

Dynamics and near-optimal control in a stochastic rumor propagation model incorporating media coverage and Lévy noise

Liang'an Huo(霍良安) and Yafang Dong(董雅芳)
Citation: Chin. Phys. B, 2022, 31 (**3**): 030202. DOI: 10.1088/1674-1056/ac2f35
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Near-optimal control of a stochastic rumor spreading model with Holling II functional response function and imprecise parameters

Liang'an Huo(霍良安) and Xiaomin Chen(陈晓敏)
Chin. Phys. B, 2021, 30 (**12**): 120205. DOI: 10.1088/1674-1056/ac0eea

Consistent Riccati expansion solvability, symmetries, and analytic solutions of a forced variable-coefficient extended Korteveg-de Vries equation in fluid dynamics of internal solitary waves

Ping Liu(刘萍), Bing Huang(黄兵), Bo Ren(任博), and Jian-Rong Yang(杨建荣)
Chin. Phys. B, 2021, 30 (**8**): 080203. DOI: 10.1088/1674-1056/ac052a

Dynamics of a stochastic rumor propagation model incorporating media coverage and driven by Lévy noise

Liang-An Huo(霍良安), Ya-Fang Dong(董雅芳), and Ting-Ting Lin(林婷婷)
Chin. Phys. B, 2021, 30 (**8**): 080201. DOI: 10.1088/1674-1056/ac0423

Closed form soliton solutions of three nonlinear fractional models through proposed improved Kudryashov method

Zillur Rahman, M Zulfikar Ali, and Harun-Or Roshid
Chin. Phys. B, 2021, 30 (**5**): 050202. DOI: 10.1088/1674-1056/abd165

Furi-Martelli-Vignoli spectrum and Feng spectrum of nonlinear block operator matrices

Xiao-Mei Dong(董小梅), De-Yu Wu(吴德玉), and Alatancang Chen(陈阿拉坦仓)
Chin. Phys. B, 2021, 30 (**4**): 040201. DOI: 10.1088/1674-1056/abd7df

---

Dynamics and near-optimal control in a stochastic rumor propagation model incorporating media coverage and Lévy noise

Liang'an Huo(霍良安)[†] and Yafang Dong(董雅芳)

Business School, University of Shanghai for Science and Technology, Shanghai 200093, *China*

(Received 4 April 2021; revised manuscript received 21 September 2021; accepted manuscript online 13 October 2021)

The appearance of rumors intensifies people's panic and affects social stability. How to control the spread of rumors has become an important issue which is worth studying. In order to more accurately reflect the actual situation in the real world, a stochastic model incorporating media coverage and Lévy noise is proposed to describe the dynamic process of rumor propagation. By introducing two control strategies of popular science education and media coverage in an emergency event, an near-optimal control problem that minimizes the influence and control cost of rumor propagation is proposed. Sufficient conditions for near-optimal control of the model are established by using a Hamiltonian function. Then the necessary conditions for near-optimal control are obtained by using the Pontryagin maximum principle. Finally, the effect of popular science education, media coverage and Lévy noise on rumor propagation process control is verified by numerical simulation.

Keywords: rumor propagation, near-optimal control, popular science education, media coverage, Lévy noise

PACS: 02.30.Jr, 02.50.Ey, 02.50.Fz, 02.30.–f **DOI:** 10.1088/1674-1056/ac2f35

1. Introduction

In recent years, with the rapid development of the Internet, information exchange between individuals has become more rapid and convenient. As a kind of information with negative influence, the spreading speed and influence range of rumors has grown exponentially. In serious cases, rumors also harm social stability, which should cause all society to be highly vigilant.[1–3] With the outbreak of COVID-19, there have been many rumors, such as "You can prevent the virus with multiple layers of masks", "A patient flees from Peking Union Medical College Hospital" and so on. These rumors spread wildly on the Internet, resulting in the inability to allocate and use material resources such as masks, and also caused strong negative effects and social panic, which largely hindered the timely prevention and control of the epidemic. Therefore, it is necessary to minimize the negative impact of rumor propagation by studying the optimal control theory of the rumor propagation process.

As for the propagation process, many scholars have done relevant studies. In 1965, Daley and Kendal proposed the classic rumor propagation model (DK model),[4] which brought a breakthrough for further research on rumor propagation. Later, Maki and Thomson further improved the DK model into the MK model.[5] On this basis, scholars have proposed and studied many extended propagation models.[6–8] With the development of science and technology, the Internet and other complex networks have penetrated into people's lives and become an important platform for rumor propagation. Therefore, many scholars began to study the influence of complex network topology on the rumor propagation process. Zanette studied the dynamics of an infectious disease-like model of rumor propagation in a small-world network,[9] while Albert *et al.* studied the tolerance of complex networks to errors.[10] Doer *et al.* analyzed how rumors spread in social networks and simulated the propagation process of rumors in different network topologies.[11] Liu *et al.* proposed a Susceptible–Hesitated–Infected–Removed (SHIR) model to study the dynamics of competitive dual information diffusion.[12] Dong *et al.* proposed a new model for the rumor spreading process in online social networks.[13] Zhao *et al.* extended the SHIR model as in epidemic disease modeling to investigate the propagation dynamics of information online.[14] Zhu *et al.* proposed a rumor propagation model with nonlinear functions and time delay in social networks.[15]

The models involved in the above articles are all deterministic, but uncertainty is inherent in most real-world systems, such as changes in national policies and the involvement of media coverage (that is, the numbers of ignorant and spreaders may change suddenly after the involvement of media coverage). The Lévy jumping process is a discontinuous function with right continuity and left limit, which can better describe this mutation. Therefore, it is of great value to study the impact of the Lévy jumping process on the dynamics of rumor propagation. The stochastic model constructed by the combination of a deterministic model and Lévy noise can better describe the propagation process of rumors in the real world. In recent years, the stochastic propagation model of rumors has also increasingly attracted scholars' attention. Some scholars have considered Lévy noise, but they mainly studied the dynamic propagation behavior of rumors.[16,17] Dauhoo *et al.*

[†]Corresponding author. E-mail: huohuolin@yeah.net

© 2022 Chinese Physical Society and IOP Publishing Ltd

converted the deterministic rumor model into a stochastic rumor model by using the stochastic coefficient representing the influence of the system, proving the existence and uniqueness of global and local solutions.[18] Jia *et al.* studied a random rumor propagation model and obtained that the threshold of the random model under the influence of white noise was less than that of the deterministic model.[19] Lv *et al.* studied a random delayed rumor propagation model.[20] Ankur *et al.* introduced a random rumor model into a social network containing expert intervention, and drew the conclusion that noise might be one of the reasons for the persistence of rumors in a stochastic environment.[21]

The purpose of studying the dynamic process of rumor transmission is to better control rumors and reduce their negative impact. Therefore, how to control the spread of rumors has become an issue of concern to an increasing number of scholars. Amine *et al.* studied a new rumor propagation model and theoretically proved the existence of optimal control.[22] Chen *et al.* discussed the control of rumor propagation in an emergency and analyzed the stability of the model.[23] The literature contains some research in this area. For example, Hui *et al.* aimed to study the rumor spreading mechanism on social network platforms during the transmission of COVID-19, and to take education as a control measure for rumor spreading.[24] Wang *et al.* studied the extent to which media reports before the crisis effectively refuted rumors, and evaluated the impact of media reports before the crisis on the effectiveness of rumor refutation in the experiment.[25] Based on the nonlinear dynamics model, Huo *et al.* analyzed the rumor propagation process and studied the influence of media coverage and government emergency strategies on the information propagation dynamics after emergencies.[26]

There are two possible reasons for the spread of rumors. First of all, as false information, rumors are extremely misleading and provocative. Some people do not have enough knowledge to distinguish between true and false information, which leads people to listen to rumors and propagate them into facts. Secondly, when a rumor is generated and propagated in an emergency, there is no corresponding real information released in time to refute the rumor, which leads to the constant spread of the rumor. In this case, popular science education and media coverage are two effective methods to prevent the crisis of rumor spreading. Popular science education can improve people's knowledge level and the ability to identify rumors, and directly transform the ignorant into the immune. In this way, the source of rumor propagation can be reasonably controlled and the path of rumor propagation can be further cut off to weaken the influence of rumors. Media coverage can refute rumors in a targeted way and reduce the probability of rumors propagating, which can not only let people know the truth in time, but also guide people to choose the right behavior and prevent overreaction. In addition, the exact solution of the equation of state and the adjoint equation is difficult to obtain and the exact control is also difficult when the stochastic factors are considered. Therefore, it is of great practical significance to study the approximate optimal control of the rumor propagation model.

Based on the above analysis and actual situation, a rumor spreading model that joins the Lévy process can describe the dynamic process of rumor propagation more accurately. Most researchers use deterministic models when conducting rumor spreading research, and rarely consider random factors. In addition, some researchers only conduct research on the process, and rarely mention control strategies. Therefore, this paper proposes a stochastic model combining media coverage and Lévy noise to describe the dynamic process of rumors spread in the real world. By introducing two control strategies, popular science education and media coverage, an optimal control problem is proposed to minimize the influence of rumors spread and control costs.

The structure of this paper is as follows. Section 2 introduces the preparatory knowledge and establishes the stochastic ignorant–spreader–recovered (ISR) model; Section 3 and 4 give the necessary and sufficient conditions for near-optimal control; Section 6 carries on the numerical simulation; and Section 7 summarizes the whole paper.

2. Model formulation

In the classical ISR model of rumor propagation,[27] nodes in social networks are divided into three categories: ignorant (I) is a node without any information about the rumor, spreader (S) is the node that accepts the rumor and propagates it, and recovered (R) is the node that finally decides to stop propagating rumors after contact with rumors. In social networks, rumors propagate among people in accordance with the rules of mass action. The ignorant node has not contacted the rumor and will not propagate it, but it may become a spreader by contacting the spreader nodes.

The propagation process of rumors can be described as follows:

(i) The main propagation model of rumors is communication after the interaction of individual opinions, that is, the information exchange between individuals through telephone, instant messaging software, face-to-face conversation and other means.

(ii) Assume that a spreader propagates rumors with a certain probability when they meet an ignorant person. Media coverage can specifically refute rumors by releasing facts, thus gradually reducing the probability of rumor propagation and finally reaching a stable value.

(iii) When the spreader meets the recovered or another spreader, they lose the motivation to spread rumors.

(iv) Rumors spread within a certain community, and the total number of people is variable. Those who initially enter the community are ignorant.

The total population is $N = I + S + R$. Due to the contact between different types of individuals, information is transmitted and diffused, and then the types of individuals are changed, as shown in Fig. 1.

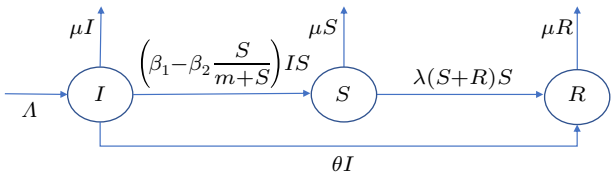


Fig. 1. Schematic diagram of a rumor under the influence of popular science education and media reports.

We develop this problem on the basis of a deterministic ISR model[28] as follows:

$$\begin{aligned}
\frac{\mathrm{d}I(t)}{\mathrm{d}t} &= \Lambda - \left(\beta_1 - \beta_2 \frac{S(t)}{m+S(t)}\right) I(t)S(t) - \theta I(t) - \mu I(t),\\
\frac{\mathrm{d}S(t)}{\mathrm{d}t} &= \left(\beta_1 - \beta_2 \frac{S(t)}{m+S(t)}\right) I(t)S(t)\\
&\quad -\lambda S(t)\left(S(t)+R(t)\right) - \mu S(t),\\
\frac{\mathrm{d}R(t)}{\mathrm{d}t} &= \lambda S(t)(S(t)+R(t)) + \theta I(t) - \mu R(t),
\end{aligned} \tag{1}$$

where Λ is the rate of migration; individuals enter a given community at a constant rate. θ is the immunity rate of ignorant individuals; under the influence of popular science education, some ignorant θI people have their own judgment of rumor, that is, they are immune to rumor and directly enter the R group. β_1 is the probability of the ignorant individual I contacting the spreader S, receiving and spreading; $\beta_1 SI$ means that some ignorant people choose to believe the information of the spreader and enter the class$S$. β_2 is the maximum influence of media on propagation probability when individual I of easy dyeing contacts with spreader S. $\beta_2 \frac{S}{m+S}$ refers to the contact between the ignorant I and the spreader S, and the probability of media obstructing the spread of the rumor. With an increase in the number of infected people, the media increases the intensity of reports, and the media obstructing the individual from believing and spreading the rumor also increases gradually, finally tending to a maximum value of β_2. m is the saturation coefficient, which measures the influence of media reports on communication. Considering that media reports can slow down the spread and diffusion of untrue information, but cannot completely stop its spread, we assume $\beta_1 > \beta_2$. μ is the rate of removal of individuals from each group for some reason. λ is the probability that the spreader S meets the recovered R, and the spreader loses the power to continue spreading; or, when two rumor spreaders S meet each other, they both lose the probability of continuing to spread rumors.

Generally, a Lévy process $\widetilde{L}(t)$ consists of the sum of a linear drift, a Brownian motion $B(t)$ and a superposition of centered independent Poisson processes with different jump sizes $\lambda(\mathrm{d}v)$ which is the rate of arrival intensity of the Poisson process with a jump of size v. According to the Lévy–Ito decomposition theorem,[29] we get

$$\widetilde{L}(t) = at + \sigma B(t) + \int_Y v\widetilde{N}(t, \mathrm{d}v), \tag{2}$$

where $a \in R$, $\sigma \geq 0, B(t)$ is standard Brownian motion. $\widetilde{N}$ is a compensated Poisson process and $\widetilde{N}(t, \mathrm{d}v) = \overline{N}(t, \mathrm{d}v) - \lambda(\mathrm{d}v)t$. $\overline{N}$ is a Poisson counting measure with characteristic measure λ on a measurable subset Y of $(0, +\infty)$ with $\lambda(Y) < \infty$.

In the process of rumor propagation in the real world, the process of rumor propagation is affected by many uncertain factors, such as national policies and individual personality factors. Therefore, in order to describe the process of rumor spreading more realistically, we assume that the transmission rate β_1 is subject to Lévy noise, that is, β_1 is replaced by $\beta_1 + \eta\,\mathrm{d}\widetilde{L}$. Then we obtain a new stochastic ISR model with Lévy noise

$$\begin{aligned}
\mathrm{d}I &= \left[\Lambda - \left(\beta_1 + \eta a - \beta_2 \frac{S}{m+S}\right) IS - \theta I - \mu I\right] \mathrm{d}t\\
&\quad -\eta\sigma IS\mathrm{d}B(t) - \int_Y \eta v IS\widetilde{N}(\mathrm{d}t, \mathrm{d}v),\\
\mathrm{d}S &= \left[\left(\beta_1 + \eta a - \beta_2 \frac{S}{m+S}\right) IS - \lambda S(S+R) - \mu S\right] \mathrm{d}t\\
&\quad +\eta\sigma IS\mathrm{d}B(t) + \int_Y \eta v IS\widetilde{N}(\mathrm{d}t, \mathrm{d}v),\\
\mathrm{d}R &= [\lambda S(S+R) + \theta I - \mu R]\mathrm{d}t.
\end{aligned} \tag{3}$$

In the above model, when rumors spread to a certain scale, they may endanger social harmony and stability, and bring huge losses to the related people. Therefore, it is necessary to study how to control the propagation of rumors. The reason why rumors are widely spread may be that some people are unable to distinguish between true and false information due to their insufficient knowledge reserve, so they take rumors as facts to propagate. Another reason may be that there is no corresponding real information about the breaking event to refute the rumor, so the rumor keeps propagating.

Under such circumstances, popular science education and media reports are the most effective ways to curb the spread of rumors. On the one hand, popular science education can improve a person's cognitive level so as to directly transform the ignorant into the immune. Popular science education stops rumors from spreading in the first place and reduces the number of the ignorant while increasing the number of the immune. On the other hand, media coverage can also be a targeted refutation of rumors. By increasing media coverage, the truth can be publicized in a timely manner and public opinion can be actively guided so as to reduce the number of spreaders and the

negative impact of rumor propagation. Reasonable coverage and time for media coverage and popular science education can reduce the number of spreaders and ignorants to a lower level. Therefore, we introduce education and media control $u(t)=(u_\theta(t),u_m(t))$ into the stochastic ISR model (3) as follows:

$$\begin{aligned}
\mathrm{d}I &= \left[\Lambda-\left(\beta_1+\eta a-\beta_2\frac{S}{(1-u_m)m+S}\right)IS\right.\\
&\quad\left.-(1+u_\theta)\theta I-\mu I\right]\mathrm{d}t-\eta\sigma IS\mathrm{d}B(t)\\
&\quad-\int_Y \eta v IS\widetilde{N}(\mathrm{d}t,\mathrm{d}v),\\
\mathrm{d}S &= \left[\left(\beta_1+\eta a-\beta_2\frac{S}{(1-u_m)m+S}\right)IS\right.\\
&\quad\left.-\lambda S(S+R)-\mu S\right]\mathrm{d}t+\eta\sigma IS\mathrm{d}B(t)\\
&\quad+\int_Y \eta v IS\widetilde{N}(\mathrm{d}t,\mathrm{d}v),\\
\mathrm{d}R &= [\lambda S(S+R)+(1+u_\theta)\theta I-\mu R]\,\mathrm{d}t, \qquad (4)
\end{aligned}$$

where u_m and u_θ represent the control variables for the influence of science popularization education and media coverage, respectively. Let $U_{\mathrm{ad}}\subset R$ be a bounded nonempty closed set. $u_\theta,u_m:[0,1]\times\Omega\to U$ is called admissible, if it is an F_t-adapted process with values in U_{ad}.

The purpose of optimal control is to minimize the negative effects caused by the propagation of rumors and maximize the positive social utility. This paper chooses the following aspects to consider the social utility of the propagation of untrue information:

(I) Minimize the number of ignorant individuals I and spreaders S of rumors.

(II) Minimize the cost of variable control in the process.

We define the associated objective function as

$$\begin{aligned}
&\mathrm{Min}\,J(I_0,S_0,R_0;u_m,u_\theta)\\
&=\int_0^T\left[B_0I(t)+B_1S(t)+\frac{1}{2}B_2\left(u_\theta^2+u_m^2\right)\right]\mathrm{d}t+h(x(T))\,\mathrm{d}t, \qquad (5)
\end{aligned}$$

where B_0,B_1 are small positive constants to balance the size of $S(t)$, which converts all social utilities into cost-oriented functions. B_2 is a positive constant. $h(x(T))=(0,S(t),0)$ is the number of spreaders at the end of rumor propagation at time T. The square of the control variable reflects the severity of side effects reported by popular science education and media coverage.[30,31]

3. Analysis of near-optimal control for stochastic ISR model

Before we study near-optimal control, we give some preliminary knowledge as follows; let $(\Omega,F,(F_t)_{0\le t\le T},P)$ be a complete filtered probability space on a finite time horizon $[0,T]$, and we assume that $(F_t)_{0\le t\le T}$ is the natural filtration and $F=F_t$. If the control process $u:[0,T]\times\Omega\to U$ is an F_t-adapted process with values in U, then it is called admissible. We use U_{ad} to denote the set that consists of all admissible controls; $|\cdot|$ denotes the form of an Euclidean space; f_x is the partial derivative of f with respect to x; χ_s denotes the indicator function of a set S; and $X+Y$ is the set $\{x+y:x\in X,y\in Y\}$ for any sets X and Y.

The control problem is to find an admissible control that minimizes or nearly minimizes the objective function $J(0,x_0;u(\cdot))$ over all $u(\cdot)\in U_{\mathrm{ad}}$. $V(0,x_0)=\min\limits_{u(\cdot)\in U_{\mathrm{ad}}}J(0,x_0;u(\cdot))$ is the value function.

Definition 1[32] If $u^*(\cdot)$ attains the minimum of $J(0,x_0;u(\cdot))$, then an admissible control $u^*(\cdot)$ is called optimal control.

Definition 2[32] For any given $\varepsilon>0$, if $|J(0,x_0;u^\varepsilon(\cdot))-V(0,x_0)|\le\varepsilon$, an admissible control $u^\varepsilon(\cdot)$ is called ε-optimal control.

Definition 3[32] A family of admissible controls $\{u^\varepsilon(\cdot)\}$ parameterized by $\varepsilon>0$ and any element $u^\varepsilon(\cdot)$ in the family are called near-optimal control, if $|J(0,x_0;u^\varepsilon(\cdot))-V(0,x_0)|\le\delta(\varepsilon)$ holds for a sufficiently small $\varepsilon>0$, where δ is a function of ε satisfying $\delta(\varepsilon)\to0$ as $\varepsilon\to0$. The estimate $\delta(\varepsilon)$ is called an error bound. If $\delta(\varepsilon)=c\varepsilon^k$ for some $c,k>0$, then $u^\varepsilon(\cdot)$ is called near-optimal with order ε^k.

Lemma 1 Ekeland's principle.[33] Let (Q,d) be a complete metric space and $F(\cdot):Q\to R$ be a lower-semicontinuous and bounded from below. For any $\varepsilon>0$, we assume that $u^\varepsilon(\cdot)\in Q$ satisfies

$$F(u^\varepsilon(\cdot))\le\inf_{u(\cdot)\in Q}F(u(\cdot))+\varepsilon.$$

Then there is $u^\varpi(\cdot)\in Q$ such that for all $\varpi>0$ and $u(\cdot)\in Q$,

$$\begin{aligned}
&F\left(u^\varpi(\cdot)\right)\le F(u^\varepsilon(\cdot)),\quad d\left(u^\varpi(\cdot),u^\varepsilon(\cdot)\right)\le\varpi,\\
&F\left(u^\varpi(\cdot)\right)\le F(u(\cdot))+\frac{\varepsilon}{\varpi}d\left(u,u^\varpi\right).
\end{aligned}$$

There are some assumptions, which will be needed in the proof of results.

(H1) For all $0\le t\le T$, the partial derivative $h_{S(t)}(x)$ is continuous. There is a positive imprecise parameter C such that $(1+|x|)^{-1}|h_{S(t)}(x)|\le C$.

(H2) Let $x(t),\widehat{x}(t)\in R_+^3$ and $h(x)$ be continuous; then there exists a positive imprecise parameter C for any $0\le t\le T$, such that

$$\begin{aligned}
&|h(x(t))-h(\widehat{x}(t))|\le C|S(t)-\widehat{S}(t)|,\\
&|h_{S(t)}(x(t))-h_{\widehat{S}(t)}(\widehat{x}(t))|\le C|S(t)-\widehat{S}(t)|.
\end{aligned}$$

(H3) The control set U is convex.

Under hypotheses (H1) and (H2), the adjoint equation admits a unique adapted solution. By Mazur's theorem,[34] we can then obtain the existence of optimal control under hypothesis (H3).

3.1. Sufficient conditions of near-optimal controls

In order to obtain sufficient conditions of near-optimal controls, we give the following lemmas for the moment estimation of the state and adjoint equation. For convenience, let $x(t) = (x_1(t), x_2(t), x_3(t))^{\mathrm{T}} = (I(t), S(t), R(t))^{\mathrm{T}}$, $u(t) = (u_m(t), u_\theta(t))$; then we get

$$\begin{aligned}
\mathrm{d}x_1(t) &= \left[\Lambda - \left(\beta_1 + \eta a - \beta_2 \frac{x_2}{(1-u_m)m + x_2}\right) x_1 x_2 \right.\\
&\quad \left. -(1+u_\theta)\theta x_1 - \mu x_1 \right] \mathrm{d}t - \eta\sigma x_1 x_2 \mathrm{d}B(t)\\
&\quad - \int_Y \eta v x_1 x_2 \widetilde{N}(\mathrm{d}t, \mathrm{d}v),\\
\mathrm{d}x_2(t) &= \left[\left(\beta_1 + \eta a - \beta_2 \frac{x_2}{(1-u_m)m + x_2}\right) x_1 x_2 \right.\\
&\quad \left. -\lambda x_2(x_2 + x_3) - \mu x_2 \right] \mathrm{d}t + \eta\sigma x_1 x_2 \mathrm{d}B(t)\\
&\quad + \int_Y \eta v x_1 x_2 \widetilde{N}(\mathrm{d}t, \mathrm{d}v),\\
\mathrm{d}x_3(t) &= [\lambda x_2(x_2 + x_3) + (1+u_\theta)\theta x_1 - \mu x_3]\,\mathrm{d}t. \qquad (6)
\end{aligned}$$

To simplify the above equation, let

$$\begin{aligned}
f_1(x(t), u(t)) &= \Lambda - \left(\beta_1 + \eta a - \beta_2 \frac{x_2}{(1-u_m)m + x_2}\right) x_1 x_2\\
&\quad -(1+u_\theta)\theta x_1 - \mu x_1,\\
f_2(x(t), u(t)) &= \left(\beta_1 + \eta a - \beta_2 \frac{x_2}{(1-u_m)m + x_2}\right) x_1 x_2\\
&\quad -\lambda x_2(x_2 + x_3) - \mu x_2,\\
f_3(x(t), u(t)) &= \lambda x_2(x_2 + x_3) + (1+u_\theta)\theta x_1 - \mu x_3.
\end{aligned}$$

Then we have

$$\begin{aligned}
\mathrm{d}x_1(t) &= f_1(x(t), u(t))\,\mathrm{d}t - \sigma_1'(x(t))\,\mathrm{d}B(t)\\
&\quad - \int_Y g_1(x(t), v)\widetilde{N}(\mathrm{d}t, \mathrm{d}v),\\
\mathrm{d}x_2(t) &= f_2(x(t), u(t))\,\mathrm{d}t + \sigma_2'(x(t))\,\mathrm{d}B(t)\\
&\quad + \int_Y g_2(x(t), v)\widetilde{N}(\mathrm{d}t, \mathrm{d}v),\\
\mathrm{d}x_3(t) &= f_3(x(t), u(t))\,\mathrm{d}t. \qquad (7)
\end{aligned}$$

3.1.1. Some priori estimates of the ignorant, spreader and recovered

Then we give some prior estimates of the ignorant, spreader and recovered. The following lemma illustrates the boundedness of ε moment estimates for the ignorant, spreader and recovered.

Lemma 2 For any $k \geq 0$ and $u(t) \in U_{\mathrm{ad}}$, we have

$$\sum_{i=1}^{3} E \sup_{0\leq t\leq T} |x_i(t)|^k \leq C, \qquad (8)$$

where C is a positive imprecise parameter that depends only on k.

Proof Let $X(t) = x_1(t) + x_2(t) + x_3(t)$; we have $\mathrm{d}X(t)/\mathrm{d}t = \Lambda - \mu x_1 - \mu x_2 - \mu x_3 = \Lambda - \mu X(t)$, which shows that $(\Lambda/\mu) - (1/\mu)\mathrm{e}^{-\mu t} \leq \liminf_{t\to\infty} X(t) \leq \limsup_{t\to\infty} X(t) \leq \Lambda/\mu$.

Therefore, all solutions $x_i(t)$ $(i = 1, 2, 3)$ of the system (6) are bounded by Λ/μ, which means Eq. (8) holds. Here we define a bounded set Γ for simplicity:

$$\Gamma = \left\{ x(t) \in R_+^3 : \frac{\Lambda}{\mu} - \frac{1}{\mu}\mathrm{e}^{-\mu t} \leq X(t) \leq \frac{\Lambda}{\mu} \right\} \subset R_+^3. \qquad (9)$$

Then we introduce the adjoint equation as follows:

$$\begin{aligned}
\mathrm{d}p_1(t) &= -b_1(x(t), u(t), p(t), q(t), r(v))\,\mathrm{d}t\\
&\quad + q_1(t)\,\mathrm{d}B(t) + \int_Y r_1(v)\widetilde{N}(\mathrm{d}t, \mathrm{d}v),\\
\mathrm{d}p_2(t) &= -b_2(x(t), u(t), p(t), q(t), r(v))\,\mathrm{d}t\\
&\quad + q_2(t)\,\mathrm{d}B(t) + \int_Y r_2(v)\widetilde{N}(\mathrm{d}t, \mathrm{d}v),\\
\mathrm{d}p_3(t) &= -b_3(x(t), u(t), p(t), q(t))\,\mathrm{d}t,\\
p_i(T) &= h_{x_i}(x(T)),\ \ i = 1, 2, 3, \qquad (10)
\end{aligned}$$

where

$$\begin{aligned}
b_1(x(t), u(t), p(t), q(t), r(v)) &= -\left[\left(\beta_1 + \eta a - \beta_2 \frac{x_2}{(1-u_m)m + x_2}\right) x_2 + (1+u_\theta)\theta + \mu\right] p_1(t) + \left(\beta_1 + \eta a - \beta_2 \frac{x_2}{(1-u_m)m + x_2}\right)\\
&\quad \times x_2 p_2(t) + (1+u_\theta)\theta p_3(t) - \eta\sigma x_2 q_1(t) + \eta\sigma x_2 q_2(t) - \int_Y \eta v x_2 r_1(v)\lambda(\mathrm{d}v) + \int_Y \eta v x_2 r_1(v)\lambda(\mathrm{d}v) + B_0,\\
b_2(x(t), u(t), p(t), q(t), r(v)) &= -\left[\beta_1 + \eta a - \frac{\beta_2 x_2(2(1-u_m)m + x_2)}{((1-u_m)m + x_2)^2}\right] x_1 p_1(t) + \left[\left(\beta_1 + \eta a - \frac{\beta_2 x_2(2(1-u_m)m + x_2)}{((1-u_m)m + x_2)^2}\right) x_1\right.\\
&\quad \left. -\lambda(2x_2 + x_3) - \mu\right] p_2(t) + \lambda(2x_2 + x_3) p_3(t) - \eta\sigma x_1 q_1(t) + \eta\sigma x_1 q_2(t) - \int_Y \eta v x_1 r_1(v)\lambda(\mathrm{d}v) + \int_Y \eta v x_1 r_1(v)\lambda(\mathrm{d}v) + B_1,\\
b_3(x(t), u(t), p(t), q(t)) &= -\lambda x_2 p_2(t) + (\lambda x_2 - \mu) p_3(t).
\end{aligned}$$

The second moment estimate of the adjoint Eq. (10) is given in the following lemma.

Lemma 3 For any $u_m(t), u_\theta(t), u'_m(t), u'_\theta(t) \in U_{\rm ad}$, we have

$$\sum_{i=1}^{3} E \sup_{0\le t\le T} |p_i(t)|^2 + \sum_{i=1}^{2} E \int_0^T |q_i(t)|^2 \mathrm{d}t + \sum_{i=1}^{2} E \int_Y |r_i(v)|^2 \mathrm{d}v \le C, \tag{11}$$

where C is a positive imprecise parameter.

Proof Integrating both sides of the first equation of Eq. (10) from t to T, we get

$$p_1(T) - p_1(t) = -\int_t^T b_1(x(s),u(s),p(s),q(s),r(v))\mathrm{d}s + \int_t^T q_1(s)\mathrm{d}B(s) + \int_t^T \int_Y r_1(v)\widetilde{N}(\mathrm{d}s,\mathrm{d}v). \tag{12}$$

Squaring Eq. (12) and making use of the second moments, where $x(t) \in \Gamma$ for all $t \ge 0$, we have

$$\begin{aligned} & E|p_1(t)|^2 + E\int_t^T |q_1(s)|^2\mathrm{d}s + (T-t)E\int_t^T |r_1(v)|^2\mathrm{d}v, \\ \le\ & CE|p_1(T)|^2 \\ & + C(T-t)E\int_t^T |b_1(x(s),u(s),p(s),q(s),r(v))|^2\mathrm{d}s \\ \le\ & CE|1+p_1(T)|^2 + C(T-t)\sum_{i=1}^{3} E\int_t^T |p_i(s)|^2\mathrm{d}s \\ & + C(T-t)\sum_{i=1}^{2} E\int_t^T |q_i(s)|^2\mathrm{d}s \\ & + C(T-t)\sum_{i=1}^{2} E\int_Y |r_i(v)|^2\mathrm{d}v. \end{aligned} \tag{13}$$

In the same way, we have

$$\begin{aligned} & E|p_2(t)|^2 + E\int_t^T |q_2(s)|^2\mathrm{d}s + (T-t)E\int_t^T |r_2(v)|^2\mathrm{d}v \\ \le\ & CE|1+p_2(T)|^2 + C(T-t)\sum_{i=1}^{3} E\int_t^T |p_i(s)|^2\mathrm{d}s \\ & + C(T-t)\sum_{i=1}^{2} E\int_t^T |q_i(s)|^2\mathrm{d}s \\ & + C(T-t)\sum_{i=1}^{2} E\int_Y |r_i(v)|^2\mathrm{d}v, \end{aligned} \tag{14}$$

$$\begin{aligned} & E|p_3(t)|^2 \\ \le\ & CE|1+p_3(T)|^2 + C(T-t)E\int_t^T |p_2(s)|^2\mathrm{d}s \\ & + C(T-t)E\int_t^T |p_3(s)|^2\mathrm{d}s. \end{aligned} \tag{15}$$

Therefore, by Eqs. (13)–(15), we can get

$$\begin{aligned} & \sum_{i=1}^{3} |p_i(t)|^2 + \sum_{i=1}^{2} E\int_t^T |q_i(s)|^2\mathrm{d}s \\ & + (T-t)\sum_{i=1}^{2} E\int_Y |r_i(v)|^2\mathrm{d}v \\ \le\ & \sum_{i=1}^{3} CE|1+p_1(T)|^2 + C(T-t)\sum_{i=1}^{3} E\int_t^T |p_i(s)|^2\mathrm{d}s \\ & + C(T-t)\sum_{i=1}^{2} E\int_t^T |q_i(s)|^2\mathrm{d}s \\ & + C(T-t)\sum_{i=1}^{2} E\int_Y |r_i(v)|^2\mathrm{d}v. \end{aligned} \tag{16}$$

Then we have

$$\begin{aligned} & \sum_{i=1}^{3} |p_i(t)|^2 + \frac{1}{2}\sum_{i=1}^{2} E\int_t^T |q_i(s)|^2\mathrm{d}s + \frac{1}{2}\sum_{i=1}^{2} E\int_Y |r_i(v)|^2\mathrm{d}v \\ \le\ & \sum_{i=1}^{3} CE|1+p_1(T)|^2 + C(T-t)\sum_{i=1}^{3} E\int_t^T |p_i(s)|^2\mathrm{d}s, \end{aligned} \tag{17}$$

where $t \in [T-\xi, T]$ with $\xi = 1/(2C)$. From Eq. (17) and Gronwall's inequality, we get

$$\begin{aligned} \sum_{i=1}^{3} E \sup_{0\le t\le T} |p_i(t)|^2 &\le C, \\ \sum_{i=1}^{2} E\int_t^T |q_i(t)|^2\mathrm{d}t &\le C, \\ \sum_{i=1}^{2} E\int_Y |r_i(v)|^2\mathrm{d}v &\le C. \end{aligned} \tag{18}$$

Using the same method in Eqs. (13)–(15) in $[T-\xi, T]$, the estimate in Eq. (17) holds for all $t \in [T-2\xi, T]$. Thus, we know that Eq. (18) holds for $t \in [0, T]$ by repeating a finite number of steps.

From Eq. (12) we can get

$$\begin{aligned} p_1(t) =\ & p_1(T) + \int_t^T b_1(x(s),u(s),p(s),q(s),r(v))\mathrm{d}s \\ & - \int_0^T q_1(s)\mathrm{d}B(s) + \int_0^t q_1(s)\mathrm{d}B(s) \\ & - \int_0^T \int_Y r_1(v)\widetilde{N}(\mathrm{d}s,\mathrm{d}v) \\ & + \int_0^t \int_Y r_1(v)\widetilde{N}(\mathrm{d}s,\mathrm{d}v). \end{aligned} \tag{19}$$

By the elementary inequality

$$|m_1+m_2+m_3+m_4|^n \le 4^n(|m_1|^n + |m_2|^n + |m_3|^n + |m_4|^n), \quad \forall n > 0,$$

then we have

$$|p_1(t)|^2 \le C\left[1 + |p_1(T)|^2 + \int_0^T \left(\sum_{i=1}^{3}|p_i(s)|^2 + \sum_{i=1}^{2}|q_i(s)|^2 + \sum_{i=1}^{2}\int_Y |r_i(v)|^2\mathrm{d}v\right)\mathrm{d}s + \left(\int_0^T q_1(s)\mathrm{d}B(s)\right)^2\right.$$

$$+\left(\int_0^t q_1(s)\,\mathrm{d}B(s)\right)^2+\left(\int_0^T\int_Y r_1(v)\widetilde{N}(\mathrm{d}s,\mathrm{d}v)\right)^2+\left(\int_0^t\int_Y r_1(v)\widetilde{N}(\mathrm{d}s,\mathrm{d}v)\right)^2\Bigg]. \tag{20}$$

Similar to Eq. (19), we can get similar results of the above inequality. Combining these inequalities, we have

$$\begin{aligned}\sum_{i=1}^3|p_i(t)|^2 \le C\Bigg[1+\sum_{i=1}^3|p_i(T)|^2+\int_0^T\left(\sum_{i=1}^3|p_i(s)|^2+\sum_{i=1}^2|q_i(s)|^2+\sum_{i=1}^2\int_Y|r_i(v)|^2\mathrm{d}v\right)\mathrm{d}s+\sum_{i=1}^2\left(\int_0^T q_i(s)\,\mathrm{d}B(s)\right)^2\\ +\sum_{i=1}^2\left(\int_0^t q_i(s)\,\mathrm{d}B(s)\right)^2+\sum_{i=1}^2\left(\int_0^T\int_Y r_1(v)\widetilde{N}(\mathrm{d}s,\mathrm{d}v)\right)^2+\sum_{i=1}^2\left(\int_0^t\int_Y r_1(v)\widetilde{N}(\mathrm{d}s,\mathrm{d}v)\right)^2\Bigg].\end{aligned} \tag{21}$$

Taking expectation on both sides and using the Burkholder–Davis–Gundy inequality (see Ref. [35], Theorem 1.7.3), we get

$$\sum_{i=1}^3 E\sup|p_i(t)|^2 \le C\left(\sum_{i=1}^3 E|p_i(T)|^2+\sum_{i=1}^3 E\int_0^T|p_i(s)|^2\mathrm{d}s+\sum_{i=1}^2 E\int_0^T|q_i(s)|^2\mathrm{d}s+\sum_{i=1}^2 E\int_Y|r_i(v)|^2\mathrm{d}v\right). \tag{22}$$

Using Gronwall's inequality, we can derive Eq. (11) from the above inequality. Then the proof is complete.

3.1.2. Sufficient conditions for near-optimal controls

In this section, we define the Hamiltonian function $H(t,x(t),u(t),p(t),q(t),r(v))$, $H:[0,T]\times R_+^3\times U_{\mathrm{ad}}\times R_+^3\times R_+^2\times R_+^2\to R$,

$$H(t,x(t),u(t),p(t),q(t),r(v))=f^T(x(t),u(t))p(t)+\sigma_*^T(x(t))q(t)+\int_Y g^T(x(t),v)r(v)\lambda(\mathrm{d}v)+L(x(t),u(t)), \tag{23}$$

where

$$f(x(t),u(t))=\begin{pmatrix}f_1(x(t),u(t))\\ f_2(x(t),u(t))\\ f_3(x(t),u(t))\end{pmatrix},\quad \sigma_*(x(t))=\begin{pmatrix}\sigma_1'(x(t))\\ \sigma_2'(x(t))\end{pmatrix},\quad g(x(t),v)=\begin{pmatrix}g_1(x(t),v)\\ g_2(x(t),v)\end{pmatrix},$$

and $L(x(t),u(t))=B_0x_1(t)+B_1x_2(t)+B_2(u_\theta^2+u_m^2)$.

Then we discuss the sufficient conditions for the existence of near-optimal control of ISR models by applying lemmas 2 and 3.

Theorem 1 Suppose (H1), (H2) and (H3) hold. Let $(x^\varepsilon(t),u^\varepsilon(t))$ be an admissible pair and $(p^\varepsilon(t),q^\varepsilon(t),r^\varepsilon(v))$ be the solutions of adjoint Eq. (10) corresponding to $(x^\varepsilon(t),u^\varepsilon(t))$. If $H(t,x(t),u(t),p(t),q(t),r(v))$ is convex, that is, for some $\varepsilon>0$, if

$$\begin{aligned}&\min_{u(t)\in U_{\mathrm{ad}}[0,T]}E\int_0^T\left(\frac{\beta_2x_1^\varepsilon(t)x_2^{2\varepsilon}(t)m}{\left((1-u_m)m+x_2^\varepsilon(t)\right)^2}\left(p_1^\varepsilon(t)-p_2^\varepsilon(t)\right)u_m+\theta x_1^\varepsilon(t)(p_3^\varepsilon(t)-p_1^\varepsilon(t))u_\theta+B_2(u_m^2+u_\theta^2)\right)\mathrm{d}t\\ &\ge E\int_0^T\left(\frac{\beta_2x_1^\varepsilon(t)x_2^{2\varepsilon}(t)m}{((1-u_m^\varepsilon)m+x_2^\varepsilon(t))^2}(p_1^\varepsilon(t)-p_2^\varepsilon(t))u_m^\varepsilon+\theta x_1^\varepsilon(t)(p_3^\varepsilon(t)-p_1^\varepsilon(t))u_\theta^\varepsilon+B_2(u_m^{2\varepsilon}+u_\theta^{2\varepsilon})\right)\mathrm{d}t-\varepsilon,\end{aligned} \tag{24}$$

then

$$J(0,x_0;u^\varepsilon(t))\le\inf_{u(t)\in U_{\mathrm{ad}}[0,T]}J(0,x_0;u(t))+C\varepsilon^{1/2}.$$

Proof From Eq. (5) and the definition of $H(t,x(t),u(t),p(t),q(t),r(v))$, we get

$$J(0,x_0;u^\varepsilon(t))-J(0,x_0;u(t))=k_1+k_2-k_3, \tag{25}$$

where

$$\begin{aligned}k_1 &= E\int_0^T\left[H(t,x^\varepsilon(t),u^\varepsilon(t),p^\varepsilon(t),q^\varepsilon(t),r^\varepsilon(v))-H(t,x(t),u(t),p^\varepsilon(t),q^\varepsilon(t),r^\varepsilon(v))\right]\mathrm{d}t,\\ k_2 &= E[h(x^\varepsilon(T))-h(x(T))],\\ k_3 &= E\int_0^T[f^T(x^\varepsilon(t),u^\varepsilon(t))-f^T(x(t),u(t))]p^\varepsilon(t)+[\sigma_*^T(x^\varepsilon(t))-\sigma_*^T(x(t))]q^\varepsilon(t)\\ &\quad+\int_Y[g^T(x^\varepsilon(t),v)-g^T(x(t),v)r^\varepsilon(v)\lambda(\mathrm{d}v)]\,\mathrm{d}t.\end{aligned} \tag{26}$$

For the convexity of $H(t,x^\varepsilon(t),u^\varepsilon(t),p^\varepsilon(t),q^\varepsilon(t),r^\varepsilon(v))$, we have

$$\begin{aligned} k_1 \le & \sum_{i=1}^{3} E\int_0^T H_{x_i}(t,x^\varepsilon(t),u^\varepsilon(t),p^\varepsilon(t),q^\varepsilon(t),r^\varepsilon(v))\,(x_i^\varepsilon(t)-x_i(t))\,\mathrm{d}t \\ & + E\int_0^T H_{u_m}(t,x^\varepsilon(t),u^\varepsilon(t),p^\varepsilon(t),q^\varepsilon(t),r^\varepsilon(v))(u_m^\varepsilon(t)-u_m(t))\,\mathrm{d}t \\ & + E\int_0^T H_{u_\theta}(t,x^\varepsilon(t),u^\varepsilon(t),p^\varepsilon(t),q^\varepsilon(t),r^\varepsilon(v))(u_\theta^\varepsilon(t)-u_\theta(t))\,\mathrm{d}t, \end{aligned} \tag{27}$$

$$k_2 \le \sum_{i=1}^{3} E\left[h_{x_i}(x^\varepsilon(T))(x_i^\varepsilon(t)-x_i(t))\right]. \tag{28}$$

Applying Ito's formula[36] to $\sum_{i=1}^{3} p_i^\varepsilon(t)(x_i^\varepsilon(t)-x_i(t))$, we obtain

$$\begin{aligned} \sum_{i=1}^{3} E[h_{x_i}(x^\varepsilon(T))(x_i^\varepsilon(t)-x_i(t))] = & -\sum_{i=1}^{3} E\int_0^T H_{x_i}(x^\varepsilon(s),u^\varepsilon(s),p^\varepsilon(s),q^\varepsilon(s),r^\varepsilon(v))(x_i^\varepsilon(s)-x_i(s))\,\mathrm{d}s \\ & + \sum_{i=1}^{3} E\int_0^T p_i^\varepsilon(s)|f_i(x^\varepsilon(s),u^\varepsilon(s))-f_i(x(s),u(s))|\,\mathrm{d}s + \sum_{i=1}^{2} E\int_0^T q_i^\varepsilon(s)|\sigma_i(x^\varepsilon(s))-\sigma_i(x(s))|\,\mathrm{d}s \\ & + \sum_{i=1}^{2} E\int_0^T\int_Y r_i^\varepsilon(v)|g_i(x^\varepsilon(s),v)-g_i(x(s),v)|\lambda(\mathrm{d}v)\,\mathrm{d}s. \end{aligned}$$

Then we have

$$k_3 = \sum_{i=1}^{3} E\int_0^T H_{x_i}(x^\varepsilon(s),u^\varepsilon(s),p^\varepsilon(s),q^\varepsilon(s),r^\varepsilon(v))(x_i^\varepsilon(s)-x_i(s))\,\mathrm{d}s + \sum_{i=1}^{3} E\left[h_{x_i}(x^\varepsilon(T))(x_i^\varepsilon(t)-x_i(t))\right]. \tag{29}$$

Substituting Eqs. (27)–(29) into Eq. (25), we obtain

$$J(0,x_0;u^\varepsilon(t))-J(0,x_0;u(t)) \le E\int_0^T H_{u(t)}(u^\varepsilon(t)-u(t))\,\mathrm{d}t. \tag{30}$$

According to Eq. (24), we can get

$$\begin{aligned} & E\int_0^T\left(\frac{\beta_2 x_1^\varepsilon(t)x_2^{2\varepsilon}(t)m}{((1-u_m)m+x_2^\varepsilon(t))^2}(p_1^\varepsilon(t)-p_2^\varepsilon(t))u_m+\theta x_1^\varepsilon(t)(p_3^\varepsilon(t)-p_1^\varepsilon(t))u_\theta+B_2(u_m^2+u_\theta^2)\right)\mathrm{d}t \\ & +E\int_0^T\left\{\left[\Lambda-\left(\beta_1+\eta a-\beta_2\frac{x_2^\varepsilon(t)}{(1-u_m)m+x_2^\varepsilon(t)}\right)x_1^\varepsilon(t)x_2^\varepsilon(t)-(1-u_\theta)\theta x_1^\varepsilon(t)-\mu x_1^\varepsilon(t)\right]p_1^\varepsilon(t)\right. \\ & +\left[\left(\beta_1+\eta a-\beta_2\frac{x_2^\varepsilon(t)}{(1-u_m)m+x_2^\varepsilon(t)}\right)x_1^\varepsilon(t)x_2^\varepsilon(t)-\lambda x_2^\varepsilon(t)(x_2^\varepsilon(t)+x_3^\varepsilon(t))-\mu x_2^\varepsilon(t)\right]p_2^\varepsilon(t) \\ & +[\lambda x_2^\varepsilon(t)(x_2^\varepsilon(t)+x_3^\varepsilon(t))+(1-u_\theta)\theta x_1^\varepsilon(t)-\mu x_3^\varepsilon(t)]\,p_3^\varepsilon(t)-\eta\sigma x_1^\varepsilon(t)x_2^\varepsilon(t)q_1^\varepsilon(t) \\ & \left.+\eta\sigma x_1^\varepsilon(t)x_2^\varepsilon(t)q_2^\varepsilon(t)-\int_Y\eta v x_1^\varepsilon(t)x_2^\varepsilon(t)r_1^\varepsilon(t)\lambda(\mathrm{d}v)+\int_Y\eta v x_1^\varepsilon(t)x_2^\varepsilon(t)r_2^\varepsilon(t)\lambda(\mathrm{d}v)\right\}\mathrm{d}t \\ \ge & \sup_{u(t)\in U[0,T]} E\int_0^T\left(\frac{\beta_2 x_1^\varepsilon(t)x_2^{2\varepsilon}(t)m}{((1-u_m^\varepsilon)m+x_2^\varepsilon(t))^2}(p_1^\varepsilon(t)-p_2^\varepsilon(t))u_m^\varepsilon+\theta x_1^\varepsilon(t)(p_3^\varepsilon(t)-p_1^\varepsilon(t))u_\theta^\varepsilon+B_2(u_m^{2\varepsilon}+u_\theta^{2\varepsilon})\right)\mathrm{d}t \\ & -\varepsilon+E\int_0^T\left\{\left[\Lambda-(\beta_1+\eta a-\beta_2\frac{x_2^\varepsilon(t)}{(1-u_m)m+x_2^\varepsilon(t)})x_1^\varepsilon(t)x_2^\varepsilon(t)-(1-u_\theta)\theta x_1^\varepsilon(t)-\mu x_1^\varepsilon(t)\right]p_1^\varepsilon(t)\right. \\ & +\left[\left(\beta_1+\eta a-\beta_2\frac{x_2^\varepsilon(t)}{(1-u_m)m+x_2^\varepsilon(t)}\right)x_1^\varepsilon(t)x_2^\varepsilon(t)-\lambda x_2^\varepsilon(t)(x_2^\varepsilon(t)+x_3^\varepsilon(t))-\mu x_2^\varepsilon(t)\right]p_2^\varepsilon(t) \\ & +[\lambda x_2^\varepsilon(t)(x_2^\varepsilon(t)+x_3^\varepsilon(t))+(1-u_\theta)\theta x_1^\varepsilon(t)-\mu x_3^\varepsilon(t)]\,p_3^\varepsilon(t)-\eta\sigma x_1^\varepsilon(t)x_2^\varepsilon(t)q_1^\varepsilon(t) \\ & \left.+\eta\sigma x_1^\varepsilon(t)x_2^\varepsilon(t)q_2^\varepsilon(t)-\int_Y\eta v x_1^\varepsilon(t)x_2^\varepsilon(t)r_1^\varepsilon(t)\lambda(\mathrm{d}v)+\int_Y\eta v x_1^\varepsilon(t)x_2^\varepsilon(t)r_2^\varepsilon(t)\lambda(\mathrm{d}v)\right\}\mathrm{d}t. \end{aligned} \tag{31}$$

According to Eqs. (31) and (23), we can obtain

$$\begin{aligned} & E\int_0^T H(t,x^\varepsilon(t),u^\varepsilon(t),p^\varepsilon(t),q^\varepsilon(t),r^\varepsilon(v))\,\mathrm{d}t \\ \geq & \sup_{u(t)\in U_{\mathrm{ad}}[0,T]} E\int_0^T H(t,x^\varepsilon(t),u(t),p^\varepsilon(t),q^\varepsilon(t),r^\varepsilon(v))\,\mathrm{d}t \\ & -\varepsilon. \end{aligned} \tag{32}$$

Then we define a function $F(\cdot): U_{\mathrm{ad}} \to R$

$$F(u(t)) = E\int_0^T H(t,x^\varepsilon(t),u(t),p^\varepsilon(t),q^\varepsilon(t),r^\varepsilon(v))\,\mathrm{d}t. \tag{33}$$

To estimate the term $H_{u(t)}(t,x^\varepsilon(t),u^\varepsilon(t),p^\varepsilon(t),q^\varepsilon(t),r^\varepsilon(v))$, we define a complete metric $\widetilde{d}$ on U_{ad}: for any $u(t)$ and $u'(t)\in U_{\mathrm{ad}}$,

$$\begin{aligned} \widetilde{d}(u(t),u'(t)) = & E\int_0^T \Big(1+\sum_{i=1}^3 |p_i^\varepsilon(t)| + \sum_{i=1}^2 |q_i^\varepsilon(t)| \\ & + \sum_{i=1}^2 |r_i^\varepsilon(v)|\Big)|u(t)-u'(t)|\,\mathrm{d}t. \end{aligned} \tag{34}$$

According to hypothesis (H2), we can obtain the fact that $F(\cdot)$ is continuous on U_{ad} with regard to the metric $\widetilde{d}$. Therefore, from Eq. (32) and lemma 1, there exists a $\widetilde{u^\varepsilon}(t)\in U_{\mathrm{ad}}$ such that

$$\begin{aligned} & \widetilde{d}(u^\varepsilon(t),\widetilde{u^\varepsilon}(t)) \leq \varepsilon^{1/2}, \\ & F(\widetilde{u^\varepsilon}(t)) \leq F(u(t)) + \varepsilon^{1/2}\widetilde{d}(u(t),\widetilde{u^\varepsilon}(t)),\ \ \forall u(t)\in U_{\mathrm{ad}}. \end{aligned} \tag{35}$$

Then we can obtain

$$\begin{aligned} & H(t,x^\varepsilon(t),\widetilde{u^\varepsilon}(t),p^\varepsilon(t),q^\varepsilon(t),r^\varepsilon(v)) \\ = & \min_{u(t)\in U_{\mathrm{ad}}}\Big[H(t,x^\varepsilon(t),u(t),p^\varepsilon(t),q^\varepsilon(t),r^\varepsilon(v)) \\ & +\varepsilon^{1/2}\Big(1+\sum_{i=1}^3 |p_i^\varepsilon(t)| + \sum_{i=1}^2 |q_i^\varepsilon(t)| \\ & + \sum_{i=1}^2 |r_i^\varepsilon(v)|\Big)|u(t)-\widetilde{u^\varepsilon}(t)|\Big]. \end{aligned} \tag{36}$$

From lemma 1 and Eq. (16), we have

$$\begin{aligned} & 0\in \partial_{u(t)} H\left(t,x^\varepsilon(t),\widetilde{u^\varepsilon}(t),p^\varepsilon(t),q^\varepsilon(t),r^\varepsilon(v)\right) \\ \subset & \ \partial_{u(t)} H\left(t,x^\varepsilon(t),\widetilde{u^\varepsilon}(t),p^\varepsilon(t),q^\varepsilon(t),r^\varepsilon(v)\right) \\ & +\left[-\varepsilon^{1/2}\omega,\varepsilon^{1/2}\omega\right], \end{aligned} \tag{37}$$

where $\omega = 1+\sum_{i=1}^3 |p_i^\varepsilon(t)| + \sum_{i=1}^2 |q_i^\varepsilon(t)| + \sum_{i=1}^2 |r_i^\varepsilon(v)|$.

Since the Hamiltonian function is differentiable in $u(t)$ by hypothesis (H2), then we can obtain that there exists an $\omega_1^\varepsilon(t)\in[-\varepsilon^{1/2}\omega,\varepsilon^{1/2}\omega]$ such that

$$H_{u(t)}\left(t,x^\varepsilon(t),\widetilde{u^\varepsilon}(t),p^\varepsilon(t),q^\varepsilon(t),r^\varepsilon(v)\right) + \omega_1^\varepsilon(t) = 0. \tag{38}$$

Therefore, from Eq. (38) and hypothesis (H2), we get

$$\begin{aligned} & \left|H_{u(t)}(t,x^\varepsilon(t),u^\varepsilon(t),p^\varepsilon(t),q^\varepsilon(t),r^\varepsilon(v))\right| \\ \leq & \Big|H_{u(t)}(t,x^\varepsilon(t),u^\varepsilon(t),p^\varepsilon(t),q^\varepsilon(t),r^\varepsilon(v)) \\ & -H_{u(t)}\left(t,x^\varepsilon(t),\widetilde{u^\varepsilon}(t),p^\varepsilon(t),q^\varepsilon(t),r^\varepsilon(v)\right)\Big| \\ & +\Big|H_{u(t)}\left(t,x^\varepsilon(t),\widetilde{u^\varepsilon}(t),p^\varepsilon(t),q^\varepsilon(t),r^\varepsilon(v)\right)\Big| \\ \leq & C\omega^\varepsilon(t)|u^\varepsilon(t)-\widetilde{u^\varepsilon}(t)| + \omega_1^\varepsilon(t) \\ \leq & C\omega^\varepsilon(t)|u^\varepsilon(t)-\widetilde{u^\varepsilon}(t)| + 2\varepsilon^{1/2}\omega^\varepsilon(t). \end{aligned} \tag{39}$$

According to lemma 3 and the definition of $\widetilde{d}$, the expected conclusion can be obtained from Eqs. (30) and (39) and Holder's inequality.

3.2. Necessary conditions for near-optimal controls

In this section, we discuss the necessary conditions of near-optimal control. First, we give the following priori estimates to the ignorant, spread and recovered persons of the system to get our results.

In order to give the following lemma, we first define a metric on the admissible control field $U_{\mathrm{ad}}[0,T]$ as follows:

$$\begin{aligned} & d(u(t),\widehat{u}(t)) = E\left[\mathrm{mes}\left\{t\in[0,T]: u(t)\neq\widehat{u}(t)\right\}\right], \\ & \forall u(t),\widehat{u}(t)\in U_{\mathrm{ad}}, \end{aligned} \tag{40}$$

where mes stands for the Lebesgue measure. Since U_{ad} is closed, it can be proven that U_{ad} is a complete metric space under d.[34] The problem of moment estimation of the distance between two state variables is discussed by using the distance of the defined control variables.

Lemma 4 For all $k\geq 0$ and $0<l<1$ satisfying $kl<1$, and $u(t),\widehat{u}(t)\in U_{\mathrm{ad}}$, along with the corresponding trajectories $x(t),x'(t)$, there exists a positive imprecise parameter $C=C(k,l)$ such that

$$\begin{aligned} & \sum_{i=1}^3 E\sup_{0\leq t\leq T}|x_i(t)-\widehat{x}_i(t)|^{2k} \\ \leq & C\left[d(u_m(t),\widehat{u}_m(t))^{lk} + d(u_\theta(t),\widehat{u}_\theta(t))^{lk}\right]. \end{aligned} \tag{41}$$

Proof We assume that $k\geq 1$. Let $\beta=\beta_1+\eta a$, $v'=\int_Y \eta v\,\mathrm{d}v$. The following equation is obtained by using the elementary inequality

$$\begin{aligned} & E\sup_{0\leq t\leq r}|x_1(t)-\widehat{x}_1(t)|^{2k} \\ \leq & CE\int_0^r \Big[(\beta^{2k}+(\eta\sigma)^{2k}+v'^{2k})|x_1(t)x_2(t)-\widehat{x}_1(t)\widehat{x}_2(t)|^{2k} \\ & +\mu^{2k}|x_1(t)-\widehat{x}_1(t)|^{2k}\Big] \\ & +\left|\frac{\beta_2 x_1(t)x_2^2(t)}{(1-u_m)m+x_2(t)} - \frac{\beta_2\widehat{x}_1(t)\widehat{x}_2{}^2(t)}{(1-u_m)m+\widehat{x}_2(t)}\right|^{2k} \\ & +|(1+u_\theta)\theta x_1(t)-(1+u'_\theta)\theta\widehat{x}_1(t)|^{2k}\,\mathrm{d}t \end{aligned}$$

$$
\begin{aligned}
&\le CE\int_0^r \sum_{i=1}^{2} |x_i(t)-\widehat{x}_i(t)|^{2k}\mathrm{d}t \\
&\quad +C\left[E\int_0^r \chi_{u(t)\neq\widehat{u}(t)}(t)\mathrm{d}t\right]^{lk} \\
&\le C\left[E\int_0^r \sum_{i=1}^{2} |x_i(t)-\widehat{x}_i(t)|^{2k}\mathrm{d}t + d(u_m(t),\widehat{u_m}(t))^{lk}\right. \\
&\quad \left. +d(u_\theta(t),\widehat{u_\theta}(t))^{lk}\right]. \qquad (42)
\end{aligned}
$$

If $0<l<1, k\ge 1$, through the Cauchy–Schwarz inequality, we get

$$
\begin{aligned}
&E\int_0^r \chi_{u(t)\neq\widehat{u}(t)}(t)\mathrm{d}t \\
&\le C\left(E\int_0^r \mathrm{d}t\right)^{1-lk}\times\left(E\int_0^r \chi_{u(t)\neq\widehat{u}(t)}(t)\mathrm{d}t\right)^{lk} \\
&\le C\left[d\left(u_m(t),\widehat{u_m}(t)\right)^{lk}+d\left(u_\theta(t),\widehat{u_\theta}(t)\right)^{lk}\right]. \qquad (43)
\end{aligned}
$$

Then, we assume $k\ge 1$. For $\forall r>0$, we get the following inequality through the elementary inequality:

$$
\begin{aligned}
&E\sup_{0\le t\le r}|x_2(t)-\widehat{x}_2(t)|^{2k} \\
&\le CE\int_0^r \left[(\beta^{2k}+(\eta\sigma)^{2k}+\nu'^{2k})\,|x_1(t)x_2(t)-\widehat{x}_1(t)\widehat{x}_2(t)|^{2k}\right. \\
&\quad \left.+\lambda^{2k}\left|x_2^2(t)-\widehat{x}_2^2(t)\right|^{2k}\right] \\
&\quad +\left|\frac{\beta_2 x_1(t)x_2^2(t)}{(1-u_m)m+x_2(t)}-\frac{\beta_2\widehat{x}_1(t)\widehat{x}_2^2(t)}{(1-\widehat{u_m})m+\widehat{x}_2(t)}\right|^{2k} \\
&\quad +\lambda^{2k}|x_2(t)x_3(t)-\widehat{x}_2(t)\widehat{x}_3(t)|^{2k} \\
&\quad -\mu^{2k}\left|x_2^2(t)-\widehat{x}_2^2(t)\right|^{2k}\mathrm{d}t \\
&\le CE\int_0^r \sum_{i=1}^{3}|x_i(t)-\widehat{x}_i(t)|^{2k}\mathrm{d}t+C\left[E\int_0^r \chi_{u(t)\neq\widehat{u}(t)}(t)\mathrm{d}t\right]^{lk} \\
&\le C\left[E\int_0^r \sum_{i=1}^{3}|x_i(t)-\widehat{x}_i(t)|^{2k}\mathrm{d}t+d(u_m(t),\widehat{u_m}(t))^{lk}\right]. \qquad (44)
\end{aligned}
$$

Using the same method to estimate to $|x_3(t)-\widehat{x}_3(t)|^{2k}$,

$$
\begin{aligned}
&E\sup_{0\le t\le r}|x_3(t)-\widehat{x}_3(t)|^{2k} \\
&\le CE\int_0^r \left[\lambda^{2k}|x_2^2(t)-\widehat{x}_2^2(t)|^{2k}\right] \\
&\quad +\lambda^{2k}|x_2(t)x_3(t)-\widehat{x}_2(t)\widehat{x}_3(t)|^{2k}+\mu^{2k}|x_2^2(t)-\widehat{x}_2^2(t)|^{2k} \\
&\quad +|(1+u_\theta)\theta x_1(t)-(1+\widehat{u_\theta})\theta\widehat{x}_1(t)|^{2k} \\
&\quad -\mu^{2k}\left|x_2^2(t)-\widehat{x}_2^2(t)\right|^{2k}\mathrm{d}t \\
&\le CE\int_0^r \sum_{i=1}^{3}|x_i(t)-\widehat{x}_i(t)|^{2k}\mathrm{d}t+C\left[E\int_0^r \chi_{u(t)\neq\widehat{u}(t)}(t)\mathrm{d}t\right]^{lk} \\
&\le C\left[E\int_0^r \sum_{i=1}^{3}|x_i(t)-\widehat{x}_i(t)|^{2k}\mathrm{d}t+d(u_\theta(t),\widehat{u_\theta}(t))^{lk}\right]. \qquad (45)
\end{aligned}
$$

From Eqs. (44) and (45), we obtain

$$
\begin{aligned}
&\sum_{i=1}^{3}E\sup_{0\le t\le r}|x_i(t)-x_i'(t)|^{2k} \\
&\le C\left[\int_0^r \sum_{i=1}^{3}E\sup_{0\le t\le r}|x_i(t)-x_i'(t)|^{2k}\mathrm{d}t+d(u(t),u'(t))^{lk}\right].
\end{aligned}
$$

We have this result with Gronwall's inequality.

Then, for $0\le k\le 1$, using the Cauchy–Schwarz inequality, we have

$$
\begin{aligned}
\sum_{i=1}^{3}E\sup_{0\le t\le r}|x_i(t)-\widehat{x}_i(t)|^{2k} &\le \sum_{i=1}^{3}E\sup_{0\le t\le r}\left[|x_i(t)-\widehat{x}_i(t)|^2\right]^k \\
&\le \left[Cd(u(t),\widehat{u}(t))^l\right]^k \\
&\le C^k\left[d(u(t),\widehat{u}(t))^{lk}\right]. \qquad (46)
\end{aligned}
$$

The proof is complete, which is the result of the estimation of the state Eq. (6).

For the result of moment estimation of the adjoint equation, we have the following lemma.

Lemma 5 Let hypotheses (H2) and (H3) hold. For any $0<l<1$ and $1<k<2$ satisfying $(1+l)k<2$, there exists a positive imprecise parameter $C=C(l,k)$ such that for any $u(t),\widehat{u}(t)\in U$ along with the corresponding trajectory $x(t),\widehat{x}(t)$ and the solution $p(t),q(t),r(v),(\widehat{p}(t),\widehat{q}(t),\widehat{r}(v))$ of the adjoint equation, we have

$$
\begin{aligned}
&\sum_{i=1}^{3}E\int_0^T |p_i(t)-\widehat{p}_i(t)|^k\mathrm{d}t+\sum_{i=1}^{2}E\int_0^T |q_i(t)-\widehat{q}_i(t)|^k\mathrm{d}t \\
&+\sum_{i=1}^{2}E\int_Y |r_i(t)-\widehat{r}_i(t)|^k\mathrm{d}t\le Cd(u(t),\widehat{u}(t))^{lk/2}.
\end{aligned}
$$

It is shown that the lemma can be proven by the similar method of lemma 4.2 in Ref. [32].

Then we discuss the necessary conditions for near-optimal controls.

Lemma 6 Let hypotheses (H1) and (H2) hold. $(p^\varepsilon(t),q^\varepsilon(t),r^\varepsilon(v))$ is the solution of the adjoint Eq. (10) under the control $u^\varepsilon(t)$. For any $l\in[0,1)$, $\varepsilon>0$ and any ε-optional pair $(x^\varepsilon(t),u^\varepsilon(t))$, there exists a positive impressive parameter C such that

$$
\begin{aligned}
&\min_{u(t)\in U_{\mathrm{ad}}[0,T]}E\int_0^T\left(\frac{\beta_2 x_1^\varepsilon(t)x_2^{2\varepsilon}(t)m}{((1-u_m)m+x_2^\varepsilon(t))^2}(p_1^\varepsilon(t)-p_2^\varepsilon(t))\right. \\
&\left.\times u_m+\theta x_1^\varepsilon(t)(p_3^\varepsilon(t)-p_1^\varepsilon(t))u_\theta+B_2(u_m^2+u_\theta^2)\right)\mathrm{d}t \\
&\ge E\int_0^T\left(\frac{\beta_2 x_1^\varepsilon(t)x_2^{2\varepsilon}(t)m}{((1-u_m^\varepsilon)m+x_2^\varepsilon(t))^2}(p_1^\varepsilon(t)-p_2^\varepsilon(t))\right. \\
&\left.\times u_m^\varepsilon+\theta x_1^\varepsilon(t)(p_3^\varepsilon(t)-p_1^\varepsilon(t))u_\theta^\varepsilon+B_2(u_m^{2\varepsilon}+u_\theta^{2\varepsilon})\right)\mathrm{d}t
\end{aligned}
$$

$$-C\varepsilon^{l/3}. \tag{47}$$

Proof The key step in the proof is to figure out that $H_{u(t)}(t,x^{\varepsilon}(t),u^{\varepsilon}(t),p^{\varepsilon}(t),q^{\varepsilon}(t),r^{\varepsilon}(v))$ is very small and estimate it with ε, defining a new metric d as follows:

$$d(u^{\varepsilon}(t),\widetilde{u^{\varepsilon}}(t)) \le \varepsilon^{2/3}, \quad \varepsilon > 0, \tag{48}$$

$$\widetilde{J}(0,x_0;\widetilde{u^{\varepsilon}}(t)) \le \widetilde{J}(0,x_0;u(t)), \quad \forall u(t) \in U_{\mathrm{ad}}[0,T], \tag{49}$$

where the cost function is

$$\widetilde{J}(0,x_0;u(t)) \le J(0,x_0;u(t)) + \varepsilon^{13} d(u(t),\widetilde{u^{\varepsilon}}(t)). \tag{50}$$

This means $(\widetilde{x^{\varepsilon}}(t),\widetilde{u^{\varepsilon}}(t))$ is optimal for system (6) with the above cost function. Then for a.e. $t \in [0,1]$, $l \in [0,1)$, a stochastic maximum principle yields

$$\begin{aligned} & H(t,\widetilde{x^{\varepsilon}}(t),\widetilde{u^{\varepsilon}}(t),\widetilde{p^{\varepsilon}}(t),\widetilde{q^{\varepsilon}}(t),\widetilde{r^{\varepsilon}}(t)) \\ = & \min_{u(t)\in U_{\mathrm{ad}}} \Big[H(t,\widetilde{x^{\varepsilon}}(t),\widetilde{u^{\varepsilon}}(t),\widetilde{p^{\varepsilon}}(t),\widetilde{q^{\varepsilon}}(t),\widetilde{r^{\varepsilon}}(t)) \\ & + \varepsilon^{l/3}\left(|u(t)-\widetilde{u^{\varepsilon}}(t)|\right)\Big], \end{aligned} \tag{51}$$

where $\widetilde{p^{\varepsilon}}(t)$, $\widetilde{q^{\varepsilon}}(t)$, $\widetilde{r^{\varepsilon}}(t)$ is the solution of Eq. (10) with $\widetilde{u^{\varepsilon}}(t)$.

Let

$$\begin{aligned} M_1(t) &= H\left(t,\widetilde{x^{\varepsilon}}(t),\widetilde{u^{\varepsilon}}(t),\widetilde{p^{\varepsilon}}(t),\widetilde{q^{\varepsilon}}(t),\widetilde{r^{\varepsilon}}(t)\right) \\ &\quad -H\left(t,x^{\varepsilon}(t),u^{\varepsilon}(t),p^{\varepsilon}(t),q^{\varepsilon}(t),r^{\varepsilon}(v)\right), \\ M_2(t) &= H\left(t,\widetilde{x^{\varepsilon}}(t),\widetilde{u^{\varepsilon}}(t),\widetilde{p^{\varepsilon}}(t),\widetilde{q^{\varepsilon}}(t),\widetilde{r^{\varepsilon}}(t)\right) \\ &\quad -H\left(t,x^{\varepsilon}(t),u(t),p^{\varepsilon}(t),q^{\varepsilon}(t),r^{\varepsilon}(v)\right). \end{aligned}$$

According to Eq. (51) and elementary inequality, there exists a constant C such that

$$\begin{aligned} & E\int_0^T H\left(t,x^{\varepsilon}(t),u^{\varepsilon}(t),p^{\varepsilon}(t),q^{\varepsilon}(t),r^{\varepsilon}(v)\right)\mathrm{d}t \\ \le & E\int_0^T H\left(t,\widetilde{x^{\varepsilon}}(t),\widetilde{u^{\varepsilon}}(t),\widetilde{p^{\varepsilon}}(t),\widetilde{q^{\varepsilon}}(t),\widetilde{r^{\varepsilon}}(t)\right)\mathrm{d}t \\ & + E\int_0^T |M_1(t)|\,\mathrm{d}t \\ = & E\int_0^T \min_{u(t)\in U_{\mathrm{ad}}} \Big[H\left(t,\widetilde{x^{\varepsilon}}(t),\widetilde{u^{\varepsilon}}(t),\widetilde{p^{\varepsilon}}(t),\widetilde{q^{\varepsilon}}(t),\widetilde{r^{\varepsilon}}(t)\right) \\ & + \varepsilon^{l/3}\left(|u(t)-\widetilde{u^{\varepsilon}}(t)|\right)\Big]\mathrm{d}t + E\int_0^T |M_1(t)|\,\mathrm{d}t \\ \le & E\int_0^T \min_{u(t)\in U_{\mathrm{ad}}} [H\left(t,x^{\varepsilon}(t),u^{\varepsilon}(t),p^{\varepsilon}(t),q^{\varepsilon}(t),r^{\varepsilon}(v)\right)]\,\mathrm{d}t \\ & + E\int_0^T |M_1(t)|\,\mathrm{d}t + C\varepsilon^{l/3} \\ \le & E\int_0^T \min_{u(t)\in U_{\mathrm{ad}}} [H\left(t,x^{\varepsilon}(t),u^{\varepsilon}(t),p^{\varepsilon}(t),q^{\varepsilon}(t),r^{\varepsilon}(v)\right)]\mathrm{d}t \\ & + E\int_0^T \min_{u(t)\in U_{\mathrm{ad}}} |M_2(t)|\,\mathrm{d}t + E\int_0^T |M_1(t)|\,\mathrm{d}t + C\varepsilon^{l/3} \\ \le & \min_{u(t)\in U_{\mathrm{ad}}} E\int_0^T [H\left(t,x^{\varepsilon}(t),u^{\varepsilon}(t),p^{\varepsilon}(t),q^{\varepsilon}(t),r^{\varepsilon}(v)\right)]\,\mathrm{d}t \\ & + \min_{u(t)\in U_{\mathrm{ad}}} E\int_0^T |M_2(t)|\,\mathrm{d}t + E\int_0^T |M_1(t)|\,\mathrm{d}t + C\varepsilon^{l/3}. \end{aligned}$$

Then, we can get

$$\begin{aligned} & \min_{u(t)\in U_{\mathrm{ad}}[0,T]} E\int_0^T \left(\frac{\beta_2 x_1^{\varepsilon}(t) x_2^{2\varepsilon}(t) m}{\left((1-u_m)m + x_2^{\varepsilon}(t)\right)^2} \left(p_1^{\varepsilon}(t) - p_2^{\varepsilon}(t)\right) \right. \\ & \left. \times u_m + \theta x_1^{\varepsilon}(t)(p_3^{\varepsilon}(t) - p_1^{\varepsilon}(t))u_\theta + B_2\left(u_m^2 + u_\theta^2\right) \right)\mathrm{d}t \\ \ge & E\int_0^T \left(\frac{\beta_2 x_1^{\varepsilon}(t) x_2^{2\varepsilon}(t) m}{\left((1-u_m^{\varepsilon})m + x_2^{\varepsilon}(t)\right)^2} \left(p_1^{\varepsilon}(t) - p_2^{\varepsilon}(t)\right) \right. \\ & \left. \times u_m^{\varepsilon} + \theta x_1^{\varepsilon}(t)\left(p_3^{\varepsilon}(t) - p_1^{\varepsilon}(t)\right)u_\theta^{\varepsilon} + B_2\left(u_m^{2\varepsilon} + u_\theta^{2\varepsilon}\right) \right)\mathrm{d}t \\ & - C\varepsilon^{l/3}. \end{aligned} \tag{52}$$

4. Numerical simulations

In this section, we use MATLAB to perform numerical simulations and use these numerical examples to illustrate the impact of popular science education, media reports and Lévy noise on the dynamics of rumor spreading. We set the parameter values as follows: $\Lambda = 0.25$, $\mu = 0.15$, $\beta_1 = 0.6$, $\beta_2 = 0.4$, $\lambda = 0.2$, $\theta = 0.2$, $\sigma = 0.009$.

According to Milstein's method,[37] we can get the corresponding discretization equation of the state Eqs. (4) and (10) as follows:

$$\begin{aligned} I_{i+1} &= I_i + \left[\Lambda - \left(\beta_1 + \eta a - \beta_2 \frac{S_i}{(1-u_m)m + S_i}\right) I_i S_i - (1+u_\theta)\theta I_i - \mu I_i\right]\Delta t - \eta\sigma I_i S_i \sqrt{\Delta t}\xi_i \\ &\quad - \eta\sigma I_i S_i(\xi_i^2 - 1)\Delta t - \sum_{j=1}^{\infty} V_j \eta v I_i S_i (a(i))^{-1/d}, \\ S_{i+1} &= S_i + \left[\left(\beta_1 + \eta a - \beta_2 \frac{S_i}{(1-u_m)m + S_i}\right) I_i S_i - \lambda S_i(S_i + R_i) - \mu S_i\right]\Delta t + \eta\sigma I_i S_i \sqrt{\Delta t}\xi_i \\ &\quad + \eta\sigma I_i S_i(\xi_i^2 - 1)\Delta t + \sum_{j=1}^{\infty} V_j \eta v I_i S_i (a(i))^{-1/d}, \\ R_{i+1} &= R_i + [\lambda S_i(S_i + R_i) + (1+u_\theta)\theta I_i - \mu R_i]\Delta t, \end{aligned} \tag{53}$$

$$
\begin{aligned}
p_{1_i} &= p_{1_{i+1}} - \Bigg[\left(\left(\beta_1 + \eta a - \beta_2 \frac{S_{i+1}}{(1-u_m)m + S_{i+1}} \right) S_{i+1} + (1+u_\theta)\theta + \mu \right) p_{1_{i+1}} \\
&\quad - \left(\beta_1 + \eta a - \beta_2 \frac{S_{i+1}}{(1-u_m)m + S_{i+1}} \right) S_{i+1} p_{2_{i+1}} - (1+u_\theta)\theta p_{3_{i+1}} + \eta\sigma S_{i+1} q_{1_{i+1}} - \eta\sigma S_{i+1} q_{2_{i+1}} \\
&\quad + \sum_{j=1}^{\infty} V_j \eta v S_{i+1} (a(i+1))^{-1/d} r_{1_{i+1}} - \sum_{j=1}^{\infty} V_j \eta v S_{i+1} (a(i+1))^{-1/d} r_{2_{i+1}} - B_0 \Bigg] \Delta t \\
&\quad - q_{1_{i+1}} \sqrt{\Delta t}\, \xi_{i+1} - \frac{q_{1_{i+1}}^2}{2} (\xi_{i+1}^2 - 1)\Delta t + \sum_{j=1}^{\infty} V_j (a(i+1))^{-1/d} r_{1_{i+1}}, \\
p_{2_i} &= p_{2_{i+1}} - \Bigg[\left(\beta_1 + \eta a - \frac{\beta_2 S_{i+1}(2(1-u_m)m + S_{i+1})}{((1-u_m)m + S_{i+1})^2} \right) I_{i+1} p_{1_{i+1}} - \Bigg(\left(\beta_1 + \eta a - \frac{\beta_2 S_{i+1}(2(1-u_m)m + S_{i+1})}{((1-u_m)m + S_{i+1})^2} \right) I_{i+1} \\
&\quad - \lambda(2S_{i+1} + R_{i+1}) - \mu \Bigg) p_{2_{i+1}} - \lambda(2S_{i+1} + R_{i+1}) p_{3_{i+1}} + \eta\sigma I_{i+1} q_{1_{i+1}} - \eta\sigma I_{i+1} q_{2_{i+1}} + \sum_{j=1}^{\infty} V_j \eta v I_{i+1} (a(i+1))^{-1/d} r_{1_{i+1}} \\
&\quad - \sum_{j=1}^{\infty} V_j \eta v I_{i+1} (a(i+1))^{-1/d} r_{2_{i+1}} - B_1 \Bigg] \Delta t - q_{2_{i+1}} \sqrt{\Delta t}\, \xi_{i+1} - \frac{q_{2_{i+1}}^2}{2} (\xi_{i+1}^2 - 1)\Delta t + \sum_{j=1}^{\infty} V_j (a(i+1))^{-1/d} r_{2_{i+1}}, \\
p_{3_i} &= p_{3_{i+1}} - \left[\lambda S_{i+1} p_{2_{i+1}} + (\lambda S_{i+1} - \mu) p_{3_{i+1}} \right] \Delta t.
\end{aligned} \tag{54}
$$

The intensity of Lévy noise can affect the rumor propagation process. Figure 2 depicts the influence of Lévy noise on the rumor propagation process. $r(v)$ stands for response intensities. As shown in the figure, the trend chart of the number of spreaders over time is different under different response intensities $r(v)$. The greater the value of $r(v)$, the faster the value of $S(t)$ decreases. Therefore, we can reduce the ultimate spread scale of rumors by strengthening external intervention. For example, when rumors spread wildly and even cause bad social impact, the government should promptly popularize relevant knowledge about rumors to guide people to correctly view the content of rumors, and relevant media should also increase their reporting efforts. In this way, the number of spreaders of rumors can be reduced by strengthening external intervention.

We also simulated the time evolution diagram of the number of people in different states during rumor propagation under four different control conditions: (1) $u_m = 0, u_\theta = 0$ represents the uncontrolled situation; (2) $u_m = 0, u_\theta \neq 0$ represents the situation where only popular science education control is added; (3) $u_m \neq 0, u_\theta = 0$ represents the situation where only media coverage control is added; and (4) $u_m \neq 0, u_\theta \neq 0$ represents the situation where both media coverage and popular science education control are added. As can be seen from Fig. 4, when media coverage and popular science education control are added at the same time, the peak value and number of spreaders drop sharply, and the influence scope of rumors becomes smaller. Compared with the other three conditions, the rumor propagation process is easier to control, indicating that this situation is nearoptimal control. For example, in the early days of the COVID-19 outbreak, many people were not familiar with the specific case of novel coronavirus. It was rumored that "symptoms such as fever and cough were signs of infection with COVID-19". The rampant propagation of this rumor was bound to cause public panic. In order to prevent this crisis, the government provided the latest research on COVID-19 in official news and educated the public through multiple media channels. Popular science education allowed people to learn more about the symptoms of COVID-19 infection, which directly reduced the number of the ignorant and enabled them to convert directly to the immune. Media coverage directly reduced the number of communicators through targeted refutation of the rumor.

In addition, it can be seen from Figs. 3–5 that the dynamic optimal control is effective. Both the control of popular science education and the control of media coverage decline faster than the situation without control. Under the control of government and media intervention, the number of spreaders declines faster than in the absence of control. For example, during the COVID-19 outbreak, the public listened to the rumor that "Shuanhuanglian could prevent the novel coronavirus", so they rushed to buy Shuanhuanglian, causing chaos in the market. The government quickly educated the public about the novel coronavirus treatment, and many media outlets debunked the rumor. The double control increased the number of people immune to the rumor. Taking a closer analysis, we can see that when the media coverage control strategy is added, the control effect is good but not great. When the government science education control strategy is added, the control effect is significantly improved. This simulation result shows that the government science education control strategy is more effective than the media coverage control strategy.

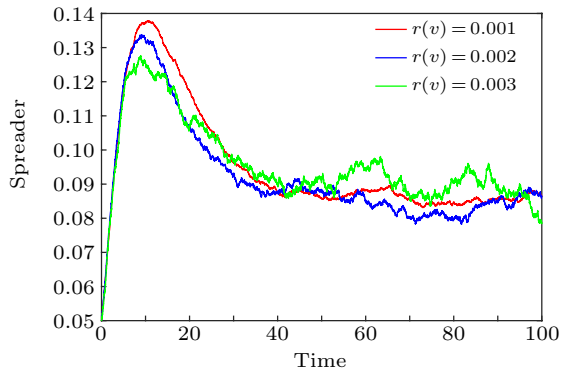


Fig. 2. Path of $S(t)$ under different response intensities $r(v)$.

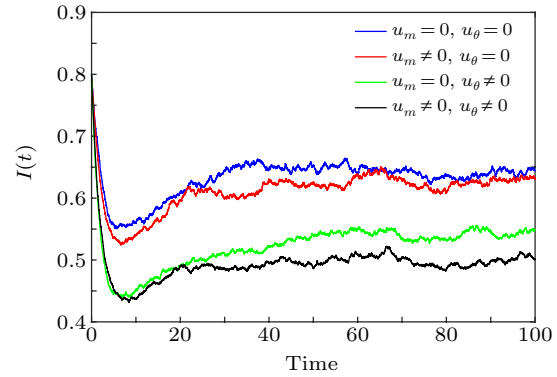


Fig. 3. Path of $I(t)$ for the stochastic ISR model with initial $(I_0,S_0,R_0)=(0.8,0.2,0.05)$.

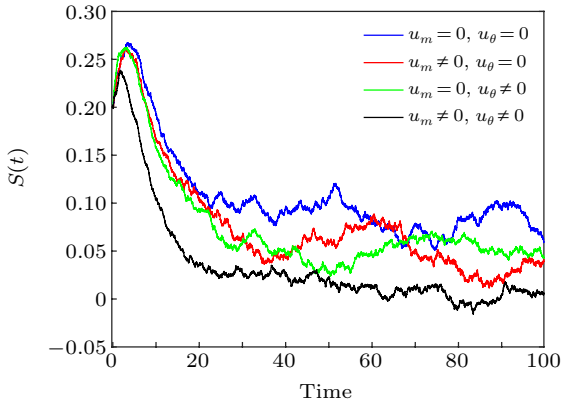


Fig. 4. Path of $S(t)$ for the stochastic ISR model with initial $(I_0,S_0,R_0)=(0.8,0.2,0.05)$.

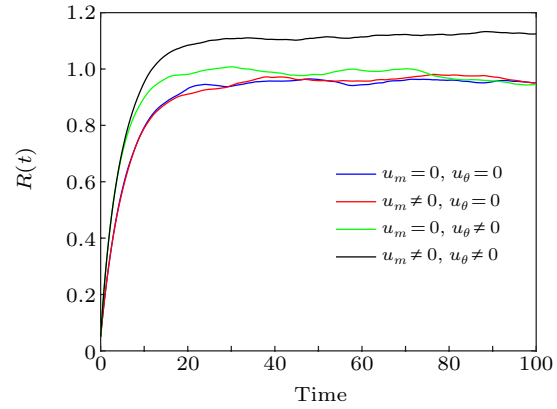


Fig. 5. Path of $R(t)$ for the stochastic ISR model with initial $(I_0,S_0,R_0)=(0.8,0.2,0.05)$.

As for the near-optimal control, it can be seen from Fig. 6 that solutions p_1, p_2, and p_3 of adjoint Eq. (10) converge to zero in the end, indicating that the objective function reaches the minimum value.

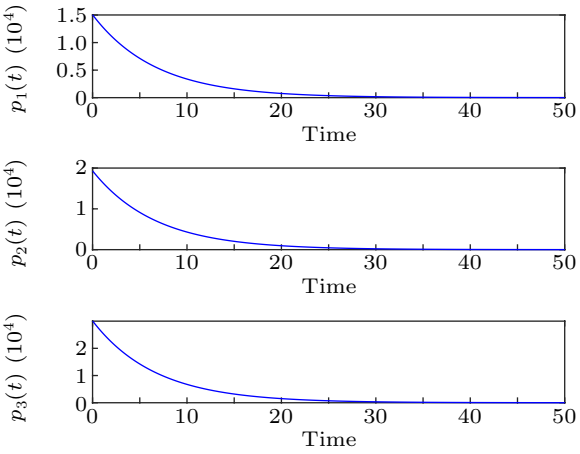


Fig. 6. Path of p_1, p_2, p_3 for the stochastic ISR model with initial $(I_0,S_0,R_0)=(0.8,0.2,0.05)$.

5. Conclusion

This paper studies the problem of near-optimal control of rumor propagation. We first construct a stochastic model that includes media reports and Lévy noise. Then, in order to minimize the impact of rumor propagation and control costs, we introduce two control strategies: popular science education and media coverage. Using the Hamiltonian function and Pontryagin maximum principle, the necessary and sufficient conditions for the model to be close to the optimal control are obtained. Finally, we verify the effects of popular science education, media report control and noise control through numerical simulations. The results show that popular science education and media coverage can effectively control the spread of rumors, and these two control strategies can reduce the cost of rumor dispelling to the lowest level. At the same time, Lévy noise will also affect the process of rumor propagation. The louder the noise, the faster the attenuation speed of the spreader. In this paper, we assume that the parameters are precisely known, but may not be true due to various uncertainties. The influence of uncertain parameters and time delay on the rumor control process is worthy of further study.

Acknowledgments

Project supported by the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning, and the Project for the Natural Science Foundation of Shanghai, China (Grant No. 21ZR1444100), and the Project for the National Natural Science Foundation of China (Grant Nos. 72174121, 71774111, 71871144, and 71804047).

References

[1] Chen X L and Wang N 2020 *Sci. Rep.* **10** 5887
[2] Huo L A, Chen S J and Zhao L J 2021 *Phys. A* **571** 125828
[3] Vardanjania H M, Heydaria S T, Dowranb B and Pasalarc M 2020 *Integr. Med. Res.* **9** 100482
[4] Daley D J and Kendall D G 1965 *IMA J. Appl. Math.* **1** 42
[5] Maki D P and Thompson M 1973 *Mathematical models and applications: With emphasis on the social, life, and management sciences* (New Jersey: Prentice-Hall) 9780135616703
[6] Huo L A, Wang L and Zhao X M 2019 *Phys. A* **517** 551
[7] Kawachi K, Seki M, Yoshida H, Otake Y, Warashina K and Ueda H 2008 *J. Theor. Biol.* **253** 55
[8] Laarabi H, Abta A, Rachik M and Bouyaghroumni J 2016 *Differ. Equ. Dyn. Syst.* **24** 407
[9] Zanette D H 2002 *Phys. Rev. E* **65** 041908
[10] Albert R, Jeong H and Baraba'si A L 2000 *Nature* **406** 378
[11] Doer B, Fouz M and Friedrich T 2012 *Commun. ACM* **55** 70
[12] Liu Y, Diao S M, Zhu Y X and Liu Q 2016 *Phys. A* **461** 543
[13] Dong S, Deng Y B and Huang Y C 2017 *Commun. Theor. Phys.* **68** 545
[14] Zhao D L, Sun J B, Tan Y J and Wu J H 2018 *Phys. A* **512** 1019
[15] Zhu L H and Huang X Y 2020 *Commun. Theor. Phys.* **72** 015002
[16] Jia F J, Lv G Y and Zou G 2018 *Math. Methods Appl. Sci.* **41** 1661
[17] Yu J Y and Liu M 2017 *Phys. A* **482** 14
[18] Dauhoo M Z, Juggurnath D and Adam N R B 2016 *Math. Soc. Sci.* **82** 85
[19] Jia F G and Lv G Y 2018 *Phys. A* **490** 613
[20] Jia F G, Lv G Y, Wang S F and Zou G A 2018 *J. Stat. Mech. Theory Exp.* **2018** 023502
[21] Jain A, Dhar J and Gupta V 2019 *Phys. A* **519** 227
[22] El Bhih A, Ghazzali R, Ben Rhila S, Rachik M and Laaroussi A E 2020 *Discrete Dyn. Nat. Soc.* **2020** 1
[23] Chen G H 2019 *Phys. A* **522** 88
[24] Hui, H W, Zhou, C C, Lu X and Li J R 2020 *Nonlinear Dyn.* **101** 1933
[25] Wang Q S and Song P J 2015 *J. Organ. Comput. Electron. Commer.* **25** 98
[26] Huo L A and Cheng Y Y 2018 *Discrete Dyn. Nat. Soc.* **2018** 4137129
[27] Nekovee M, Moreno Y, Bianconi G and Marsili M 2007 *Phys. A* **374** 457
[28] Huo L A and Dong Y F 2020 *Math. Methods Appl. Sci.* **43** 6903
[29] Applebaum D and Wang W 2004 *Lévy Processes and Stochastic Calculus* (New York: Cambridge University Press) 0521738652
[30] Mu X J and Zhang Q M 2019 *Math. Methods Appl. Sci.* **42** 767
[31] Mu X J, Zhang Q M and Rong L B 2019 *J. Franklin Inst.* **356** 11385
[32] Guo W J, Zhang Q M and Rong L B 2018 *Inf. Sci.* **467** 670
[33] Ekeland I 1979 *Bull. New. Ser. Am. Math. Soc.* **1** 443
[34] Menaldi J L 2001 *SIAM Rev.* **43** 727
[35] Zhou Y L, Yuan S L and Zhao D L 2016 *Appl. Math. Comput.* **275** 255
[36] Mao X R 2007 *Stochastic Differential Equations and Applications 2nd ed.* (Chichester: Horwood Publishing Ltd) 9781904275343
[37] Higham D J 2001 *SIAM Rev.* **43** 525