\tau_1, \tau_2) \to (0, 0)$, so

$$\overline{\lambda} = \frac{a+\delta-e}{2} + \sqrt{\left(\frac{a+\delta-e}{2}\right)^2 + \alpha n} - a - \delta.$$

It implies that $\overline{\lambda} = 0$ iff $R_0 = 1$, $\overline{\lambda} < 0$ iff $R_0 < 1$, and $\overline{\lambda} > 0$ iff $R_0 > 1$. This completes the proof. □

References

Arnold L (1998) Random dynamical systems. Springer, New York

Arnold L, Wihstutz V, Eckmann JP, Edited (1990) Lyapunov exponents. In: Proceedings of a conference held in Oberwolfach, May 28, 1990; Lecture notes in mathematics, vol 1486. Springer

Arnold L, Wihstutz V, Edited (1984) Lyapunov exponents. In: Proceedings of a workshop held in Brmen; lecture notes in mathematics, vol 1186. Springer, Berlin, November 12–15, 1984

Asih Tri Sri Noor, Lenhart Suzanne, Wise Steven, Aryati Lina, Adi-Kusumo F, Hardianti Mardiah S, Forde Jonathan (2016) The dynamics of HPV infection and cervical cancer cells. Bull Math Biol 78:4–20

Barnabas RV, Laukkanen P, Koskela P, Kontula O, Lehtinen M, Garnett GP (2006) Epidemiology of HPV 16 and cervical cancer in Finland and the potential impact of vaccination: mathematical modelling analyses. PLOS Med 3:0624–0632. https://doi.org/10.1371/journal.pmed.0030138

Bellet LR (2006) Ergodic properties of Markov processes. In: Open quantum systems II. Springer, Berlin, pp 1–39

Bosch FX, Manos MM, Munoz N, Sherman M, Jansen AM, Peto J et al (1995) Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. J Nat Cancer Inst 87:796–802

Bouvard V et al (2009) A review of human carcinogens-part B: biological agents. Lancet Oncol 10:321–322

Brown VI, White KAJ (2011) The role of optimal control in assessing themost cost-effective implementation of a vaccination programme: HPV as a case study. Math Biosci 231:126–134. https://doi.org/10.1016/j.mbs.2011.02.009123

Bumrungthai S et al (2023) Mathematical modelling of cervical precancerous lesion grade risk scores: linear regression analysis of cellular protein biomarkers and human papillomavirus E6/E7 RNA staining patterns. Diagnostics 13:1084. https://doi.org/10.3390/diagnostics13061084

Butel JS (2000) Viral carcinogenesis: revelation of molecular mechanisms and etiology of human disease. Carcinogenesis 21(3):405–26. https://doi.org/10.1093/carcin/21.3.405

Clifford GM, Smith JS, Aguado T, Franceschi S (2003) Comparison of HPV type distribution in high-grade cervical lesions and cervical cancer: a meta-analysis. Br J Cancer 89(101–105):6601024. https://doi.org/10.1038/sj.bjc

de Villiers EM (1994) Human pathogenic papillomavirus types: an update. Curr Top Microbiol Immunol 184:1–12

Dieu NT, Du NH, Nguyen HD, Yin G (2016) Protection zones for survival of species in random environment. SIAM J Appl Math 76(4):1382–1402

Du NH, Nguyen HD, Yin G (2016) Conditions for permanence and ergodicity of certain stochastic predator-prey models. J Appl Probab 53:187–202

Duan J (2015) An introduction to stochastic dynamics. Cambridge University Press, Cambridge

Elbasha EH (2008) Global stability of equilibria in a two-sex HPV vaccination model. Bull Math Biol 70:894–909. https://doi.org/10.1007/s11538-007-9283-0

Frazer IH, Cox JT, Mayezux EJ, Franco EL, Moscicki AB, Palefsky JM et al (2006) Advances in prevention of cervical cancer and other human papillomavirus-related diseases. Pediat Infect Dis J 25:S65–S81

Goldhaber-Fiebert JD, Stout NK, Salomon JA, Kuntz KM, Goldie SJ (2008) Cost-effectiveness of cervical cancer screening with human papillomavirus DNA testing and HPV-16,18 vaccination. J Natl Cancer Inst 100:308–320. https://doi.org/10.1093/jnci/djn019

Graham SV (2017) The human papillomavirus replication cycle, and its links to cancer progression: a comprehensive review. Clin Sci 131:2201–2221. https://doi.org/10.1042/CS20160786

Hassard BD, Kazarinoff ND, Wan Y-H (1981) Theory and applications of Hopf Bifurcation. Cambridge University Press, Cambridge

Hausen ZH (2001) Oncogenic DNA viruses. Oncogene 20:7820–7823

Hening A, Nguyen HD (2018) Coexistence and extinction for stochastic Kolmogorov systems. Ann Appl Probab 28:1893–1942

Herfs M et al (2012) A discrete population of squamocolumnar junction cells implicated in the pathogenesis of cervical cancer. PNAS 109(26):10516–10521

Hörmander L (1967) Hypoelliptic second order differential equations. Acta Math 119:147–171

Ikeda N, Watanabe S (1989) Stochastic differential equations and diffusion processes, 2nd edn. North-Holland, Amsterdam

Iskandar R, Taghavi K, Low N, Bramer WM, Egger M, Rohner E (2022) Mathematical models for evaluating effectiveness and cost-effectiveness of cervical cancer control policies in populations including women living with human immunodeficiency virus: a scoping review. Value Health Reg Issues 32:39–46

Jorgensen B (1982) Statistical property of the generalized inverse Gaussian distribution. Springer, New York
Jurdjevic V (1996) Geometric control theory. Cambridge University Press, Cambridge
Keller H (1996) Random attractors and and bifurcations of the stochastic Lorenz system. Technical Report 389. Institut für Dynamische Systeme, Universität Bremen
Khasminskii R (2012) Stochastic stability of differential equations. In: Stochastic modeling and applied probability, 2nd edn
Kim JJ, Brisson M, Edmunds WJ, Goldie SJ (2008) Modeling cervical cancer prevention in developed countries. Vaccine 26:K76–K86. https://doi.org/10.1016/j.vaccine.2008.06.009
Lee C, Laimins LA (2007) The differentiation-dependent life cycle of human papillomaviruses in keratinocytes. In: Garcea R, Di Maio D (eds) The papillomaviruses. Springer, New York, pp 45–67
Lee SL, Tameru AM (2012) A mathematical model of human papillomavirus (HPV) in the United States and its impact on cervical cancer. J Cancer 3:262–268. https://doi.org/10.7150/jca.4161
Mao X (1997) Stochastic differential equations and their applications. Horwood Publishing, Chichester
Meyn SP, Tweedie RL (1993) Stability of Markovian processes II: continuous-time processes and sampled chains. Adv Appl Probab 25:487–517
Meyn SP, Tweedie RL (1993) Stability of Markovian processes III: Foster–Lyapunov criteria for continuous-time processes. Adv Appl Probab 25:518–548
Munoz N, Castellsague X, de Gonzalez AB, Gissmann L (2006) Chapter 1: HPV in the etiology of human cancer. Vaccine 24:S1–S10
Murtono M, Ndii MZ, Sugiyanto S (2019) Mathematical model of cervical cancer treatment using chemotherapy drug. Biol Med Nat Prod Chem 8(1):11–15. https://doi.org/10.14421/biomedich.2019.81.11-15
Nualart D (2006) The Malliavin calculus and related topics. Springer, Berlin
Ostor AG (1993) Natural history of cervical intraepithelial neoplasia: a critical review. Int J Gynecol Pathol 12:186–192
Parkin DM (2006) The global health burden of infection-associated cancers in the year 2002. Int J Cancer 118:3030–3044
Phan TA, Tian JP (2020) Basic stochastic model for tumor virotherapy. Math Biosci Eng 17(4):4271–4294
Phan TA, Tian JP (2022) Hopf bifurcation without parameters in deterministic and stochastic modeling of cancer virotherapy, part II. J Math Anal Appl 515:126444. https://doi.org/10.1016/j.jmaa.2022.126444
Phan TA, Nguyen HD, Tian JP (2021) Deterministic and stochastic modeling for PDGF-driven gliomas reveals a classification of gliomas. J Math Biol 83:22
Phan TA, Tian JP, Wang B (2021) Dynamics of cholera epidemic models in fluctuating environments. Stoch Dyn 21(02):2150011. https://doi.org/10.1142/S0219493721500118
Plummer M et al (2016) Global burden of cancers attributable to infections in 2012: a synthetic analysis. Lancet Glob Health 4:e609–e616
Reingold AL (2000) Infectious Disease Epidemiology in the 21st Century: Will It Be Eradicated or Will It Reemerge? Epidemiol Rev 22:57–63
Ryser MD, Gravittc PE, Myersd ER (2017) Mechanistic mathematical models: an underused platform for HPV research. Papillomavirus Res 3:46–49
Shah KV, Howley PM (1996) Papillomavirus. In: Fields BN, Knipe DM, Howley PM (eds) Virology, 3rd edn. Lippincott-Raven Press Ltd, New York, pp 2077–2109
Sierra-Rojas JC, Reyes-Carreto R, Vargas-De-Leon C, Camacho JF (2022) Modeling and mathematical analysis of the dynamics of HPV in cervical epithelial cells: transient, acute, latency, and chronic infections. Comput Math Methods Med Article ID 8650071. https://doi.org/10.1155/2022/8650071
White MK, Pagano JS, Khalili K (2014) Viruses and human cancers: a long road of discovery of molecular paradigms. Clin Microbiol Rev 27(3):463–81. https://doi.org/10.1128/CMR.00124-13
Wright TC, Ferenczy A (2002) Anatomy and histology of the cervix, Blaustein's pathology of the female genital tract, 5th edn. Springer, New York, pp 207–224
Zapatka M, Borozan I, Brewer DS, Iskar M, Grundhoff A, Alawi M et al (2020) The landscape of viral associations in human cancers. Nat Genet 52:320–330
Ziyadi N (2017) A male–female mathematical model of human papillomavirus (HPV) in African American population. Math Biol Eng 14(1):339–358. https://doi.org/10.3934/mbe.2017022