



Global Dynamics of a Susceptible-Infectious-Recovered Epidemic Model with a Generalized Nonmonotone Incidence Rate

Min Lu¹ · Jicai Huang¹ · Shigui Ruan² · Pei Yu³

Received: 2 April 2019 / Revised: 27 May 2020 / Published online: 29 June 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

A susceptible-infectious-recovered (SIRS) epidemic model with a generalized nonmonotone incidence rate $\frac{kIS}{1+\beta I+\alpha I^2}$ ($\beta > -2\sqrt{\alpha}$ such that $1 + \beta I + \alpha I^2 > 0$ for all $I \geq 0$) is considered in this paper. It is shown that the basic reproduction number R_0 does not act as a threshold value for the disease spread anymore, and there exists a sub-threshold value R_* (< 1) such that: (i) if $R_0 < R_*$, then the disease-free equilibrium is globally asymptotically stable; (ii) if $R_0 = R_*$, then there is a unique endemic equilibrium which is a nilpotent cusp of codimension at most three; (iii) if $R_* < R_0 < 1$, then there are two endemic equilibria, one is a weak focus of multiplicity at least three, the other is a saddle; (iv) if $R_0 \geq 1$, then there is again a unique endemic equilibrium which is a weak focus of multiplicity at least three. As parameters vary, the model undergoes saddle-node bifurcation, backward bifurcation, Bogdanov–Takens bifurcation of codimension three, Hopf bifurcation, and degenerate Hopf bifurcation of codimension three. Moreover, it is shown that there exists a critical value α_0 for the psychological effect α , a critical value k_0 for the infection rate k , and two critical values β_0, β_1 ($\beta_1 < \beta_0$) for β that will determine whether the disease dies out or persists in the form of positive periodic coexistent oscillations or coexistent steady states under different initial populations. Numerical simulations are given to demonstrate the existence of one, two or three limit cycles.

Keywords SIRS epidemic model · Generalized nonmonotone incidence rate · Saddle-node bifurcation · Backward bifurcation · Bogdanov–Takens bifurcation of codimension three · Degenerate Hopf bifurcation of codimension three

Research was partially supported by NSFC (Nos. 11871235, 11771168), the Fundamental Research Funds for the Central Universities (CCNU19TS030) and NSERC (No. R2686A02).

✉ Pei Yu
pyu@uwo.ca

Extended author information available on the last page of the article

1 Introduction

Consider an infectious disease being transmitted person to person by direct or indirect contact in a population. The host population is divided into three classes $S(t)$, $I(t)$ and $R(t)$, where $S(t)$ denotes the number of individuals who are susceptible to the disease, that is, who are not yet infected at time t , $I(t)$ denotes the number of infectious individuals at time t who have been infected and are able to spread the disease by contact with susceptible individuals, $R(t)$ denotes the number of individuals who have been recovered from the infection and removed from the possibility of being infected again or of spreading at time t . Bifurcations in susceptible-infectious-recovered (SIRS) epidemic models with various incidence rates have been studied by many researchers, see for example, Alexander and Moghadas [1, 2], Capasso and Serio [4], Derrick and van den Driessche [7], Hethcote and van den Driessche [11], Hu et al. [12], Liu et al. [18], Li et al. [17], Ruan and Wang [22], Tang et al. [24], Xiao and Ruan [26], Xiao and Zhou [27], Zhou et al. [30]. In this paper we consider the following SIRS model with a generalized nonmonotone incidence rate:

$$\begin{aligned}\frac{dS}{dt} &= b - dS - \frac{kIS}{1 + \beta I + \alpha I^2} + \delta R, \\ \frac{dI}{dt} &= \frac{kIS}{1 + \beta I + \alpha I^2} - (d + \mu)I, \\ \frac{dR}{dt} &= \mu I - (d + \delta)R,\end{aligned}\tag{1.1}$$

under the initial conditions $S(0) \geq 0$, $I(0) \geq 0$, $R(0) \geq 0$, where $S(t)$, $I(t)$ and $R(t)$ denote the numbers of susceptible, infective, and recovered individuals at time t , respectively, $b > 0$ is the recruitment rate of the population, $d > 0$ is the natural death rate of the population, $\mu > 0$ is the natural recovery rate of the infected individuals, $\delta > 0$ is the rate at which recovered individuals lose immunity and return to the susceptible class. In the incidence rate function $\frac{kIS}{1 + \beta I + \alpha I^2}$, $k > 0$ is the infection rate and kI measures the infection force of disease, $\frac{1}{1 + \beta I + \alpha I^2}$ describes the inhibit effect from the behavioral change of the susceptible individuals when the number of infectious individuals increases, $\alpha > 0$ is a parameter which measures the psychological or inhibitory effect, and β is a parameter satisfying $\beta > -2\sqrt{\alpha}$ such that $1 + \beta I + \alpha I^2 > 0$ for all $I \geq 0$.

When studying the data on the cholera epidemic spread in Bari, Italy, in 1973, Capasso et al. [3] and Capasso and Serio [4] proposed a nonlinear incidence rate function taking psychological effects into account, which exhibits patterns showing in Fig. 1b, where the infection rate increases firstly and then decreases as the number of infected individuals increases. In fact, when a new infectious disease emerges, both the contact rate and the infection probability increase since people have very little knowledge about the disease. However, when number of infected individuals is getting larger and the disease becomes more serious, psychological factor leads people to modify their behavior and implement measures to reduce the opportunities of contact and the probability of infection. For instance, after the outbreaks of the avian influenza A (H5N1) virus, to determine the knowledge, attitudes and practices relating to avian influenza in the general populations, cross-sectional surveys conducted in Thailand and China show a high degree of awareness of human avian influenza in both urban and rural populations and a higher level of proper hygienic practice among urban residents by reducing their visits to live markets [19]. Also, during the outbreak of coronavirus disease 2019 (COVID-19), the aggressive measures and policies, such as border screening, mask wearing, quarantine, and isolation, were proved to

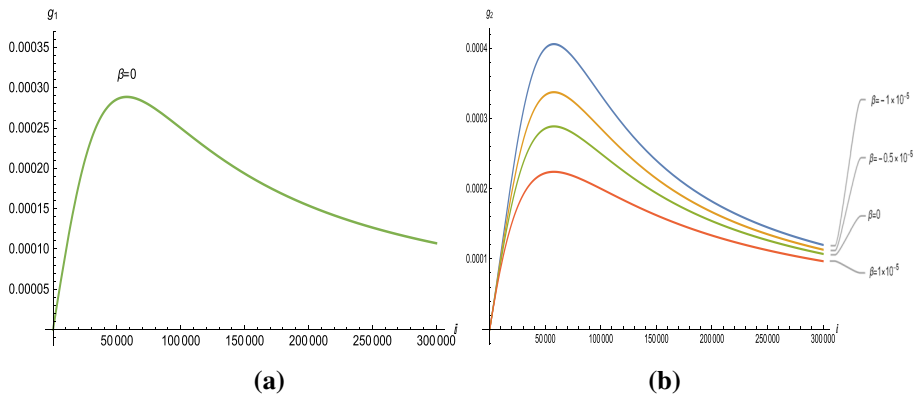


Fig. 1 **a** The graph of the nonmonotone infection force $g_1(I) = \frac{kI}{1+\alpha I^2}$; **b** the graphs of the generalized nonmonotone infection force $g_2(I) = \frac{kI}{1+\beta I+\alpha I^2}$ when $\alpha = 3 \times 10^{-10}$, $k = 1 \times 10^{-8}$

be very effective in reducing the further transmission of the disease. So with these behavioral changes the infection force decreases when the number of infected individuals becomes larger. But Capasso and Serio [4] did not give any specific function to describe such an incidence. Some scholars have proposed some functions to describe this kind of incidence rate, for example, Xiao and Ruan [26] proposed a nonmonotone incidence rate $\frac{kIS}{1+\alpha I^2}$ to take psychological effects into account.

The special case of model (1.1) when $\beta = 0$ was proposed and studied by Xiao and Ruan [26] who found that system (1.1) with $\beta = 0$ has an important threshold, i.e., the basic reproduction number R_0 , such that the disease-free equilibrium is globally asymptotically stable if $R_0 \leq 1$, and the unique endemic equilibrium is globally asymptotically stable if $R_0 > 1$. Tang et al. [24] conjectured that system (1.1) with $\beta = 0$ does not exhibit complex dynamics and bifurcations.

Model (1.1) with $\beta \neq 0$ was first considered by Xiao and Zhou [27]. Comparing the nonmonotone infection force $g_1(I) = \frac{kI}{1+\alpha I^2}$ (Fig. 1a) with the generalized nonmonotone infection force $g_2(I) = \frac{kI}{1+\beta I+\alpha I^2}$ (Fig. 1b) we can see that, when the infection number is small, $g_2(I)$ increases faster when $\beta < 0$ than when $\beta = 0$ (see Fig. 1). Moreover, for the same values of k and α , the maximum of $g_2(I)$ is bigger when $\beta < 0$ than when $\beta = 0$. We can also see that the infection force $g_2(I)$ is stronger when $\beta < 0$. Xiao and Zhou [27] presented qualitative analysis of model (1.1) and showed the existence of a cusp of codimension two, bistable phenomenon and periodic oscillations. Later, Zhou et al. [30] further studied the existence of different kinds of bifurcations in model (1.1), such as Bogdanov–Takens bifurcation of codimension two and Hopf bifurcation of codimension one, by choosing several sets of specific parameter values. Their results showed that model (1.1) with a generalized nonmonotone incidence rate can exhibit complex dynamics and bifurcations. However, several sets of parameter values that they chose to unfold the bifurcations may be not biologically meaningful.

In this paper, we continue to consider model (1.1) with a generalized nonmonotone incidence rate function. By considering general parameter conditions, not specific parameter values as chosen in Zhou et al. [30], we show that the basic reproduction number R_0 in model (1.1) does not act as a threshold value for the disease spread anymore, and there exists a sub-threshold value R_* (< 1) such that: (i) if $R_0 < R_*$, then the disease-free

equilibrium is globally asymptotically stable; (ii) if $R_0 = R_*$, then there is a unique endemic equilibrium which is a nilpotent cusp of codimension at most three; (iii) if $R_* < R_0 < 1$, then there are two endemic equilibria, one is a weak focus of multiplicity at least three, the other is a saddle; (iv) if $R_0 \geq 1$, then there is again a unique endemic equilibrium which a weak focus of multiplicity at least three. In case (iv), sufficient conditions to guarantee the globally asymptotically stability or the uniqueness of a limit cycle for the model are given. As the parameters vary, the model undergoes a sequence of bifurcations including saddle-node bifurcation, backward bifurcation, Bogdanov–Takens bifurcation of codimension three, Hopf bifurcation, and degenerate Hopf bifurcation of codimension three. Moreover, it is shown that there exists a critical value α_0 for the psychological effect α , a critical value k_0 for the infection rate k , and two critical values β_0, β_1 ($\beta_1 < \beta_0$) for β that will determine whether the disease dies out or persists in the form of positive periodic coexistent oscillations or coexistent steady states under different initial populations. More specifically, it is shown that: (i) when $\alpha > \alpha_0$, or $\alpha \leq \alpha_0$, $k < k_0$ and $\beta \geq \beta_0$, or $k = k_0$ and $\beta \geq \beta_0$, the disease will die out for all positive initial populations; (ii) when $\alpha = \alpha_0$ and $\beta_1 \leq \beta < \beta_0$, the disease will die out for almost all positive initial populations; (iii) when $\alpha = \alpha_0$ and $\beta < \beta_1$, the disease will persist in the form of a positive coexistent steady state for some positive initial populations; and (iv) when $\alpha < \alpha_0$, $k < k_0$ and $\beta < \beta_0$, or $k = k_0$ and $\beta < \beta_0$, or $k > k_0$, the disease will persist in the form of multiple positive periodic coexistent oscillations and coexistent steady states for some positive initial populations. Numerical simulations are presented to illustrate the theoretical results, including the existence of one, two or three limit cycles, bifurcation diagrams and corresponding phase portraits.

The rest of the paper is organized as follows. In Sect. 2, we summarize the qualitative analysis of the model, including the types and stability of equilibria. In Sect. 3, we present sufficient conditions to guarantee the nonexistence, existence and uniqueness of limit cycles of the model. In Sect. 4, we show that the model undergoes saddle-node bifurcation, backward bifurcation, Bogdanov–Takens bifurcation of codimensions two and three, Hopf bifurcation and degenerate Hopf bifurcation of codimension three. Numerical simulations are given to demonstrate the existence of one, two or three limit cycles in Sect. 5. A brief discussion is given in the last section.

2 Basic Properties of the Model

Adding up the three equations of (1.1), we have the following lemma.

Lemma 2.1 *The plane $S + I + R = \frac{b}{d}$ is an invariant manifold of system (1.1) which is attracting in the first octant.*

The limit set of system (1.1) is on the plane $S + I + R = \frac{b}{d}$ on which model (1.1) can be reduced to a two-dimensional system:

$$\begin{aligned} \frac{dI}{dt} &= \frac{kI}{1 + \beta I + \alpha I^2} \left(\frac{b}{d} - I - R \right) - (d + \mu)I, \\ \frac{dR}{dt} &= \mu I - (d + \delta)R. \end{aligned} \quad (2.1)$$

Letting

$$I = \frac{d + \delta}{k} x, \quad R = \frac{d + \delta}{k} y, \quad t = \frac{1}{d + \delta} \tau,$$

then system (2.1) can be rescaled as follows (we still denote τ by t)

$$\begin{aligned} \frac{dx}{dt} &= \frac{x}{1 + mx + nx^2} (A - x - y) - px \triangleq P(x, y), \\ \frac{dy}{dt} &= qx - y \triangleq Q(x, y), \end{aligned} \tag{2.2}$$

where

$$m = \frac{\beta(d + \delta)}{k}, \quad n = \frac{\alpha(d + \delta)^2}{k^2}, \quad A = \frac{bk}{d(d + \delta)}, \quad p = \frac{d + \mu}{d + \delta}, \quad q = \frac{\mu}{d + \delta}. \tag{2.3}$$

Moreover, $\mu, \delta, k, \alpha, \beta$ can be expressed by A, m, p, q, n, b and d as follow

$$\mu = \frac{dq}{p - q}, \quad \delta = \frac{d(1 - p + q)}{p - q}, \quad k = \frac{Ad^2}{b(p - q)}, \quad \alpha = \frac{A^2 d^2 n}{b^2}, \quad \beta = \frac{Adm}{b}. \tag{2.4}$$

Denote

$$\Gamma_1 = \{(A, m, p, q, n) : q < p < q + 1, m > -2\sqrt{n}, A, p, q, n > 0\}. \tag{2.5}$$

We can see that

$$\Omega = \{(x, y) | 0 \leq x \leq A, 0 \leq y \leq qA\}$$

is a positively invariant and bounded region for system (2.2) when $(A, m, p, q, n) \in \Gamma_1$.

System (2.2) always has an equilibrium $E_0 = (0, 0)$ which corresponds to the disease-free equilibrium $(\frac{b}{a}, 0, 0)$ of system (1.1). To find the endemic equilibria, we set

$$\frac{x}{1 + mx + nx^2} (A - x - y) - px = 0, \quad qx - y = 0, \tag{2.6}$$

which yield

$$pnx^2 + (1 + q + pm)x + p - A = 0. \tag{2.7}$$

The discriminant of (2.7) is

$$\Delta = (1 + q + pm)^2 - 4pn(p - A).$$

We can see that (2.7) has at most two roots x^- and x^+ , which may coalesce into a unique root x_1 , where

$$x^- = \frac{-(1 + q + pm) - \sqrt{\Delta}}{2pn}, \quad x^+ = \frac{-(1 + q + pm) + \sqrt{\Delta}}{2pn}, \quad x_1 = -\frac{1 + q + pm}{2pn}. \tag{2.8}$$

By using the procedure in van den Driessche and Watmough [25] and the existence of positive equilibria, we find the basic reproduction number and a sub-threshold value as follows:

$$R_0 = \frac{A}{p}, \quad R_* = 1 - \frac{(1 + q + pm)^2}{4np^2}. \tag{2.9}$$

Clearly, $R_* < 1$ and $\Delta \geq 0$ is equivalent to $R_0 \geq R_*$.

By analyzing the existence of positive roots to (2.7), we obtain the following results.

Theorem 2.2 *Model (2.2) always has a boundary equilibrium $E_0(0, 0)$. Moreover,*

- (I) *when $R_0 < 1$, we have*
 - (i.1) *if $R_0 < R_*$, then system (2.2) has no positive equilibria;*
 - (i.2) *if $R_0 = R_*$ and $m < -\frac{1+q}{p}$, then system (2.2) has a unique positive equilibrium $E_1(x_1, y_1)$, where $y_1 = qx_1$;*
 - (i.3) *if $R_0 > R_*$ and $m < -\frac{1+q}{p}$, then system (2.2) has two positive equilibria $E_2(x_2, y_2)$ and $E_3(x_3, y_3)$, where $x_2 = x^-$, $y_2 = qx_2$; $x_3 = x^+$, $y_3 = qx_3$;*
- (II) *when $R_0 = 1$ and $m < -\frac{1+q}{p}$, then system (2.2) has a unique positive equilibrium $E_4(x_4, y_4)$, where $x_4 = x^+$, $y_4 = qx_4$;*
- (III) *when $R_0 > 1$, then system (2.2) has a unique positive equilibrium $E_5(x_5, y_5)$, where $x_5 = x^+$, $y_5 = qx_5$.*

Next we study the locally stability of equilibria of system (2.2). Combining the results in Xiao and Zhou [27], we have the following lemma.

Lemma 2.3 *The boundary equilibrium $E_0(0, 0)$ of system (2.2) is*

- (i.1) *a hyperbolic stable node if $R_0 < 1$;*
- (i.2) *a saddle-node if $R_0 = 1$ and $m \neq -\frac{1+q}{p}$. Moreover, a stable parabolic sector lies in the right (or left) half plane of \mathbb{R}^2 if $m > -\frac{1+q}{p}$ (or $m < -\frac{1+q}{p}$);*
- (i.3) *a degenerate stable node if $R_0 = 1$ and $m = -\frac{1+q}{p}$;*
- (i.4) *a hyperbolic saddle if $R_0 > 1$.*

To determine when the disease cannot persist, we have to study the global stability of the equilibrium $(\frac{d}{b}, 0, 0)$. Next, we improve Theorem 2.5 in Xiao and Zhou [27] for the global stability of $(\frac{d}{b}, 0, 0)$ in the interior \mathbb{R}_+^3 .

Theorem 2.4 *The disease-free equilibrium $(\frac{d}{b}, 0, 0)$ of model (1.1) is globally asymptotically stable in the interior \mathbb{R}_+^3 , and the disease cannot invade the population if one of the following conditions holds:*

- (i.1) $R_0 < R_*$;
- (i.2) $R_* \leq R_0 < 1$ and $m \geq -\frac{1+q}{p}$;
- (i.3) $R_0 = 1$ and $m \geq -\frac{1+q}{p}$.

Proof Since Lemma 2.1 implies that the stability of the disease-free equilibrium $(\frac{d}{b}, 0, 0)$ of system (1.1) in the interior \mathbb{R}_+^3 is equivalent to that of equilibrium $E_0(0, 0)$ of system (2.2) in \mathbb{R}_+^2 , we only discuss the stability of equilibrium $E_0(0, 0)$ of system (2.2) in \mathbb{R}_+^2 . The phase portraits are shown in Fig. 2.

Since Ω is a positively invariant region, $x = 0$ is an invariant line and by index theory, we can obtain that system (2.2) has no nontrivial periodic orbits in \mathbb{R}_+^2 when system (2.2) has no positive equilibria. From Theorem 2.2 and Lemma 2.3, we know that system (2.2) only has a hyperbolic stable node $E_0(0, 0)$ in \mathbb{R}_+^2 when the condition (i.1) or (i.2) is satisfied (see Fig. 2a). Thus, $E_0(0, 0)$ is a global attractor in \mathbb{R}_+^2 . Hence, the condition (i.1) or (i.2) guarantees that $(\frac{d}{b}, 0, 0)$ of model (1.1) is globally asymptotically stable in the interior \mathbb{R}_+^3 .

When $R_0 = 1$ and $m = -\frac{1+q}{p}$, from Theorem 2.2 and Lemma 2.3, we know that system (2.2) only has a degenerate stable node $E_0(0, 0)$ in \mathbb{R}_+^2 . When $R_0 = 1$ and $m > -\frac{1+q}{p}$, it follows from Theorem 2.2 and Lemma 2.3 that system (2.2) has a unique equilibrium $E_0(0, 0)$ in \mathbb{R}_+^2 , and $E_0(0, 0)$ is a saddle-node, which has a stable parabolic sector in the right half plane of \mathbb{R}^2 (see Fig. 2b). We obtain the conclusion by the same arguments as above. \square

Remark 2.5 We can see that $R_0 \leq 1$ is equivalent to $k \leq k_0$, $m \geq -\frac{1+q}{p}$ is equivalent to $\beta \geq \beta_0$, and $R_0 \leq R_*$ is equivalent to $\alpha \geq \alpha_0$, where

$$\begin{aligned} k_0 &\triangleq \frac{d(\mu + d)}{b}, \\ \beta_0 &\triangleq -k \frac{\mu + d + \delta}{(d + \delta)(d + \mu)}, \\ \alpha_0 &\triangleq \frac{d[k(\mu + \delta + d) + \beta(d + \mu)(d + \delta)]^2}{4(d + \delta)^2(d + \mu)(d(d + \mu) - bk)} = \frac{d(\beta - \beta_0)^2(d + \mu)}{4b(k_0 - k)}. \end{aligned} \tag{2.10}$$

Lemma 2.3 and Theorem 2.4 indicate that the disease will die out for all positive initial populations if one of the following cases holds

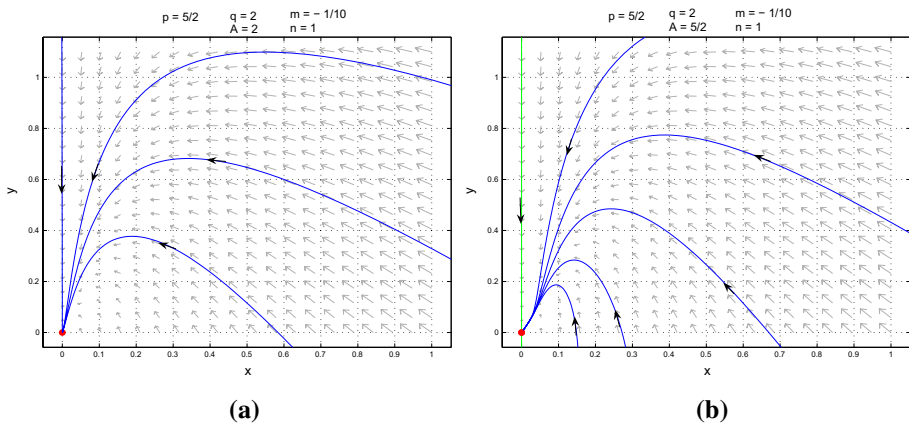


Fig. 2 The phase portraits of system (2.2) with no positive equilibria. **a** A hyperbolic stable node $E_0(0, 0)$ when $R_* < R_0 < 1$ and $m > -\frac{1+q}{p}$; **b** a saddle-node $E_0(0, 0)$ with a stable parabolic sector in the right half plane when $R_0 = 1$ and $m > -\frac{1+q}{p}$

- (i.1) $\alpha > \alpha_0$, i.e., the psychological or inhibitory effect α is greater than a critical value α_0 ;
- (i.2) $\alpha \leq \alpha_0, k < k_0, \beta \geq \beta_0$, i.e., the infection rate k is smaller than the critical value k_0 and the psychological effect α is smaller than or equal to the critical value α_0 , but β is greater than or equal to a critical value β_0 ;
- (i.3) $k = k_0, \beta \geq \beta_0$, i.e., the infection rate k is equal to the critical value k_0 , but β is greater than or equal to the critical value β_0 .

Next we consider the stability of the positive equilibria of system (2.2). The Jacobian matrix of system (2.2) at a positive equilibria $E(x, y)$ is given by

$$J(E) = \begin{pmatrix} -\frac{x(1+(m+2nx)p)}{1+mx+nx^2} & \frac{-x}{1+mx+nx^2} \\ q & -1 \end{pmatrix}.$$

The determinant of $J(E)$ is

$$\det(J(E)) = \frac{x(1 + q + pm + 2npx)}{1 + mx + nx^2},$$

its sign is determined by

$$S_D = 1 + q + pm + 2npx. \tag{2.11}$$

The trace of $J(E)$ is

$$\text{tr}(J(E)) = \frac{-n(1 + 2p)x^2 - (1 + m + mp)x - 1}{1 + mx + nx^2},$$

its sign is determined by

$$S_T = -n(1 + 2p)x^2 - (1 + m + mp)x - 1. \tag{2.12}$$

In the following, we classify topological types of the positive equilibria of system (2.2). Define

$$A_* = \frac{2p(1 + q + pq)}{1 - mp + q + 2pq}, \quad m_* = -\frac{1 + p + q + 2pq}{p^2}. \tag{2.13}$$

Theorem 2.6 *When $(A, m, p, q, n) \in \Gamma_1, R_0 = R_*$ and $m < -\frac{1+q}{p}$, system (2.2) has a unique positive equilibrium $E_1(x_1, y_1)$ and no closed orbits in Ω . Moreover,*

- (I) *if $A \neq A_*$, then $E_1(x_1, y_1)$ is a saddle-node, which includes a stable parabolic sector (or an unstable parabolic sector) if $A > A_*$ (or $A < A_*$);*
- (II) *if $A = A_*$, then $E_1(x_1, y_1)$ is a cusp. Moreover,*

- (i.1) *if $m \neq m_*$, then $E_1(x_1, y_1)$ is a cusp of codimension two;*
- (i.2) *if $m = m_*$, then $E_1(x_1, y_1)$ is a cusp of codimension three.*

The phase portraits are shown in Fig. 3.

Proof Substituting the value of x_1 and $R_0 = R_*$ into S_D and S_T , which are given in (2.11) and (2.12), respectively, then by a direct calculation, we can deduce that $S_D(x_1) = 0$, and

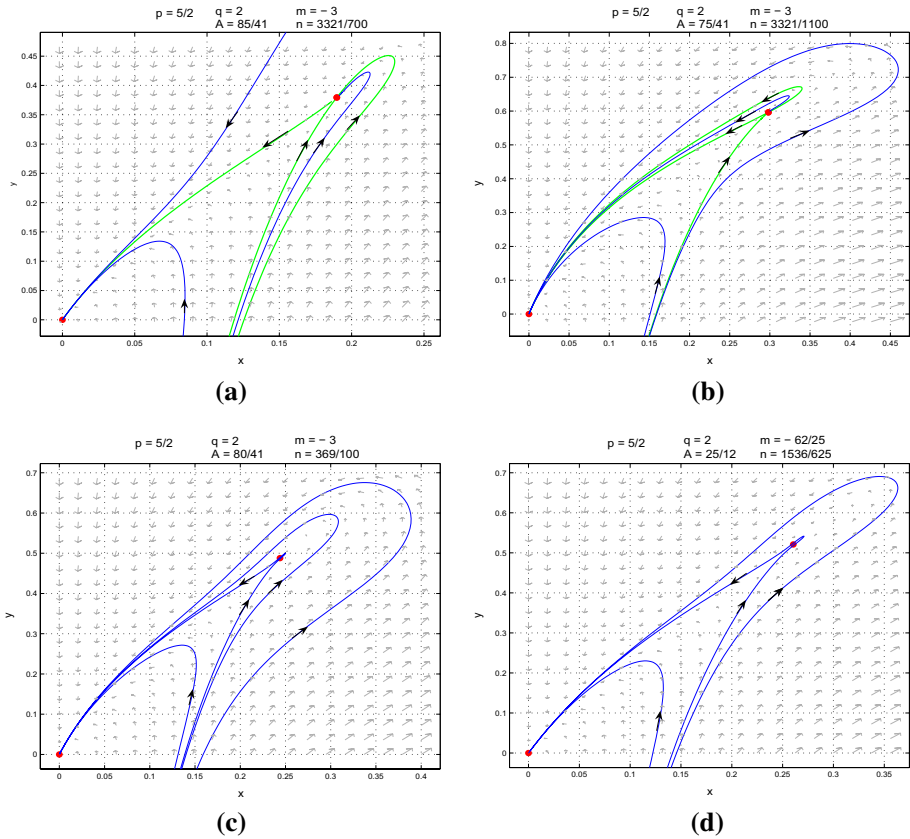


Fig. 3 Types of the unique positive equilibrium $E_1(x_1, y_1)$ of system (2.2): **a** A saddle-node with a stable parabolic sector when $A > A_*$; **b** a saddle-node with an unstable parabolic sector when $A < A_*$; **c** a cusp of codimension two when $A = A_*$ and $m \neq m_*$; **d** a cusp of codimension three when $A = A_*$ and $m = m_*$.

$$S_T(x_1) = \frac{-2p(1 + q + pq) + A(1 - mp + q + 2pq)}{p(1 + q + pm)}.$$

Letting $S_T(x_1) = 0$, we have

$$A = \frac{2p(1 + q + pq)}{1 - mp + q + 2pq}.$$

The assertions (I) and (i.1) of (II) are proved in [27] (see Lemma 2.7).

Next we prove the assertion (i.2) of (II). Let $X = x - x_1$, $Y = y - y_1$, and use Taylor expansion, system (2.2) can be rewritten as (for simplicity, we still denote X, Y by x, y , respectively)

$$\begin{aligned} \frac{dx}{dt} &= x - \frac{1}{q}y + a_{20}x^2 + a_{11}xy + a_{30}x^3 + a_{21}x^2y + a_{40}x^4 + a_{31}x^3y + o(|x, y|^4), \\ \frac{dy}{dt} &= qx - y, \end{aligned} \tag{2.14}$$

where

$$\begin{aligned} a_{20} &= \frac{(1 + 2p)(1 + q + pq)^2}{4p^3q}, & a_{11} &= -\frac{(1 + 2p)(1 + q + pq)^2}{2p^3q^2}, \\ a_{30} &= -\frac{(1 + 2p)^2(pq - 1)(1 + q + pq)^3}{8p^6q^2}, & a_{21} &= -\frac{(1 + 2p)^2(2 + q)(1 + q + pq)^3}{8p^6q^3}, \\ a_{40} &= -\frac{(1 + 2p)^3(1 + q + pq)^4[-2 + (-1 + 2p)q + (1 + 3p)q^2]}{32p^9q^3}, \\ a_{31} &= \frac{(1 + 2p)^3(1 + q + pq)^4(-2 - 2q + pq^2)}{16p^9q^4}, \end{aligned}$$

where A , n and m have been eliminated by using the equations $A = A_*$, $R_0 = R_*$ and $m = m_*$, respectively.

The following transformation

$$\begin{aligned} X &= x, \\ Y &= x - \frac{1}{q}y + a_{20}x^2 + a_{11}xy + a_{30}x^3 + a_{21}x^2y \\ &\quad + a_{40}x^4 + a_{31}x^3y + o(|x, y|^4), \end{aligned}$$

brings (2.14) into (we still denote X, Y by x, y , respectively)

$$\begin{aligned} \frac{dx}{dt} &= y, \\ \frac{dy}{dt} &= b_{20}x^2 + b_{02}y^2 + b_{30}x^3 + b_{21}x^2y \\ &\quad + b_{12}xy^2 + b_{40}x^4 + b_{31}x^3y + b_{22}x^2y^2 + o(|x, y|^4), \end{aligned} \tag{2.15}$$

where

$$\begin{aligned} b_{20} &= -\frac{(1 + 2p)(1 + q + pq)^2}{4p^3q}, & b_{02} &= \frac{(1 + 2p)(1 + q + pq)^2}{2p^3q}, \\ b_{30} &= -\frac{(1 + 2p)^2(1 + q + pq)^4}{8p^6q^2}, \\ b_{21} &= -\frac{(1 + 2p)^3(1 + q + pq)^3}{8p^6q}, & b_{12} &= -\frac{(1 + 2p)^2(pq - 1)(1 + q + pq)^3}{4p^6q^2}, \\ b_{40} &= -\frac{(1 + 2p)^3(2 + q)(1 + q + pq)^5}{32p^9q^3}, & b_{31} &= -\frac{(1 + 2p)^4(1 + q)(1 + q + pq)^4}{16p^9q^2}, \\ b_{22} &= \frac{(1 + 2p)^3(1 + q + pq)^4[2 + q - 2pq + (-1 - 2p + 2p^2)q^2]}{16p^9q^3}. \end{aligned}$$

In addition, let $dt = (1 - b_{02}x)d\tau$. Then system (2.15) becomes (we still denote τ by t)

$$\begin{aligned} \frac{dx}{dt} &= y(1 - b_{02}x), \\ \frac{dy}{dt} &= (1 - b_{02}x)(b_{20}x^2 + b_{02}y^2 + b_{30}x^3 + b_{21}x^2y + b_{12}xy^2 + b_{40}x^4 + b_{31}x^3y + b_{22}x^2y^2 \\ &\quad + o(|x, y|^4)). \end{aligned} \tag{2.16}$$

Next let $X = x, Y = y(1 - b_{02}x)$. Then system (2.16) is transformed into (we still denote X, Y by x, y , respectively)

$$\begin{aligned} \frac{dx}{dt} &= y, \\ \frac{dy}{dt} &= b_{20}x^2 + (b_{30} - 2b_{20}b_{02})x^3 + (b_{20}b_{02}^2 - 2b_{02}b_{30} + b_{40})x^4 + b_{21}x^2y \\ &\quad + (b_{31} - b_{02}b_{21})x^3y + (b_{12} - b_{02}^2)xy^2 + (b_{22} - b_{02}^3)x^2y^2 + o(|x, y|^4). \end{aligned} \tag{2.17}$$

Finally, noticing $b_{20} < 0$ and letting $X = -x, Y = -\frac{y}{\sqrt{-b_{20}}}$ and $\tau = \sqrt{-b_{20}}t$, we obtain (we still denote X, Y and τ by x, y and t , respectively)

$$\begin{aligned} \frac{dx}{dt} &= y, \\ \frac{dy}{dt} &= x^2 - \frac{b_{30} - 2b_{20}b_{02}}{b_{20}}x^3 + \frac{b_{20}b_{02}^2 - 2b_{02}b_{30} + b_{40}}{b_{20}}x^4 + y\left(\frac{b_{21}}{\sqrt{-b_{20}}}x^2 - \frac{b_{31} - b_{02}b_{21}}{\sqrt{-b_{20}}}x^3\right) \\ &\quad + y^2[(b_{12} - b_{02}^2)x - (b_{22} - b_{02}^3)x^2] + o(|x, y|^4). \end{aligned} \tag{2.18}$$

By Proposition 5.3 in Lamontagne et al. [16] (see also Lemma 2 in Huang et al. [13]), we obtain the equivalent system of (2.18) as follows:

$$\begin{aligned} \frac{dx}{dt} &= y, \\ \frac{dy}{dt} &= x^2 + Gx^3y + o(|x, y|^4). \end{aligned} \tag{2.19}$$

where

$$G = \frac{(1 + 2p)^3(1 + q)(1 + q + pq)^3\sqrt{(1 + 2p)pq}}{8p^8q^2}.$$

Because $p, q > 0$, we have $G > 0$. So $E_1(x_1, y_1)$ is a cusp of codimension three.

Nonexistence of limit cycles in Ω comes from the index theory. More precisely, from the previous arguments, system (2.2) has a unique equilibrium $E_1(x_1, y_1)$ in Ω which is a saddle-node or a cusp whose index is not one. Hence, it is impossible to have any limit cycle in Ω when $(A, m, p, q, n) \in \Gamma_1$ and $R_0 = R_*, m < -\frac{1+q}{p}$. This completes the proof. \square

Remark 2.7 Zhou et al. [30] showed that system (2.2) has a cusp of codimension three by choosing a set of specific parameter values, but their chosen parameter values are not biologically meaningful since they do not satisfy $p > q$.

Remark 2.8 When the psychological effect α is equal to α_0 given by (2.10), Theorem 2.6 implies that whether the disease persists or dies out will depend on β and the initial populations. More precisely, when $\alpha = \alpha_0$ and $\beta < \beta_0$ given by (2.10), i.e., $R_0 = R_*$ and $m < -\frac{1+q}{p}$, then system (1.1) has two equilibria, a disease-free equilibrium and an endemic equilibrium. Moreover, the disease will persist in the form of a positive steady state for some positive initial populations if β is smaller than a smaller critical value ($\beta < \beta_1$, i.e. $A > A_*$), given by

$$\beta_1 \triangleq \beta_0 - \frac{2(k_0 - k) [(d + \delta)^2 + \mu(\mu + \delta + 2d)]}{(d + \delta)^2(d + \mu)}, \tag{2.20}$$

(see Fig. 3a). Otherwise, the disease will die out for almost all positive initial populations if $\beta \geq \beta_1$ (i.e., $A \leq A_*$) (see Fig. 3b–d).

Let

$$A_i = \frac{2p^2 + (1 + p + mp^2 + q + 2pq)x_i}{1 + 2p} \quad (i = 2, 3, 4, 5). \tag{2.21}$$

Theorem 2.9 *If equilibria $E_i(x_i, y_i)$ ($i=2,3,4,5$) exist, then $E_2(x_2, y_2)$ is a hyperbolic saddle. Moreover,*

- (i.1) $E_i(x_i, y_i)$ ($i=3,4,5$) is a hyperbolic unstable node or focus if $A < A_i$;
- (i.2) $E_i(x_i, y_i)$ ($i=3,4,5$) is a hyperbolic stable node or focus if $A > A_i$;
- (i.3) $E_i(x_i, y_i)$ ($i=3,4,5$) is a weak focus or center if $A = A_i$.

Proof To investigate the type of $E_i(x_i, y_i)$, it suffices to consider the signs of S_D and S_T , where S_D and S_T are given in (2.11) and (2.12), respectively, from which we have

$$S_D(x_i) = 1 + q + pm + 2npx_i.$$

Since x^- is the smaller root of (2.7), x^+ is the larger root of (2.7), we obtain

$$2pnx^- + 1 + q + pm < 0, \quad 2pnx^+ + 1 + q + pm > 0.$$

Then it is easy to get $S_D(x_2) < 0, S_D(x_i) > 0, i = 3, 4, 5$. On the other hand,

$$\begin{aligned} S_T(x_i) &= -n(1 + 2p)x_i^2 - (1 + m + mp)x_i - 1 \\ &= \frac{2p^2 - (1 + 2p)A + (1 + p + mp^2 + q + 2pq)x_i}{p}, \end{aligned}$$

it is easy to see that $S_T(x_i) > 0, S_T(x_i) = 0$ and $S_T(x_i) < 0$ if $A < A_i, A = A_i$, and $A > A_i$, respectively, leading to the conclusions. □

3 Nonexistence, Existence and Uniqueness of Limit Cycles

Now we consider the nonexistence, existence and uniqueness of limit cycles of system (2.2).

Theorem 3.1 *Suppose $(A, m, n, p, q) \in \Gamma_1$ and $m > m_1$, then system (2.2) does not have non-trivial periodic orbits in the interior of \mathbb{R}_+^2 , where*

$$m_1 = \frac{-1 - 2\sqrt{n(1+2p)}}{1+p}. \quad (3.1)$$

Proof Taking a Dulac function

$$D(x, y) = \frac{1 + mx + nx^2}{x},$$

we have

$$\frac{\partial(DP)}{\partial x} + \frac{\partial(DQ)}{\partial x} = \frac{-n(1+2p)x^2 - (1+m+mp)x - 1}{x},$$

and the discriminant for $-n(1+2p)x^2 - (1+m+mp)x - 1$ is

$$\Delta_1 = (1+m+mp)^2 - 4n(1+2p) = m^2(1+p)^2 + 2m(1+p) + 1 - 4n(1+2p).$$

The discriminant for $\Delta_1 = 0$ is

$$\Delta_2 = 16n(1+2p)(1+p)^2.$$

Clearly, $\Delta_2 > 0$, and $\Delta_1 < 0$ if $m_1 < m < m_2$, where

$$m_1 = \frac{-1 - 2\sqrt{n(1+2p)}}{1+p}, \quad m_2 = \frac{-1 + 2\sqrt{n(1+2p)}}{1+p}.$$

The above discussions imply that

$$\frac{\partial(DP)}{\partial x} + \frac{\partial(DQ)}{\partial x} < 0$$

for $x > 0$ if

$$m \geq -\frac{1}{1+p} \quad \text{or} \quad m_1 < m < m_2,$$

that is, $m > m_1$ since $m_1 < -\frac{1}{1+p} < m_2$. \square

Remark 3.2 When $m = 0$ in system (2.2) (i.e., $\beta = 0$ in system (2.1)), from Theorems 2.2, 2.4 and 3.1, we can see that: **(I)** if $R_0 \leq 1$, then system (2.2) has no positive equilibria, and the unique boundary equilibrium $E_0(0, 0)$ (i.e., the disease-free equilibrium $(\frac{d}{b}, 0, 0)$ of system (1.1)) is globally asymptotically stable (by Theorem 2.4); **(II)** if $R_0 > 1$, then system (2.2) has a unique positive equilibrium E_5 which is globally asymptotically stable (by Theorem 3.1) since the unique boundary equilibrium E_0 is a hyperbolic saddle. The above two results when $\beta = 0$ in system (1.1) are shown in Xiao and Ruan [26], thus our

results when $\beta > -2\sqrt{\alpha}$ in system (1.1) can be seen as a generalization of those in Xiao and Ruan [26] by considering a more general system.

Remark 3.3 From Theorems 2.2, 3.1 and Lemma 2.3, we can also see that: **(I)** if $R_0 = 1$ and $m_1 < m < -\frac{1+q}{p}$, then the unique positive equilibrium E_4 of system (2.2) is globally asymptotically stable since the unique boundary equilibrium E_0 is a saddle-node with a stable parabolic sector lies in the left half plane of \mathbb{R}^2 and Ω is a positive invariant and bounded region; **(II)** if $R_0 > 1$ and $m > m_1$, then the unique positive equilibrium E_5 is globally asymptotically stable since the unique boundary equilibrium E_0 is a hyperbolic saddle and Ω is a positive invariant and bounded region. We use simulations to illustrate the results. We take a set of parameters values: $p = \frac{5}{2}$, $A = \frac{5}{2}$, $q = 2$, $n = 3$, $m = -2$ so that $R_0 = 1$ and $m > m_1$ for case **(I)**, as shown in Fig. 4a: E_4 is globally asymptotically stable; we take another set of parameters values: $p = \frac{5}{2}$, $A = 3$, $q = 2$, $n = 3$, $m = -2$ so that $R_0 > 1$ and $m > m_1$ for case **(II)**, as shown in Fig. 4b: E_5 is globally asymptotically stable.

We next investigate the existence and uniqueness of a limit cycle of system (2.2). Our method is to convert system (2.2) into a Liénard system and then apply Theorem 1.1 in Kooij and Zegeling [15], which is a modification of Z.-F. Zhang’s Theorem [28]. Firstly, we rewrite system (2.2) as

$$\frac{dx}{dt} = g_0(x) - g_1(x)y, \quad \frac{dy}{dt} = qx - y, \tag{3.2}$$

where

$$g_0(x) = \frac{x(A - p - (1 + mp)x - np x^2)}{1 + mx + nx^2}, \quad g_1(x) = \frac{x}{1 + mx + nx^2}. \tag{3.3}$$

Note that $g_1(x) > 0$ for all $x > 0$. Let $dt = -\frac{1}{g_1(x)} d\tau$ (we still denote τ by t). Then system (3.2) can be rewritten as

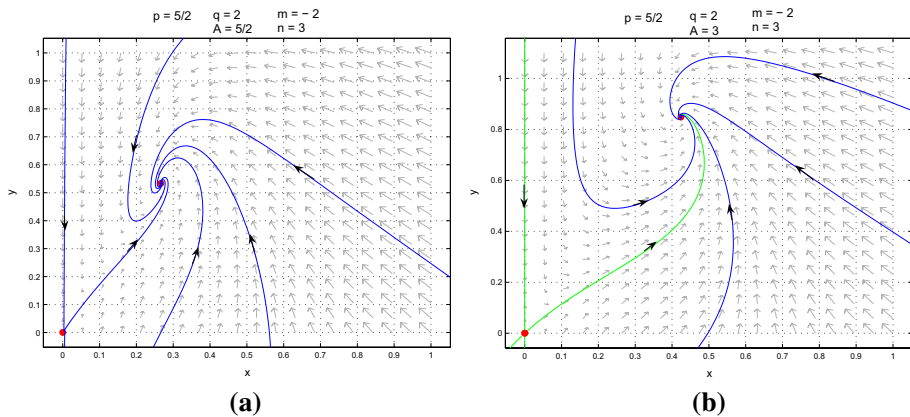


Fig. 4 Global stability of system (2.2): **a** a unique positive equilibrium E_4 is globally asymptotically stable if $R_0 = 1$ and $m_1 < m < -\frac{1+q}{p}$; **b** a unique positive equilibrium E_5 is globally asymptotically stable if $R_0 > 1$ and $m_1 < m < -\frac{1+q}{p}$

$$\frac{dx}{dt} = y - \frac{g_0(x)}{g_1(x)}, \quad \frac{dy}{dt} = \frac{y - qx}{g_1(x)}. \tag{3.4}$$

System (3.4) has the same qualitative properties as that of system (3.2) except the directions of trajectories. With the following transformation,

$$X = x, \quad Y = y - \int_{x^+}^x \frac{1}{g_1(u)} du, \tag{3.5}$$

system (3.4) is reduced to the following Liénard system (we still denote X, Y as x, y , respectively)

$$\frac{dx}{dt} = y - F(x), \quad \frac{dy}{dt} = -g(x), \tag{3.6}$$

where

$$\begin{aligned} F(x) &= \frac{g_0(x)}{g_1(x)} - \int_{x^+}^x \frac{1}{g_1(u)} du, \quad f(x) = F'(x) = \frac{-n(1 + 2p)x^2 - (1 + m + mp)x - 1}{x}, \\ g(x) &= \frac{qxg_1(x) - g_0(x)}{g_1^2(x)} = \frac{1 + mx + nx^2}{x} [npx^2 + (1 + q + mp)x + p - A], \\ \left(\frac{f(x)}{g(x)}\right)' &= \frac{h(x)}{[npx^2 + (1 + q + mp)x + p - A]^2 (1 + mx + nx^2)^2}, \end{aligned} \tag{3.7}$$

in which

$$\begin{aligned} h(x) &= l_0 + l_1x + l_2x^2 + l_3x^3 + l_4x^4 + l_5x^5, \\ l_0 &= 1 + q + mp + (1 + mp)(A - p), \\ l_1 &= 4np(A - p) + 2m(1 + q + mp) + 2np, \\ l_2 &= (1 + q + 2mp)[m(1 + m + mp) + n(2 - p)] + n[mp(A - p + 4) + p - A - pq], \\ l_3 &= 2n(1 + q + mp)(1 + m + mp) + 4n^2p, \\ l_4 &= n^2[(2p + 1)(1 + q + mp) + 3p(1 + m + mp) + mp(1 + 2p)], \\ l_5 &= 2n^3p(1 + 2p). \end{aligned} \tag{3.8}$$

Now we state and prove the existence and uniqueness of a limit cycle in system (2.2).

Theorem 3.4 *Suppose that $(A, m, p, q, n) \in \Gamma_1$ and $m \leq m_1$. Then system (2.2) has at most one closed orbit in \mathbb{R}_+^2 if $S_T(x^+)h(x) < 0$ (i.e., $f(x^+)\left(\frac{f(x)}{g(x)}\right)' < 0$) for $x \neq x^+$. Moreover, the closed orbit is hyperbolic if it exists. Here, $S_T(x)$, x^- and x^+ are given in (2.12) and (2.8), respectively.*

Proof By Theorem 2.2, we only need to consider the following two cases: (i) $1 > R_0 > R_*$ and $m < -\frac{1+q}{p}$; (ii) $R_0 \geq 1$.

When $1 > R_0 > R_*$ and $m < -\frac{1+q}{p}$, the limit cycles of system (2.2) (if exist) must lie in the stripe region between the vertical lines $x = x^-$ and $x = A$, since on $x = x^-$ the derivative

$$\begin{aligned} \left. \frac{dx}{dt} \right|_{x=x^-} &= \frac{x^-}{1 + mx^- + n(x^-)^2} [(A - p - (1 + mp)x^- - np x^{-2}) - y] \\ &= \frac{x^-}{1 + mx^- + n(x^-)^2} (qx^- - y) \neq 0 \end{aligned}$$

except at the saddle $E_2(x_2, y_2) = E_2(x^-, qx^-)$, and the positive invariant region is $\Omega = \{(x, y) \mid 0 \leq x \leq A, 0 \leq y \leq qA\}$. Let $S(x^-, A)$ denote the stripe region, since the transformation between (3.4) and (3.6) does not change x , it suffices to discuss (3.6) in $S(x^-, A)$ (i.e., $x^- < x < A$).

Transformation (3.5) is one-to-one for $(x, y) \in \mathbb{R}_+^2$, so it is equivalent to discuss the uniqueness of closed orbits for system (3.6) where $x > 0$. Corresponding to equilibrium $E(x^+, qx^+)$ (i.e., $E_i(x_i, y_i)$, $i = 3, 4, 5$, of system (2.2)), system (3.6) has an equilibrium (x^+, qx^+) . It is easy to see that

$$g(x^+) = \frac{1 + mx^+ + nx^{+2}}{x^+} [np x^{+2} + (1 + q + mp)x^+ + p - A] = 0, \quad (x - x^+)g(x) > 0$$

in $S(x^-, A)$. The other conditions given in Kooij and Zegeling [15] (see Theorem 1.1) can be directly verified.

When $R_0 \geq 1$, we simply replace $x = x^-$ by $x = 0$ and then use a similar argument. The proof is complete. □

Remark 3.5 From Theorems 2.2 and 3.4, we can see that system (2.2) has a unique limit cycle if one of the following conditions holds: (i) $R_0 = 1$, $m < \max\{m_1, -\frac{1+q}{p}\}$, $S_T(x^+) > 0$ and $h(x) < 0$ for $0 < x < A$; (ii) $R_0 > 1$, $S_T(x^+) > 0$, $m \leq m_1$ and $h(x) < 0$ for $0 < x < A$.

4 Bifurcation Analysis

From Theorems 2.2, 2.6 and 2.9, we know that system (2.2) may exhibit saddle-node bifurcation, backward bifurcation, Bogdanov–Takens bifurcation around $E_1(x_1, y_1)$, and Hopf bifurcation around $E_i(x_i, y_i)$ ($i = 3, 4, 5$). In this section, we investigate various possible bifurcations in system (2.2).

4.1 Saddle-Node Bifurcation

From Theorems 2.2, 2.6 and 2.9, we know that the surface

$$SN = \left\{ (A, m, n, p, q) : R_0 = R_*, A \neq A_*, q < p < q + 1, -2\sqrt{n} < m < -\frac{1+q}{p}, A, n, p, q > 0 \right\},$$

is the saddle-node bifurcation surface. When the parameters are varied to cross the surface from one side to the other side, the number of positive equilibria of system (2.2) changes from zero to two and the saddle-node bifurcation yields two positive equilibria. This implies that there exists a critical psychological effect value α_0 such that the disease cannot invade the population when $\alpha > \alpha_0$ (i.e., $R_0 < R_*$), and the disease will persist for some positive initial populations when $\alpha \leq \alpha_0$ (i.e., $R_0 \geq R_*$).

4.2 Backward Bifurcation

The basic reproduction number R_0 is an expected average number of new infected individuals produced by a single infective individual in a completely susceptible population. In most classic epidemic models, the disease-free equilibrium loses its stability when R_0 crosses one, which results in a bifurcation where a curve of endemic equilibria emerges. The direction of this bifurcation is forward if the bifurcating equilibrium occurs when R_0 is slightly above 1 and there is no positive equilibria near the disease-free equilibrium for $R_0 < 1$. In contrast, this bifurcation is backward if the endemic curve occurs when $R_0 < 1$ (see Dushoff et al. [9]), then R_0 does not act as a threshold value of disease invasion anymore. In order to control the disease, one must further reduce R_0 so that it passes a sub-threshold value R_* (< 1), where a saddle-node bifurcation occurs, no endemic equilibrium exists and the disease-free equilibrium is globally asymptotically stable when $R_0 < R_*$. By Theorem 2.2, we can see that system (2.2) exhibits a backward bifurcation.

Theorem 4.1 *If $(A, m, p, q, n) \in \Gamma_1$, then system (2.2) admits a backward bifurcation as R_0 crosses one if*

$$m < -\frac{1 + q}{p},$$

and admits a forward bifurcation if

$$m \geq -\frac{1 + q}{p}.$$

The bifurcation diagrams are shown in Fig. 5.

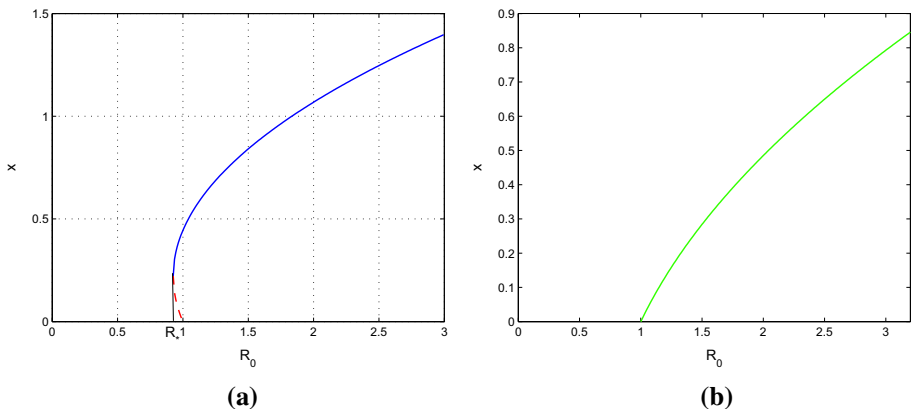


Fig. 5 **a** Backward bifurcation diagram in (R_0, x) -plane for system (2.2) when $p = \frac{3}{2}, q = 1, n = \frac{3}{2}, m = -2$, where $R_* = \frac{25}{27}$; **b** forward bifurcation diagram in (R_0, x) -plane for system (2.2) when $p = \frac{3}{2}, q = 1, n = \frac{3}{2}, m = 0$

Remark 4.2 Theorem 4.1 implies that there exists a critical value β_0 given by (2.10) such that when $\beta < \beta_0$, (i.e., $m < -\frac{1+q}{p}$) system (2.2) exhibits backward bifurcation with rich dynamics.

4.3 Bogdanov–Takens Bifurcation of Codimension Three

From Theorem 2.6, we can see that system (2.2) may undergo Bogdanov–Takens bifurcation of codimension three around $E_1(x_1, y_1)$ if the bifurcation parameters are chosen appropriately. Let

$$\Gamma_2 = \{(A, m, p, q, n) : R_0 = R_*, A = A_*, m = m_*, q < p < q + 1, -2\sqrt{n} < m < -\frac{1+q}{p}, A, p, q, n > 0\}, \quad (4.1)$$

where R_* and A_* are given in (2.9) and (2.13), respectively. To explore the existence of Bogdanov–Takens bifurcation of codimension three, we first take the definition from Perko [21, pages 484–485] (see also Huang et al. [14]) for Bogdanov–Takens bifurcation of codimension three as follows.

Definition 4.3 The bifurcation that results from unfolding the following normal form of a cusp of codimension three,

$$\frac{dx}{dt} = y, \quad \frac{dy}{dt} = x^2 \pm x^3y, \quad (4.2)$$

is called a *Bogdanov–Takens bifurcation of codimension three* or a *cusp bifurcation of codimension three*. A universal unfolding of the above normal form is given by

$$\frac{dx}{dt} = y, \quad \frac{dy}{dt} = \mu_1 + \mu_2y + \mu_3xy + x^2 \pm x^3y. \quad (4.3)$$

Theorem 2.6 indicates that system (2.2) may exhibit a Bogdanov–Takens bifurcation of codimension three. In order to make sure if such a bifurcation can be fully unfolded inside the class of system (2.2), we choose A , p and q as bifurcation parameters and obtain the following unfolding system

$$\begin{aligned} \frac{dx}{dt} &= \frac{x}{1 + mx + nx^2} (A + r_1 - x - y) - (p + r_2)x, \\ \frac{dy}{dt} &= (q + r_3)x - y, \end{aligned} \quad (4.4)$$

where $(A, m, p, q, n) \in \Gamma_2$ and $(r_1, r_2, r_3) \sim (0, 0, 0)$. If we can transform the unfolding system (4.4) into the following versal unfolding of a Bogdanov–Takens singularity (cusp case) of codimension three by a series of near-identity transformations,

$$\begin{aligned} \frac{dx}{dt} &= y, \\ \frac{dy}{dt} &= \gamma_1 + \gamma_2y + \gamma_3xy + x^2 + x^3y + R(x, y, r), \end{aligned} \quad (4.5)$$

where

$$R(x, y, r) = y^2 O(|x, y|^2) + O(|x, y|^5) + O(r)(O(y^2) + O(|x, y|^3)) + O(r^2)O(|x, y|), \tag{4.6}$$

and $\left. \frac{\partial(\gamma_1, \gamma_2, \gamma_3)}{\partial(r_1, r_2, r_3)} \right|_{r=0} \neq 0$, then we can claim that system (4.4) (i.e., system (2.2)) undergoes a Bogdanov–Takens bifurcation of codimension three (see the works of Dumortier et al. [8], and Chow et al. [5]). In fact, we have the following theorem.

Theorem 4.4 *When $(A, m, p, q, n) \in \Gamma_2$, system (2.2) has a nilpotent cusp $E_1(x_1, y_1)$ of codimension three. If we choose A, p and q as bifurcation parameters, then system (2.2) undergoes a Bogdanov–Takens bifurcation of codimension three in a small neighborhood of E_1 . Hence, system (2.2) exhibits the coexistence of an unstable homoclinic loop and a stable limit cycle, coexistence of two limit cycles (the inner one stable and the outer unstable), and a semi-stable limit cycle for different sets of parameters.*

Proof Firstly, we translate the equilibrium $E_1(x_1, y_1)$ of system (4.4) when $r = 0$ into the origin. Let $X = x - x_1, Y = y - y_1$ and use Taylor expansion. Then system (4.4) becomes (we still denote X, Y by x, y , respectively)

$$\frac{dx}{dt} = a_{00} + a_{10}x + a_{01}y + a_{20}x^2 + a_{11}xy + a_{30}x^3 + a_{21}x^2y + a_{40}x^4 + a_{31}x^3y + O(|x, y|^5), \tag{4.7}$$

$$\frac{dy}{dt} = r_3x_1 + (q + r_3)x - y,$$

where

$$\begin{aligned} a_{00} &= \frac{r_1}{q} - \frac{2p^2r_2}{(1+2p)q(1+p+pq)}, & a_{10} &= 1 - r_2 + \frac{(1+2p)(1+q+pq)^2r_1}{2p^3q^2}, \\ a_{01} &= -\frac{1}{q}, \\ a_{20} &= \frac{(1+2p)(1+q+pq)^2}{8p^6q^3} (2p^3q^2 + (1+2p)(2+q)(1+q+pq)r_1), \\ a_{11} &= -\frac{(1+2p)(1+q+pq)^2}{2p^3q^2}, \\ a_{30} &= -\frac{(1+2p)^2(1+q+pq)^3}{16p^9q^4} \{2p^3q^2(pq-1) + [2p^3q^3 - 2(1+q)^2 + p^2q(-4-2q+3q^2) \\ &\quad + p(-4-10q-5q^2+q^3)]r_1\}, \\ a_{21} &= -\frac{(2+q)(1+2p)^2(1+q+pq)^3}{8p^6q^3}, \\ a_{40} &= -\frac{(1+2p)^3(1+q+pq)^4}{64p^{12}q^5} \{2p^3q^2[-2+(-1+2p)q+(1+3p)q^2] + [2p^3q^3(4+3q) \\ &\quad + (1+q)^2(-4-2q+q^2) + p^2q(-8-4q+16q^2+11q^3) \\ &\quad + 2p(-4-12q-8q^2+3q^3+3q^4)]r_1\}, \\ a_{31} &= \frac{(1+2p)^3(1+q+pq)^4(-2-2q+pq^2)}{16p^9q^4}, \end{aligned}$$

in which A , n and m have been eliminated by using the equations: $A = A_*$, $R_0 = R_*$ and $m = m_*$, respectively. Note that system (4.7) is reduced to system (2.14) when $r = 0$.

Next let

$$\begin{aligned} X &= x, \\ Y &= a_{00} + a_{10}x + a_{01}y + a_{20}x^2 + a_{11}xy + a_{30}x^3 \\ &\quad + a_{21}x^2y + a_{40}x^4 + a_{31}x^3y + O(|x, y|^5). \end{aligned}$$

Then system (4.7) can be put in the form of (we still denote X, Y by x, y , respectively)

$$\begin{aligned} \frac{dx}{dt} &= y, \\ \frac{dy}{dt} &= b_{00} + b_{10}x + b_{01}y + b_{20}x^2 + b_{11}xy \\ &\quad + b_{02}y^2 + b_{30}x^3 + b_{21}x^2y + b_{12}xy^2 + b_{40}x^4 \\ &\quad + b_{31}x^3y + b_{22}x^2y^2 + O(|x, y|^5), \end{aligned} \tag{4.8}$$

where

$$\begin{aligned} b_{00} &= a_{00} + a_{01}r_3x_1, & b_{10} &= a_{10} + a_{01}(q + r_3) + a_{11}r_3x_1, & b_{01} &= \frac{a_{01}(a_{10} - 1) - a_{00}a_{11}}{a_{01}}, \\ b_{20} &= a_{20} + a_{11}(q + r_3) + a_{21}r_3x_1, & b_{11} &= \frac{a_{00}a_{11}^2 + 2a_{01}^2a_{20} - a_{01}a_{10}a_{11} - 2a_{00}a_{01}a_{21}}{a_{01}^2}, \\ b_{02} &= \frac{a_{11}}{a_{01}}, \\ b_{30} &= \frac{1}{a_{01}^4} \{ a_{00}^2(a_{11}^4 - 4a_{01}a_{11}^2a_{21} + 2a_{01}^2a_{21}^2 + 4a_{01}^2a_{11}a_{31}) \\ &\quad + a_{01}^4[a_{30} + a_{21}(q + r_3) + a_{31}r_3x_1] \}, \\ b_{21} &= \frac{1}{a_{01}^3} [a_{01}a_{11}(a_{10}a_{11} + 3a_{00}a_{21}) + 3a_{01}^3a_{30} - a_{00}a_{11}^3 - a_{01}^2(a_{11}a_{20} \\ &\quad + 2a_{10}a_{21} + 3a_{00}a_{31})], \\ b_{12} &= \frac{2a_{01}a_{21} - a_{11}^2}{a_{01}^2}, & b_{22} &= \frac{a_{11}^3 - 3a_{01}a_{11}a_{21} + 3a_{01}^2a_{31}}{a_{01}^3}, \\ b_{40} &= \frac{1}{a_{01}^4} \{ a_{00} [a_{10}(a_{11}^4 - 4a_{01}a_{11}^2a_{21} + 2a_{01}^2a_{21}^2 + 4a_{01}^2a_{11}a_{31}) - a_{01}(a_{11}^3a_{20} - a_{01}a_{11}^2a_{30} \\ &\quad + a_{01}^2(2a_{21}a_{30} + 3a_{20}a_{31}) + a_{01}a_{11}(-3a_{20}a_{21} + a_{01}a_{40})] + a_{01}^4[a_{40} + a_{31}(q + r_3)] \}, \\ b_{31} &= \frac{1}{a_{01}^4} [-a_{00}a_{11}^4 + a_{01}a_{11}^2(4a_{00}a_{21} - a_{10}a_{11}) \\ &\quad - a_{01}^3(2a_{20}a_{21} + a_{11}a_{30} + 3a_{10}a_{31}) + a_{01}^2(a_{11}^2a_{20} \\ &\quad + 3a_{10}a_{11}a_{21} - 2a_{00}a_{21}^2 - 4a_{00}a_{11}a_{31}) + 4a_{01}^4a_{40}]. \end{aligned}$$

Again note that system (4.8) is reduced to system (2.15) when $r = 0$.

Secondly, following the procedure given by Li et al. [17] (see also [14]), we use seven steps to transform system (4.8) to the versal unfolding of the Bogdanov–Takens singularity (cusp case) of codimension three, that is system (4.5).

(I) Removing the y^2 -term from system (4.8). In order to remove the y^2 -term, we let $x = X + \frac{b_{02}}{2}X^2$, $y = Y + b_{02}XY$, which is a near-identity transformation for (x, y) near $(0, 0)$. Then system (4.8) is changed into (we still denote X, Y by x, y , respectively)

$$\begin{aligned} \frac{dx}{dt} &= y, \\ \frac{dy}{dt} &= c_{00} + c_{10}x + c_{01}y + c_{20}x^2 + c_{11}xy + c_{30}x^3 + c_{21}x^2y + c_{12}xy^2 + c_{40}x^4 + c_{31}x^3y \\ &\quad + c_{22}x^2y^2 + O(|x, y|^5), \end{aligned} \tag{4.9}$$

where

$$\begin{aligned} c_{00} &= b_{00}, \quad c_{10} = -b_{00}b_{02} + b_{10}, \quad c_{01} = b_{01}, \\ c_{20} &= \frac{1}{2}(2b_{00}b_{02}^2 - b_{02}b_{10} + 2b_{20}), \quad c_{11} = b_{11}, \\ c_{30} &= \frac{1}{2}(-2b_{00}b_{02}^3 + b_{02}^2b_{10} + 2b_{30}), \\ c_{21} &= \frac{1}{2}(b_{02}b_{11} + 2b_{21}), \quad c_{12} = 2b_{02}^2 + b_{12}, \\ c_{31} &= b_{02}b_{21} + b_{31}, \\ c_{40} &= \frac{1}{4}(4b_{00}b_{02}^4 - 2b_{02}^3b_{10} + b_{02}^2b_{20} + 2b_{02}b_{30} + 4b_{40}), \\ c_{22} &= \frac{1}{2}(-2b_{02}^3 + 3b_{02}b_{12} + 2b_{22}). \end{aligned}$$

Note that $c_{00} = c_{10} = c_{01} = c_{11} = 0$ when $r = 0$.

(II) Eliminating the xy^2 -term in system (4.9). Let $x = X + \frac{c_{12}}{6}X^3$, $y = Y + \frac{c_{12}}{2}X^2Y$. Then we obtain the following system (we still denote X, Y by x, y , respectively)

$$\begin{aligned} \frac{dx}{dt} &= y, \\ \frac{dy}{dt} &= d_{00} + d_{10}x + d_{01}y + d_{20}x^2 + d_{11}xy + d_{30}x^3 + d_{21}x^2y + d_{40}x^4 + d_{31}x^3y + d_{22}x^2y^2 \\ &\quad + O(|x, y|^5), \end{aligned} \tag{4.10}$$

where

$$\begin{aligned} d_{00} &= c_{00}, \quad d_{10} = c_{10}, \quad d_{01} = c_{01}, \quad d_{20} = c_{20} - \frac{c_{12}c_{00}}{2}, \quad d_{11} = c_{11}, \quad d_{30} = c_{30} - \frac{c_{12}c_{10}}{3}, \\ d_{21} &= c_{21}, \quad d_{40} = c_{40} + \frac{c_{00}c_{12}^2}{4} - \frac{c_{12}c_{20}}{6}, \quad d_{31} = c_{31} + \frac{c_{11}c_{12}}{6}, \quad d_{22} = c_{22}. \end{aligned}$$

Note that $d_{00} = d_{10} = d_{01} = d_{11} = 0$ when $r = 0$.

(III) Removing the x^2y^2 -term in system (4.10). Let $x = X + \frac{d_{22}}{12}X^4$, $y = Y + \frac{d_{22}}{3}X^3Y$. Then we have the following system (we still denote X, Y by x, y , respectively)

$$\begin{aligned} \frac{dx}{dt} &= y, \\ \frac{dy}{dt} &= e_{00} + e_{10}x + e_{01}y + e_{20}x^2 + e_{11}xy \\ &\quad + e_{30}x^3 + e_{21}x^2y + e_{40}x^4 + e_{31}x^3y + O(|x, y|^5), \end{aligned} \tag{4.11}$$

where

$$\begin{aligned} e_{00} &= d_{00}, \quad e_{10} = d_{10}, \quad e_{01} = d_{01}, \quad e_{20} = d_{20}, \quad e_{11} = d_{11}, \quad e_{30} = d_{30} - \frac{d_{22}d_{00}}{3}, \\ e_{21} &= d_{21}, \quad e_{40} = d_{40} - \frac{d_{22}d_{10}}{4}, \quad e_{31} = d_{31}. \end{aligned}$$

Note that $e_{00} = e_{10} = e_{01} = e_{11} = 0$ when $r = 0$.

(IV) Removing the x^3 and x^4 -terms in system (4.11). Note that $e_{20} = -\frac{(1+2p)(1+q+pq)^2}{4p^3q} + O(r)$, $e_{20} \neq 0$ for small r since $p, q > 0$. We let

$$x = X - \frac{e_{30}}{4e_{20}}X^2 + \frac{15e_{30}^2 - 16e_{20}e_{40}}{80e_{20}^2}X^3, \quad y = Y, \quad t = \left(1 - \frac{e_{30}}{2e_{20}}X + \frac{45e_{30}^2 - 48e_{20}e_{40}}{80e_{20}^2}X^2\right)\tau,$$

and then obtain the following system from system (4.11) (we still denote X, Y, τ by x, y, t , respectively)

$$\begin{aligned} \frac{dx}{dt} &= y, \\ \frac{dy}{dt} &= f_{00} + f_{10}x + f_{01}y + f_{20}x^2 + f_{11}xy \\ &\quad + f_{30}x^3 + f_{21}x^2y + f_{40}x^4 + f_{31}x^3y + O(|x, y|^5), \end{aligned} \tag{4.12}$$

where

$$\begin{aligned} f_{00} &= e_{00}, \quad f_{10} = e_{10} - \frac{e_{00}e_{30}}{2e_{20}}, \quad f_{01} = e_{01}, \\ f_{20} &= e_{20} - \frac{60e_{10}e_{20}e_{30} - 45e_{00}e_{30}^2 + 48e_{00}e_{20}e_{40}}{80e_{20}^2}, \\ f_{11} &= e_{11} - \frac{e_{01}e_{30}}{2e_{20}}, \quad f_{30} = \frac{e_{10}(35e_{30}^2 - 32e_{20}e_{40})}{40e_{20}^2}, \\ f_{21} &= e_{21} - \frac{60e_{11}e_{20}e_{30} - 45e_{01}e_{30}^2 + 48e_{01}e_{20}e_{40}}{80e_{20}^2}, \\ f_{40} &= \frac{e_{10}(-15e_{30}^3 + 16e_{20}e_{40}e_{40})}{64e_{20}^3}, \quad f_{31} = e_{31} + \frac{7e_{11}e_{30}^2}{8e_{20}^2} - \frac{5e_{21}e_{30} + 4e_{11}e_{40}}{5e_{20}}. \end{aligned}$$

Note that $f_{00} = f_{10} = f_{01} = f_{11} = f_{30} = f_{40} = 0$ when $r = 0$.

(V) Removing the x^2y -term in system (4.12). Since $f_{20} = -\frac{(1+2p)(1+q+pq)^2}{4p^3q} + O(r)$, $f_{20} \neq 0$ for small r because $p, q > 0$, we can make the following transformation

$$x = X, \quad y = Y + \frac{f_{21}}{3f_{20}}Y^2 + \frac{f_{21}^2}{36f_{20}^2}Y^3, \quad \tau = \left(1 + \frac{f_{21}}{3f_{20}}Y + \frac{f_{21}^2}{36f_{20}^2}Y^2\right)t,$$

under which system (4.12) becomes (we still denote X, Y, τ by x, y, t , respectively)

$$\begin{aligned} \frac{dx}{dt} &= y, \\ \frac{dy}{dt} &= g_{00} + g_{10}x + g_{01}y + g_{20}x^2 + g_{11}xy \\ &\quad + g_{31}x^3y + R_1(x, y, r), \end{aligned} \tag{4.13}$$

where

$$g_{00} = f_{00}, \quad g_{10} = f_{10}, \quad g_{01} = f_{01} - \frac{f_{00}f_{21}}{f_{20}}, \quad g_{11} = f_{11} - \frac{f_{10}f_{21}}{f_{20}}, \quad g_{20} = f_{20}, \quad g_{31} = f_{31} - \frac{f_{21}f_{30}}{f_{20}}.$$

Note that $g_{00} = g_{10} = g_{01} = g_{11} = 0$ when $r = 0$, and $R_1(x, y, r)$ has the property of (4.6).

(VI) Changing g_{20} and g_{31} to 1 in system (4.13). It is seen that $g_{20} = -\frac{(1+2p)(1+p+pq)^2}{4p^3q} + O(r) < 0$ and $g_{31} = -\frac{(1+2p)^4(1+p+pq)^4(1+q)}{16p^9q^2} + O(r) < 0$ for small r since $p, q > 0$. By making the following changes of variables and time rescaling:

$$x = g_{20}^{\frac{1}{5}}g_{31}^{-\frac{2}{5}}X, \quad y = g_{20}^{\frac{4}{5}}g_{31}^{-\frac{3}{5}}Y, \quad t = g_{20}^{-\frac{3}{5}}g_{31}^{\frac{1}{5}}\tau,$$

we can transform system (4.13) to (we still denote X, Y, τ by x, y, t , respectively)

$$\begin{aligned} \frac{dx}{dt} &= y, \\ \frac{dy}{dt} &= h_{00} + h_{10}x + h_{01}y + x^2 + h_{11}xy + x^3y + R_2(x, y, r), \end{aligned} \tag{4.14}$$

where

$$h_{00} = g_{00}g_{31}^{\frac{4}{5}}g_{20}^{-\frac{7}{5}}, \quad h_{10} = g_{10}g_{31}^{\frac{2}{5}}g_{20}^{-\frac{6}{5}}, \quad h_{01} = g_{01}g_{31}^{\frac{1}{5}}g_{20}^{-\frac{3}{5}}, \quad h_{11} = g_{11}g_{31}^{-\frac{1}{5}}g_{20}^{-\frac{2}{5}}.$$

Note that $h_{00} = h_{10} = h_{01} = h_{11} = 0$, when $r = 0$ and $R_2(x, y, r)$ has the property of (4.6).

(VII) Removing the x_{10} -term in system (4.14). Finally, let $x = X - \frac{h_{10}}{2}$, $y = Y$. Then system (4.14) becomes (we still denote X, Y by x, y , respectively)

$$\begin{aligned} \frac{dx}{d\tau} &= y, \\ \frac{dy}{d\tau} &= \gamma_1 + \gamma_2y + \gamma_3xy + x^2 + x^3y + R_3(x, y, r), \end{aligned} \tag{4.15}$$

where $\gamma_1 = h_{00} - \frac{1}{4}h_{10}^2$, $\gamma_2 = h_{01} - \frac{1}{8}(h_{10}^3 + 4h_{10}h_{11})$, $\gamma_3 = h_{11} + \frac{3}{4}h_{10}^2$, and $R_3(x, y, r)$ has the property of (4.6).

A direct computation with the help of *Mathematica* shows that

$$\left| \frac{\partial(\gamma_1, \gamma_2, \gamma_3)}{\partial(r_1, r_2, r_3)} \right|_{r=0} = \frac{2^{\frac{3}{5}}(1+q)^{\frac{9}{5}}(1+2p)^{\frac{14}{5}}}{p^4 q^{\frac{11}{5}}(1+q+pq)^{\frac{3}{5}}} > 0$$

due to $p, q > 0$. System (4.15) is exactly in the form of system (4.5). Therefore, by the results of Dumortier et al. [8] and Chow et al. [5], we know that system (4.15) is the versal unfolding of Bogdanov–Takens singularity (cusp case) of codimension three. The remainder term $R_3(x_3, y_3, r)$ with the property (4.6) has no influence on the bifurcation analysis. □

Now we describe the bifurcation diagram of system (4.15) following the bifurcation diagram given in Figure 4.2 of Dumortier et al. [8] (see also [14]). System (4.15) has no equilibria for $\gamma_1 > 0$. $\gamma_1 = 0$ is a saddle-node bifurcation plane in a neighborhood of the origin. Crossing this plane in the direction of decreasing γ_1 results in two equilibria: a saddle and a node or focus. The other surfaces of bifurcations are located in the half space $\gamma_1 < 0$. The bifurcation diagram has the conical structure in \mathbb{R}^3 , starting from $(\gamma_1, \gamma_2, \gamma_3) = (0, 0, 0)$. It can be best shown by drawing its intersection with the half sphere

$$S = \{(\gamma_1, \gamma_2, \gamma_3) | \gamma_1^2 + \gamma_2^2 + \gamma_3^2 = \lambda^2, \gamma_1 \leq 0, \lambda > 0, \text{ sufficiently small}\}.$$

To clearly see the traces of the intersections, we draw the projections of traces onto the (γ_2, γ_3) plane, as shown in Fig. 6.

In the following, we summarize the bifurcation phenomena of system (4.15), which are equivalent to that of the original system (2.2). There are three bifurcation curves on S as shown in Fig. 6:

- C : homoclinic bifurcation curve;
- H : Hopf bifurcation curve;
- L : saddle-node bifurcation curve of limit cycles.

The curve L is tangent to H at a point h_2 and tangent to C at a point c_2 . The curves H and C have first order contact with the boundary of S at the points b_1 and b_2 . In the neighborhoods of b_1 and b_2 , system (4.15) is an unfolding of the cusp singularity of codimension two.

Along the curve C , except at the point c_2 , a homoclinic bifurcation of codimension one occurs. When crossing the arc b_1c_2 of C from left to right, the two separatrices of the saddle point coincide and an unstable limit cycle appears. A similar phenomenon gives rise to a stable limit cycle when crossing the arc c_2b_2 of C from right to left. The point c_2 corresponds to a homoclinic bifurcation of codimension two.

Along the arc b_1h_2 of the curve H , a subcritical Hopf bifurcation occurs with an unstable limit cycle appearing when crossing the arc b_1h_2 of H from right to left. Along the arc h_2b_2 of the curve H , a supercritical Hopf bifurcation occurs with a stable limit cycle appearing when crossing the arc h_2b_2 of H from left to right. The point h_2 is a degenerate Hopf bifurcation point, i.e., a Hopf bifurcation point of codimension two.

The curves H and C intersect transversally at a unique point d representing a parameter value of simultaneous Hopf bifurcation and homoclinic bifurcation.

For parameter values in the triangle dh_2c_2 , there exist exactly two limit cycles: the inner one is stable and the outer one is unstable. These two limit cycles coalesce in a generic way in a saddle-node bifurcation of limit cycles when the curve L is crossed from right to left. On the arc L itself, there exists a unique semistable limit cycle.

Remark 4.5 Zhou et al. [30] showed that system (2.2) exhibits an attracting Bogdanov–Takens bifurcation of codimension two by unfolding a set of specific parameter values, but their chosen parameter values do not satisfy $p > q$. Our results obtained in this paper about the existence of a Bogdanov–Takens bifurcation of codimension three imply the existence of repelling and attracting Bogdanov–Takens bifurcation of codimension two for general parameter conditions.

4.4 Hopf Bifurcation of Codimension Three

From Theorem 2.9, we can see that system (2.2) may exhibit Hopf bifurcation around $E_i(x_i, y_i)$ ($i = 3, 4, 5$). In this section we discuss Hopf bifurcation around $E_i(x_i, y_i)$ ($i = 3, 4, 5$).

Firstly, letting $dt = (1 + mx + nx^2)d\tau$, we obtain (we still denote τ by t)

$$\begin{aligned} \frac{dx}{dt} &= x(A - x - y) - p(1 + mx + nx^2)x, \\ \frac{dy}{dt} &= (qx - y)(1 + mx + nx^2). \end{aligned} \tag{4.16}$$

Obviously, system (4.16) has the same topological structure as that of system (2.2), since we consider system (2.2) in $\mathbb{R}_2^+ = \{(x, y) : x \geq 0, y \geq 0\}$ and $1 + mx + nx^2 > 0$ holds for all $x \geq 0$.

Following the technique in Dai et al. [6] and noticing that $x_i = x^+$ and $y_i = qx^+$ ($i = 3, 4, 5$), where x^+ is given in (2.8), we use the following scalings of the coordinates,

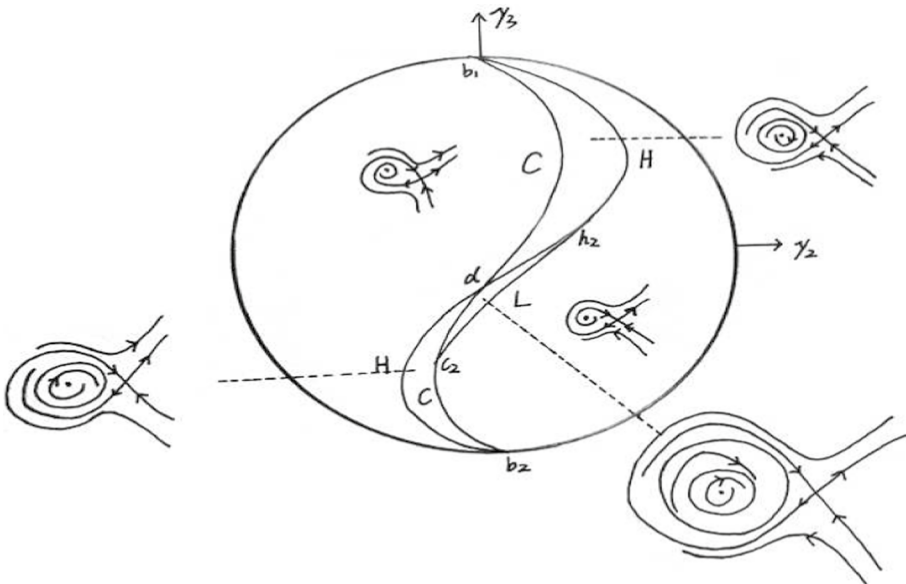


Fig. 6 Bifurcation diagram for system (4.15) on S

$$\bar{x} = \frac{x}{x^+}, \quad \bar{y} = \frac{y}{qx^+}, \quad \tau = x^+t, \quad (4.17)$$

to transform system (4.16) to an equivalent polynomial differential system (we still denote τ by t)

$$\begin{aligned} \frac{d\bar{x}}{dt} &= \bar{x} \left(\frac{A}{x^+} - \bar{x} - q\bar{y} \right) - \frac{P}{x^+} (1 + mx^+\bar{x} + nx^{+2}\bar{x}^2)\bar{x}, \\ \frac{d\bar{y}}{dt} &= \frac{1}{x^+} (\bar{x} - \bar{y})(1 + mx^+\bar{x} + nx^{+2}\bar{x}^2). \end{aligned} \quad (4.18)$$

Further, letting

$$\bar{A} = \frac{A}{x^+}, \quad \bar{m} = mx^+, \quad \bar{n} = nx^{+2}, \quad \bar{p} = \frac{P}{x^+}, \quad \bar{q} = q, \quad a = \frac{1}{x^+}, \quad (4.19)$$

in system (4.18) and dropping the bars, we obtain a new system,

$$\begin{aligned} \frac{dx}{dt} &= x(A - x - qy) - px(1 + mx + nx^2), \\ \frac{dy}{dt} &= a(x - y)(1 + mx + nx^2). \end{aligned} \quad (4.20)$$

Since system (4.20) has an equilibrium $\tilde{E}(1, 1)$ (corresponding to $E_i(x_i, y_i)$ in system (2.2), $i = 3, 4, 5$), we have

$$A = p(1 + m + n) + q + 1,$$

which is then substituted into (4.20) to finally yield the following system

$$\begin{aligned} \frac{dx}{dt} &= x(p(1 + m + n) + q + 1 - x - qy) - px(1 + mx + nx^2), \\ \frac{dy}{dt} &= a(x - y)(1 + mx + nx^2), \end{aligned} \quad (4.21)$$

where the parameters satisfy the following conditions:

$$p, q, n, a > 0, \quad m > -2\sqrt{n}, \quad aq < p < a(q + 1). \quad (4.22)$$

Since the transformation (4.17) is a linear sign-reserving transformation, system (4.21) and system (2.2) have the same qualitative property. In order to get the necessary conditions for Hopf bifurcation around $\tilde{E}(1, 1)$, we firstly let

$$a_* = \frac{-(mp + 2np + 1)}{m + n + 1}, \quad (4.23)$$

and have the following results.

Theorem 4.6 *When the conditions in (4.22) are satisfied, system (4.21) has an equilibrium at $\tilde{E}(1, 1)$. Moreover,*

- (I) *when $a < a_*$, $\tilde{E}(1, 1)$ is an unstable hyperbolic node or focus;*
- (II) *when $a > a_*$, $\tilde{E}(1, 1)$ is a stable hyperbolic node or focus;*
- (III) *when $a = a_*$, $\tilde{E}(1, 1)$ is a weak focus or center.*

Proof The Jacobian matrix of system (4.21) at $\tilde{E}(1, 1)$ is

$$J(\tilde{E}(1, 1)) = \begin{pmatrix} -1 - mp - 2np & -q \\ a(m + n + 1) & -a(m + n + 1) \end{pmatrix},$$

then the determinant of $J(\tilde{E}_i(1, 1))$ is

$$\det(J(\tilde{E})) = a(m + n + 1)(1 + q + mp + 2np),$$

and the trace of $J(\tilde{E}_i(1, 1))$ is

$$\text{tr}(J(\tilde{E})) = -1 - mp - 2np - a(m + n + 1).$$

Since $m + n + 1 > 0$, and $2np x^+ + 1 + q + mp > 0$ in system (2.3) $\Leftrightarrow 1 + q + mp + 2np > 0$, we can see that $\det(J(\tilde{E})) > 0$, and $\text{tr}(J(\tilde{E})) = 0$ ($> 0, < 0$) if $a = a_*$ ($a < a_*, a > a_*$), then the conclusions hold. \square

Next we continue to consider the case (III) of Theorem 4.6 and study the existence of Hopf bifurcation around $\tilde{E}(1, 1)$ of system (4.21), which corresponds the Hopf bifurcation around $E_i(x_i, y_i)$ ($i = 3, 4, 5$) of system (2.2), respectively. We shall investigate the nondegenerate condition and stability of the bifurcating periodic orbits at the positive equilibrium $\tilde{E}(1, 1)$ of system (4.21) by calculating the first Lyapunov coefficient.

Let $X = x - 1, Y = y - 1$, and $a = a_*$, system (4.21) can be written as (we still denote X, Y by x, y , respectively)

$$\begin{aligned} \frac{dx}{dt} &= \tilde{b}x - qy + (\tilde{b} - np)x^2 - qxy - np x^3, \\ \frac{dy}{dt} &= \tilde{b}x - \tilde{b}y + a_*(m + 2n)x^2 - a_*(m + 2n)xy + a_*nx^3 - a_*nx^2y, \end{aligned} \tag{4.24}$$

where $\tilde{b} = -(1 + mp + 2np) > 0$ since $a_* > 0$. Let $\omega = \sqrt{\tilde{b}q - \tilde{b}^2}$ and make a transformation of $x = qX, y = \tilde{b}X - \omega Y$ and $dt = \frac{1}{\omega}d\tau$, system (4.24) becomes (we still denote X, Y, τ by x, y, t , respectively)

$$\begin{aligned} \frac{dx}{dt} &= y + f(x, y), \\ \frac{dy}{dt} &= -x + g(x, y), \end{aligned} \tag{4.25}$$

where

$$\begin{aligned} f(x, y) &= a_{20}x^2 + a_{11}xy + a_{02}y^2 + a_{30}x^3 + a_{21}x^2y + a_{12}xy^2 + a_{03}y^3, \\ g(x, y) &= b_{20}x^2 + b_{11}xy + b_{02}y^2 + b_{30}x^3 + b_{21}x^2y + b_{12}xy^2 + b_{03}y^3, \end{aligned}$$

and

$$\begin{aligned} a_{20} &= -\frac{npq}{\omega}, \quad a_{11} = q, \quad a_{02} = 0, \quad a_{30} = -\frac{npq^2}{\omega}, \quad a_{21} = 0, \quad a_{12} = 0, \quad a_{03} = 0, \\ b_{20} &= -\frac{q}{\omega^2}(n\tilde{b}p + a_*(m + 2n)(q - \tilde{b})), \quad b_{11} = \frac{q(\tilde{b} - a_*(m + 2n))}{\omega}, \quad b_{02} = 0, \\ b_{30} &= -\frac{nq^2}{\omega^2}(\tilde{b}p + a_*(q - \tilde{b})), \quad b_{21} = \frac{-naq^2}{\omega}, \quad b_{12} = 0, \quad b_{03} = 0. \end{aligned}$$

Using the formula in Zhang et al. [29] and carrying out lengthy calculations by MATLAB, we obtain the first Lyapunov coefficient as follows

$$\sigma_1 = \frac{q^2 (b_0 + b_1p + b_2p^2) + (c_0 + c_1p)q}{4(1 + m + n)^2(1 + q + mp + 2np)\sqrt{-(1 + mp + 2np)(1 + q + mp + 2np)}}$$

where

$$b_0 = c_0 = m + 3n - n^2, \quad b_1 = 2m^2 + n(9 - n)m - n(3n^2 - 12n + 1),$$

$$b_2 = m^3 + m^2n(6 + 2n) + mn(3n^2 + 18n - 1) + 16n^3, \quad c_1 = m^2 + mn(3 + n) + 2n(3n - 1).$$

Since $1 + q + mp + 2np > 0$, $1 + mp + 2np < 0$, and $1 + m + n > 0$, the sign of σ_1 is same as

$$\sigma_{11} = (b_0 + b_1p + b_2p^2) + (c_0 + c_1p)q. \tag{4.26}$$

Thus we have the following results.

Theorem 4.7 *When $a = a_*$ and the conditions in (4.22) are satisfied, we have*

- (I) *if $\sigma_{11} < 0$, then $\tilde{E}(1, 1)$ is a stable weak focus with multiplicity one and one stable limit cycle bifurcates from $\tilde{E}(1, 1)$ by the supercritical Hopf bifurcation;*
- (II) *if $\sigma_{11} > 0$, then $\tilde{E}(1, 1)$ is an unstable weak focus with multiplicity one and one unstable limit cycle bifurcates from $\tilde{E}(1, 1)$ by the subcritical Hopf bifurcation;*
- (III) *if $\sigma_{11} = 0$, then $\tilde{E}(1, 1)$ is a weak focus with multiplicity at least two and system (4.21) may exhibit degenerate Hopf bifurcation.*

Next we continue to consider the case (III) in Theorem 4.7, in fact, we can see that $\sigma_{11} = 0$ if $b_0 + b_1p + b_2p^2 = c_0 + c_1p = 0$, or $(c_0 + c_1p) \neq 0$ and $q = -\frac{b_0 + b_1p + b_2p^2}{c_0 + c_1p}$.

Define

$$q_1 = -\frac{b_0 + b_1p + b_2p^2}{c_0 + c_1p}. \tag{4.27}$$

From the case (III) of Theorem 4.7, we know that system (4.21) may exhibit degenerate Hopf bifurcation (i.e., Hopf bifurcation of codimension 2) when $q = q_1$, $a = a_*$ and the conditions in (4.22) are satisfied. Using the formal series method in [29] and MATLAB software, we get the second Lyapunov coefficient as follows

$$\sigma_2 = \frac{q_1^4(1 + mp + 2np)(e_0 + e_1p + e_2p^2)}{24(m + n + 1)^{\frac{7}{2}}(1 - n + (m + mn + 4n)p)\sqrt{\frac{np(1 + mp + 2np)(1 - n + (m + mn + 4n)p)}{c_0 + c_1p}}},$$

where

$$e_0 = -(2 + 3n)m^2 - n(11 + 9n)m + 2n(1 - 9n - 2n^2),$$

$$e_1 = -4m^3 - 20nm^2 + 8n(1 - 5n)m + 4n^2(5 - 10n + n^2),$$

$$e_2 = (m + 2n)[(3n - 2)m^3 + n(3n - 5)m^2 - 2n(n^2 + 12n - 3)m - 4n^2(7n - 1)].$$

Since $1 + mp + 2np < 0, 1 + m + n > 0$ and $\frac{1-n+(m+mn+4n)p}{c_0+c_1p} < 0$, the sign of σ_2 is determined by

$$\sigma_{22} = e_0 + e_1p + e_2p^2. \tag{4.28}$$

Summarizing the above analysis, we have the following results.

Theorem 4.8 *When $a = a_*, q = q_1$, and the conditions in (4.22) are satisfied, we have*

- (I) *if $\sigma_{22} > 0$, then $\tilde{E}(1, 1)$ is an unstable weak focus with multiplicity 2, system (4.21) exhibits a degenerate Hopf bifurcation of codimension 2 and there are two limit cycles surrounding $\tilde{E}(1, 1)$, the outer one is unstable;*
- (II) *if $\sigma_{22} < 0$, then $\tilde{E}(1, 1)$ is a stable weak focus with multiplicity 2, system (4.21) exhibits a degenerate Hopf bifurcation of codimension 2 and there are two limit cycle surrounding $\tilde{E}(1, 1)$, the outer one is stable;*
- (III) *if $\sigma_{22} = 0$, then $\tilde{E}(1, 1)$ is a weak focus with multiplicity at least 3, and system (4.21) may exhibit degenerate Hopf bifurcations of codimension at least 3.*

When $\sigma_2 = 0$ (or $\sigma_{22} = 0$), i.e.,

$$p = p_{\pm} \triangleq \frac{-e_1 \pm \sqrt{e_1^2 - 4e_0e_2}}{2e_2}, \tag{4.29}$$

system (4.21) may exhibit degenerate Hopf bifurcations of codimension at least 3. In order to understand the exact codimension of the Hopf bifurcation around $\tilde{E}(1, 1)$, we need to calculate the third Lyapunov coefficient, which has a very lengthy expression about m and n . For simplification, we first let $m = -\frac{3}{20}$ and $n = \frac{3}{500}$, and get $p = p_+ = \frac{20(7860009+214\sqrt{574600561})}{8474511}$ from $\sigma_2 = 0$; then get $q = q_1 = \frac{611412311729+25777159\sqrt{574600561}}{169631461850}$ from $\sigma_1 = 0$; finally get $a = a_* = \frac{10(22381+\sqrt{574600561})}{122819}$ from $\text{tr}(J(\tilde{E})) = 0$. With this set of parameters values, we obtain the value of the third Lyapunov coefficient as follows:

$$\sigma_3 \approx -3.726014431455.$$

Thus, $\tilde{E}(1, 1)$ is a stable weak focus with multiplicity 3 and system (4.21) exhibits degenerate Hopf bifurcation of codimension 3.

Theorem 4.9 *If*

$$m = -\frac{3}{20}, n = \frac{3}{500}, p = \frac{20(7860009 + 214\sqrt{574600561})}{8474511},$$

$$a = \frac{10(22381 + \sqrt{574600561})}{122819}, q = \frac{611412311729 + 25777159\sqrt{574600561}}{169631461850},$$

then $\tilde{E}(1, 1)$ is a stable weak focus with multiplicity 3, system (4.21) exhibits a degenerate Hopf bifurcation of codimension 3, and three small amplitude limit cycles bifurcate from $\tilde{E}(1, 1)$ when the bifurcation parameters p, q and a are chosen properly.

Remark 4.10 From Theorem 4.9 we can see that system (4.21) may exhibit degenerate Hopf bifurcations of codimension larger than 3 if m and n are arbitrary parameters.

5 Numerical Simulations

In this section, we present simulations to illustrate the existence of one, two and three limit cycles arising from the subcritical and supercritical Hopf bifurcations, Hopf bifurcation of codimension two, and Hopf bifurcation of codimension three around $\tilde{E}(1, 1)$ in system (4.21), respectively, which correspond to the Hopf bifurcation around $E_i(x_i, y_i)$ ($i = 3, 4, 5$) in system (2.2) for the following five classes:

(i) Hopf bifurcation of codimension one around $E_3(x_3, y_3)$ for $R_* < R_0 < 1$. The existence of one stable (or unstable) limit cycle arising from the supercritical (or subcritical) Hopf bifurcation around $\tilde{E}(1, 1)$ in system (4.21) is given in Fig. 7a (or b), which corresponds to the supercritical (or subcritical) Hopf bifurcation around $E_3(x_3, y_3)$ for $R_* < R_0 < 1$ in system (2.2). Moreover, from the following parameter values for system (4.21)

$$(n, m, p, q, a) = \left(\frac{1}{2}, -\frac{6}{5}, \frac{17}{2}, \frac{20837}{6550}, \frac{7}{3} - 0.02 \right)$$

in Fig. 7a, we can get the original parameter values for system (1.1)

$$(\mu, \delta, k, \alpha, \beta) = \left(\frac{7230439d}{1120811}, \frac{11520392d}{1120811}, \frac{6613425d^2}{1120811b}, \frac{7775536041d^2}{343220000b^2}, -\frac{264537d}{32750b} \right)$$

by (4.19) and (2.3). On the other hand, we select

$$b = 1.64 \times 10^7, \quad d = 0.006988$$

following our previous study Lu et al. [20]. Then the original parameter values for system (1.1) are as follows

$$(\mu, \delta, k, \alpha, \beta) = (0.0450801, 0.0071827, 1.75694 \times 10^{-11}, 4.11316 \times 10^{-18}, -3.44179 \times 10^{-9}).$$

Similarly, from the following parameter values for system (4.21)

$$(n, m, p, q, a) = \left(\frac{1}{2}, -\frac{6}{5}, \frac{17}{2}, \frac{3591}{1310}, \frac{7}{3} + 0.05 \right)$$

in Fig. 7b, we obtain the original parameter values for system (1.1)

$$(\mu, \delta, k, \alpha, \beta) = (0.023213, 0.00148013, 9.52459 \times 10^{-12}, 3.59301 \times 10^{-18}, -3.21681 \times 10^{-9}).$$

(ii) Hopf bifurcation of codimension one around $E_4(x_4, y_4)$ for $R_0 = 1$. The existence of one stable limit cycle arising from the supercritical Hopf bifurcation around $\tilde{E}(1, 1)$ in system (4.21) is given in Fig. 7c, which corresponds to the supercritical Hopf bifurcation around $E_4(x_4, y_4)$ for $R_0 = 1$ in system (2.2). Moreover, as in case (i), from the following parameter values for system (4.21)

$$(n, m, p, q, a) = \left(\frac{1}{2}, -\frac{6}{5}, \frac{15}{2}, \frac{17}{4}, \frac{5}{3} - 0.1 \right)$$

in Fig. 7c, we obtain the original parameter values for system (1.1)

$$(\mu, \delta, k, \alpha, \beta) = (0.0552813, 0.00601937, 2.65328 \times 10^{-11}, 5.10635 \times 10^{-18}, -3.83488 \times 10^{-9}).$$

(iii) Hopf bifurcation of codimension one around $E_5(x_5, y_5)$ for $R_0 > 1$. The existence of one stable limit cycle arising from the supercritical Hopf bifurcation around $\tilde{E}(1, 1)$ in system (4.21) is given in Fig. 7d, which corresponds to the supercritical Hopf bifurcation around $E_5(x_5, y_5)$ for $R_0 > 1$ in system (2.2). Moreover, as in case (i), from the following parameter values for system (4.21)

$$(n, m, p, q, a) = \left(\frac{1}{2}, -\frac{6}{5}, \frac{15}{2}, \frac{22}{5}, \frac{5}{3} - 0.1 \right)$$

in Fig. 7d, we get the original parameter values for system (1.1)

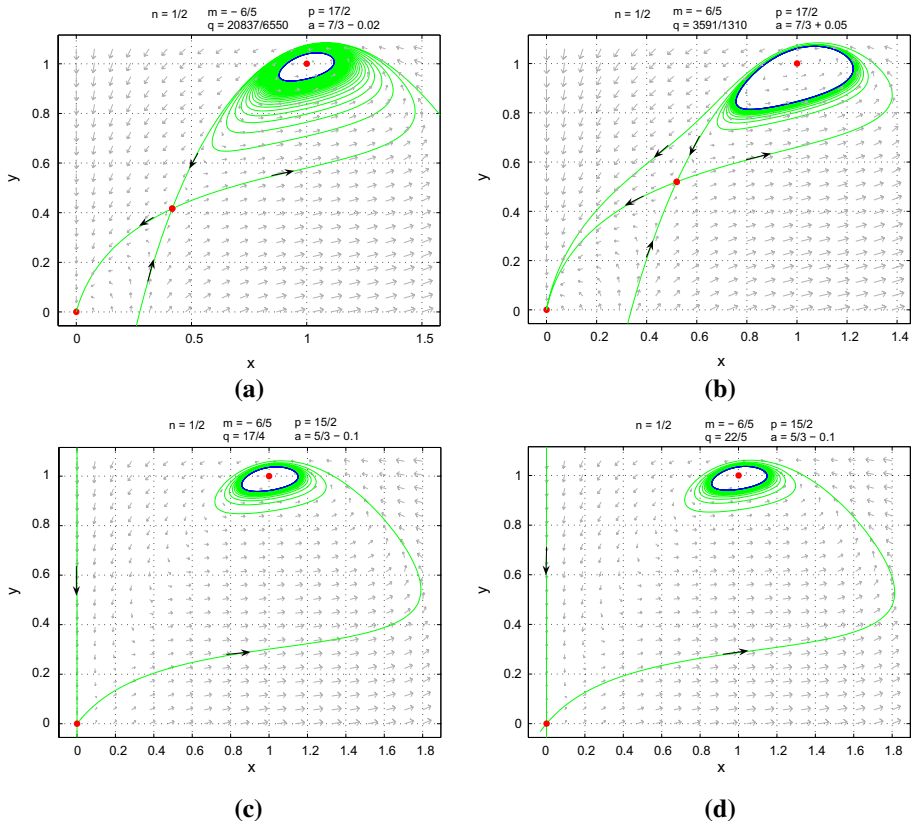


Fig. 7 Hopf bifurcations of codimension one around $\tilde{E}(1, 1)$ for system (4.21) corresponds to: **a** supercritical Hopf bifurcation around $E_3(x_3, y_3)$ for $R_* < R_0 < 1$ in system (2.2); **b** subcritical Hopf bifurcation around $E_3(x_3, y_3)$ for $R_* < R_0 < 1$ in system (2.2); **c** supercritical Hopf bifurcation around $E_4(x_4, y_4)$ for $R_0 = 1$ in system (2.2); **d** supercritical Hopf bifurcation around $E_4(x_4, y_4)$ for $R_0 > 1$ in system (2.2)

$$(\mu, \delta, k, \alpha, \beta) = (0.0794021, 0.0110579, 3.75468 \times 10^{-11}, 5.31265 \times 10^{-18}, -3.91158 \times 10^{-9}).$$

Remark 5.1 When $R_* < R_0 < 1$ or $R_0 > 1$, Zhou et al. [30] showed that system (2.2) exhibits subcritical Hopf bifurcation around E_3 or E_5 by choosing several sets of specific parameter values, but their chosen parameter values do not satisfy $p > q$.

Remark 5.2 In Fig. 7a, model (4.21) has three equilibria, a disease-free equilibrium which is a stable hyperbolic node, two endemic equilibria (a hyperbolic saddle and an unstable focus), and a stable limit cycle. If the initial populations lie in the right side of the two stable manifolds of the saddle, then the disease will tend to periodic fluctuations. If the initial populations lie in the left side of the two stable manifolds of the saddle, then the disease will die out.

Remark 5.3 In Fig. 7b, model (4.21) has three equilibria, a disease-free equilibrium which is a stable hyperbolic node, two endemic equilibria (a hyperbolic saddle and a stable focus), and an unstable limit cycle. If the initial populations lie on the limit cycle, then the disease will persist in the form of periodic oscillations. If the initial populations lie inside the limit cycle, then the disease will tend to a positive coexistence steady state. If the initial populations lie outside the limit cycle, the disease will die out for almost all positive initial populations.

Remark 5.4 When $k > k_0$, or $k = k_0$ and $\beta < \beta_0$ (i.e., $R_0 > 1$, or $R_0 = 1$ and $m < -\frac{1+q}{p}$), model (4.21) has only two equilibria. In Fig. 7c, d, there exist a disease-free equilibrium which is a saddle, one endemic equilibrium which is an unstable focus, and a stable limit cycle. We can see that the disease will tend to periodic outbreaks or persist in the form of a positive coexistent steady state for all positive initial populations.

(iv) Hopf bifurcation of codimension two around $E_3(x_3, y_3)$ for $R_* < R_0 < 1$. The existence of two limit cycles arising from the Hopf bifurcation of codimension two around $\tilde{E}(1, 1)$ in system (4.21) is given in Fig. 8, which corresponds to the Hopf bifurcation of codimension two around $E_3(x_3, y_3)$ for $R_* < R_0 < 1$ in system (2.2). We firstly fix $n = \frac{1}{2}$, $m = -\frac{6}{5}$ and $p = \frac{17}{2}$, then get $q = \frac{1861}{655}$ from $\sigma_1 = 0$, and get $a = \frac{7}{3}$ from $\text{tr}(J(\tilde{E})) = 0$, finally get $\sigma_2 = 135.766$. For this set of parameter values, $\tilde{E}(1, 1)$ is an unstable multiple focus with multiplicity 2. Next we first perturb q such that it increases to $\frac{1861}{655} + 0.15$, then $\tilde{E}(1, 1)$ becomes a stable multiple focus with multiplicity 1, an unstable limit cycle occurs around $\tilde{E}(1, 1)$ which is the outer limit cycle in Fig. 8. Secondly, we perturb a such that it reduces to $\frac{7}{3} - 0.005$, then $\tilde{E}(1, 1)$ becomes an unstable hyperbolic focus, another stable limit cycle occurs around $\tilde{E}(1, 1)$, which is the inner limit cycle in Fig. 8. Moreover, as in case (i), from the following parameter values for system (4.21)

$$(n, m, p, q, a) = \left(\frac{1}{2}, -\frac{6}{5}, \frac{17}{2}, \frac{1861}{655} + 0.15, \frac{7}{3} - 0.005 \right)$$

in Fig. 8, we obtain the original parameter values for system (1.1)

$$(\mu, \delta, k, \alpha, \beta) = (0.0316967, 0.00360857, 1.26849 \times 10^{-11}, 3.70814 \times 10^{-18}, -3.26794 \times 10^{-9}).$$

Remark 5.5 From Fig. 8, we can see the existence of two periodic coexistent oscillations and coexistence steady states when the infection rate k is smaller than the critical value k_0 given by (2.10), the psychological effect α is smaller than the critical value α_0 , and β is smaller than the critical value β_0 , i.e., $k < k_0, \alpha < \alpha_0$ and $\beta < \beta_0$ (corresponding to $R_* < R_0 < 1$ and $m < -\frac{1+q}{p}$). The disease will die out for almost all positive initial populations outside the outer unstable periodic orbit, will tend to periodic outbreaks for almost all positive initial populations on or inside the outer unstable periodic orbit, and will persist in the form of positive steady states when the initial populations lie on the positive equilibria or their stable manifolds.

(v) Hopf bifurcation of codimension three around $E_5(x_5, y_5)$ for $R_0 > 1$. The simulation of three limit cycles arising from the Hopf bifurcation of codimension three around $\tilde{E}(1, 1)$ in system (4.21) is shown in Fig. 9, which corresponds to the Hopf bifurcation of codimension three around $E_5(x_5, y_5)$ for $R_0 > 1$ in system (2.2). In order to numerically simulate the three limit cycles in system (4.21), we use the normal form to determine the parameter values. We take $(m, n) = (-0.15, 0.006)$ and obtain the critical values $p = 36.65608396 \dots$ and $q = 7.24695834 \dots$ from (4.29) and (4.27), respectively. At these critical values, $tr(J(\tilde{E})) = \sigma_1 = \sigma_2 = 0$ but $\sigma_3 = -3.72601443$. Next, we perturb the parameters p, q and a such that $0 < tr(J(\tilde{E})) \ll -\sigma_1 \ll \sigma_2 \ll -\sigma_3$, yielding three small-amplitude limit cycles around $\tilde{E}(1, 1)$. We apply the 4th-order Runge–Kutta method to run the simulations on a

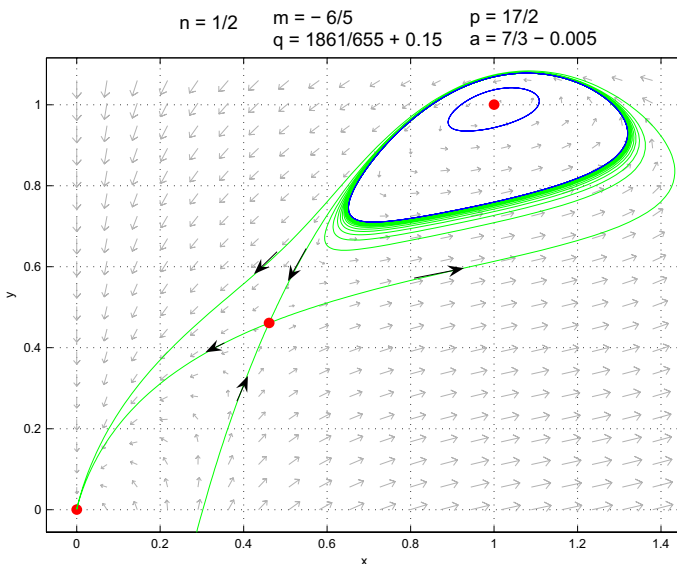


Fig. 8 Two limit cycles enclosing an unstable hyperbolic focus $\tilde{E}(1, 1)$ for system (4.21), which corresponds to two limit cycles around $E_3(x_3, y_3)$ for $R_* < R_0 < 1$ in system (2.2)

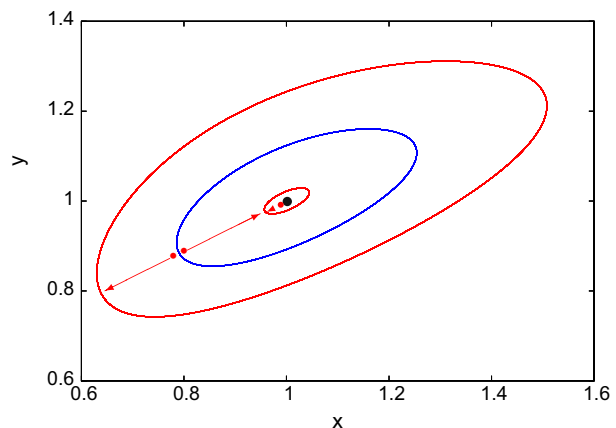
PC machine. Since the model is a two-dimensional differential system, we can use negative time steps in the integration scheme to simulate the unstable limit cycle. Since $-\sigma_1 < 0$, $\sigma_2 > 0$, and $\sigma_3 < 0$, the innermost and outermost limit cycles are stable while the middle one is unstable. All the three limit cycles enclose the equilibrium $\tilde{E}(1, 1)$ which is an unstable hyperbolic focus since $\text{tr}(J(\tilde{E})) > 0$. Our simulations show that the convergent speed is extremely slow and the process is very time consuming. For each limit cycle, we choose two initial points, one lying outside the limit cycle and one lying inside the cycle, and have trajectories initiated from both points converging to the limit cycle. (Note that convergence also appears for the unstable limit cycle since negative time steps are used.) The simulations of three limit cycles are shown in Fig. 9, where we only present the very last portion of each trajectory in order to avoid massive data plotting. The red and blue curves represent stable and unstable limit cycles, respectively.

Remark 5.6 When the infection rate k is larger than the critical value k_0 given by (2.10), i.e., $k > k_0$ (corresponding to $R_0 > 1$), we can see that the existence of three periodic coexistent solutions and a coexistence steady state from Fig. 9, where model (4.21) has only two equilibria: a disease-free equilibrium which is a saddle and an endemic equilibrium which is an unstable hyperbolic focus. The disease will tend to periodic outbreaks for almost all positive initial populations outside or inside the outermost stable periodic orbit, and will persist in the form of positive steady states or unstable periodic solutions when the initial populations lie on the unique positive equilibrium or the middle unstable periodic orbit.

6 Discussion

In this paper, we studied the global dynamics of the SIRS model (1.1) with a generalized nonmonotone incidence rate. By considering general parameter conditions, instead of specific parameter values as in Zhou et al. [30], we have shown that the basic reproduction number R_0 in model (1.1) does not act as a threshold value for disease spread anymore, and there exists a sub-threshold value $R_* (< 1)$ such that: (i) if $R_0 < R_*$, then the disease-free equilibrium is globally asymptotically stable; (ii) if $R_0 = R_*$, then the model has a

Fig. 9 Three limit cycles enclosing an unstable hyperbolic focus $\tilde{E}(1, 1)$ for system (4.21), which corresponds to three limit cycles around $E_5(x_5, y_5)$ for $R_0 > 1$ in system (2.2). The innermost and outermost limit cycles are stable and the middle limit cycle is unstable. The initial points for the simulation are also shown with positive time steps for convergence to the stable limit cycle (in red color) and negative time steps for convergence to the unstable limit cycles (in blue color) (Color figure online)



unique endemic equilibrium which is a nilpotent cusp of codimension at most three; (iii) if $R_* < R_0 < 1$, then the model has two endemic equilibria, one is a weak focus of multiplicity at least three and the other is a saddle; (iv) if $R_0 \geq 1$, then the unique endemic equilibrium is a weak focus of multiplicity at least three. As parameters vary, the model undergoes a sequence of bifurcations, including saddle-node bifurcation, backward bifurcation, Bogdanov–Takens bifurcation of codimension three, Hopf bifurcation, and degenerate Hopf bifurcation of codimension three.

Moreover, it is shown that there exist a critical value $\alpha = \alpha_0$ for the psychological effect, a critical value $k = k_0$ for the infection rate, and two critical values $\beta = \beta_0, \beta_1$ ($\beta_1 < \beta_0$) such that: (i) when $\alpha > \alpha_0$, or $\alpha \leq \alpha_0, k < k_0$ and $\beta \geq \beta_0$, or $k = k_0$ and $\beta \geq \beta_0$, the disease will die out for all positive initial populations; (ii) when $\alpha = \alpha_0$ and $\beta_1 \leq \beta < \beta_0$, the disease will die out for almost all positive initial populations; (iii) when $\alpha = \alpha_0$ and $\beta < \beta_1$, the disease will persist in the form of a positive coexistent steady state for some positive initial populations; and (iv) when $\alpha < \alpha_0, k < k_0$ and $\beta < \beta_0$, or $k = k_0$ and $\beta < \beta_0$, or $k > k_0$, the disease will persist in the form of multiple positive periodic coexistent oscillations and coexistent steady states for some positive initial populations. With different choices of parameter values, numerical simulations demonstrate that model (1.1) has one, two or three limit cycles due to these bifurcations.

The existence of limit cycles (isolated periodic orbits) for epidemic models is interesting and significant both in mathematics and applications since the existence of stable limit cycles provides a satisfactory explanation for disease recurrence and break out in a rather reproducible periodic manner, which may have profound implications for the control, prevention and prediction of disease transmission. For example, from the data of some diseases such as measles, researchers observed some complex periodic patterns, see Stone et al. [23].

On the other hand, the existence of multiple limit cycles implies diseases establish or break out in different periodic manners with different periods. Pattern changes of epidemics have been observed in some childhood infectious diseases such as measles, the major transitions are between regular cycles and irregular cycles, and from regionally synchronized oscillations to complex, spatially incoherent epidemics, etc. Measles is a natural ecological system that exhibits different dynamical transitions at different times and places, yet all of these transitions can be predicted as bifurcations of a single nonlinear model ([10]).

Notice that in a recent paper, we (Lu et al. [20]) studied model (1.1) with a different incidence rate

$$\frac{kI^2S}{1 + \beta I + \alpha I^2}, \quad (6.1)$$

in which the infection function first increases to a maximum when a new infectious disease emerges, then decreases due to psychological effect, and eventually tends to a saturation level due to crowding effect. Recall that the incidence rate considered in this paper takes the following form:

$$\frac{kIS}{1 + \beta I + \alpha I^2}, \quad (6.2)$$

which first increases to a maximum, then decreases and tends to zero when the number of infectious individuals become larger and larger. These two functions are certainly different (see Fig. 1.2 in [20] and Fig. 1b). Then it is necessary and interesting to compare

the dynamics of these two cases. In [20] we found that model (1.1) with incidence rate (6.1) has a weak focus of multiplicity at most two and a cusp of codimension at most two, undergoes saddle-node bifurcation, Bogdanov–Takens bifurcation of codimension two, Hopf bifurcation, and degenerate Hopf bifurcation of codimension two, and possesses one or two limit cycles for various parameter values. In this paper we showed that model (1.1) with incidence rate (6.2) has a weak focus of multiplicity at least three and a nilpotent cusp of codimension at most three, exhibits saddle-node bifurcation, backward bifurcation, Bogdanov–Takens bifurcation of codimension three, Hopf bifurcation, and degenerate Hopf bifurcation of codimension three, and has one, two or three limit cycles as parameters vary. This demonstrates that the dynamics of model (1.1) with incidence rate (6.2) are much more complex than that of model (1.1) with incidence rate (6.1).

From Theorem 4.9, we can see that system (4.21) may exhibit degenerate Hopf bifurcations of codimension larger than three if we let m and n be arbitrary parameters. It will be very interesting to study these bifurcations and we leave it for future consideration.

Acknowledgements We would like to thank the two anonymous reviewers for their helpful comments and suggestions which really helped us to improve the manuscript.

References

- Alexander, M.E., Moghadas, S.M.: Periodicity in an epidemic model with a generalized nonlinear incidence. *Math. Biosci.* **189**, 75–96 (2004)
- Alexander, M.E., Moghadas, S.M.: Bifurcations analysis of an SIRS epidemic model with generalized incidence. *SIAM J. Appl. Math.* **65**, 1794–1816 (2005)
- Capasso, V., Crosso, E., Serio, G.: I modelli matematici nella indagine epidemiologica. I Applicazione all'epidemia di colera verificatasi in Bari nel 1973. *Annali Sclavo* **19**, 193–208 (1977)
- Capasso, V., Serio, G.: A generalization of the Kermack–Mckendrick deterministic epidemic model. *Math. Biosci.* **42**, 43–61 (1978)
- Chow, S.-N., Li, C., Wang, D.: *Normal Forms and Bifurcations of Planar Vector Fields*. Cambridge University Press, Cambridge (1994)
- Dai, Y., Zhao, Y., Sang, B.: Four limit cycles in a predator–prey system of Leslie type with generalized Holling type III functional response. *Nonlinear Anal. Real World Appl.* **50**, 218–239 (2019)
- Derrick, W.R., van den Driessche, P.: A disease transmission in a nonconstant population. *J. Math. Biol.* **31**, 495–512 (1993)
- Dumortier, F., Roussarie, R., Sotomayor, J.: Generic 3-parameter families of vector fields on the plane, unfolding a singularity with nilpotent linear part. The cusp case of codimension 3, *Ergodic Theor. Dyn. Syst.* **7**(3), 375–413 (1987)
- Dushoff, J., Huang, W., Castillo-chavez, C.: Backward bifurcations and catastrophe in simple models of fatal disease. *J. Math. Biol.* **36**, 227–248 (1998)
- Earn, D.J.D., Rohani, P., Bolker, B.M., Grenfell, B.T.: A simple model for complex dynamical transitions in epidemics. *Science* **287**, 667–670 (2000)
- Hethcote, H.W., van den Driessche, P.: Some epidemiological models with nonlinear incidence. *J. Math. Biol.* **29**, 271–287 (1991)
- Hu, Z., Bi, P., Ma, W., Ruan, S.: Bifurcations of an SIRS epidemic model with nonlinear incidence rate. *Discrete Contin. Dyn. Syst. Ser. B* **15**, 93–112 (2011)
- Huang, J., Gong, Y., Chen, J.: Multiple bifurcations in a predator–prey system of Holling and leslie type with constant–yield prey harvesting. *Int. J. Bifurc. Chaos* **23**, 1350164 (2013). 24 pp
- Huang, J., Liu, S., Ruan, S., Zhang, X.: Bogdanov–Tankens bifurcation of codimension 3 in a predator–prey model with constant–yield predator harvesting. *Commun. Pure Appl. Anal.* **15**, 1041–1055 (2016)
- Kooij, R.E., Zegeling, A.: A predator–prey model with Ivlev's functional response. *J. Math. Anal. Appl.* **198**, 473–489 (1996)
- Lamontagne, Y., Coutu, C., Rousseau, C.: Bifurcation analysis of a predator–prey system with generalized Holling type III functional response. *J. Differ. Equ.* **79**, 132–167 (2008)

17. Li, C., Li, J., Ma, Z.: Codimension 3 B-T bifurcation in an epidemic model with a nonlinear incidence. *Discrete Contin. Dyn. Syst. Ser. B* **20**, 1107–1116 (2015)
18. Liu, W.-M., Levin, S., Iwasa, Y.: Influence of nonlinear incidence rates upon the behavior of SIRS epidemiological models. *J. Math. Biol.* **23**, 187–204 (1986)
19. Liu, S., Pang, L., Ruan, S., Zhang, X.: Global dynamics of avian influenza epidemic models with psychological effect. *Comput. Math. Methods Med.* (2015). <https://doi.org/10.1155/2015/913726>
20. Lu, M., Huang, J., Ruan, S., Yu, P.: Bifurcation analysis of an SIRS epidemic model with a generalized nonmonotone and saturated incidence rate. *J. Differ. Equ.* **267**, 1859–1898 (2019)
21. Perko, L.: *Differential Equations and Dynamical Systems*, 3rd edn. Springer, New York (2001)
22. Ruan, S., Wang, W.: Dynamical behavior of an epidemical model with with a nonlinear incidence rate. *J. Differ. Equ.* **188**, 135–163 (2003)
23. Stone, L., Olinky, R., Huppert, A.: Seasonal dynamics of recurrent epidemic. *Nature* **446**, 533–536 (2007)
24. Tang, Y., Huang, D., Ruan, S., Zhang, W.: Coexistence of limit cycles and homoclinic loops in a SIRS model with a nonlinear incidence rate. *SIAM J. Appl. Math.* **69**, 621–639 (2008)
25. van den Driessche, P., Watmough, J.: Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease of disease transmission. *Math. Biosci.* **180**, 29–48 (2002)
26. Xiao, D., Ruan, S.: Global analysis of an epidemic model with a nonlinear incidence rate. *Math. Biosci.* **208**, 419–429 (2007)
27. Xiao, D., Zhou, Y.: Qualitative analysis of an epidemic model. *Can. Appl. Math. Q.* **14**, 469–492 (2006)
28. Zhang, Z.: Proof of the uniqueness theorem of limit cycles of generalized Liénard equations. *Appl. Anal.* **23**, 63–76 (1986)
29. Zhang, Z., Ding, T., Huang, W., Dong, Z.: *Qualitative Theory of Differential Equations*. Translations of Mathematical Monographs, vol. 101. American Mathematical Society, Providence, RI (1992)
30. Zhou, Y., Xiao, D., Li, Y.: Bifurcations of an epidemic model with non-monotonic incidence rate of saturated mass action. *Chaos Solitons Fractals* **32**, 1903–1915 (2007)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Min Lu¹ · Jicai Huang¹ · Shigui Ruan² · Pei Yu³

Min Lu
lumin@mails.cnu.edu.cn

Jicai Huang
hjc@mail.cnu.edu.cn

Shigui Ruan
ruan@math.miami.edu

¹ School of Mathematics and Statistics, Central China Normal University, Wuhan 430079, Hubei, People's Republic of China

² Department of Mathematics, University of Miami, Coral Gables, FL 33124-4250, USA

³ Department of Applied Mathematics, Western University, London, ON N6A 5B7, Canada