Journal of Mathematical Biology (2024) 88:43
https://doi.org/10.1007/s00285-024-02065-0

Mathematical Biology

Viral infection dynamics with immune chemokines and CTL mobility modulated by the infected cell density

Hongying Shu[1] · Hai-Yang Jin[2] · Xiang-Sheng Wang[3] · Jianhong Wu[4]

Received: 4 July 2023 / Revised: 5 December 2023 / Accepted: 9 February 2024

Abstract

We study a viral infection model incorporating both cell-to-cell infection and immune chemokines. Based on experimental results in the literature, we make a standing assumption that the cytotoxic T lymphocytes (CTL) will move toward the location with more infected cells, while the diffusion rate of CTL is a decreasing function of the density of infected cells. We first establish the global existence and ultimate boundedness of the solution via a priori energy estimates. We then define the basic reproduction number of viral infection R_0 and prove (by the uniform persistence theory, Lyapunov function technique and LaSalle invariance principle) that the infection-free steady state E_0 is globally asymptotically stable if $R_0 < 1$. When $R_0 > 1$, then E_0 becomes unstable, and another basic reproduction number of CTL response R_1 becomes the dynamic threshold in the sense that if $R_1 < 1$, then the CTL-inactivated steady state E_1 is globally asymptotically stable; and if $R_1 > 1$, then the immune response is uniform persistent and, under an additional technical condition the CTL-activated steady state E_2 is globally asymptotically stable. To establish the global stability results, we need to prove point dissipativity, obtain uniform persistence, construct suitable Lyapunov functions, and apply the LaSalle invariance principle.

✉ Jianhong Wu
wujh@mathstat.yorku.ca

Hongying Shu
hshu@snnu.edu.cn

Hai-Yang Jin
mahyjin@scut.edu.cn

Xiang-Sheng Wang
xswang@louisiana.edu

1 School of Mathematics and Information Science, Shaanxi Normal University, Xi'an 710062, China

2 Department of Mathematics, South China University of Technology, Guangzhou 510640, China

3 Department of Mathematics, University of Louisiana at Lafayette, Lafayette, LA 70503, USA

4 Department of Mathematics and Statistics, York University, Toronto, ON M3J 1P3, Canada

Published online: 15 March 2024

Keywords Viral infection model · Immune chemokines · Cell-to-cell infection · Density-suppressed motility · Global dynamics

Mathematics Subject Classification 35A01 · 35K57 · 37B25 · 92C17

1 Introduction

Mathematical models have been playing an important role in understanding viral dynamics (Perelson and Nelson 1999). A simple three-compartment ordinary differential system was proposed in (Nowak et al. 1996) to study the interaction of uninfected target cells, infected cells, and free viruses. This model was further extended in (Nowak and Bangham 1996) to describe the crucial role of cytotoxic T lymphocytes (CTL) in antiviral defense. A general result was obtained in (Shu et al. 2013) in determining sharp conditions for global stability in equilibria of viral infection models with CTL immune responses and intracellular delays.

Chemokines compose a complicated communication system among all kinds of cell types, including immune cells (Rot and von Andrian 2004). However, the understanding of this mechanism is far from clear (Bromley et al. 2008). A simple model of the immune system was proposed by (Lee et al. 2017) to characterize the dynamics of the immune cells, the chemokines, and the antigens. It was shown that strong chemotaxis may induce nonlinear instability of the system. Recently, Zheng and Shan (Zheng and Shan 2023) modified the three-compartmental model of the immune system with general kinetic functions. They proved the global existence and boundedness of classical solutions under certain conditions. Stability and instability of the system were also investigated via the energy estimate and bootstrap method. As commented in (Dyson et al. 2008), the solutions of many existing models of chemotaxis may blow up at a finite time. However, in biological settings, the cell densities and the viral load shall be bounded. It is thus important to develop a viral infection model with immune chemokines so that the solutions exist globally.

Another important feature of viral dynamics is the cell-to-cell transmission mode, which allows viral particles to be transferred directly from an infected cell to an uninfected target cell through the formation of virological synapses (Galloway et al. 2015; Hübner et al. 2009; Martin and Sattentau 2009; Sattentau 2008). It has been revealed that over half of viral infections are due to the cell-to-cell infection mode (Iwami et al. 2015). Moreover, the cell-to-cell infection of HIV may permit ongoing replication even during antiretroviral therapy (Sigal 2011). Some earlier works on the cell-to-cell infection mode were given in (Gummuluru et al. 2000; Dixit and Perelson 2004). Lai and Zou (Lai and Zou 2014, 2015) proposed viral infection models with both virus-to-cell and cell-to-cell infection modes. They demonstrated that the basic reproduction number is the sum of the basic reproduction numbers for virus-to-cell and cell-to-cell infections, respectively. A similar result was obtained in (Deng et al. 2023) for a general model with immune responses and intracellular delays.

To the best of our knowledge, there is no existing model to couple the immune chemokines with both viral infection modes. In this paper, we will incorporate the viral population model with chemokines and cell-to-cell infection. According to the study in (Halle et al. 2016), CTL tends to migrate more slowly while recognizing infected cells. Hence, we will assume in our model that the diffusion rate of CTL is a decreasing function of the density of infected cells. More specifically, we propose to study the following reaction-diffusion system with general incidence functions:

$$\begin{cases} \partial_t T = d_T \Delta T + b(T) - f(T, I) - g(T, V), & x \in \Omega, t > 0, \\ \partial_t I = d_I \Delta I + f(T, I) + g(T, V) - r_1 I Z - \mu_I I, & x \in \Omega, t > 0, \\ \partial_t V = d_V \Delta V + kI - \mu_V V, & x \in \Omega, t > 0, \\ \partial_t Z = \Delta[d_Z(I) Z] + r_2 I Z - \mu_Z Z, & x \in \Omega, t > 0, \\ \partial_\nu T = \partial_\nu I = \partial_\nu V = \partial_\nu Z = 0, & x \in \partial\Omega, t > 0, \\ \{T, I, V, Z\}(x, 0) = \{T_0, I_0, V_0, Z_0\}(x), & x \in \Omega, \end{cases} \tag{1.1}$$

where $\Omega \subset \mathbb{R}^2$ is a bounded domain with smooth boundary $\partial\Omega$. $T(x,t)$, $I(x,t)$, $V(x,t)$ and $Z(x,t)$ denote the densities of uninfected targeted cells, infected targeted cells, viruses and cytotoxic T lymphocytes (CTL), respectively. d_T, d_I and d_V are positive diffusion constants for T, I and V, respectively. The growth function for uninfected targeted cells is given by $b(T)$. The cell-to-cell and virus-to-cell transmissions are characterized by the nonlinear functions $f(T, I)$ and $g(T, V)$, respectively. $r_1 IZ$ stands for the removal rate of infected cells by CTL and kI is the virion production rate. The recruitment rate of CTL is $r_2 IZ$. Upon scaling $\widetilde{Z} = (r_1/r_2)Z$, we may assume without loss of generality that $r_1 = r_2 = r$. The constant per capital death rates of I, V and Z are denoted by μ_I, μ_V and μ_Z, respectively. Finally, the function $d_Z(I) > 0$ characterizes the dependence of CTL motility on the density of infected cells. Note that $\Delta[d_Z(I)Z] = \nabla \cdot [d_Z(I)\nabla Z] + \nabla \cdot [d'_Z(I)Z\nabla I]$. The first term $\nabla \cdot [d_Z(I)\nabla Z]$ implies that the CTLs diffuse more slowly if the density of infected cells is higher, while the second term $\nabla \cdot [d'_Z(I)Z\nabla I]$ together with the assumption $d'_Z(I) \le 0$ indicates that the CTLs tend to move toward the location with higher density of infected cells. The term $\Delta[d_Z(I)Z]$ is also related to the density-suppressed motility, which may produce more complex dynamics and patterns (Fu et al. 2012; Liu 2011). The density-suppressed motility has been studied extensively in the context of chemotaxis (Keller and Segel 1970, 1971) and prey-taxis (Kareiva and Odell 1987). More precisely, it has been proved that the density-suppressed motility can prevent the blow-up of solution (Tao and Winkler 2017; Jin et al. 2018; Jiang et al. 2022; Fujie and Jiang 2020, 2021), enhance the diffusion (Winkler 2020), and trigger pattern formation (Ma et al. 2020; Jin and Wang 2021; Li and Wang 2021).

Notations: Fix $p_0 > 2$ and let $\mathbb{X} = W^{1,p_0}(\Omega, \mathbb{R})$ be the phase space for initial conditions. Moreover, we use a plus symbol in the subscript to denote the non-negative cone of the corresponding Banach space; for instance, $\mathbb{R}_+ = [0, \infty)$ and $\mathbb{X}_+ = W^{1,p_0}(\Omega, \mathbb{R}_+)$. For $1 \le p \le \infty$, $\|\cdot\|_p$ denotes the L^p-norm in $L^p(\bar{\Omega})$. Given a linear operator L, the spectral bound and spectral radius of L are denoted by $s(L)$ and $\rho(L)$, respectively. We say a linear dynamical system $u'(t) = Lu$ (or the linear operator L) is stable if $s(L) < 0$ and unstable if $s(L) > 0$. We use ∂_ν and dS to denote the normal derivative and differential form on $\partial\Omega$. The symbols ∇, ∇^2, and Δ stand for the gradient vector, Hessian matrix, and Laplacian, respectively. Especially, Δu

is the trace of $\nabla^2 u$. The notation $|\cdot|$ indicates the Euclidean norm of a vector, while $|\cdot|_E$ is the Frobenius norm of a matrix. The solution semiflow of (1.1) is denoted as $\Theta(t) : \mathbb{X}_+^4 \to \mathbb{X}_+^4$ for $t \geq 0$. For any initial condition $\phi \in \mathbb{X}_+^4$, the omega limit set of ϕ is $\omega(\phi) = \cap_{s\geq 0}\overline{\cup_{t\geq s}\Theta(t)\phi}$. For any subset $D \subset \mathbb{X}_+^4$, the omega limit set of D is $\omega(D) = \cup_{\phi\in D}\omega(\phi)$. We say the subset D is positively invariant if $\Theta(t)D \subset D$. Obviously, if a positively invariant subset D is closed then $\omega(D) \subset D$. Given a steady state $E \in \mathbb{X}_+^4$ such that $\Theta(t)E = E$ for all $t \geq 0$, the stable set of E is $W^s(E) = \{\phi \in \mathbb{X}_+^4 : \lim_{t\to\infty} \Theta(t)\phi = E\}$.

Basic assumptions: Throughout this paper, we set $r_1 = r_2 = r$, and make the following biologically relevant assumptions.

(H1) $b \in C^2(\mathbb{R}_+)$ with $b' < 0$ on $\mathbb{R}_+$. Moreover, there exists (a unique) $T_0 > 0$ such that $b(T_0) = 0$.
(H2) $f, g \in C^2(\mathbb{R}_+ \times \mathbb{R}_+)$ such that f and g vanish on the boundary of first quadrant, the first-order partial derivatives $\partial_T f$, $\partial_I f$, $\partial_T g$, and $\partial_V g$ are strictly positive in the interior of the first quadrant, the second-order partial derivatives $\partial_{II} f$ and $\partial_{VV} g$ are non-positive in the first quadrant, and there exists $K > 0$ such that $f(T, I) \leq KTI$ and $g(T, V) \leq KTV$ for all $T, I, V \in \mathbb{R}_+$.
(H3) $d_Z \in C^2(\mathbb{R}_+)$ such that $d_Z > 0$ and $d_Z' \leq 0$ on $\mathbb{R}_+$.

Main results: Our main results can be summarized in the following two theorems.

Theorem 1.1 (Global existence and point dissipativity) *Let $\Omega \subset \mathbb{R}^2$ and $\mathbb{X}_+ = W^{1,p_0}(\Omega, \mathbb{R}_+)$. Assume (H1)-(H3). For any initial value in $[\mathbb{X}_+ \setminus \{0\}]^4$, the system (1.1) has a unique nonnegative classical solution (T, I, V, Z) satisfying*

$$(T, V, Z) \in [C([0,\infty) \times \bar{\Omega}) \cap C^{2,1}((0,\infty) \times \bar{\Omega})]^3$$

and

$$I \in C([0,\infty) \times \bar{\Omega}) \cap C^{2,1}((0,\infty) \times \bar{\Omega}) \cap L^\infty_{loc}([0,\infty); W^{1,\infty}(\Omega)).$$

Moreover, there exists $C > 0$ independent of initial conditions such that

$$\limsup_{t\to\infty}(\|T(\cdot,t)\|_{\mathbb{X}} + \|I(\cdot,t)\|_{\mathbb{X}} + \|V(\cdot,t)\|_{\mathbb{X}} + \|Z(\cdot,t)\|_{\mathbb{X}}) \leq C.$$

Here, $\mathbb{X} = W^{1,p_0}(\Omega, \mathbb{R})$ and $\|u\|_{\mathbb{X}} = \|u\|_{p_0} + \|\nabla u\|_{p_0}$ for $u \in \mathbb{X}$.

Theorem 1.2 *(Global dynamics) Assume (H1)-(H3) and define*

$$R_0 := \frac{\partial_I f(T_0, 0)}{\mu_I} + \frac{k\partial_V g(T_0, 0)}{\mu_V \mu_I}.$$

Let $(T(t), I(t), V(t), V(t))$ be any solution to (1.1) with initial value in $[\mathbb{X}_+ \setminus \{0\}]^4$.

- *If $R_0 \leq 1$, then the solution will converge to the infection-free steady state $E_0 = (T_0, 0, 0, 0)$ as $t \to \infty$.*

- *If* $R_0 > 1$*, then there exists a unique CTL-inactivated steady state* $E_1 = (T_1, I_1, V_1, 0)$*. We define* $R_1 := rI_1/(\mu Z_1)$ *and assume that* $f(T, I) = h_0(T)h_d(I)$ *and* $g(T, V) = h_0(T)h_i(V)$*, where* $h_0, h_d, h_i \in C^2(\mathbb{R}_+)$*.*
  - *If* $R_0 > 1$ *and* $R_1 \le 1$*, then the solution will converge to* E_1*.*
  - *If* $R_0 > 1$ *and* $R_1 > 1$*, then there exists a unique CTL-activated steady state* $E_2 = (T_2, I_2, V_2, Z_2)$*. Moreover, the solution converges to* E_2 *provided*

$$d_I \ge \frac{Z_2}{4I_2} \max_{0 \le I \le \infty} \frac{|Id_Z'(I)|^2}{d_Z(I)}.$$

Outline of the paper: In Sect. 2, we establish the boundedness and global existence of solutions. The key step is to use the fourth equation of (1.1) to obtain the boundedness of $\int_t^{t+\tau} \int_\Omega Z^2 dxds$ for some fixed $\tau \in (0, 1]$. With this in hand, we can further show that $\int_t^{t+\tau} \|\Delta I\|_2^2 ds$ is uniformly bounded. A direct application of L^2 estimate on Z gives

$$\frac{d}{dt}\|Z\|_{L^2}^2 \le K_1 \|Z\|_{L^2}^2 (\|\Delta I\|_{L^2}^2 + 1).$$

Making use of a lemma that generalizes Gronwall's inequality, we obtain $\|Z\|_{L^2} \le K_2$, which can be further improved by the standard bootstrap argument and Moser iteration method to $\|Z\|_{L^\infty} \le K_3$. Note that the constants K_1, K_2, and K_3 may depend on initial conditions. In this section, we shall also prove a stronger result that the solution is ultimately bounded by a constant that is independent of initial conditions. This stronger result is also called the point dissipativity of the system. Once we have proved that the system is point dissipative, it immediately follows from (Hale 1988, Theorem 3.4.8) that the system possesses a nonempty global attractor in $\mathbb{X}_+^4$. In Sect. 3, we first define two basic reproduction numbers and establish the existence of three steady states. A local stability analysis of these steady states will also be conducted. Furthermore, we will use the Lyapunov function technique and the LaSalle invariance principle to obtain global asymptotic stability of the steady states under certain conditions. Due to the non-closedness of the positively invariant sets on which some Lyapunov functions are defined, we will need to prove the uniform persistence of the solution before applying the LaSalle invariance principle. In Sect. 4, we will conduct numerical simulations to illustrate the theoretical results and conclude the paper with a brief discussion.

2 Global existence and point dissipativity

In this section, we first prove that the solution exists globally and is uniformly bounded by a constant that depends on the initial conditions. As we shall see later, we will need a stronger result on the point dissipativity of the system to prove the uniform persistence of the solution. Hence, we will further prove that the solution is ultimately bounded by a constant independent of the initial conditions. The key ingredients in the proof are Gronwall-type inequality, Gagliardo-Nirenberg inequality, and Moser iteration.

2.1 Global existence

First, we state the following lemma which is a generalization of (Stinner et al. 2014, Lemma 3.4) and Gronwall's inequality.

Lemma 2.1 *Let $t_m > 0$, $\tau \in (0, t_m)$, $c_1 > 0$ and $c_2 > 0$. Assume that y is a non-negative and continuously differentiable function on $[0, t_m)$ satisfying $y'(t) + c_1 y(t) \le h(t)$ for all $t \in (0, t_m)$, where h is a non-negative and locally integrable function on $(0, t_m)$ such that $\int_{t-\tau}^{t} h(s)ds \le c_2$ for all $t \in [\tau, t_m)$. Denote $c_3 = \max\{y(0), c_2/(c_1\tau) + c_2\}$. We then have $y(j\tau) \le c_3$ for all non-negative integer $j < t_m/\tau$, and $y(t) \le c_3 + c_2$ for all $t \in [0, t_m)$.*

Proof We only need to prove the first assertion because the second one follows immediately from the non-negativity of h and y together with an integration of the inequality $y' + c_1 y \le h$. Obviously, $y(0) \le c_3$. Assume $y(j\tau) \le c_3$ and $j + 1 < t_m/\tau$. We consider two cases: (i) $y(s) \ge c_2/(c_1\tau)$ for all $s \in [j\tau, (j+1)\tau]$; and (ii) $y(s_0) < c_2/(c_1\tau)$ for some $s_0 \in [j\tau, (j + 1)\tau]$. In the first case, we integrate $y' + c_1 y \le h$ to obtain $y((j + 1)\tau) \le y(j\tau) \le c_3$; while in the second case, an integration of $y' + c_1 y \le h$ gives $y((j + 1)\tau) \le y(s_0) + c_2 \le c_3$. This completes the proof. □

Now, we are ready to prove the global existence of classical solutions as follows.

Theorem 2.2 *Let $\Omega \subset \mathbb{R}^2$ and $\mathbb{X}_+ = W^{1,p_0}(\Omega, \mathbb{R}_+)$. Assume (H1)-(H3). For any initial value in $[\mathbb{X}_+ \setminus \{0\}]^4$, the system (1.1) has a unique nonnegative classical solution (T, I, V, Z) satisfying*

$$(T, V, Z) \in [C([0, \infty) \times \bar{\Omega}) \cap C^{2,1}((0, \infty) \times \bar{\Omega})]^3$$

and

$$I \in C([0, \infty) \times \bar{\Omega}) \cap C^{2,1}((0, \infty) \times \bar{\Omega}) \cap L^\infty_{loc}([0, \infty); W^{1,\infty}(\Omega)).$$

Proof By adopting the idea in (Tao and Winkler 2011, Lemma 2.1) and (Jin et al. 2018), we can prove the existence of local solutions by using the Schauder fixed point theorem and the parabolic regularity theory; namely, there exists $t_m \in (0, \infty]$ such that the problem (1.1) has a unique nonnegative classical solution (T, I, V, Z) satisfying

$$(T, V, Z) \in [C([0, t_m) \times \bar{\Omega}) \cap C^{2,1}((0, t_m) \times \bar{\Omega})]^3$$

and

$$I \in C([0, t_m) \times \bar{\Omega}) \cap C^{2,1}((0, t_m) \times \bar{\Omega}) \cap L^\infty_{loc}([0, t_m); W^{1,\infty}(\Omega)).$$

Moreover, we have either $t_m = \infty$ or

$$\limsup_{t \to t_m^-} (\|T(\cdot, t)\|_{L^\infty} + \|V(\cdot, t)\|_{L^\infty} + \|Z(\cdot, t)\|_{L^\infty} + \|I(\cdot, t)\|_{W^{1,\infty}}) = \infty.$$

Note that $[0, t_m)$ is the maximum interval for the existence of classical solutions. It suffices to prove that the solution does not blow up as $t \to t_m^-$. Denote $\tau = \min\{1, t_m/2\}$. For convenience, we will use c to denote a generic (large) positive constant that is independent of t. We proceed in the following nine steps.

Step 1. $\|T(\cdot, t)\|_\infty \le c$.

In view of the non-negativity of the solution and (H2), we obtain from the first equation of (1.1) that $\partial_t T \le d_T \Delta T + b(T)$. By (H1) and maximum principle, we have $T(x, t) \le \max\{\|T(\cdot, 0)\|_\infty, T_0\}$.

Step 2. $\|I(\cdot, t)\|_1 + \|V(\cdot, t)\|_1 + \|Z(\cdot, t)\|_1 \le c$.

From (1.1), we obtain

$$\begin{aligned}\frac{d}{dt}\int_\Omega (T + I + Z)dx &= \int_\Omega [b(T) - \mu_I I - \mu_Z Z]dx\\ &\le \int_\Omega [b(0) + c - T - \mu_I I - \mu_Z Z]dx.\end{aligned}$$

It then follows from Gronwall's inequality that $\|I(\cdot, t)\|_1 + \|Z(\cdot, t)\|_1 \le c$. Integrating the fourth equation of (1.1) and applying Gronwall's inequality yield $\|V(\cdot, t)\|_1 \le c$.

Step 3. For any $p > 1$, $\|I(\cdot, t)\|_p + \|V(\cdot, t)\|_p \le c$.

Multiplying the second equation of (1.1) by I^{p-1} and integrating over the domain Ω, we obtain

$$\begin{aligned}\frac{1}{p}\frac{d}{dt}\int_\Omega I^p dx = d_I \int_\Omega (\Delta I)I^{p-1}dx + \int_\Omega [f(T, I) + g(T, V)]I^{p-1}dx\\ - \int_\Omega (r_1 Z + \mu_I)I^p dx.\end{aligned}$$

The Green's identity gives

$$\int_\Omega (\Delta I)I^{p-1}dx = -\int_\Omega (\nabla I)\cdot((p-1)I^{p-2}\nabla I)dx = -\frac{4(p-1)}{p^2}\int_\Omega |\nabla I^{p/2}|^2 dx.$$

On account of (H2), we have

$$\int_\Omega [f(T, I) + g(T, V)]I^{p-1}dx \le \int_\Omega KT(I^p + I^{p-1}V)dx \le c\int_\Omega (I^p + V^p)dx.$$

Hence,

$$\frac{1}{p}\frac{d}{dt}\int_\Omega I^p dx + \frac{4(p-1)d_I}{p^2}\int_\Omega |\nabla I^{p/2}|^2 dx \le c\int_\Omega (I^p + V^p)dx. \tag{2.1}$$

Similarly, we multiply the third equation of (1.1) by V^{p-1} and then apply Green's identity to have

$$\frac{1}{p}\frac{d}{dt}\int_\Omega V^p dx + \frac{4(p-1)d_V}{p^2}\int_\Omega |\nabla V^{p/2}|^2 dx \le c\int_\Omega (I^p + V^p)dx. \tag{2.2}$$

Now, we use the Gagliardo-Nirenberg inequality and boundedness of $\|I\|_q + \|V\|_q$ to find

$$\|I^{p/2}\|_2^2 \le c(\|\nabla I^{p/2}\|_2^{2(1-1/p)}\|I^{p/2}\|_{2/p}^{2/p} + \|I^{p/2}\|_{2/p}^2) \le \frac{1}{c}\|\nabla I^{p/2}\|_2^2 + c, \quad (2.3)$$

and similarly,

$$\|V^{p/2}\|_2^2 \le \frac{1}{c}\|\nabla V^{p/2}\|_2^2 + c, \quad (2.4)$$

Taking a weighted sum of inequalities (2.1), (2.2), (2.3), and (2.4), and then applying Gronwall's inequality, we obtain $\|I(\cdot,t)\|_p^p + \|V(\cdot,t)\|_p^p \le c$.

Step 4. $\|I(\cdot,t)\|_\infty + \|V(\cdot,t)\|_\infty \le c$.
It follows from the third equation of (1.1) that

$$V(\cdot,t) = e^{(d_V\Delta-\mu_V)t}V(\cdot,0) + k\int_0^t e^{(d_V\Delta-\mu_V)(t-s)}I(\cdot,s)ds.$$

By estimates of Neumann heat semigroup (Horstmann and Winkler 2005; Kowalczyk and Szymańska 2008) and uniform boundedness of $\|I\|_3$, we find $\varepsilon > 0$ such that

$$\|V(\cdot,t)\|_\infty \le \|V(\cdot,0)\|_\infty + c\int_0^t [1+(t-s)^{-1/3}]e^{-\varepsilon(t-s)}ds \le c.$$

Similarly, using the assumptions in (H2), we obtain from the equation of (1.1) and the uniform boundedness of $\|I\|_3 + \|V\|_3$ that $\|I(\cdot,t)\|_\infty \le c$.

Step 5. $\int_{t-\tau}^t \int_\Omega Z^2(x,s)dxds \le c$ for all $t \in [\tau, t_m)$.
To this end, we shall use the duality arguments based on some nice ideas developed in (Tao and Winkler 2017). Denote $\delta_1 = \mu_Z/[2d_Z(0)]$ and let $\mathcal{A}$ be the self-adjoint positive operator $-\Delta + \delta_1$ with Neumann boundary condition on $L^2(\Omega)$. Then $\mathcal{A}^{-1}$ is a bounded and positive operator on $L^2(\Omega)$. Adding the second and the fourth equation of (1.1) gives (recalling that $r_1 = r_2 = r$)

$$\begin{aligned}\partial_t(I+Z) + \mathcal{A}[d_I I + d_Z(I)Z] = \delta_1[d_I I + d_Z(I)Z]& \\ + f(T,I) + g(T,V) - \mu_I I - \mu_Z Z \le c&,\end{aligned}$$

where we have made use of (H2), $\delta_1 < \mu_Z/d_Z(0) \le \mu_Z/d_Z(I)$, and the uniform boundedness of $\|T\|_\infty + \|I\|_\infty + \|V\|_\infty$. Multiplying the above inequality by $2\mathcal{A}^{-1}(I+Z) \ge 0$ gives

$$\frac{d}{dt}\int_\Omega |\mathcal{A}^{-1/2}(I+Z)|^2dx + 2\int_\Omega (I+Z)[d_I I + d_Z(I)Z]dx \le c\int_\Omega \mathcal{A}^{-1}(I+Z)dx. \quad (2.5)$$

Since $\|I\|_\infty$ is uniformly bounded, it follows from (H1) that $d_Z(I) \geq d_Z(c_2) > 0$. Let $\delta_2 > 0$ be sufficiently small such that

$$d_I I + d_Z(I)Z \geq \delta_2(I + Z). \tag{2.6}$$

It follows from boundedness of $\mathcal{A}^{-1}$ and $\mathcal{A}^{-1/2}$ on $L^2(\Omega)$ that

$$\|\mathcal{A}^{-1}(I+Z)\|_1 \leq c\|\mathcal{A}^{-1}(I+Z)\|_2 \leq c\|I+Z\|_2 \leq \delta_3\|I+Z\|_2^2 + c, \tag{2.7}$$

and

$$\|\mathcal{A}^{-1/2}(I+Z)\|_2^2 \leq c\|I+Z\|_2^2, \tag{2.8}$$

where $\delta_3 > 0$ is sufficiently small. A combination of the above four inequalities (2.5–2.8) yields

$$\frac{d}{dt}\int_\Omega |\mathcal{A}^{-1/2}(I+Z)|^2 dx + \frac{1}{c}\int_\Omega |\mathcal{A}^{-1/2}(I+Z)|^2 dx + \frac{1}{c}\int_\Omega (I+Z)^2 dx \leq c.$$

Gronwall's inequality implies that $\|\mathcal{A}^{-1/2}(I+Z)\|_2^2 \leq c$, which together with an integration of the above inequality gives $\int_{t-\tau}^t \int_\Omega (I+Z)^2 dxds \leq c$. Since I and Z are positive, we have $\int_{t-\tau}^t \int_\Omega Z^2(x,s)dxds \leq c$ for all $t \in [\tau, t_m)$.

Step 6. $\|\nabla I(\cdot,t)\|_2 \leq c$ for all $t \in [0, t_m)$ and $\int_{t-\tau}^t \|\Delta I(\cdot,s)\|_2^2 ds \leq c$ for all $t \in [\tau, t_m)$.

We multiple the second equation of (1.1) by $-2\Delta I$ and apply Green's identity to find

$$\frac{d}{dt}\int_\Omega |\nabla I|^2 dx + 2d_I\int_\Omega |\Delta I|^2 dx \leq d_I\int_\Omega |\Delta I|^2 dx + c\left(\int_\Omega Z^2 dx + 1\right). \tag{2.9}$$

By Gagliardo-Nirenberg inequality and the uniform boundedness of $\|I\|_2$, we have

$$\|\nabla I\|_2^2 \leq c(\|\Delta I\|_2\|I\|_2 + \|I\|_2^2) \leq \frac{d_I}{2}\|\Delta I\|_2^2 + c. \tag{2.10}$$

Adding the above two inequalities (2.9) and (2.10) yields

$$\frac{d}{dt}\int_\Omega |\nabla I|^2 dx + \int_\Omega |\nabla I|^2 dx + \frac{d_I}{2}\int_\Omega |\Delta I|^2 dx \leq c\int_\Omega Z^2 dx + c.$$

By Lemma 2.1, we have $\|\nabla I\|_2 \leq c$. An integration of the above inequality then gives $\int_{t-\tau}^t \|\Delta I\|_2^2 ds \leq c$ for all $t \in [\tau, t_m)$.

Step 7. $\|Z(\cdot,t)\|_2 \leq c$.

Multiplying the fourth equation of (1.1) by $2Z$ and applying Green's identity give

$$\frac{d}{dt}\int_\Omega Z^2 dx + 2\int_\Omega d_Z(I)|\nabla Z|^2 dx \leq c\int_\Omega (Z|\nabla Z|\cdot|\nabla I| + Z^2)dx.$$

By (H1) and uniform boundedness of $\|I\|_\infty$, we can find a small $\delta_4 > 0$ such that $d_Z(I) \geq \delta_4$. It then follows from the above inequality and Cauchy's inequality that

$$\frac{d}{dt}\int_\Omega Z^2 dx + 2\delta_4 \int_\Omega |\nabla Z|^2 dx \leq \delta_4 \int_\Omega |\nabla Z|^2 dx + c\int_\Omega (Z^2|\nabla I|^2 + Z^2)dx.$$

By Hölder's inequality, we obtain

$$\frac{d}{dt}\|Z\|_2^2 + \delta_4\|\nabla Z\|_2^2 \leq c(\|Z\|_4^2\|\nabla I\|_4^2 + \|Z\|_2^2).$$

The right-hand side of the above inequality can be further estimated via the Gagliardo-Nirenberg inequality and elliptic regularity theory. Hence, we obtain

$$\begin{aligned}\frac{d}{dt}\|Z\|_2^2 + \delta_4\|\nabla Z\|_2^2 &\leq c(\|\nabla Z\|_2\|Z\|_2 + \|Z\|_2^2)(\|\Delta I\|_2\|\nabla I\|_2 + \|\nabla I\|_2^2) + c\|Z\|_2^2\\ &\leq c(\|\nabla Z\|_2\|Z\|_2 + \|Z\|_2^2)(\|\Delta I\|_2 + c)\\ &\leq c\|\nabla Z\|_2\|Z\|_2\|\Delta I\|_2 + c\|\nabla Z\|_2\|Z\|_2 + c\|Z\|_2^2\|\Delta I\|_2 + c\|Z\|_2^2\\ &\leq \delta_4\|\nabla Z\|_2^2 + c\|Z\|_2^2(\|\Delta I\|_2^2 + 1).\end{aligned}$$

This together with the uniform boundedness of $\int_{t-\tau}^t \|\Delta I\|_2^2 ds$ and Lemma 2.1 implies that $\|Z(\cdot,t)\|_2^2 \leq c$.

Step 8. For any $p > 1$, $\|Z(\cdot,t)\|_p \leq c$.
Since $\|Z\|_2$ is uniformly bounded, an application of parabolic regularity (Kowalczyk and Szymańska 2008, Lemma 1) to the second equation of (1.1) gives the uniform boundedness of $\|\nabla I\|_4$. Recall that $d_Z(I) \geq \delta_4 > 0$. We multiple the fourth equation of (1.1) by Z^{p-1} and apply Green's identity to find

$$\frac{1}{p}\frac{d}{dt}\int_\Omega Z^p dx + (p-1)\delta_4\int_\Omega Z^{p-2}|\nabla Z|^2 dx \leq c\int_\Omega Z^{p-1}|\nabla Z||\nabla I| dx + c\int_\Omega Z^p dx.$$

Applying Cauchy's inequality, Hölder's inequality, and the Gagliardo-Nirenberg inequality, we obtain

$$\begin{aligned}&\frac{1}{p}\frac{d}{dt}\int_\Omega Z^p dx + \frac{1}{c}\int_\Omega Z^{p-2}|\nabla Z|^2 dx + \int_\Omega Z^p dx \leq c\int_\Omega Z^p(|\nabla I|^2 + 1)dx\\ \leq& c\|Z^{p/2}\|_4^2 \leq c(\|\nabla Z^{p/2}\|_2^{2(1-1/p)}\|Z^{p/2}\|_{4/p}^{2/p} + \|Z^{p/2}\|_{4/p}^2)\\ \leq& c(\|\nabla Z^{p/2}\|_2^{2(1-1/p)} + 1) \leq \frac{1}{c}\|\nabla Z^{p/2}\|_2^2 + c.\end{aligned}$$

By Gronwall's inequality, we have $\|Z(\cdot,t)\|_p \leq c$.

Step 9. $\|Z(\cdot,t)\|_\infty \leq c$.
Since $\|Z\|_3$ is uniformly bounded, a simple application of parabolic regularity (Kowalczyk and Szymańska 2008, Lemma 1) to the second equation of (1.1) gives the uniform boundedness of $\|\nabla I\|_\infty$. Based on a modified Moser iteration as in (Tao and Winkler (2012), Lemma A.1), we can obtain $\|Z(\cdot,t)\|_\infty \leq c$ directly.

Since $\|T\|_\infty + \|I\|_\infty + \|\nabla I\|_\infty + \|V\|_\infty + \|Z\|_\infty$ is uniformly bounded, the solution cannot blow up as $t \to t_m^-$. Thus, $t_m = \infty$; namely, the solution exists globally. This completes the proof. $\square$

2.2 Point dissipativity

In this subsection, we will show that the solution is ultimately bounded by a constant independent of initial conditions. However, the constants $0 \le t_0 \le t_1 \le \cdots$ may depend on the initial condition. We first state a lemma that is analogous to Lemma 2.1.

Lemma 2.3 *Let $y(t) \ge 0$ be continuously differentiable, $h(t) \ge 0$ be locally integrable on $[t_0, \infty)$, and $y'(t) + C_1 y(t) \le h(t)$ for all $t \ge t_0$, where $C_1 > 0$ and $t_0 \ge 0$. Assume $\limsup_{t\to\infty} \int_{t-1}^{t} h(s)ds \le C_2$, then we have $\limsup_{t\to\infty} y(t) \le C_2/C_1 + 2C_2$.*

Proof For any sufficiently small $\varepsilon > 0$, there exists $t_1 \ge t_0$ such that $\int_{t-1}^{t} h(s)ds \le C_2 + \varepsilon$ for all $t \ge t_1$. We claim that there exists $t_2 \ge t_1$ such that $y(t_2) \le (C_2 + 2\varepsilon)/C_1$; otherwise, $y(t) > (C_2 + 2\varepsilon)/C_1$ for all $t \ge t_1$, which implies that $y'(t) < h(t) - (C_2 + 2\varepsilon)$ and $y(t+1) < y(t) - \varepsilon$ for all $t \ge t_1$, a contradiction. So, we have $y(t_2) \le (C_2 + 2\varepsilon)/C_1$ for some $t_2 \ge t_1$, which together with Lemma 2.1 implies that $y(t) \le \max\{(C_2 + 2\varepsilon)/C_1 + C_2, C_2/C_1 + 2C_2\}$ for all $t \ge t_2$. Especially, $\limsup_{t\to\infty} y(t) \le \max\{(C_2 + 2\varepsilon)/C_1 + C_2, C_2/C_1 + 2C_2\}$. Since $\varepsilon > 0$ is arbitrarily small, we may choose $\varepsilon < C_1 C_2/2$ and obtain $\limsup_{t\to\infty} y(t) \le C_2/C_1 + 2C_2$. This completes that proof. $\square$

Now, we are ready to show that the system is point dissipative.

Theorem 2.4 *Let $\Omega \subset \mathbb{R}^2$ and $\mathbb{X}_+ = W^{1,p_0}(\Omega, \mathbb{R}_+)$. Assume (H1)-(H3). Let (T, I, V, Z) be the classical solution of system (1.1) obtained in Theorem 2.2. There exists $C > 0$ independent of initial conditions such that*

$$\limsup_{t\to\infty}(\|T(\cdot,t)\|_{\mathbb{X}} + \|I(\cdot,t)\|_{\mathbb{X}} + \|V(\cdot,t)\|_{\mathbb{X}} + \|Z(\cdot,t)\|_{\mathbb{X}}) \le C. \tag{2.11}$$

Here, $\mathbb{X} = W^{1,p_0}(\Omega, \mathbb{R})$ and $\|u\|_{\mathbb{X}} = \|u\|_{p_0} + \|\nabla u\|_{p_0}$ for $u \in \mathbb{X}$.

Proof We proceed in the following 12 steps. The first nine steps are similar to the proof of Theorem 2.2 and the details are omitted. As we shall see in the last step, the main difficulty lies in the estimate of $\|\nabla Z\|_{2p}$ with $p > 1$. For convenience, we use C_i (with $i = 1, 2, \cdots$) to denote large positive constants that are independent of the initial condition.

Step 1. From the first equation of (1.1), we have $\partial_t T \le d_T \Delta T + b(T)$. By (H1), we obtain from the comparison principle that $\limsup_{t\to\infty} \|T\|_\infty \le T_0$.

Step 2. There exist $C_1 > 0$ and $t_0 \ge 0$ such that $\|T\|_\infty \le C_1$ for all $t \ge t_0$. From the system (1.1), we have

$$\frac{d}{dt}\int_\Omega (T + I + Z)dx \le \int_\Omega [b(0) + C_1 - T - \mu_I I - \mu_Z Z]dx, \quad t \ge t_0$$

which implies that $\limsup_{t\to\infty}(\|I\|_1 + \|Z\|_1) \le C_2$. An integration of the third equation of (1.1) gives $\limsup_{t\to\infty}\|V\|_1 \le C_3$.

Step 3. There exist $C_4 > 0$ and $t_1 \ge t_0$ such that $\|I\|_1 + \|V\|_1 + \|Z\|_1 \le C_4$ for all $t \ge t_1$. For any $p > 1$, it follows from the second and third equation of (1.1) and Gagliardo-Nirenberg inequality that

$$\frac{d}{dt}\int_\Omega (I^p + V^p)dx + C_5 \int_\Omega (I^p + V^p)dx \le C_6, \quad t \ge t_1$$

which implies that $\limsup_{t\to\infty}(\|I\|_p + \|V\|_p) \le C_7$. Note that C_7 depends on p.

Step 4. There exist $C_8 > 0$ and $t_2 \ge t_1$ such that $\|I\|_3 + \|V\|_3 \le C_8$ for all $t \ge t_2$. Since we have fixed $p = 3$, the constant C_8 is independent of p. By the third equation of (1.1), we have

$$V(\cdot, t) = e^{(d_V\Delta - \mu_V)(t-t_2)}V(\cdot, t_2) + k\int_{t_2}^t e^{(d_V\Delta - \mu_V)(t-s)}I(\cdot, s)ds, \quad t \ge t_2,$$

which together with the estimate of Neumann heat semigroup (Horstmann and Winkler 2005; Kowalczyk and Szymańska 2008) implies that $\limsup_{t\to\infty}\|V\|_\infty \le C_9$. Similarly, we obtain from the second equation of (1.1) that $\limsup_{t\to\infty}\|I\|_\infty \le C_{10}$.

Step 5. There exist $C_{11} > 0$ and $t_3 \ge t_2$ such that $\|T\|_\infty + \|I\|_\infty + \|V\|_\infty \le C_{11}$ for all $t \ge t_3$. Let $\mathcal{A}$ be the self-adjoint positive operator $-\Delta + \mu_Z/[2d_Z(0)]$ with Neumann boundary condition on $\partial\Omega$. It follows from the second and third equations of (1.1) that

$$\frac{d}{dt}\int_\Omega |\mathcal{A}^{-1/2}(I + Z)|^2dx + \frac{1}{C_{12}}\int_\Omega |\mathcal{A}^{-1/2}(I + Z)|^2dx + \frac{1}{C_{12}}\int_\Omega (I + Z)^2dx \le C_{13},$$

for $t \ge t_3$. Hence, we obtain $\limsup_{t\to\infty}\int_{t-1}^t\int_\Omega (I + Z)^2dxds \le C_{14}$.

Step 6. From the second equation of (1.1) and Gagliardo-Nirenberg inequality, we have

$$\frac{d}{dt}\int_\Omega |\nabla I|^2dx + \int_\Omega |\nabla I|^2dx + \frac{d_I}{2}\int_\Omega |\Delta I|^2dx \le C_{15}\int_\Omega Z^2dx + C_{15}, \quad t \ge t_3,$$

which together with Lemma 2.3 implies $\limsup_{t\to\infty}(\|\nabla I\|_2 + \int_{t-1}^t \|\Delta I\|_2^2ds) \le C_{16}$.

Step 7. There exist $C_{17} > 0$ and $t_4 \ge t_3$ such that $\|\nabla I\|_2 + \int_{t-1}^t(\|\Delta I\|_2^2 + \|Z\|_2^2)ds \le C_{17}$ for all $t \ge t_4 \ge 1$. It then follows from the fourth equation of (1.1) and Gagliardo-Nirenberg inequality that

$$\frac{d}{dt}\|Z\|_2^2 \le C_{18}\|Z\|_2^2(\|\Delta I\|_2^2 + 1), \quad t \ge t_4.$$

For any $t \geq t_4 + 1$, there exists $s \in [t-1, t]$ such that $\|Z(\cdot, s)\|_2^2 \leq C_{17}$. It then follows from the above inequality that $\|Z(\cdot, t)\|_2^2 \leq C_{17} e^{C_{18}(C_{17}+1)}$ for all $t \geq t_4 + 1$. Especially, $\limsup_{t\to\infty} \|Z\|_2 \leq C_{19}$.

Step 8. There exist $C_{20} > 0$ and $t_5 \geq t_4$ such that $\|Z\|_2 \leq C_{20}$ for all $t \geq t_5$. An application of (Kowalczyk and Szymańska 2008, Lemma 1) to the second equation of (1.1) gives $\|\nabla I\|_4 \leq C_{21}$ for $t \geq t_5$. For any $p > 1$, by the fourth equation of (1.1) and Gagliardo-Nirenberg inequality, we obtain

$$\frac{d}{dt}\int_\Omega Z^p dx + \int_\Omega Z^p dx \leq C_{22}, \quad t \geq t_5,$$

which implies that $\limsup_{t\to\infty} \|Z\|_p \leq C_{23}$.

Step 9. There exist $C_{24} > 0$ and $t_6 \geq t_5$ such that $\|Z\|_3 \leq C_{24}$ for all $t \geq t_6$. Applying (Kowalczyk and Szymańska 2008, Lemma 1), we can derive that $\|\nabla T\|_\infty + \|\nabla I\|_\infty + \|\nabla V\|_\infty \leq C_{25}$ for $t \geq t_6$. Based on a modified Moser iteration as in (Tao and Winkler (2012), Lemma A.1), we can show that $\|Z(\cdot, t)\|_p$ is ultimately bounded for any finite $p > 0$. This together with (Dung 1997, Theorem 1) then gives $\limsup_{t\to\infty} \|Z\|_\infty \leq C_{26}$.

Step 10. There exist $C_{27} > 0$ and $t_7 \geq t_6$ such that $\|Z\|_\infty \leq C_{27}$ for all $t \geq t_7$. It then follows from the fourth equation of (1.1) and the Gagliardo-Nirenberg inequality that

$$\frac{d}{dt}\int_\Omega |\nabla Z|^2 dx + \int_\Omega |\nabla Z|^2 dx \leq C_{28}, \quad t \geq t_7,$$

which implies that $\limsup_{t\to\infty} \|\nabla Z\|_2^2 \leq C_{28}$.

Step 11. There exist $C_{29} > 0$ and $t_8 \geq t_7$ such that $\|\nabla Z\|_2 \leq C_{29}$ for all $t \geq t_8$. Denote $J := \nabla I$. It follows from the second equation of (1.1) that

$$\partial_t J = d_I \Delta J - \mu_I J + \Psi(T, I, V, Z, \nabla T, J, \nabla Z),$$

where

$$\Psi = (\partial_T f + \partial_T g)\nabla T + \partial_I f J + \partial_Z g \nabla Z - rJZ - rI\nabla Z.$$

It is easily seen that $\|\Psi\|_2 \leq C_{30}$ for $t \geq t_8$. For each $p \in [1, \infty)$, we apply the method of variations of constants to the equation for J and use the estimates of Neumann heat semigroup (Horstmann and Winkler 2005; Kowalczyk and Szymańska 2008) to obtain $\|\nabla J\|_p \leq C_{31}$ for all $t \geq t_8$. Especially, $\limsup_{t\to\infty} \|\Delta I\|_p \leq C_{32}$.

Step 12. For each $p > 1$, there exist $C_{33} > 0$ and $t_9 \geq t_8$ such that $\|\Delta I\|_{2p} \leq C_{33}$ for all $t \geq t_9$. It follows from the fourth equation of (1.1) and Green's identity that

$$\frac{1}{2p}\frac{d}{dt}\int_\Omega |\nabla Z|^{2p} dx = \int_\Omega |\nabla Z|^{2p-2}\nabla Z \cdot \nabla(\partial_t Z) dx = K_1 + K_2 + K_3, \tag{2.12}$$

where

$$
\begin{aligned}
K_1 &:= -\int_\Omega \nabla\cdot(|\nabla Z|^{2p-2}\nabla Z)\nabla\cdot[d_Z(I)\nabla Z]dx,\\
K_2 &:= -\int_\Omega \nabla\cdot(|\nabla Z|^{2p-2}\nabla Z)\nabla\cdot[Zd_Z'(I)\nabla I]dx,\\
K_3 &:= \int_\Omega |\nabla Z|^{2p-2}\nabla Z\cdot(rI\nabla Z + rZ\nabla I - \mu_Z\nabla Z)dx.
\end{aligned}
$$

For $t \geq t_9$, we have $\|I\|_\infty \leq C_{11}$, $\|\nabla I\|_\infty \leq C_{25}$ and $\|Z\|_\infty \leq C_{27}$. Consequently,

$$
K_3 \leq C_{34}\left(\int_\Omega |\nabla Z|^{2p}dx + 1\right), \tag{2.13}
$$

and

$$
K_2 \leq C_{35}\int_\Omega [|\nabla Z|^{2p-3}|\nabla(|\nabla Z|^2)| + |\nabla Z|^{2p-2}|\Delta Z|](|\nabla Z| + |\Delta I| + 1)dx.
$$

To estimate K_1, we apply Green's identity again and use the identity $\nabla Z\cdot\nabla(\Delta Z) = \frac{1}{2}\Delta(|\nabla Z|^2) - |\nabla^2 Z|_E^2$, where $|\nabla^2 Z|_E$ is the Frobenius norm of the Hessian matrix $\nabla^2 Z$. It is readily seen that

$$
\begin{aligned}
K_1 =& -\int_\Omega [\nabla(|\nabla Z|^{2p-2})\cdot\nabla Z]\nabla\cdot[d_Z(I)\nabla Z]dx\\
&+ \int_\Omega \nabla(|\nabla Z|^{2p-2}\Delta Z)\cdot[d_Z(I)\nabla Z]dx\\
=& K_{11} + K_{12} + K_{13},
\end{aligned} \tag{2.14}
$$

where

$$
\begin{aligned}
K_{11} =& -\int_\Omega [\nabla(|\nabla Z|^{2p-2})\cdot\nabla Z][d_Z'(I)\nabla I\cdot\nabla Z]dx\\
\leq& C_{36}\int_\Omega |\nabla Z|^{2p-2}|\nabla(|\nabla Z|^2)|dx,
\end{aligned} \tag{2.15}
$$

$$
\begin{aligned}
K_{12} =& \frac{1}{2}\int_\Omega d_Z(I)|\nabla Z|^{2p-2}\Delta(|\nabla Z|^2)dx\\
=& \frac{1}{2}\int_{\partial\Omega} d_Z(I)|\nabla Z|^{2p-2}\partial_\nu(|\nabla Z|^2)dS - \frac{1}{2}\int_\Omega [d_Z'(I)|\nabla Z|^{2p-2}\nabla I]\cdot\nabla(|\nabla Z|^2)dx\\
&- \frac{p-1}{2}\int_\Omega d_Z(I)|\nabla Z|^{2p-4}|\nabla(|\nabla Z|^2)|^2dx,
\end{aligned}
$$

$$
K_{13} = -\int_\Omega d_Z(I)|\nabla Z|^{2p-2}|\nabla^2 Z|_E^2dx \leq -d_Z(C_{11})\int_\Omega |\nabla Z|^{2p-2}|\nabla^2 Z|_E^2dx. \tag{2.16}
$$

By (Mizoguchi and Souplet 2014, Lemma 4.2), we have $|\partial_\nu(|\nabla Z|^2)| \le 2\kappa(\partial\Omega)|\nabla Z|^2$, where $\kappa(\partial\Omega)$ is the maximum of curvatures on $\partial\Omega$. It follows that

$$\begin{aligned} K_{12} \le &\kappa(\partial\Omega) d_Z(0) \int_{\partial\Omega} |\nabla Z|^{2p} dS + C_{37} \int_\Omega |\nabla Z|^{2p-2} |\nabla(|\nabla Z|^2)| dx \\ &- \frac{d_Z(C_{11})(p-1)}{2} \int_\Omega |\nabla Z|^{2p-4} |\nabla(|\nabla Z|^2)|^2 dx. \end{aligned}$$

In view of the trace inequality (Souplet and Quittner 2007, Remark 52.9):

$$\|u\|_{L^2(\partial\Omega)} \le \varepsilon \|\nabla u\|_{L^2(\Omega)} + C(\varepsilon)\|u\|_{L^2(\Omega)}$$

and Cauchy's inequality: $2uv \le \varepsilon u^2 + (1/\varepsilon)v^2$ with any small $\varepsilon > 0$, we further obtain

$$K_{12} \le C_{38} \int_\Omega |\nabla Z|^{2p} dx - \frac{d_Z(C_{11})(p-1)}{3} \int_\Omega |\nabla Z|^{2p-4} |\nabla(|\nabla Z|^2)|^2 dx. \tag{2.17}$$

Another application of Cauchy's inequality to (2.15) gives

$$K_{11} \le C_{39} \int_\Omega |\nabla Z|^{2p} dx + \frac{d_Z(C_{11})(p-1)}{9} \int_\Omega |\nabla Z|^{2p-4} |\nabla(|\nabla Z|^2)|^2 dx. \tag{2.18}$$

Note that $|\Delta Z| \le \sqrt{2}|\nabla^2 Z|_E$. We apply Cauchy's inequality again to K_2 and make use of $\|\Delta I\|_{2p} \le C_{33}$ to find

$$\begin{aligned} K_2 \le &\frac{d_Z(C_{11})(p-1)}{9} \int_\Omega |\nabla Z|^{2p-4} |\nabla(|\nabla Z|^2)|^2 dx \\ &+ \frac{d_Z(C_{11})}{2} \int_\Omega |\nabla Z|^{2p-2} |\nabla^2 Z|_E^2 dx + C_{40} \int_\Omega |\nabla Z|^{2p} dx + C_{40}. \end{aligned} \tag{2.19}$$

A combination of (2.12), (2.13), (2.14), (2.16), (2.17), (2.18), and (2.19) yields

$$\begin{aligned} &\frac{1}{2p}\frac{d}{dt} \int_\Omega |\nabla Z|^{2p} dx + \frac{d_Z(C_{11})(p-1)}{9} \int_\Omega |\nabla Z|^{2p-4} |\nabla(|\nabla Z|^2)|^2 dx \\ &\qquad + \frac{d_Z(C_{11})}{2} \int_\Omega |\nabla Z|^{2p-2} |\nabla^2 Z|_E^2 dx \le C_{41} \int_\Omega |\nabla Z|^{2p} dx + C_{41}. \end{aligned}$$

Finally, we use Green's identity, $|\Delta Z| \le \sqrt{2}|\nabla^2 Z|_E$ and Cauchy's inequality to obtain

$$\begin{aligned}
&(C_{41}+2)\int_\Omega |\nabla Z|^{2p}dx \\
&= (C_{41}+2)\int_\Omega (|\nabla Z|^{2p-2}\nabla Z)\cdot \nabla Z dx \\
&= -(C_{41}+2)\int_\Omega Z[|\nabla Z|^{2p-2}\Delta Z + (p-1)|\nabla Z|^{2p-4}\nabla(|\nabla Z|^2)\cdot \nabla Z]dx \\
&\le \frac{d_Z(C_{11})(p-1)}{9}\int_\Omega |\nabla Z|^{2p-4}|\nabla(|\nabla Z|^2)|^2 dx \\
&\quad + \frac{d_Z(C_{11})}{2}\int_\Omega |\nabla Z|^{2p-2}|\nabla^2 Z|_E^2 dx + C_{42}\int_\Omega |\nabla Z|^{2p-2}dx,
\end{aligned}$$

and

$$C_{42}\int_\Omega |\nabla Z|^{2p-2}dx \le \int_\Omega |\nabla Z|^{2p}dx + C_{43}.$$

Consequently, we have

$$\frac{1}{2p}\frac{d}{dt}\int_\Omega |\nabla Z|^{2p}dx + \int_\Omega |\nabla Z|^{2p}dx \le C_{44},$$

which implies $\limsup_{t\to\infty} \|\nabla Z\|_{2p}^{2p} \le C_{44}$.

Therefore, we have proved the point dissipativity of the system. □

Coupling Theorem 2.2 and Theorem 2.4 gives Theorem 1.1.

3 Global dynamics

In this section, we investigate the long-time behaviors of the solutions to the system (1.1). First, we study the constant steady state $E = (T, I, V, Z)$ that satisfies the following equilibrium equations

$$b(T) = f(T, I) + g(T, V) = (rZ + \mu_I)I, \quad V = kI/\mu_V, \quad (rI - \mu_Z)Z = 0. \tag{3.1}$$

In view of (H1) and (H2), one can easily check that (3.1) possibly has three types of homogeneous steady states: the infection-free steady state $E_0 := (T_0, 0, 0, 0)$, the CTL-inactivated steady state $E_1 := (T_1, I_1, V_1, 0)$, and the CTL-inactivated steady state $E_2 := (T_2, I_2, V_2, Z_2)$. In the following subsections, we will analyze the existence and stability of these steady states.

3.1 Basic reproduction numbers and local analyses

In this subsection, we will establish the existence and local stability results for constant steady states. In view of (H1) and (H2), there exists a unique infection-free steady state $E_0 = (T_0, 0, 0, 0)$ of (1.1). Linearizing the system (1.1) about E_0 gives two equations

$$\partial_t T = d_T \Delta T + [b'(T) - \partial_T f(T_0, 0) - \partial_T g(T_0, 0)]T - \partial_I f(T_0, 0)I - \partial_V g(T_0, 0)V$$

and $\partial_t Z = d_Z(0)\Delta Z - \mu_Z Z$, which do not affect the stability of E_0, and a decoupled system

$$\begin{cases} \partial_t I = d_I \Delta I + \partial_I f(T_0, 0)I + \partial_V g(T_0, 0)V - \mu_I I, \\ \partial_t V = d_V \Delta V + kI - \mu_V V. \end{cases}$$

The linear operator on the right-hand side of the above equation can be decomposed as $\mathcal{F} - \mathcal{V}$, where

$$\mathcal{F} = \begin{pmatrix} \partial_I f(T_0, 0) & \partial_V g(T_0, 0) \\ 0 & 0 \end{pmatrix}, \quad \mathcal{V} = \begin{pmatrix} -d_I \Delta + \mu_I & 0 \\ -k & -d_V \Delta + \mu_V \end{pmatrix}.$$

Now, we define the basic reproduction number of viral infection as $R_0 = \rho(\mathcal{F}\mathcal{V}^{-1})$, the spectral radius of the next generation operator $\mathcal{F}\mathcal{V}^{-1}$ on the Banach space $\mathbb{X}^2 = W^{1,p_0}(\Omega, \mathbb{R}^2)$. Let $\mathcal{B}_d = \partial_I f(T_0, 0)(-d_I \Delta + \mu_I)^{-1}$ be the next generation operator of direct transmission and $\mathcal{B}_i = k\partial_V g(T_0, 0)(-d_V \Delta + \mu_V)^{-1}(-d_I \Delta + \mu_I)^{-1}$ be the next generation operator of indirect transmission. A simple calculation gives $R_0 = \rho(\mathcal{B}_d + \mathcal{B}_i)$. Note that the constant function $1 \in \mathbb{X}$ is an eigenfunction of $\mathcal{B}_d + \mathcal{B}_i$. It follows from Krein-Rutman theorem that the corresponding eigenvalue is a principal eigenvalue; namely,

$$R_0 = \frac{\partial_I f(T_0, 0)}{\mu_I} + \frac{k\partial_V g(T_0, 0)}{\mu_V \mu_I}. \tag{3.2}$$

Remark 3.1 The biological interpretation of the above formula is clear. The first fraction gives the average number of newly infected cells through cell-to-cell transmission during the life span of an infected cell which is introduced to the uninfected target cells with density T_0. The second fraction counts the average number of new viruses produced during the indirect transmission cycle when a virus is introduced to the uninfected target cells with density T_0. These two fractions, denoted by R_d and R_i, are also referred to as the basic reproduction numbers of direct and indirect transmissions, respectively. The above formula coincides with the earlier work on viral models without spatial diffusion (Lai and Zou 2014, 2015; Pourbashash et al. 2014); see also (Magal et al. 2019) for a similar formula of R_0 for an HIV model without immune response. Since $R_0 = R_d + R_i$, it is concluded that the ignorance of either transmission mechanism will underestimate the basic reproduction number of viral infection.

Proposition 3.2 *Assume (H1)-(H3). If $R_0 < 1$, then the infection-free steady state E_0 is linearly stable. If $R_0 > 1$, then E_0 is linearly unstable.*

Proof Since the linearized equations for T and Z are stable, it suffices to show that the decoupled linear system for I and V is stable when $R_0 < 1$ and unstable when $R_0 > 1$.

If $R_0 < 1$, we claim that all eigenvalues of $\mathcal{F} - \mathcal{V}$ have negative real parts. Assume to the contrary, there exist $\lambda \in \mathbb{C}$ with $\mathrm{Re}\lambda \geq 0$ and an eigenvalue $\xi \geq 0$ of $-\Delta$ with Neumann boundary condition on $\partial\Omega$ such that

$$\det \begin{pmatrix} \lambda + d_I\xi + \mu_I - \partial_I f(T_0, 0) & -\partial_V g(T_0, 0) \\ -k & \lambda + d_V\xi + \mu_V \end{pmatrix} = 0.$$

A simple calculation gives

$$\frac{\lambda + d_I\xi}{\mu_I} + 1 = \frac{\partial_I f(T_0, 0)}{\mu_I} + \frac{k\partial_V g(T_0, 0)}{\mu_I(\lambda + d_V\xi + \mu_V)}.$$

The modulus of the left-hand side is bounded below by one, while the modulus of the right-hand side is bounded above by $R_0 < 1$. This leads to a contradiction. Hence, all eigenvalues of $\mathcal{F} - \mathcal{V}$ have negative real parts, which implies that E_0 is linearly stable.

If $R_0 > 1$, we have $-\mu_V(\partial_I f(T_0, 0) - \mu_I) - k\partial_V g(T_0, 0) < 0$; namely, the determinant of

$$A = \begin{pmatrix} \partial_I f(T_0, 0) - \mu_I & \partial_V g(T_0, 0) \\ k & -\mu_V \end{pmatrix}$$

is negative. Hence, there exist an eigenvalue $\lambda \in \mathbb{C}$ with $\mathrm{Re}\lambda > 0$ and an eigenvector $w \in \mathbb{R}^2 \setminus \{0\}$ such that $Aw = \lambda w$. We regard w as a constant vector-valued function on Ω. It turns out that λ is also an eigenvalue of the linear operator $\mathcal{F} - \mathcal{V}$ with eigenfunction w. Therefore, E_0 is linearly unstable. $\square$

Next, we shall study the existence and linear stability of CTL-inactivated steady state $E_1 = (T_1, I_1, V_1, 0)$. We first show the CTL-inactivated steady state E_1 exists uniquely if and only if $R_0 > 1$. For $T \in \mathbb{R}_+$, define $i(T) = b(T)/\mu_I$, $v(T) = ki(T)/\mu_V = kb(T)/(\mu_I\mu_V)$, and

$$R(T) = \frac{f(T, i(T))}{\mu_I i(T)} + \frac{kg(T, v(T))}{\mu_I\mu_V v(T)}, \quad T \neq T_0. \tag{3.3}$$

When $T = T_0$, we define

$$R(T_0) := \lim_{T \to T_0} R(T).$$

Since $i(T_0) = v(T_0) = 0$, we obtain from L'Hôpital's rule that

$$\lim_{T \to T_0} \frac{f(T, i(T))}{\mu_I i(T)} = \frac{\partial_T f(T_0, 0) + \partial_I f(T_0, 0)i'(T_0)}{\mu_I i'(T_0)} = \frac{\partial_I f(T_0, 0)}{\mu_I},$$

and

$$\lim_{T\to T_0}\frac{kg(T,v(T))}{\mu_I\mu_V v(T)}=\frac{k\partial_T g(T_0,0)+\partial_V g(T_0,0)v'(T_0)}{\mu_I\mu_V v'(T_0)}=\frac{k\partial_V g(T_0,0)}{\mu_V\mu_I}.$$

In view of (3.2), we obtain $R(T_0) = R_0$. On account of (H1) and (H2), we have $i'(T) < 0$, $v'(T) < 0$ and $R'(T) > 0$ for $T \in \mathbb{R}_+$. Next, we shall show that if $R_0 > 1$, there exists a unique CTL-inactivated steady state E_1. Noting steady state E_1 satisfies (3.1), we only need to check that $R(T) = 1$ has a unique solution. One can easily check that $R(0) = 0$, which combined with the facts $R'(T) > 0$ and $R(T_0) = R_0 > 1$ implies that there exists a unique $T_1 \in (0, T_0)$ such that $R(T_1) = 1$. On the other hand, we can show that $R_0 > 1$ if E_1 exists. In fact, if E_1 exists, one has $R(T_1) = 1$, which together with the facts $R(T_0) = R_0$ and $R'(T) > 0$ as well as $0 < T_1 < T_0$, gives $1 = R(T_1) < R(T_0) = R_0$.

Assuming $R_0 > 1$, we linearize (1.1) about E_1 to find

$$\begin{cases}\partial_t T = d_T\Delta T + [b'(T_1) - \partial_T f - \partial_T g]T - \partial_I f I - \partial_V g V,\\ \partial_t I = d_I\Delta I + (\partial_T f + \partial_T g)T + (\partial_I f - \mu_I)I + \partial_V g V - rI_1 Z,\\ \partial_t V = d_V\Delta V + kI - \mu_V V,\end{cases}\tag{3.4}$$

and a decoupled equation

$$\partial_t Z = d_Z(I_1)\Delta Z + (rI_1 - \mu_Z)Z.\tag{3.5}$$

Here, the variables of $\partial_T f$ and $\partial_I f$ are (T_1, I_1), while the variables of $\partial_T g$ and $\partial_V g$ are (T_1, V_1). We first claim that the system (3.4) (with $Z = 0$) is stable. Assume to the contrary, then there exist $\lambda \in \mathbb{C}$ with $\mathrm{Re}\lambda \geq 0$ and an eigenvalue $\xi \geq 0$ of $-\Delta$ with Neumann boundary condition on $\partial\Omega$ such that

$$\det\begin{pmatrix}\lambda + d_T\xi - b'(T_1) + \partial_T f + \partial_T g & \partial_I f & \partial_V g\\ -\partial_T f - \partial_T g & \lambda + d_I\xi + \mu_I - \partial_I f & -\partial_V g\\ 0 & -k & \lambda + d_V\xi + \mu_V\end{pmatrix} = 0.$$

A simplification of the above equation gives

$$\left[1+\frac{\partial_T f+\partial_T g}{\lambda+d_T\xi-b'(T_1)}\right]\left(1+\frac{\lambda+d_I\xi}{\mu_I}\right)=\frac{\partial_I f}{\mu_I}+\frac{k\partial_V g}{\mu_I(\lambda+d_V\xi+\mu_V)}.$$

Taking modulus on both sides yields

$$1<\frac{\partial_I f}{\mu_I}+\frac{k\partial_V g}{\mu_I\mu_V}\leq\frac{f}{\mu_I I_1}+\frac{kg}{\mu_I\mu_V V_1}=1,$$

a contradiction. Therefore, the linearized system (3.4) for T, I, V (with $Z = 0$) is stable. The local stability of E_1 is then determined by the stability of (3.5); namely, by the sign of $R_1 - 1$, where

$$R_1 := \frac{rI_1}{\mu_Z} \tag{3.6}$$

is the basic reproduction number of the immune response. We summarize the above results in the following proposition.

Proposition 3.3 *Assume (H1)-(H3). The CTL-inactivated steady state E_1 exists uniquely (i.e., $0 < T_1 < T_0$, $I_1 > 0$ and $V_1 > 0$) if and only if $R_0 > 1$. Moreover, if $R_0 > 1$, then E_1 is linearly stable when $R_1 < 1$ and linearly unstable when $R_1 > 1$.*

Finally, we shall study the existence and locally asymptotic stability of the CTL-activated steady state $E_2 = (T_2, I_2, V_2, Z_2)$. In fact, we have the following proposition.

Proposition 3.4 *Assume (H1)–(H3). The CTL-activated steady state E_2 exists uniquely (i.e., $0 < T_2 < T_0$, $I_2 > 0$, $V_2 > 0$ and $Z_2 > 0$) if and only if $R_0 > 1$ and $R_1 > 1$. Moreover, E_2 is linearly stable whenever it exists.*

Proof From (3.1), we have $I_2 = \mu_Z/r$ and $V_2 = k\mu_Z/(r\mu_V)$. To prove the existence of E_2, it suffices to show that $H(T, I_2, V_2) = 1$ has a unique solution $T_2 \in (0, T_0)$ such that $b(T_2) > \mu_I I_2$ and then $Z_2 = [b(T_2)/I_2 - \mu_I]/r$, where

$$H(T, I, V) = \frac{f(T, I) + g(T, V)}{b(T)}$$

is strictly increasing in T, I, V due to (H1)-(H2). To achieve this, we observe that $R_1 > 1$ gives $I_1 > \mu_Z/r$, which entails $I_1 > \mu_Z/r = I_2$ and $V_1 = kI_1/\mu_V > V_2$. Consequently, $H(T_1, I_2, V_2) < H(T_1, I_1, V_1) = R(T_1) = 1$. Since $H(T, I_2, V_2) \to \infty$ as T approaches T_0 from the left, the equation $H(T, I_2, V_2) = 1$ has a unique solution $T_2 \in (T_1, T_0)$. Moreover, noting the fact that $b'(T) < 0$ in (H1), we obtain $b(T_2) > b(T_1) = \mu_I I_1 > \mu_I I_2$.

Next, we show that $R_0 > 1$ and $R_1 > 1$ are necessary conditions for the existence of E_2. In fact, if E_2 exists, then $Z_2 > 0$ implies $T_2 < T_0$, $I_2 < b(T_2)/\mu_I = i(T_2)$ and $V_2 = kI_2/\mu_V < kb(T_2)/(\mu_I\mu_V) = v(T_2)$. Consequently, using the definition of $R(T)$ in (3.3), one has $1 = H(T_2, I_2, V_2) < R(T_2)$. It then follows from $T_2 < T_0$ and (H1), as well as the fact $R'(T) > 0$, one has $R_0 = R(T_0) > R(T_2) > 1$. We now claim $R_1 > 1$. Otherwise, $I_1 \le \mu_Z/r = I_2$, which implies $b(T_2) > \mu_I I_2 \ge \mu_I I_1 = b(T_1)$. By (H1), we have $T_1 > T_2$, which gives $R(T_2) < R(T_1) = 1$, a contradiction.

From the above argument, we also note that there exists at most one CTL-activated steady state. Linearizing (1.1) about E_2 gives

$$\begin{cases} \partial_t T = & d_T \Delta T + [b'(T_2) - \partial_T f - \partial_T g]T - \partial_I f I - \partial_V g V, \\ \partial_t I = & d_I \Delta I + [\partial_T f + \partial_T g]T + [\partial_I f - rZ_2 - \mu_I]I + \partial_V g V - rI_2 Z, \\ \partial_t V = & d_V \Delta V + kI - \mu_V V, \\ \partial_t Z = & d_Z(I_2)\Delta Z + d_Z'(I_2)Z_2 \Delta I + rZ_2 I, \end{cases}$$

where the variables of $\partial_T f$ and $\partial_I f$ are (T_2, I_2), while the variables of $\partial_T g$ and $\partial_V g$ are (T_2, V_2). We will prove by contradiction that all eigenvalues of the linear operator corresponding to the above linear system have negative real parts. Assume to the contrary that there exist $\lambda \in \mathbb{C}$ with $\mathrm{Re}\lambda \geq 0$ and $\xi \geq 0$ such that $\det(\lambda + B) = 0$, where

$$B = \begin{pmatrix} d_T\xi - b'(T_2) + \partial_T f + \partial_T g & \partial_I f & \partial_V g & 0 \\ -\partial_T f - \partial_T g & d_I\xi + \mu_I + rZ_2 - \partial_I f & -\partial_V g & rI_2 \\ 0 & -k & d_V\xi + \mu_V & 0 \\ 0 & -[r - d'_Z(I_2)\xi]Z_2 & 0 & d_Z(I_2)\xi \end{pmatrix}.$$

A simple calculation gives

$$\begin{aligned} &\left(1 + \frac{\partial_T f + \partial_T g}{\lambda + d_T\xi}\right)\left(1 + \frac{\lambda + d_I\xi}{\mu_I + rZ_2} + \frac{rI_2[r - d'_Z(I_2)\xi]}{(\mu_I + rZ_2)[\lambda + d_Z(I_2)\xi]}\right) \\ &= \frac{\partial_I f}{\mu_I + rZ_2} + \frac{k\partial_V g}{(\mu_I + rZ_2)(\lambda + d_V\xi + \mu_V)}. \end{aligned}$$

Taking modulus on both sides gives

$$1 < \frac{\partial_I f}{\mu_I + rZ_2} + \frac{k\partial_V g}{\mu_V(\mu_I + rZ_2)} \leq \frac{f}{(\mu_I + rZ_2)I_2} + \frac{kg}{\mu_V(\mu_I + rZ_2)V_2} = 1,$$

a contradiction. This completes the proof. □

3.2 Global dynamics: global attractivity and uniform persistence

In this subsection, we will establish global dynamics of the solution semiflow $\Theta(t)$ for system (1.1) with an initial condition in $\mathbb{X}_+^4$. To obtain the compactness of the solution semiflow, we first improve the regularity of the solution obtained in Theorem 2.2. More precisely, we have the following results:

Lemma 3.5 *Assume (H1)-(H3). Let (T, I, V, Z) be the non-negative global classical solution of the system (1.1) obtained in Theorem 2.2. Then there exist $\sigma \in (0, 1)$ and $c > 0$ such that*

$$\|(T, I, V, Z)\|_{C^{2+\sigma, 1+\frac{\sigma}{2}}(\bar{\Omega}\times[1,\infty))} \leq c.$$

Proof From Theorem 2.2, we can find two positive constants c_1, c_2 such that

$$0 < T(x,t), I(x,t), V(x,t), Z(x,t) \leq c_1, \ |\nabla I(x,t)| \leq c_2 \text{ for all } x \in \Omega \text{ and } t > 0.$$

We rewrite the fourth equation of (1.1) as $\partial_t Z = \nabla \cdot A + B$ where $A = d_Z(I)\nabla Z + d'_Z(I)Z\nabla I$ and $B = rIZ - \mu_Z Z$. Obviously, $|B| \leq c_3$ and $|A| \leq d_Z(0)|\nabla Z| + c_4$.

By (H3) and Cauchy's inequality, we also have

$$A \cdot \nabla Z = d_Z(I)|\nabla Z|^2 + d_Z'(I) Z \nabla I \cdot \nabla Z \geq \frac{d_Z(c_1)}{2}|\nabla Z|^2 - c_5.$$

Thus, we apply the Hölder regularity for quasilinear parabolic equations (Porzio and Vespri 1993, Theorem 1.3 and Remark 1.4) to obtain $\|Z\|_{C^{\sigma,\frac{\sigma}{2}}(\bar{\Omega}\times[1,\infty))} \leq c_6$. Moreover, applying the standard parabolic Schauder theory (Ladyzhenskaya et al. 1968) to the system (1.1), one has

$$\|(T, I, V, Z)\|_{C^{2+\sigma,1+\frac{\sigma}{2}}(\bar{\Omega}\times[1,\infty))} \leq c_7,$$

which completes the proof. □

From Lemma 3.5, we observe that, for any $t > 1$ and a bounded set B in $\mathbb{X}_+^4$, the set $\Theta(t)B$ is bounded in $C^2(\Omega, \mathbb{R}_+^4)$ and thus precompact in $\mathbb{X}_+^4 = W^{1,p_0}(\Omega, \mathbb{R}_+^4)$ for $p_0 > 2$. This implies that the solution map $\Theta(t) : \mathbb{X}_+^4 \to \mathbb{X}_+^4$ is compact for all $t > 1$. Recall from Theorem 2.4 that $\Theta(t)$ is point dissipative. It then follows from (Hale 1988, Theorem 3.4.8) that the system admits a nonempty global attractor in $\mathbb{X}_+^4$. Next, we obtain the uniform persistence of solution which will be used to prove the globally asymptotic stability of steady states.

Proposition 3.6 *Assume (H1)-(H3). Let (T, I, V, Z) be the non-negative global classical solution of the system (1.1) with an initial value in $[\mathbb{X}_+ \setminus \{0\}]^4$. There exists $\delta_0 > 0$ such that*

$$\liminf_{t\to\infty} \min_{x\in\bar{\Omega}} T(x,t) \geq \delta_0.$$

If $R_0 > 1$, then there exists $\delta_1 > 0$ such that

$$\liminf_{t\to\infty} \min_{x\in\bar{\Omega}} \min\{I(x,t), V(x,t)\} \geq \delta_1.$$

Proof In view of the first equation of (1.1), (H2) and (2.11), we find $C_0 > 0$ and $t_0 > 0$ such that $\partial_t T \geq d_T \Delta T + b(T) - C_0 T$ for all $t > t_0$. By (H1), there exists $\delta_0 \in (0, T_0)$ such that $b(\delta_0) - C_0\delta_0 = 0$. Comparison principle gives $\liminf_{t\to\infty} T(x,t) \geq \delta_0$ for all $x \in \Omega$.

When $R_0 > 1$, we denote $\mathcal{X}_1 = \mathbb{X}_+ \times [\mathbb{X}_+ \setminus \{0\}]^2 \times \mathbb{X}_+$ and $\partial\mathcal{X}_1 = (\mathbb{X}_+ \times \{0\} \times \mathbb{X}_+^2) \cup (\mathbb{X}_+^2 \times \{0\} \times \mathbb{X}_+)$. By strong maximum principle, the largest positively invariant set in $\partial\mathcal{X}_1$ is $M_1 = \mathbb{X}_+ \times \{0\} \times \{0\} \times \mathbb{X}_+$. On M_1, the system (1.1) reduces to two decoupled equations $\partial_t T = d_T \Delta T + b(T)$ and $\partial_t Z = d_Z(0)\Delta Z - \mu_Z Z$. It is readily seen that $E_0 = (T_0, 0, 0, 0)$ is globally attractive on M_1. Following the idea in (Smith and Zhao 2001), we introduce a generalized distance function $\eta_1(u) = \min_{x\in\bar{\Omega}, i=2,3} u_i(x)$ for $u \in \mathbb{X}_+^4$. It follows from strong maximum principle that $\eta_1(\Theta(t)\phi) > 0$ for all $\phi \in \mathcal{X}_1$. Note that $\eta_1^{-1}(0,\infty) \subset \mathcal{X}_1$. Hence, the condition (P) in (Smith and Zhao 2001, Section 3) is verified. Now, we claim $W^s(E_0) \cap \eta_1^{-1}(0,\infty) = \emptyset$. Assume to the contrary that there exists a solution (T, I, V, Z) with initial condition in $\mathcal{X}_1$ such that $\lim_{t\to\infty}(T, I, V, Z) = (T_0, 0, 0, 0)$. There exists $t_1 > 0$ such that

$$\begin{aligned}
&\partial_t I \geq d_I \Delta I + (1-\varepsilon)[\partial_I f(T_0, 0)I + \partial_V g(T_0, 0)V] - \mu_I I,\\
&\partial_t V = d_V \Delta V + kI - \mu_V V,
\end{aligned}$$

where $\varepsilon = (1 - 1/R_0)/2 > 0$. On account of Perron-Frobenius theorem, we let $(I_\varepsilon, V_\varepsilon) \in \mathbb{R}^2_+$ be a positive eigenvector of the matrix

$$\begin{pmatrix} (1-\varepsilon)\partial_I f(T_0, 0) - \mu_I & (1-\varepsilon)\partial_V g(T_0, 0)\\ k & -\mu_V \end{pmatrix}$$

corresponding to the principal eigenvalue $\lambda_1 > 0$. There exists $c_1 > 0$ such that $c_1 I(x, t_1) \geq I_\varepsilon$ and $c_1 V(x, t_1) \geq V_\varepsilon$ for all $x \in \Omega$. It then follows from comparison principle that $I(x, t) \geq e^{\lambda_1(t-t_1)} I_\varepsilon/c_1$ and $V(x, t) \geq e^{\lambda_1(t-t_1)} V_\varepsilon/c_1$ for all $t > t_1$. This contradicts to the assumption that $I(x, t) \to 0$ and $V(x, t) \to 0$ as $t \to \infty$. Thus, the stable set of E_0 does not intersect $\eta_1^{-1}(0, \infty)$. By (Smith and Zhao 2001, Theorem 3), there exists $\delta_1 > 0$ such that $\liminf_{t\to\infty} p(\Theta(t)\phi) \geq \delta_1$ for any $\phi \in \mathcal{X}_1$. □

Now, we are ready to prove the global attractiveness of steady states via the Lyapunov function technique and LaSalle invariance principle. Together with locally asymptotic stability obtained in Sect. 3.1, we will actually obtain globally asymptotic stability of steady states (under certain conditions). First, we shall establish global attractiveness of the infection-free steady state $E_0 = (T_0, 0, 0, 0)$ based on the following Lyapunov function:

$$\mathcal{L}_0(T, I, V, Z) := \frac{1}{2}\int_\Omega [kI^2 + \partial_V g(T_0, 0)V^2]dx. \tag{3.7}$$

Here, and in the forthcoming content, a Lyapunov function is a differentiable and non-negative functional on a subset of $\mathbb{X}^4_+$ which is positively invariant with respect to the solution map $\Theta(t)$. For instance, we define $\mathcal{L}_0$ on the subset

$$D_0 = \{\phi = (\phi_1, \phi_2, \phi_3, \phi_4) \in \mathbb{X}^4_+ : \phi_1(x) \leq T_0\}. \tag{3.8}$$

The positive invariance of D_0 follows from the non-negativity of f and g, (H1) and a direct application of the comparison principle on the first equation of (1.1). Moreover, the omega limit set $\omega(\phi) \subset D_0$ for any $\phi \in \mathbb{X}^4_+$. The main step in the Lyapunov function technique is to construct a suitable Lyapunov function that is non-increasing along the solution trajectory. When restricted along a solution (T, I, V, Z), the Lyapunov function can be always regarded as a single-variable function of $t \in \mathbb{R}_+$.

Theorem 3.7 *Assume (H1)–(H3). If $R_0 \leq 1$, then the infection-free steady state E_0 is globally attractive. If $R_0 < 1$, then E_0 is globally asymptotically stable.*

Proof On account of the second and third equations of system (1.1), we obtain from (3.7) and Green's identity

$$\begin{aligned}\frac{d\mathcal{L}_0}{dt} &= -kd_I\int_\Omega |\nabla I|^2dx - \partial_V g(T_0,0)d_V\int_\Omega |\nabla V|^2dx - rk\int_\Omega I^2Zdx - k\mu_I\int_\Omega I^2dx\\ &\quad -\mu_V\partial_V g(T_0,0)\int_\Omega V^2dx + k\int_\Omega If(T,I)dx + k\int_\Omega I(g(T,V)+\partial_V g(T_0,0)V)dx,\end{aligned}$$

Since D_0 in (3.8) is positively invariant, we have $\|T\|_\infty \le T_0$. It then follows from (H2) that $f(T,I) \le f(T_0,I) \le \partial_I f(T_0,0)I$ and $g(T,V) \le g(T_0,V) \le \partial_T g(T_0,0)V$. Substituting these into the above equation gives

$$\begin{aligned}\frac{d\mathcal{L}_0}{dt} &\le \int_\Omega k[\partial_I f(T_0,0)-\mu_I]I^2dx + \int_\Omega 2k\partial_V g(T_0,0)IVdx - \int_\Omega \mu_V\partial_V g(T_0,0)V^2dx\\ &= -\mu_V\partial_V g(T_0,0)\int_\Omega (V-\frac{k}{\mu_I}I)^2dx + k\mu_I(R_0-1)\int_\Omega I^2dx \le 0,\end{aligned}$$

because $R_0 \le 1$. The largest positively invariance set of $d\mathcal{L}_0/dt = 0$ in D_0 is a singleton $\{E_0\}$. Obviously, $\omega(\phi) \in D_0$ for any $\phi \in D_0$. By LaSalle invariance principle, E_0 is globally attractive in D_0; namely, D_0 is a subset of $W^s(E_0)$, the stable set of E_0. Now, for any initial value $\phi \in \mathbb{X}_+^4$, the omega limit set $\omega(\phi)$ is internally chain transitive (Zhao 2017, Lemma 1.2.1) and $\omega(\phi) \subset D_0$. On account of (Zhao 2017, Theorem 1.2.1), E_0 is globally attractive in $\mathbb{X}_+^4$. If further, $R_0 < 1$, then Proposition 3.2 implies that E_0 is actually globally asymptotically stable. □

If $R_0 < 1$, one may use a similar argument as in (Bai and Winkler 2016) to prove the global attractiveness of E_0. However, for the critical case $R_0 = 1$, we shall apply the LaSalle invariance principle as in the above proof.

Next, we will explore the conditions for the global stability of $E_1 = (T_1, I_1, V_1, 0)$ and $E_2 = (T_2, I_2, V_2, Z_2)$. To this end, we assume that the incidence functions take the following forms:

(H4) $f(T,I) = h_0(T)h_d(I)$ and $g(T,V) = h_0(T)h_i(V)$, where $h_0, h_d, h_i \in C^2(\mathbb{R}_+)$.

Remark 3.8 We should point out that hypothesis (H4) is not stringent, because it can be satisfied by many forms of $f(T,I)$ and $g(T,V)$ including the mass action incidence $f(T,I) = \beta_d TI,\ \ g(T,V) = \beta_i TV$ used in simulations, see Sect. 4.1 for details. In fact, hypothesis (H4) is only used to prove the global stability of the homogeneous steady states E_1 and E_2. For the linear stability of E_0, E_1, E_2 and the nonlinear stability of E_0, hypothesis (H4) can be removed. Hence, it is natural to study the global stability of the CTL-inactivated/activated steady state without assumption (H4) in our future work.

Similar as in (Magal et al. 2010), we make use of the function $u-1-\ln u$ to define the following Lyapunov function:

$$\mathcal{L}_1(T,I,V,Z) := \int_\Omega \left(T - T_1 - \int_{T_1}^{T} \frac{h_0(T_1)}{h_0(\theta)} d\theta\right) dx + \int_\Omega \left(I - I_1 - I_1 \ln \frac{I}{I_1}\right) dx \\ + \frac{h_0(T_1)h_i(V_1)}{\mu_V V_1} \int_\Omega \left(V - V_1 - V_1 \ln \frac{V}{V_1}\right) dx + \int_\Omega Z dx \tag{3.9}$$

on the positively invariant subset

$$D_1 = \{\phi = (\phi_1, \phi_2, \phi_3, \phi_4) \in \mathbb{X}_+^4 : \phi_i(x) > 0 \text{ for all } i = 1, 2, 3 \text{ and } x \in \bar{\Omega}\}. \tag{3.10}$$

For any increasing function $h \in C(\mathbb{R}_+, \mathbb{R}_+)$ and any $u_0 > 0$, the function $1 - h(u_0)/h(u)$ is increasing on $(0, \infty)$ with a unique zero at u_0. Consequently, the function

$$u - u_0 - \int_{u_0}^{u} \frac{h(u_0)}{h(\theta)} d\theta = \int_{u_0}^{u} \left(1 - \frac{h(u_0)}{h(\theta)}\right) d\theta$$

is non-negative on $(0, \infty)$ with a unique minimum at u_0. Especially, when $h(\theta) = \theta$, we have $u - u_0 - u_0 \ln(u/u_0) \geq 0$ for all $u > 0$. Therefore, the Laypunov function in (3.9) is non-negative and achieves its unique minimum value 0 at $E_1 = (T_1, I_1, V_1, 0)$. Under the assumption (H4), we can rewrite the basic reproduction numbers R_0 in (3.2) as

$$R_0 = \frac{h_0(T_0)h_d'(0)}{\mu_I} + \frac{kh_0(T_0)h_i'(0)}{\mu_I \mu_V}.$$

Theorem 3.9 *Assume (H1)-(H4). If $R_0 > 1$ and $R_1 < 1$, then the CTL-inactivated steady state E_1 is globally asymptotically stable. If $R_0 > 1$ and $R_1 \leq 1$, then E_1 is globally attractive.*

Proof On account of (1.1) and (H4), we have

$$\frac{d\mathcal{L}_1}{dt} = \int_\Omega \left[1 - \frac{h_0(T_1)}{h_0(T)}\right] (d_T \Delta T) dx + \int_\Omega \left(1 - \frac{I_1}{I}\right) (d_I \Delta I) dx \\ + \int_\Omega \frac{h_0(T_1)h_i(V_1)}{\mu_V V_1} \left(1 - \frac{V_1}{V}\right) (d_V \Delta V) dx + \int_\Omega \Delta[d_Z(I)Z] dx + \int_\Omega \mathcal{M}_1 dx,$$

where

$$\mathcal{M}_1 = \left[1 - \frac{h_0(T_1)}{h_0(T)}\right] b(T) + h_0(T_1)[h_d(I) + h_i(V)] - \frac{I_1}{I} h_0(T)[h_d(I) + h_i(V)] \\ + \mu_I(I_1 - I) + \frac{h_0(T_1)h_i(V_1)}{\mu_V V_1} \left(kI - \mu_V V - \frac{kV_1 I}{V} + \mu_V V_1\right) + (rI_1 - \mu_Z)Z.$$

Note from (H2) and (H4) that $h_0' > 0$. This together with Green's identity implies that the first integral in the expression of $d\mathcal{L}_1/dt$ is no more than zero. Similarly, it follows from Green's identity that the second and third integrals are non-positive, while the fourth integral equals zero. Therefore, $d\mathcal{L}_1/dt \le \int_\Omega \mathcal{M}_1 dx$.

Next, we will show that $\mathcal{M}_1 \le 0$. Since $E_1 = (T_1, I_1, V_1, 0)$ satisfies the equations (3.1) and $R_1 = rI_1/\mu_Z$, we can rewrite $b(T) = b(T) - b(T_1) + h_0(T_1)[h_d(I_1) + h_i(V_1)]$, $\mu_I = h_0(T_1)[h_d(I_1)+h_i(V_1)]/I_1$, $k = \mu_V V_1/I_1$, and $rI_1 - \mu_Z = \mu_Z(R_1 - 1)$. Consequently,

$$\begin{aligned}\mathcal{M}_1 = &\left[1 - \frac{h_0(T_1)}{h_0(T)}\right][b(T) - b(T_1)] + \mu_Z(R_1 - 1)Z \\ &+ h_0(T_1)[h_d(I_1)\mathcal{M}_{11} + h_i(V_1)\mathcal{M}_{12}],\end{aligned} \tag{3.11}$$

where

$$\begin{aligned}\mathcal{M}_{11} =& 1 - \frac{h_0(T_1)}{h_0(T)} + \frac{h_d(I)}{h_d(I_1)} - \frac{I_1 h_0(T) h_d(I)}{I h_0(T_1) h_d(I_1)} + 1 - \frac{I}{I_1},\\ \mathcal{M}_{12} =& 1 - \frac{h_0(T_1)}{h_0(T)} + \frac{h_i(V)}{h_i(V_1)} - \frac{I_1 h_0(T) h_i(V)}{I h_0(T_1) h_i(V_1)} + 1 - \frac{I}{I_1} + \frac{I}{I_1} - \frac{V}{V_1} - \frac{V_1 I}{V I_1} + 1.\end{aligned}$$

By adding and subtracting $2 - \frac{I h_d(I_1)}{I_1 h_d(I)}$ we obtain

$$\begin{aligned}\mathcal{M}_{11} =& \left[3 - \frac{h_0(T_1)}{h_0(T)} - \frac{I_1 h_0(T) h_d(I)}{I h_0(T_1) h_d(I_1)} - \frac{I h_d(I_1)}{I_1 h_d(I)}\right] + \left[\frac{h_d(I)}{h_d(I_1)} - 1 - \frac{I}{I_1} + \frac{I h_d(I_1)}{I_1 h_d(I)}\right],\\ =& \left[3 - \frac{h_0(T_1)}{h_0(T)} - \frac{I_1 h_0(T) h_d(I)}{I h_0(T_1) h_d(I_1)} - \frac{I h_d(I_1)}{I_1 h_d(I)}\right] + \left[\frac{h_d(I)}{h_d(I_1)} - 1\right]\left[1 - \frac{I h_d(I_1)}{I_1 h_d(I)}\right].\end{aligned}$$

Similarly, by adding and subtracting $3 - \frac{V h_i(V_1)}{V_1 h_i(V)}$ we obtain

$$\begin{aligned}\mathcal{M}_{12} =& \left[4 - \frac{h_0(T_1)}{h_0(T)} - \frac{I_1 h_0(T) h_i(V)}{I h_0(T_1) h_i(V_1)} - \frac{V h_i(V_1)}{V_1 h_i(V)} - \frac{I V_1}{I_1 V}\right]\\ &+ \left[\frac{h_i(V)}{h_i(V_1)} - 1 - \frac{V}{V_1} + \frac{V h_i(V_1)}{V_1 h_i(V)}\right]\\ =& \left[4 - \frac{h_0(T_1)}{h_0(T)} - \frac{I_1 h_0(T) h_i(V)}{I h_0(T_1) h_i(V_1)} - \frac{V h_i(V_1)}{V_1 h_i(V)} - \frac{I V_1}{I_1 V}\right]\\ &+ \left[\frac{h_i(V)}{h_i(V_1)} - 1\right]\left[1 - \frac{V h_i(V_1)}{V_1 h_i(V)}\right].\end{aligned}$$

On account of (H2) and (H4), we have $h_d' > 0$ and $h_d'' \le 0$. This together with Cauchy's inequality implies $\mathcal{M}_{11} \le 0$. Similarly, we obtain from (H2), (H4) and Cauchy's inequality that $\mathcal{M}_{12} \le 0$. Moreover, in view of (H1), (H2) and (H4), we have $b' < 0$ and $h_0' > 0$. Hence, the first term in the expression of $\mathcal{M}_1$ in (3.11) is non-positive. The second term is also no more than zero when $R_1 \le 1$. Combining these with non-positiveness of $\mathcal{M}_{11}$ and $\mathcal{M}_{12}$ yields $\mathcal{M}_1 \le 0$ and $d\mathcal{L}_1/dt \le 0$.

The largest positively invariant set of $d\mathcal{L}_1/dt = 0$ in D_1 is a singleton $\{E_1\}$. Since $\omega(D_1) \subset D_1$ by Proposition 3.6, we apply LaSalle invariance principle to find $D_1 \subset W^s(E_1)$. Now, we consider the initial condition $\phi \in [\mathbb{X}_+ \setminus \{0\}]^4$. It again follows from Proposition 3.6 that $\omega(\phi) \subset D_1$. By (Zhao (2017), Lemma 1.2.1), $\omega(\phi)$ is internally chain transitive. An application of (Zhao 2017, Theorem 1.2.1) yields $\omega(\phi) = E_1$. This proves global attractiveness of E_1 in $[\mathbb{X}_+ \setminus \{0\}]^4$ when $R_0 > 1$ and $R_1 \leq 1$. If further, $R_1 < 1$, then we obtain from Proposition 3.3 that E_1 is globally asymptotically stable in $[\mathbb{X}_+ \setminus \{0\}]^4$. □

The following result gives uniform persistence in Z when both $R_0 > 1$ and $R_1 > 1$.

Proposition 3.10 *Assume (H1)-(H4). Let (T, I, V, Z) be the solution of (1.1) with initial condition in $[\mathbb{X}_+ \setminus \{0\}]^4$. If $R_0 > 1$ and $R_1 > 1$, then there exists $\delta_2 > 0$, independent of initial condition, such that*

$$\liminf_{t\to\infty} \min_{x\in\bar{\Omega}} Z(x, t) \geq \delta_2.$$

Proof Define $\mathcal{X}_2 = \mathbb{X}_+ \times [\mathbb{X}_+ \setminus \{0\}]^3$ and $\partial\mathcal{X}_2 = \mathbb{X}_+ \times [\mathbb{X}_+\setminus\{0\}]^2 \times \{0\}$. Let M_2 be the largest positively invariant set in $\partial\mathcal{X}_2$. Using a similar argument as in the proof of Proposition 3.6, we find that $E_1 = (T_1, I_1, V_1, 0)$ is globally attractive on M_2. Now, we introduce a generalized distance function $\eta_2(u) = \min_{x\in\bar{\Omega}} u_4(x)$ for $u \in \mathbb{X}_+^4$. It follows from strong maximum principle that $\eta_1(\Theta(t)\phi) > 0$ for all $\phi \in \mathcal{X}_2$. Note that $\eta_2^{-1}(0, \infty) \subset \mathcal{X}_2$. Hence, the condition (P) in (Smith and Zhao 2001, Section 3) is verified. We are left to show that $W^s(E_1)\cap\eta_2^{-1}(0, \infty) = \emptyset$. Assume to the contrary that for some initial condition in $\mathcal{X}_2$, the solution $(T, I, V, Z) \to (T_1, I_1, V_1, 0)$ as $t \to \infty$. There exists $t_2 > 0$ such that $\partial_t Z \geq \Delta[d_Z(I)Z]+\varepsilon_2 Z$, where $\varepsilon_2 = (rI_2-\mu_Z)/2 > 0$. By comparison principle, we have $\int_\Omega Z(x, t)dx \geq e^{\varepsilon_2(t-t_2)} \int_\Omega Z(x, t_2)dx$, which contradicts to the assumption $Z \to 0$ as $t \to \infty$. Therefore, we obtain from (Smith and Zhao 2001, Theorem 3) that $\liminf_{t\to\infty} p_2(\Theta(t)\phi) \geq \delta_2$ for some $\delta_2 > 0$ independent of the initial condition $\phi \in \mathcal{X}_2$. □

Next, we shall prove the global attractiveness (as well as globally asymptotic stability) of $E_2 = (T_2, I_2, V_2, Z_2)$ based on the following Lyapunov function:

$$\begin{aligned}\mathcal{L}_2(T, I, V, Z) := &\int_\Omega \left(T - T_2 - \int_{T_2}^{T} \frac{h_0(T_2)}{h_0(\theta)} d\theta\right) dx + \int_\Omega \left(I - I_2 - I_2 \ln \frac{I}{I_2}\right) dx \\ &+ \frac{h_0(T_2)h_i(V_2)}{\mu_V V_2} \int_\Omega \left(V - V_2 - V_2 \ln \frac{V}{V_2}\right) dx \\ &+ \int_\Omega \left(Z - Z_2 - Z_2 \ln \frac{Z}{Z_2}\right) dx\end{aligned}$$

defined on the positively invariant set

$$D_2 = \{\phi = (\phi_1, \phi_2, \phi_3, \phi_4) \in \mathbb{X}_+^4 : \phi_i(x) > 0 \text{ for all } i = 1, 2, 3, 4 \text{ and } x \in \bar{\Omega}\}.$$

Similar to D_1 defined in (3.10), the set D_2 is positively invariant but not closed. For any $\phi \in D_2$, we have $\Theta(t)\phi \in D_2$. However, to show that $\omega(\phi) \in D_2$, we need to make use of the persistence results in Propositions 3.6 and 3.10.

Theorem 3.11 *Assume (H1)–(H4). Let (T, I, V, Z) be the solution of (1.1) with initial condition in $[\mathbb{X}_+ \setminus \{0\}]^4$. If $R_0 > 1$ and $R_1 > 1$, then the CTL-activated steady state E_2 is globally asymptotically stable provided*

$$d_I \geq \frac{Z_2}{4I_2} \max_{0 \leq I \leq \infty} \frac{|I d_Z'(I)|^2}{d_Z(I)}. \tag{3.12}$$

Proof Using a similar argument as in the proof of Theorem 3.9, we obtain after a tedious calculation and an application of Green's identity that $d\mathcal{L}_2/dt = \int_\Omega \mathcal{M}_2 dx$, where

$$\begin{aligned}
\mathcal{M}_2 = &- \frac{d_T h_0(T_2) h_0'(T)}{h_0^2(T)} |\nabla T|^2 - \frac{d_V h_0(T_2) h_i(V_2)}{\mu_V V^2} |\nabla V|^2 \\
&- \frac{d_I I_2}{I^2} |\nabla I|^2 - \frac{Z_2 d_Z(I)}{Z^2} |\nabla Z|^2 - \frac{Z_2 d_Z'(I)}{Z} (\nabla I \cdot \nabla Z) \\
&+ h_0(T_2) h_d(I_2) \left(\frac{h_d(I)}{h_d(I_2)} - 1 \right) \left(1 - \frac{I h_d(I_2)}{I_2 h_d(I)} \right) \\
&+ h_0(T_2) h_i(V_2) \left(\frac{h_i(V)}{h_i(V_2)} - 1 \right) \left(1 - \frac{V h_i(V_2)}{V_2 h_i(V)} \right) \\
&+ \left(1 - \frac{h_0(T_2)}{h_0(T)} \right) [b(T) - b(T_2)] \\
&+ h_0(T_2) h_d(I_2) \left(3 - \frac{h_0(T_2)}{h_0(T)} - \frac{I_2 h_0(T) h_d(I)}{I h_0(T_2) h_d(I_2)} - \frac{I h_d(I_2)}{I_2 h_d(I)} \right) \\
&+ h_0(T_2) h_i(V_2) \left(4 - \frac{h_0(T_2)}{h_0(T)} - \frac{I_2 h_0(T) h_i(V)}{I h_0(T_2) h_i(V_2)} - \frac{V h_i(V_2)}{V_2 h_i(V)} - \frac{I V_2}{I_2 V} \right).
\end{aligned}$$

The first line in the expression of $\mathcal{M}_2$ is obviously no more than zero. The second line is bounded by zero due to (3.12). The last five lines are also no more than zero because $h_0' > 0$, $h_d' > 0$, $h_i' > 0$ and $b' < 0$. Consequently, we have $\mathcal{M}_2 \leq 0$ and $d\mathcal{L}_2/dt \leq 0$. The largest positively invariant set of $d\mathcal{L}_2/dt = 0$ in D_2 is a singleton $\{E_2\}$. For any $\phi \in D_2$, it follows from Propositions 3.6 and 3.10 that $\omega(\phi) \in D_2$. By LaSalle invariance principle, $D_2 \subset W^s(E_2)$. Now, for any $\phi \in [\mathbb{X}_+ \setminus \{0\}]^4$, thanks to Propositions 3.6 and 3.10, we still have $\omega(\phi) \subset D_2$. Moreover, (Zhao 2017, Lemma 1.2.1) implies that $\omega(\phi)$ is internally chain transitive. By (Zhao 2017, Theorem 1.2.1), E_2 is globally attractive. This together with Proposition 3.4 gives globally asymptotic stability of E_2 in $[\mathbb{X}_+ \setminus \{0\}]^4$. □

A combination of Theorem 3.7, Theorem 3.9, and Theorem 3.11 gives Theorem 1.2.

4 Simulations and discussions

4.1 Simulations

In this subsection, we conduct some numerical simulations of our model with specific functions

$$\begin{aligned} b(T) &= \lambda_c - \mu_T T + \lambda_l T(1 - T/K_T), \quad f(T, I) = \beta_d T I, \quad g(T, V) \\ &= \beta_i T V, \quad d_Z(I) = d_0 e^{-I}. \end{aligned}$$

The parameter values are chosen to be consistent with the studies in the literature (Komarova et al. 2003; Li and Shu 2012; Shu et al. 2014):

$$\begin{aligned} &\lambda_c = 10, \ \mu_T = 0.02, \ \lambda_l = 0.005, \ K_T = 1500, \ \beta_d = 0.003, \ \beta_i = 0.0027, \\ &\mu_I = 3, \ \mu_V = 2.4, \ \mu_Z = 0.3, \ r_1 = r_2 = r = 0.3, \ k = 3, \\ &d_T = 0.01, \ d_I = 0.01, \ d_V = 0.1, \ d_0 = 0.01. \end{aligned}$$

One of the simple positive and decreasing functions is the exponential function, so we choose $d_Z(I) = d_0 e^{-I}$ with $d_0 = 0.001$. The values of the diffusion coefficients d_T, d_I, and d_V are arbitrarily chosen. Based on the theoretical results, the diffusion coefficients do not affect the global dynamics. The constant growth rate λ_c and the logistic growth rate λ_l of the targeted cells as well as the cell-to-cell infection rate β_d are also arbitrarily chosen. The removal rate of infected cells r_1 and the recruitment rate of CTL r_2 will be varying from 0.3 to 1 to simulate two cases $R_1 < 1$ and $R_1 > 1$, respectively; noting that $r_1 = 1$ in (Li and Shu 2012; Shu et al. 2014). The descriptions and references for other parameters are listed in Table 1.

It can be calculated that $T_0 \approx 589$. The domain is set to be $\Omega = (0, 1) \times (0, 1)$. We choose the initial condition as a perturbation of infection-free steady state $E_0 = (T_0, 0, 0, 0)$:

$$T(x, 0) = T_0, \ I(x, 0) = V(x, 0) = Z(x, 0) = e^{-100[(x_1-0.5)^2+(x_2-0.5)^2]}.$$

Table 1 Parameter values and descriptions

Parameter value	Description	Reference
$\mu_T = 0.02$	Death rate of targeted cells	(Li and Shu 2012)
$K_T = 1500$	Carrying capacity of targeted cells	(Li and Shu 2012)
$\beta_i = 0.0027$	Virus-to-cell transmission rate	(Li and Shu 2012)
$\mu_I = 3$	Death rate of infected cells	(Li and Shu 2012)
$\mu_V = 2.4$	Death rate of virus	(Li and Shu 2012)
$k = 3$	Production rate of virus	(Li and Shu 2012)
$\mu_Z = 0.3$	Death rate of CTL	(Komarova et al. 2003)

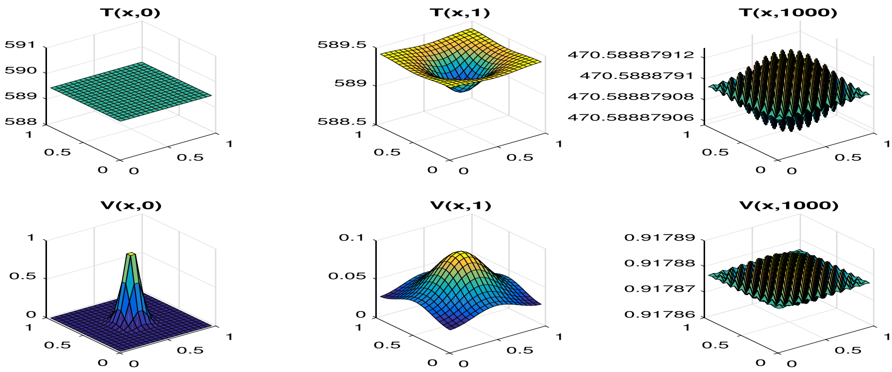

Fig. 1 The CTL-inactivated steady state E_1 is globally asymptotically stable when $R_0 > 1$ and $R_1 < 1$

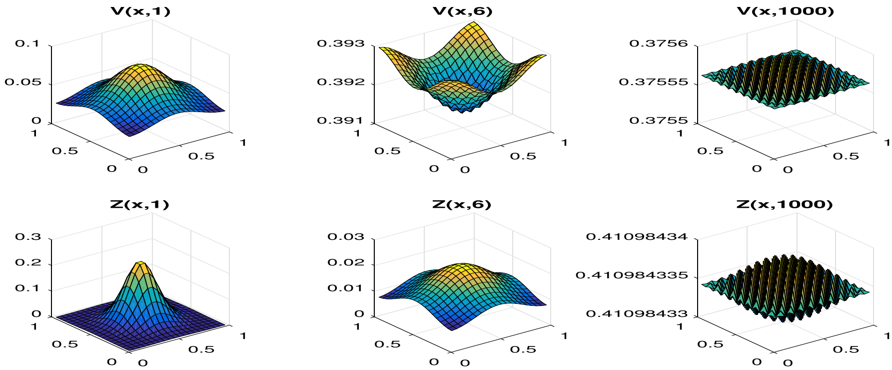

Fig. 2 The CTL-activated steady state E_2 is globally asymptotically stable when $R_0 > 1$ and $R_1 > 1$

From (3.2) and (3.6), we calculate $R_0 \approx 1.25$ and $R_1 \approx 0.73$. By Theorem 3.9, the CTL-inactivated steady state $E_1 = (T_1, I_1, V_1, 0)$ is globally asymptotically stable, where $T_1 \approx 471$, $I_1 \approx 0.73$, and $V_1 \approx 0.92$. The global dynamics of the model are illustrated in Fig. 1.

Next, we increase the parameter value $r = 1$ such that the basic reproduction number of immune response becomes $R_1 \approx 2.45$. By Proposition 3.4, there exists a unique CTL-activated steady state $E_2 = (T_2, I_2, V_2, Z_2)$. It can be calculated that $T_2 \approx 535$, $I_2 \approx 0.30$, $V_2 \approx 0.38$ and $Z_2 \approx 0.41$. Moreover, (3.12) can be verified:

$$\frac{Z_2}{4I_2} \max_{0 \le I \le \infty} \frac{[I d_Z'(I)]^2}{d_Z(I)} = \frac{Z_2}{4I_2} \max_{0 \le I \le \infty} d_0 I^2 e^{-I} = \frac{d_0 Z_2}{I_2 e^2} \approx 0.0018 < 0.01 = d_I.$$

An application of Theorem 3.11 gives global asymptotic stability of E_2; see Fig. 2.

We have conducted several numerical simulations (not shown here) with different choices of the parameter values including the diffusion rates, and selected different types of the positive and decreasing function $d_Z(I)$. The simulation results all coincide with our theoretical conclusions.

4.2 Discussions

We considered a general viral infection model with cell-to-cell transmission and immune chemokines. The global existence and ultimately boundedness of the solution were obtained via a priori energy estimate. We introduced the basic reproduction number of infection R_0 and proved that it is the threshold parameter to determine extinction and persistence of viral infection; namely, the unique infection-free steady state E_0 is globally asymptotically stable if $R_0 < 1$ and globally attractive if $R_0 = 1$, while the viral infection will be uniformly persistent if $R_0 > 1$. When $R_0 > 1$, there exists a unique CTL-inactivated steady state E_1. We defined another basic reproduction number of the CTL immune response R_1 which determines the threshold dynamics of CTL. To be more specific, if $R_1 < 1$ then E_1 is globally asymptotically stable; if $R_1 = 1$ then E_1 is globally attractive; and if $R_1 > 1$ then there exists a unique CTL-activated steady state E_2 which is globally asymptotically stable.

The assumptions (H1)–(H4) are biologically relevant and mathematically broad. For instance, the nonlinear infection rates $f(T, I)$ and $g(T, V)$ generalize most commonly used functional responses of Holling's type I and II; and the growth function $b(T)$ includes a linear function $b_l(T) = \lambda_c - \mu_T T$ plus a possible nonlinear mitosis rate $b_m(T)$ as long as $b_m'(T) < \mu_T$ for all $T > 0$. In the numerical simulation, we chose a logistic growth $b_m(T) = \lambda_l T(1 - T/K_T)$ with $\lambda_l < \mu_T$. It is worth to remark that Hopf bifurcation may occur if $\lambda_l > \mu_T$; see (Li and Shu 2012). Thus, the monotonicity condition $b'(T) < 0$ is important to establish the stability results of steady states. In the proof of Theorem 3.11, we made a technical assumption (3.12) so as to construct a suitable Lyapunov function which is non-increasing along the solution semiflow. Numerical simulations (not shown here) suggest that this assumption could be removed; namely, the statement of Theorem 3.11 remains valid without the technical assumption (3.12). We leave this problem for further investigation.

There are some possible generalizations of our model. One extension is to incorporate the intracellular delays during viral infection (Lai and Zou 2014; Li and Shu 2012; Shu et al. 2013) and time lags in the process of immune response (Fenton et al. 2006; Shu et al. 2014). Another extension is to replace the term $\Delta[d_Z(I)Z] = \nabla \cdot [d_Z(I)\nabla Z + d_Z'(I)Z\nabla I]$ with a more general term of diffusion and chemotaxis $\nabla \cdot [d_Z(I)\nabla Z - \chi(I)Z\nabla I]$; see (Jin and Wang 2021). It is expected that the extended model with time delays and more general chemotaxis will exhibit more rich dynamics. We will provide a detailed study of this extended model in a forthcoming paper. On the other hand, for the two-component system with degenerate motility, the global existence of weak and generalized solutions for a large class of $d_Z(I)$ with various decay behavior (Winkler 2023a, b, c; Li and Winkler 2023; Li and Wang 2023) have been established, hence it is an interesting issue to study the solution behavior for the system (1.1) with the degenerate motility, that is $d_Z(I) = 0$ as $I = 0$.

Acknowledgements We would like to express our gratitude to the anonymous referees for careful reading and valuable suggestions which help to improve the presentation of this paper. H. Shu was partially supported by the National Natural Science Foundation of China (No. 12371495), the Fundamental Research Funds for the Central Universities (No. GK202201002), and the Natural Science Basic Research Program of Shaanxi (No. 2023-JC-JQ-03). H. Y. Jin was partially supported by the National Natural Science Foundation of China (No. 12371203), and Guangdong Basic and Applied Basic Research Foundation (No. 2022B1515020032).

X.-S. Wang was partially supported by the Louisiana Board of Regents Support Fund under contract No. LEQSF(2022–25)-RD-A-26. J. Wu was partially supported by Canada Research Chair Program (No. 230720).

References

Bai X, Winkler M (2016) Equilibration in a fully parabolic two-species chemotaxis system with competitive kinetics. Indiana Univ Math J 65:553–583

Bromley SK, Mempel TR, Luster AD (2008) Orchestrating the orchestrators: chemokines in control of T cell traffic. Nat Immunol 9:970–980

Deng J, Shu H, Wang L, Wang X-S (2023) Viral dynamics with immune responses: Effects of distributed delays and Filippov antiretroviral therapy. J Math Biol 86:37

Dixit NM, Perelson AS (2004) Multiplicity of human immunodeficiency virus infections in lymphoid tissue. J Virol 78:8942–8945

Dung L (1997) Dissipativity and global attractors for a class of quasilinear parabolic systems. Commun Partial Differ Equ 22:413–433

Dyson J, Villella-Bressan R, Webb GF (2008) Global existence and boundedness of solutions to a model of chemotaxis. Math Model Nat Phenom 3:17–35

Fenton A, Lello J, Bonsall MB (2006) Pathogen responses to host immunity: the impact of time delays and memory on the evolution of virulence. Proc R Soc B 273:2083–2090

Fu X, Tang L-H, Liu C, Huang J-D, Hwa T, Lenz P (2012) Stripe formation in bacterial system with density-suppressed motility. Phys Rev Lett 108:198102

Fujie K, Jiang J (2020) Global existence for a kinetic model of pattern formation with density-suppressed motilities. J Differ Equ 269:5338–5378

Fujie K, Jiang J (2021) Comparison methods for a Keller-Segel-type model of pattern formations with density-suppressed motilities. Calc Var Partial Differ Equ 60:92

Galloway NLK, Doitsh G, Monroe KM, Yang Z, Muñoz-Arias I, Levy DN, Greene WC (2015) Cell-to-cell transmission of HIV-1 is required to trigger pyroptotic death of lymphoid-tissue-derived CD4 T cells. Cell Rep 12:1555–1563

Gummuluru S, Kinsey CM, Emerman M (2000) An in vitro rapid-turnover assay for human immunodeficiency virus type 1 replication selects for cell-to-cell spread of virus. J Virol 74:10882–10891

Hale JK (1988) Asymptotic behavior of dissipative systems. American Mathematical Society, Providence

Halle S, Keyser KA, Stahl FR et al (2016) In vivo killing capacity of cytotoxic T cells is limited and involves dynamic interactions and T cell cooperativity. Immunity 44:233–245

Horstmann D, Winkler M (2005) Boundedness vs. blow-up in a chemotaxis system. J Differ Equ 215:52–107

Hübner W, McNerney GP, Chen P, Dale BM, Gordan RE, Chuang FYS, Li XD, Asmuth DM, Huser T, Chen BK (2009) Quantitative 3D video microscopy of HIV transfer across T cell virological synapses. Science 323:1743–1747

Iwami S, Takeuchi JS, Nakaoka S, Mammano F, Clavel F, Inaba H, Kobayashi T, Misawa N, Aihara K, Koyanagi Y, Sato K (2015) Cell-to-cell infection by HIV contributes over half virus infection. eLife 4:08150

Jiang J, Laurençot P, Zhang Y (2022) Global existence, uniform boundedness, and stabilization in a chemotaxis system with density-suppressed motility and nutrient consumption. Comm Partial Diff Equ 47:1024–1069

Jin HY, Kim YJ, Wang ZA (2018) Boundedness, stabilization and pattern formation driven by density-suppressed motility. SIAM J Appl Math 78:1632–1657

Jin HY, Wang ZA (2021) Global dynamics and spatio-temporal patterns of predator-prey systems with density-dependent motion. Eur J Appl Math 32:652–682

Kareiva P, Odell GT (1987) Swarms of predators exhibit "preytaxis" if individual predators use area-restricted search. Am Nat 130:233–270

Keller EF, Segel LA (1970) Initiation of slime mold aggregation viewed as an instability. J Theor Biol 26:399–415

Keller EF, Segel LA (1971) Model for chemotaxis. J Theor Biol 30:225–234

Komarova NL, Barnes E, Klenerman P, Wodarz D (2003) Boosting immunity by antiviral drug therapy: a simple relationship among timing, efficacy, and success. Proc Natl Acad Sci USA 100:1855–1860

Kowalczyk R, Szymańska Z (2008) On the global existence of solutions to an aggregation model. J Math Anal Appl 343:379–398
Ladyzhenskaya O, Solonnikov V, Uralceva N (1968) Linear and quasilinear equations of parabolic type. American Mathematical Society, Providence, RI
Lai X, Zou X (2014) Modeling HIV-1 virus dynamics with both virus-to-cell infection and cell-to-cell transmission. SIAM J Appl Math 74:898–917
Lai X, Zou X (2015) Modeling cell-to-cell spread of HIV-1 with logistic target cell growth. J Math Anal Appl 426:563–584
Lee S, Kim S, Oh Y, Hwang HJ (2017) Mathematical modeling and its analysis for instability of the immune system induced by chemotaxis. J Math Biol 75:1101–1131
Li G, Wang L (2023) Boundedness in a taxis-consumption system involving signal-dependent motilities and concurrent enhancement of density-determined diffusion and cross-diffusion. Z Angew Math Phys 74:92
Li G, Winkler M (2023) Relaxation in a Keller-Segel-consumption system involving signal-dependent motilities. Commun Math Sci 21:299–322
Li J, Wang ZA (2021) Traveling wave solutions to the density-suppressed motility model. J Diff Equ 301:1–36
Li MY, Shu H (2012) Joint effects of mitosis and intracellular delay on viral dynamics: two-parameter bifurcation analysis. J Math Biol 64:1005–1020
Liu C et al (2011) Sequential establishment of stripe patterns in an expanding cell population. Science 334:238
Ma M, Peng R, Wang ZA (2020) Stationary and non-stationary patterns of the density-suppressed motility model. Phys D 402:132259
Magal P, McCluskey CC, Webb GF (2010) Lyapunov functional and global asymptotic stability for an infection-age model. Appl Anal 89:1109–1140
Magal P, Webb GF, Wu Y (2019) On the basic reproduction number of reaction-diffusion epidemic models. SIAM J Appl Math 79:284–304
Martin N, Sattentau Q (2009) Cell-to-cell HIV-1 spread and its implications for immune evasion. Curr Opin HIV AIDS 4:143–149
Mizoguchi N, Souplet P (2014) Nondegeneracy of blow-up points for the parabolic Keller-Segel system. Ann Inst H Poincaré Anal Non Linéaire 31:851–875
Nowak MA, Bangham CRM (1996) Population dynamics of immune responses to persistent viruses. Science 272:74–79
Nowak MA, Bonhoeffer S, Hill AM, Boehme R, Thomas HC (1996) Viral dynamics in hepatitis B virus infection. Proc Natl Acad Sci USA 93:4398–4402
Perelson AS, Nelson PW (1999) Mathematical analysis of HIV-1 dynamics in vivo. SIAM Rev 41:3–44
Porzio MM, Vespri V (1993) Hölder estimates for local solutions of some doubly nonlinear degenerate parabolic equations. J Diff Equ 103:146–178
Pourbashash H, Pilyugin SS, Leenheer PD, McCluskey C (2014) Global analysis of within host virus models with cell-to-cell viral transmission. Discr Contin Dyn Syst Ser B 19:3341–3357
Rot A, von Andrian UH (2004) Chemokines in innate and adaptive host defense: basic chemokinese grammar for immune cells. Ann Rev Immunol 22:891–928
Sattentau Q (2008) Avoiding the void: cell-to-cell spread of human viruses. Nat Rev Microbiol 6:28–41
Shu H, Wang L, Watmough J (2013) Global stability of a nonlinear viral infection model with infinitely distributed intracellular delays and CTL immune responses. SIAM J Appl Math 73:1280–1302
Shu H, Wang L, Watmough J (2014) Sustained and transient oscillations and chaos induced by delayed antiviral immune response in an immunosuppressive infection model. J Math Biol 68:477–503
Sigal A, Kim JT, Balazs AB, Dekel E, Mayo A, Milo R, Baltimore D (2011) Cell-to-cell spread of HIV permits ongoing replication despite antiretroviral therapy. Nature 477:95–98
Smith HL, Zhao X-Q (2001) Robust persistence for semidynamical systems. Nonlinear Anal 47:6169–6179
Souplet P, Quittner P (2007) Superlinear parabolic problems: blow-up and global existence and steady states. Birkhäuser Advanced Texts, Basel/Boston/Berlin
Stinner C, Surulescu C, Winkler M (2014) Global weak solutions in a PDE-ODE system modeling multiscale cancer cell invasion. SIAM J Math Anal 46:1969–2007
Tao Y, Winkler M (2011) A chemotaxis-haptotaxis model: the roles of nonlinear diffusion and logistic source. SIAM J Math Anal 43:685–704

Tao Y, Winkler M (2012) Boundedness in a quasilinear parabolic-parabolic Keller-Segel system with subcritical sensitivity. J Diff Equ 252:692–715

Tao Y, Winkler M (2017) Effects of signal-dependent motilities in a Keller-Segel-type reaction-diffusion system. Math Models Methods Appl Sci 27:1645–1683

Winkler M (2020) Can simultaneous density-determined enhancement of diffusion and cross-diffusion foster boundedness in KellerCSegel type systems involving signal-dependent motilities? Nonlinearity 33:6590

Winkler M (2023) Global generalized solvability in a strongly degenerate taxis-type parabolic system modeling migration-consumption interaction. Z Angew Math Phys 74:32

Winkler M (2023) Stabilization despite pervasive strong cross-degeneracies in a nonlinear diffusion model for migration-consumption interaction. Nonlinearity 36:4438–4469

Winkler M (2023) A quantitative strong parabolic maximum principle and application to a taxis-type migration-consumption model involving signal-dependent degenerate diffusion. Inst. H. Poincaré-ANL, accepted, Ann. https://doi.org/10.4171/AIHPC/73

Zhao X-Q (2017) Dynamical systems in population biology. CMS Books in Mathematics. Springer, Cham, 2 edition

Zheng P, Shan W (2023) Global boundedness and stability analysis of the quasilinear immune chemotaxis system. J Diff Equ 344:556–607

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