Epidemic control
via stochastic optimal control

Andrew Lesniewski
Department of Mathematics
Baruch College
55 Lexington Avenue
New York, NY 10010

Draft of April 15, 2020
First draft: April 14, 2020

Abstract

We study the problem of optimal control of the stochastic SIR model. Models of this type are used in mathematical epidemiology to capture the time evolution of highly infectious diseases such as COVID-19. Our approach relies on reformulating the Hamilton-Jacobi-Bellman equation as a stochastic minimum principle. This results in a system of forward backward stochastic differential equations, which is amenable to numerical solution via Monte Carlo simulations. We present a number of numerical solutions of the system under a variety of scenarios.

1   Introduction

The classic SIR model, originally proposed in [15], is the standard tool of mathematical epidemiology [4] for quantitative analysis of the spread of an epidemic. It describes the state of the affected population in terms of three state variables, traditionally denoted by S, I, and R:

  (i) $0 \leq S \leq 1$, the fraction of individuals who are *susceptible* to the disease.

 (ii) $0 \leq I \leq 1$, the fraction of individuals who are *infected* with the disease.

(iii) $0 \leq R \leq 1$, the fraction of individuals who have been *removed* and are immune to the disease. Note that, in this simplified model, R includes the individuals who have left the population through mortality.

The model dynamics is given as the following dynamical system:

$$\begin{aligned}
\frac{\dot{S}(t)}{S(t)} &= -\beta I(t), \\
\frac{\dot{I}(t)}{I(t)} &= \beta S(t) - \gamma, \\
\frac{\dot{R}(t)}{I(t)} &= \gamma,
\end{aligned} \tag{1}$$

where the constant parameters $\beta > 0$ and $\gamma > 0$ are called the *infection rate* and *recovery rate*, respectively. The evolution above is subject to the initial conditions:

$$\begin{aligned}
S(0) &= S_0, \\
I(0) &= I_0, \\
R(0) &= 1 - S_0 - I_0.
\end{aligned} \tag{2}$$

Notice that this dynamics obeys the conservation law

$$S(t) + I(t) + R(t) = 1, \tag{3}$$

consistent with the assumption that the variables S, I, and R represent population fractions. This means that the variable R is, in a way, redundant, as its current value does not affect the dynamics of S and I, and it can be computed in a straightforward manner from (3). It is thus natural to consider the two dimensional system defined by S, I only.

Eventually every epidemic comes to a natural halt, but its impact on the population may be very serious. The overall objective of epidemic control is to slow down the spread of infection in a way that it does not overwhelm the healthcare system and to allow the the economy to function. All of this should be done within the limits of available resources.

In this note we study the problem of optimal control of an epidemic modeled by means of a stochastic extension of the SIR model (see Section 2 for definition). We assume that the controlling agent ("government") has the ability to impact the spread of the disease through one of the policies (or the combination of both):

(i) Vaccination of susceptible individuals, with rate v. This makes the fraction vS of the susceptible population immune to the disease.

(ii) Isolation of infected individuals, with rate i. This removes the fraction iI of the infected population and prevents it from spreading the disease.

The controlled dynamics of the SIR model reads [1], [13]:

$$\begin{aligned}
\frac{\dot{S}(t)}{S(t)} &= -\beta I(t) - v(t), \\
\frac{\dot{I}(t)}{I(t)} &= \beta S(t) - \gamma - i(t), \\
\dot{R}(t) &= v(t)S(t) + (\gamma + i(t))I(t).
\end{aligned} \tag{4}$$

For efficiency, we will be using the notation $X_1 = S$ and $X_2 = I$ throughout the remainder of the paper.

Mathematical models are only as good as (i) their analytic specifications, and (ii) the data that fuel them. During initial phases of an epidemic the available data tend to be of limited usefulness: because of the lack of reliable large scale testing, it is not really known what fractions of the population fall into the different compartments S, I, and R. This may lead to a panic reaction of the population and a chaotic and economically devastating public health response to the epidemic. In the absence of an effective vaccine (which would allow to immunize a portion of the population) the optimal policy is to isolate at least a significant fraction of the infected individuals so that the basic reproduction ration R_0 can be brought significantly below one. Since we lack the knowledge who is infected and who is not, the public health response is to try to isolate everyone, whether susceptible, infected or immune.

These circumstances impose a serious limitation on practical applicability of the approach to optimal epidemic control discussed in this paper, as well as other quantitative approaches. Unless the inputs to the model (β, γ, and the current snapshots of S and I) can reliably be estimated from the available data, the model's output is unreliable[1].

The paper is organized as follows. In Section 2 we review the formulation of the continuous time stochastic SIR model. The optimal control problem for the stochastic SIR model is formulated in Section 3. The optimal control problem is recast as the stochastic minimum principle problem and formulated in terms of a system of backward forward stochastic differential equations (FBSDE). We present an algorithm for solving this system in Section 5. This section presents also the results of a number of numerical experiments involving the three mitigation policies within various cost function regimes.

---

[1]"This is indeed a mystery," Watson remarked. "What do you imagine that it means?"
"I have no data yet," Holmes replied. "It is a capital mistake to theorise before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts." [8]

2   Stochastic SIR model

We consider a continuous time stochastic extension of the deterministic SIR model, see [3] and references therein. Let W_t denote the standard Brownian motion, and let $\dot{W}_t$ denote the white noise process. We assume that the infection rate β is subject to random shocks, rather than being constant, namely

$$\beta_t = \beta + \sigma \dot{W}_t, \tag{5}$$

while the recovery rate γ remains constant. Here $\sigma > 0$ is a constant volatility parameter. This leads to the following system of stochastic differential equations (SDE), driven by W_t

$$\begin{aligned}
\frac{dX_{1,t}}{X_{1,t}} &= -\beta X_{2,t} dt - \sigma X_{2,t} dW_t, \\
\frac{dX_{2,t}}{X_{2,t}} &= (\beta X_{1,t} - \gamma)dt + \sigma X_{1,t} dW_t,
\end{aligned} \tag{6}$$

with the initial conditions

$$\begin{aligned}
X_{1,0} &= S_0, \\
X_{2,0} &= I_0.
\end{aligned} \tag{7}$$

The third component of the process, $X_3 = R$, follows the dynamics

$$\frac{dX_{3,t}}{X_{2,t}} = \gamma dt, \tag{8}$$

which implies that the conservation law

$$X_t + Y_t + Z_t = 1 \tag{9}$$

continues to hold in the stochastic model.

Notice that (6) is a one factor model, driven by a single source of randomness. There is a natural two-factor version of the stochastic SIR model [17], in which also the recovery rate γ is allowed to be subject to white noise shocks. For simplicity, our analysis will focus on the one-factor model (6).

3   HJB equation

We frame the problem of epidemic control as a stochastic control problem [2]. We denote by $u = (u_1, u_2) \equiv (v, i)$ the vaccination and isolation controls, and we denote the controlled process by X_t^u. Generalizing the deterministic specification (4) to the stochastic case, we assume that the dynamics of X_t^u is given by:

$$\begin{aligned}
\frac{dX_{1,t}^u}{X_{1,t}^u} &= -(\beta X_{2,t}^u + u_{1,t})dt - \sigma X_{2,t}^u dW_t, \\
\frac{dX_{2,t}^u}{X_{2,t}^u} &= (\beta X_{1,t}^u - \gamma - u_{2,t})dt + \sigma X_{1,t}^u dW_t,
\end{aligned} \tag{10}$$

subject to initial conditions (7).

Two special cases of the controlled process are of interest. If a vaccine against the disease is unavailable, we set $u_1 = 0$ in the equation above, which yields the following controlled process:

$$
\begin{aligned}
\frac{dX_{1,t}^u}{X_{1,t}^u} &= -\beta X_{2,t}^u dt - \sigma X_{2,t}^u dW_t, \\
\frac{dX_{2,t}^u}{X_{2,t}^u} &= (\beta X_{1,t}^u - \gamma - u_{2,t})dt + \sigma X_{1,t}^u dW_t.
\end{aligned}
\tag{11}
$$

We will refer to this policy as an isolation policy.

Similarly, we can consider a vaccination policy, for which $u_2 = 0$. In this case the controlled dynamics reads

$$
\begin{aligned}
\frac{dX_{1,t}^u}{X_{1,t}^u} &= -(\beta X_{2,t}^u + u_{1,t})dt - \sigma X_{2,t}^u dW_t, \\
\frac{dX_{2,t}^u}{X_{2,t}^u} &= (\beta X_{1,t}^u - \gamma)dt + \sigma X_{1,t}^u dW_t.
\end{aligned}
\tag{12}
$$

We assume a finite time horizon $T < \infty$. The controlling agent's objective is to minimize a running cost function $c(X_t, u_t)$ and the terminal value function $G(X_T)$. In other words, we are seeking a policy u^* such that

$$
u^* = \arg\min_u \mathsf{E}\Big(\int_0^T c(X_t^u, u_t)dt + G(X_T) \Big).
\tag{13}
$$

We consider the following cost function:

$$
c(x, u) = c_1(x_1, u_1) + c_2(x_2, u_2),
\tag{14}
$$

where

$$
c_i(x_i, u_i) = \big(\frac{1}{2} L_i u_i^2 + M_i u_i + N_i \big) x_i,
\tag{15}
$$

for $i = 1, 2$. In other words, the running cost of vaccination is assumed to be proportional to the number of susceptible individuals, while the cost of isolation is assumed to be proportional to the number of infected individuals. The coefficients $L_i > 0, M_i, N_i$ are determined by the overall cost of following the mitigation policy. In particular, they should be selected so that the running cost functions are strictly positive.

As the terminal value function we take the transmission rate of the infection [13], namely

$$
G(x) = \beta x_1 x_2.
\tag{16}
$$

We now invoke stochastic dynamic programming, see eg. [10], [20]. The key element of this approach is the value function $J(t, x, y)$. It is determined by two requirements:

(B1) It satisfies Bellman's principle of optimality,

$$J(t, X_t) = \min_u \mathsf{E}_t\big(J(t + dt, X_{t+dt}) + c(X_t, u_t)dt\big), \tag{17}$$

for all $0 \leq t < T$, where E_t denotes conditional expectation with respect to the information set at time t, and where the minimum is taken over all admissible controls u_t, [10].

(B2) It satisfies the terminal condition,

$$J(T, X_T) = G(X_T). \tag{18}$$

Using Ito's lemma, we verify that these two conditions lead to the following nonlinear partial differential equation for the value function, namely the stochastic Hamilton-Jacobi-Bellman equation:

$$\dot{J} + \min_u \Big\{ -x_1(\beta x_2 + u_1)J_{x_1} + (\beta x_1 - \gamma - u_2)x_2 J_{x_2}$$
$$+ \frac{1}{2}\sigma^2 x_1^2 x_2^2\big(J_{x_1 x_1} + J_{x_1 x_2} - 2J_{x_1 x_2}\big) + c(x, u)\Big\} = 0, \tag{19}$$

subject to the terminal condition

$$J(T, x) = G(x). \tag{20}$$

As the first step towards solving the HJB equation (19), we let $u = u^*$ denote the minimizer of the expression inside the curly parentheses in (19). In other words, u^* satisfies

$$-x_i J_{x_i} + (L_i u_i + M_i)x_i = 0,$$

which leads to the following first order condition on u^*:

$$u_i^* = \frac{1}{L_i}\left(J_{x_i} - M_i\right). \tag{21}$$

Substituting u^* back to the HJB equation yields the equation

$$\dot{J} - \beta x_1 x_2 J_{x_1} + (\beta x_1 - \gamma)x_2 J_{x_2} + \frac{1}{2}\sigma^2 x_1^2 x_2^2\big(J_{x_1 x_1} + J_{x_1 x_2} - 2J_{x_1 x_2}\big)$$
$$\left(-\frac{1}{2}\frac{(J_{x_1} - M_1)^2}{L_1} + N_1\right)x_1 + \left(-\frac{1}{2}\frac{(J_{x_2} - M_2)^2}{L_2} + N_2\right)x_1 = 0. \tag{22}$$

We do not believe that the solution to this equation can be explicitly represented in terms of standard functions. This is a three dimensional partial differential equation, and solving it numerically may pose challenges. Rather than following this path, we shall invoke the stochastic minimum principle and reformulate the problem as a system of FBSDEs. Among the advantages of this approach is that it might be amenable to a deep learning approach via the method advocated in [14].

4 Stochastic minimum principle

The stochastic minimum principle, see [20] and references therein, offers an alternative approach to stochastic optimal control. It is a stochastic version of Pontryagin's minimum principle introduced in the context of deterministic optimal control [21]. It also offers an effective numerical methodology to solving the HJB equation (22) via Monte Carlo simulations. In this approach, the key object is the Hamiltonian function $\mathcal{H} = \mathcal{H}(x, u, y, z)$ of four arguments and a system of stochastic differential equations, both forward and backward, determined by $\mathcal{H}$.

Specifically, for the case of the controlled stochastic SIR model (12), we have $x, u, y, z \in \mathbb{R}^2$, and the Hamiltonian function reads:

$$
\begin{aligned}
\mathcal{H}(x, u, y, z) = & - x_1(\beta x_2 + u_1)y_1 + (\beta x_1 - \gamma - u_2)x_2 y_2 \\
& + \sigma x_1 x_2 (z_2 - z_1) + c(x, u).
\end{aligned}
\tag{23}
$$

We consider the following system of stochastic Hamilton's equations:

$$
\begin{aligned}
dX_t &= \nabla_y \mathcal{H}(X_t, u_t, Y_t, Z_t)dt + \sigma(X_t)dW_t, \\
X_0 &= \begin{pmatrix} S_0 \\ I_0 \end{pmatrix},
\end{aligned}
\tag{24}
$$

and

$$
\begin{aligned}
-dY_t &= \nabla_x \mathcal{H}(X_t, u_t, Y_t, Z_t)dt - Z_t dW_t, \\
Y_T &= \nabla_x G(X_T),
\end{aligned}
\tag{25}
$$

where

$$
\sigma(X_t) = \sigma X_{1,t} X_{2,t} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.
\tag{26}
$$

Equation (24) is merely an alternative way of writing the underlying controlled diffusion process (12), while equation (25) reflects the dynamics of the control variables given the running cost function. Note that while the first of the equations (24) is a standard (forward) SDE, the second one is a backward stochastic differential equation (BSDE), see e.g. [9].

Explicitly, these four equations can be stated as

$$
\begin{aligned}
dX_{1,t} &= -X_{1,t}(\beta X_{2,t} + u_{1,t})dt - \sigma X_{1,t}X_{2,t}dW_t, \\
dX_{2,t} &= (\beta X_{1,t} - \gamma - u_{2,t})X_{2,t}dt + \sigma X_{1,t}X_{2,t}dW_t, \\
-dY_{1,t} &= \big(\beta(Y_{2,t} - Y_{1,t}) + \sigma(Z_{2,t} - Z_{1,t}))X_{2,t} - u_{1,t}Y_{1,t} \\
&\qquad\qquad + c_{x_1}(X_t, u_t)\big)dt - Z_{1,t}dW_t, \\
-dY_{2,t} &= \big(\beta(Y_{2,t} - Y_{1,t}) + \sigma(Z_{2,t} - Z_{1,t}))X_{1,t} - (\gamma + u_{2,t})Y_{2,t} \\
&\qquad\qquad + c_{x_2}(X_t, u_t)\big)dt - Z_{2,t}dW_t,
\end{aligned}
\tag{27}
$$

subject to the boundary conditions in (24).

Let now choose u^* to be the policy such that $X_t^* \equiv X_t^{u^*}$ is a solution to (27) and

$$
\mathcal{H}(X_t^*, u_t^*, Y_t^*, Z_t^*) = \min_u \mathcal{H}(X_t^u, u_t, Y_t^u, Z_t^u),
\tag{28}
$$

Then u^* is an optimal control, i.e. it satisfies the optimality condition (13).

In fact, there is a direct link between the Hamilton function $\mathcal{H}$ and the value function J (also known in classical dynamics as the action function). Namely, we set [20], [22]:

$$J(t, X_t) = \int_0^t Y_s^{\mathrm{T}} dX_s - \mathcal{H}(X_s, u_s, Y_s, Z_s) ds. \tag{29}$$

If u^* is an optimal control and X^* denotes the corresponding optimal diffusion process, then the pair

$$\begin{aligned} Y_t &= \nabla_x J(t, X_t^*), \\ Z_t &= \nabla_x^2 J(t, X_t^*)\, \sigma(X_t^*) \end{aligned} \tag{30}$$

is the solution to the BSDE (25).

Going back to the main line of reasoning, we find that u^* has to satisfy

$$u_i^* = \frac{y_i - M_i}{L_i}\,. \tag{31}$$

From equations (30) we see that, up to a simple linear transformation, Y_t is essentially the optimal policy. Furthermore, the process Z_t can be thought of as the sensitivity of the optimal policy to the underlying process X_t (multiplied by the instantaneous volatility of that process).

Substituting this expression into (27), we find that the optimal process has to follow the dynamics:

$$\begin{aligned} dX_{1,t}^* &= -X_{1,t}^*\big(\beta X_{2,t}^* + \frac{1}{L_1}\,(Y_{1,t}^* - M_1)\big)dt - \sigma X_{1,t}^* X_{2,t}^* dW_t, \\ dX_{2,t}^* &= \big(\beta X_{1,t}^* - \gamma - \frac{1}{L_2}\,(Y_{2,t}^* - M_2)\big)X_{2,t}^* dt + \sigma X_{1,t}^* X_{2,t}^* dW_t, \\ -dY_{1,t}^* &= \big(\beta(Y_{2,t}^* - Y_{1,t}^*) + \sigma(Z_{2,t}^* - Z_{1,t}^*)\big)X_{2,t}^* dt \\ &\qquad\qquad + \big(N_1 - \frac{1}{2L_1}\,(Y_{1,t}^* - M_1)^2\big)dt - Z_{1,t}^* dW_t, \\ -dY_{2,t}^* &= \big(\beta(Y_{2,t}^* - Y_{1,t}^*) + \sigma(Z_{2,t}^* - Z_{1,t}^*)\big)X_{1,t}^* dt \\ &\qquad\qquad + \big(-\gamma Y_{2,t}^* + N_2 - \frac{1}{2L_2}\,(Y_{2,t}^* - M_2)^2\big)dt - Z_{2,t}^* dW_t, \end{aligned} \tag{32}$$

subject to the boundary conditions in (24) and (25). In particular, the isolation only and vaccination only policies are given by the following systems of FBSDEs:

$$\begin{aligned} dX_{1,t}^* &= -\beta X_{1,t}^* X_{2,t}^* dt - \sigma X_{1,t}^* X_{2,t}^* dW_t, \\ dX_{2,t}^* &= \big(\beta X_{1,t}^* - \gamma - \frac{1}{L_2}\,(Y_{2,t}^* - M_2)\big)X_{2,t}^* dt + \sigma X_{1,t}^* X_{2,t}^* dW_t, \\ -dY_{1,t}^* &= \big(\beta(Y_{2,t}^* - Y_{1,t}^*) + \sigma(Z_{2,t}^* - Z_{1,t}^*)\big)X_{2,t}^* dt - Z_{1,t}^* dW_t, \\ -dY_{2,t}^* &= \big(\beta(Y_{2,t}^* - Y_{1,t}^*) + \sigma(Z_{2,t}^* - Z_{1,t}^*)\big)X_{1,t}^* dt \\ &\qquad\qquad + \big(-\gamma Y_{2,t}^* + N_2 - \frac{1}{2L_2}\,(Y_{2,t}^* - M_2)^2\big)dt - Z_{2,t}^* dW_t, \end{aligned} \tag{33}$$

and

$$
\begin{aligned}
dX_{1,t}^* &= -X_{1,t}^*\big(\beta X_{2,t}^* + \frac{1}{L_1}\left(Y_{1,t}^* - M_1\right)\big)dt - \sigma X_{1,t}^* X_{2,t}^* dW_t, \\
dX_{2,t}^* &= \big(\beta X_{1,t}^* - \gamma\big)X_{2,t}^* dt + \sigma X_{1,t}^* X_{2,t}^* dW_t, \\
-dY_{1,t}^* &= \big(\beta(Y_{2,t}^* - Y_{1,t}^*) + \sigma(Z_{2,t}^* - Z_{1,t}^*)\big)X_{2,t}^* dt \\
&\qquad + \big(N_1 - \frac{1}{2L_1}\left(Y_{1,t}^* - M_1\right)^2\big)dt - Z_{1,t}^* dW_t, \\
-dY_{2,t}^* &= \big(\beta(Y_{2,t}^* - Y_{1,t}^*) + \sigma(Z_{2,t}^* - Z_{1,t}^*)\big)X_{1,t}^* dt - \gamma Y_{2,t}^* dt - Z_{2,t}^* dW_t,
\end{aligned}
\tag{34}
$$

respectively.

Even though derived in the context of a meaningful underlying dynamics, there is no *a priori* reason why these equations should have solutions. The drivers of the backward equations above contain terms quadratic in Y, and so the standard existence theorems [9] do not apply. We proceed in the following assuming that these systems do, in fact, have solutions.

5 Numerical solution

In this section we discuss a numerical algorithm for solving the system (32). Applying this methodology in a number of numerical experiments, we present compelling evidence that solutions to (32) exist and are meaningful over a wide range of model parameters.

5.1 Simulating the optimal FBSDE

We start by describing a method for discretization of the basic FBSDE system (32). We notice first that the two dimensional system defined by the state variables S, I is in fact Hamiltonian [16]. Namely, we define new (canonical) variables

$$
\begin{aligned}
q &= -\log S, \\
p &= -\log I,
\end{aligned}
\tag{35}
$$

and set

$$
H(p,q) = -\beta e^{-p} - \gamma q - \beta e^{-q}.
\tag{36}
$$

It is easy to verify that

$$
\frac{d}{dt} H = 0,
\tag{37}
$$

under the dynamics (1). The system (1) can be explicitly written in the Hamilton form

$$
\begin{aligned}
\dot{p} &= -\frac{\partial H}{\partial q}, \\
\dot{q} &= \frac{\partial H}{\partial p}.
\end{aligned}
\tag{38}
$$

Notice also that the Hamiltonian function is separable, i.e. it is of the form $H(p,q) = T(p) + V(q)^2$.

A convenient discretization to (6) can be formulated in terms of the canonical variables $(p_t, q_t) = (-\log X_{2,t}, -\log X_{1,t})$ as follows [17]. Choose the number of steps $\mathfrak{n}$, define the time step $\delta = T/\mathfrak{n}$, set $t_n = n\delta$ for $n = 0, 1, \ldots, \mathfrak{n}$, and use simplified $(p_n, q_n) \equiv (p_{t_n}, q_{t_n})$. This yields the following Euler scheme:

$$
\begin{aligned}
q_{n+1} &= q_n + (\beta\delta + \sigma\Delta W_n)e^{-p_n} + \frac{1}{2}\sigma^2\delta e^{-2p_n}, \\
p_{n+1} &= p_n + \gamma\delta - (\beta\delta + \sigma\Delta W_n)e^{-q_n} + \frac{1}{2}\sigma^2\delta e^{-2q_n},
\end{aligned}
\tag{39}
$$

for $n = 0, 1, \ldots, \mathfrak{n} - 1$. Here, the Brownian motion increments $\Delta W = W_{(n+1)\delta} - W_{n\delta}$ are independent variates drawn from the normal distribution $N(0, \delta)$. At each iteration step also have to to floor q_{n+1} and p_{n+1} at 0, $q_{n+1} = \max(q_{n+1}, 0)$, $p_{n+1} = \max(p_{n+1}, 0)$, so that $e^{-q_n}, e^{-p_n} \leq 1$.

Approximating the backward equations of the system (32) leads to the following backward Euler scheme:

$$
\begin{aligned}
Y_{1,n} &= Y_{1,n+1} + f_1(X_n, Y_n, Z_n)\delta - Z_{1,n}\Delta W_n, \\
Y_{2,n} &= Y_{2,n+1} + f_2(X_n, Y_n, Z_n)\delta - Z_{2,n}\Delta W_n,
\end{aligned}
\tag{40}
$$

where

$$
\begin{aligned}
f_1(x, y, z) &= \big(\beta(y_2 - y_1) + \sigma(z_2 - z_1)\big)x_2 + N_1 - \frac{1}{2L_1}(y_1 - M_1)^2, \\
f_2(x, y, z) &= \big(\beta(y_2 - y_1) + \sigma(z_2 - z_1)\big)x_1 - \gamma y_2 + N_2 - \frac{1}{2L_2}(y_2 - M_2)^2
\end{aligned}
\tag{41}
$$

denote the generators of the two BSDEs. Starting with the terminal condition

$$
Y_{\mathfrak{n}} = \nabla G(X_{\mathfrak{n}}),
\tag{42}
$$

we will move backward in time with computing Y_n and Z_n, for $n = \mathfrak{n} - 1, \ldots, 0$.

Notice that two difficulties arise while doing so: (i) the Y_n's in (40) are not necessarily adapted, and (ii) they depend on Z_n. These two problems can be solved by taking conditional expectations $\mathsf{E}_n(\cdot) = \mathsf{E}(\cdot \mid X_0, X_1, \ldots, X_n)$ on time t_n. This leads to the following condition:

$$
\begin{aligned}
Y_{i,n} &= \mathsf{E}_n(Y_{i,n}) \\
&= \mathsf{E}_n(Y_{i,n+1}) + f_i(X_n, Y_n, Z_n)\delta.
\end{aligned}
$$

This is an implicit scheme, which may slow down the computations. However, we can easily replace it with an explicit scheme of the same order:

$$
Y_{i,n} = \mathsf{E}_n\big(Y_{i,n+1} + f_i(X_n, Y_{i,n+1}, Z_n)\delta\big).
$$

---

[2]There is a practical dimension to these theoretical observations The fact that the SIR model is a Hamiltonian allows us to solve it numerically by means of the extremely efficient Störmer-Verlet method [12], [16]. We will not make use of this method, as it does not seem to directly extend to the stochastic case

In order to determine $Z_{i,n}$, $i = 1, 2$, we multiply (40) by an increment ΔW_n and take conditional expectations. This yields

$$
\begin{aligned}
0 &= \mathsf{E}_n(Y_{i,n}\Delta W_n) \\
&= \mathsf{E}_n(Y_{i,n+1}\Delta W_n) - Z_{i,n}\delta,
\end{aligned}
$$

and hence we obtain the following expression for $Z_{i,n}$:

$$
Z_{i,n} = \frac{1}{\delta}\,\mathsf{E}_n(Y_{i,n+1}\Delta W_n).
$$

In summary, we are led to the following discrete time scheme for solving (40):

$$
\begin{aligned}
Y_{\mathfrak{n}} &= \nabla G(X_{\mathfrak{n}}), \\
Z_{i,n} &= \frac{1}{\delta}\,\mathsf{E}_n(Y_{i,n+1}\Delta W_n), \\
Y_{i,n} &= \mathsf{E}_n\big(Y_{i,n+1} + f_i(X_n, Y_{n+1}, Z_n)\delta\big),
\end{aligned}
\tag{43}
$$

for $n = \mathfrak{n} - 1, \ldots, 0$. Note that simulating this system requires numerical estimation of the conditional expected values $\mathsf{E}_n(\,\cdot\,)$ in the formulas above. We discuss this issue in the following section.

5.2 Computing the conditional expectations

A practical and powerful method of computing the conditional expected values in (43) is the Longstaff-Schwartz regression method originally developed for modeling American options [19]. We use a variant of this method that involves the Hermite polynomials, and that was used for a similar purpose in [18]. This choice is natural, as conditional expectations of Hermite polynomials of a Gaussian random variable lead to simple closed form expressions.

Let $\mathrm{He}_k(x)$, $k = 0, 1, \ldots$, denote the k-th normalized Hermite polynomial corresponding to the standard Gaussian measure $d\mu(x) = (2\pi)^{-1/2}\,e^{-x^2/2}dx$. These functions form an orthonormal basis for the Hilbert space $L^2(\mathbb{R}, d\mu)$.

The key property of $\mathrm{He}_k(x)$ is the following addition formula for $\chi \in [0, 1]$ and $w, x \in \mathbb{R}$:

$$
\mathrm{He}_k(\sqrt{\chi}w + \sqrt{1-\chi}x) = \sum_{j=0}^{k}\binom{k}{j}^{1/2}\chi^{j/2}(1-\chi)^{(k-j)/2}\mathrm{He}_j(w)\mathrm{He}_{k-j}(x). \tag{44}
$$

Consequently, integrating over x with respect to to the measure μ yields the following conditioning rule:

$$
\mathsf{E}\big(\mathrm{He}_k(\sqrt{\chi}w + \sqrt{1-\chi}x)\,|\,w\big) = \chi^{k/2}\mathrm{He}_k(w). \tag{45}
$$

Here, w, x are independent standard normal random variables. We shall use this relation in order to estimate the conditional expected values in (43).

We set $W_{t_i} = \sqrt{t_i}\, w_i$, for $i = 1, \ldots, m$, where w_i is an n-dimensional standard normal random variable. We notice that

$$w_{i+1} = \sqrt{\chi_i}\, w_i + \sqrt{1 - \chi_i}\, X_i, \tag{46}$$

where $\chi_i = t_i/t_{i+1}$, and where X_i is standard normal and independent of w_i. In the following, we shall use this decomposition in conjunction with (45).

Now, we assume the following linear architecture:

$$Y_{i+1} = \sum_{k=0}^{K} g_{k,i+1} \mathrm{He}_k(w_{i+1}), \tag{47}$$

where K is the cutoff value of the order of the Hermite polynomials. This is simply a truncated expansion of the random variable Y_{i+1} in terms of the orthonormal basis $\mathrm{He}_k(w_{i+1})$. The values of the Fourier coefficients are *estimated by means of ordinary least square regression*. Then, as a consequence of the conditioning rule (45),

$$\mathsf{E}_i(Y_{i+1}) = \sum_{k=0}^{K} g_{k,i+1}\, \chi_i^{k/2}\, \mathrm{He}_k(w_i). \tag{48}$$

In other words, conditioning Y_{i+1} on w_i is equivalent to multiplying its Fourier coefficients g_k by the factor $\chi_i^{k/2}$. This allows us to calculate the first term on the right hand side of the third equation in (43). In order to calculate the second term, we substitute the explicit formula for Z_n into the generators (41) and repeat the calculations in (47) and (48) with Y_{n+1} replaced by $Y_{n+1} + f(X_n, Y_n, Z_n)\delta$.

5.3   Numerical experiments

For numerical experiments within the framework developed above, we assume a time horizon of 1 year (365 days), and choose the SIR model parameters as follows:

$$\begin{aligned}
\beta &= 38.0, \\
\gamma &= 11.5, \\
S_0 &= 0.999, \\
I_0 &= 0.001,
\end{aligned} \tag{49}$$

These parameters are purely hypothetical, and they do not arise from the calibration to any actually reported data. The corresponding value of the basic reproduction ratio $R_0 = S_0\beta/\gamma$ is 3.30 and it indicates a highly infectious disease such as COVID-19. We choose the diffusion parameter

$$\sigma = 3.1. \tag{50}$$

A typical scenario generated by this model is graphed in Figure 1.

For solving the FBSDE system (32) we generate $2,000$ scenarios (Monte Carlo paths) using the low variance spectral decomposition method. Each scenario is based
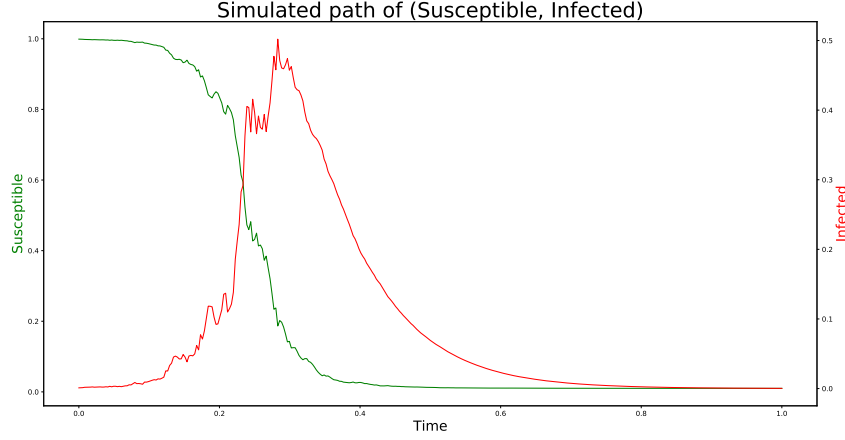
Figure 1: A single Monte Carlo path of the raw process (6).

on daily sampling, i.e. $\mathfrak{n} = 365$. For calculating the conditional expectations in the Longstaff-Schwartz regression we use the cutoff value of $K = 4$. The expectation behind this choice is that we obtain the accuracy of four sigma.

We attempt to construct a numerical solution to (32) through the following iterative procedure. We start by generating initial Monte Carlo paths $X_t^{(0)}$ using the (discretized version of the) raw process (6). For the initial guess of Y_t we take $Y_t^{(0)} = \nabla G(X_t^{(0)})$. Notice that this is *not* at solution to the backward equations in (32), it merely satisfies the terminal condition. After this, we iterate

$$
\begin{aligned}
X_t^{(k+1)} &= \text{ solution of the discretized forward equation with } Y_t = Y_t^{(k)}, \\
Y_t^{(k+1)} &= \text{ solution of the discretized backward equation with } X_t = X_t^{(k+1)},
\end{aligned}
\tag{51}
$$

for $k = 0, 1, \ldots$, until the stopping criterion is satisfied or the maximum number of iterations is reached. For the stopping criterion we choose the condition that the average L^2-norm change of a Monte Carlo path falls below a tolerance threshold of 10^{-8}.

We make various choices of the coefficients L, M, N defining the running cost functions. Depending on the values of these coefficients, the iterative process described above converges to a meaningful solution or it diverges. At this point, it is unclear what choices of the coefficients lead to what outcomes, but it appears that there are well defined basins of convergence in the space of the parameters.

We first consider a high cost policy. Figures 2 - 5 show the graphs of the solutions assuming the following parameters of the quadratic polynomial in the running cost function $c_2(u_2, x_2)$:

$$
\begin{aligned}
L_2 &= 1.0, \\
M_2 &= 0.0, \\
N_2 &= 120.0.
\end{aligned}
\tag{52}
$$

Unlike Figure 1, the curves in all graphs below show the averages of the corresponding quantities over 2,000 Monte Carlo paths.
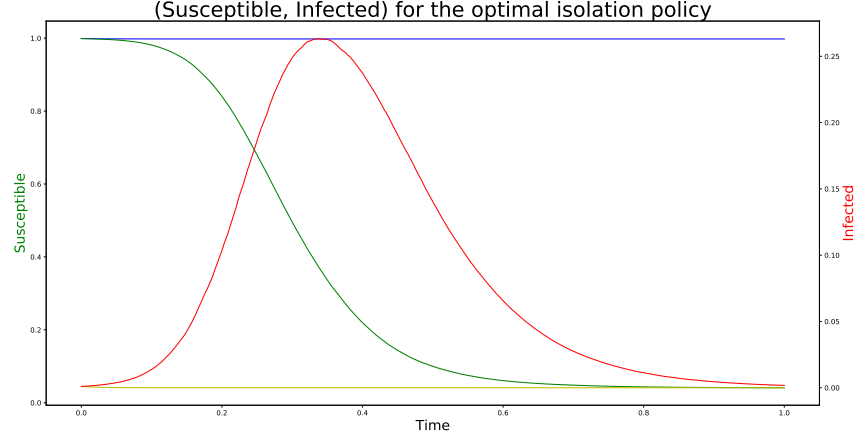
Figure 2: Raw (Susceptible, Infected) versus optimal (Susceptible, Infected) under the isolation policy assuming high cost running cost function. The read and green lines are the average values of the raw Susceptible (primary axis) and Infected (secondary axis) fractions, respectively, while the yellow and blue lines represent the averages of the corresponding optimal values.

Under this running cost function, the optimal policy is to implement a draconian isolation regime, which leads to a rapid drop in infections, while keeping the susceptible fraction of the population at a very high level.

On the other hand, Figures 6 - 9 show the plots of the solutions assuming a low cost policy, with the following parameters of the quadratic polynomial in the running cost function $c_2(u_2, x_2)$:

$$
\begin{aligned}
L_2 &= 5.0, \\
M_2 &= 0.0, \\
N_2 &= 10.0.
\end{aligned}
\tag{53}
$$

Under this running cost function, the optimal policy is a moderate isolation regime. Following this policy, the isolation rate is high, as the infections are low, and then it declines over time while the epidemic develops. Unlike the policy above, this leads to a gradual decline in both the infected and susceptible fractions of the population.

Consider now the case of an optimal vaccination strategy. Again, we make two choices of the running cost function: high cost and low cost.

For the high cost case we choose the parameters as follows:

$$
\begin{aligned}
L_1 &= 10.0, \\
M_1 &= 0.0, \\
N_1 &= 500.0.
\end{aligned}
\tag{54}
$$
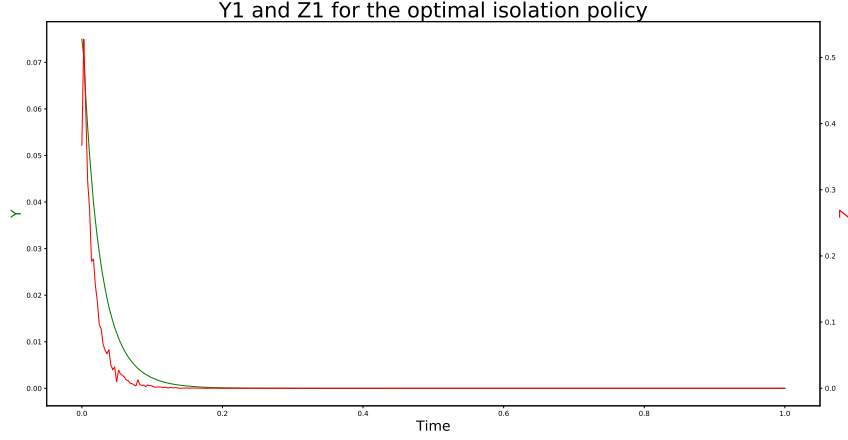
Figure 3: Plot of $Y_{1,t}$ (primary axis) and $Z_{1,t}$ (secondary axis) for equation (33).

The results of Monte Carlo simulations for this cost function are plotted in Figures 10 - 13. They parallel the results presented above in the case of high cost isolation mitigation. The optimal policy is a massive vaccination campaign that dramatically reduces the susceptible fraction of the population and leads to significantly lower infections.

The low cost running cost function is parameterized as follows:

$$
\begin{aligned}
L_2 &= 0.01, \\
M_2 &= 0.0, \\
N_2 &= 0.4.
\end{aligned}
\tag{55}
$$

The results of Monte Carlo simulations for this cost function are plotted in Figures 14 - 17.

Finally, we consider the case of the vaccination / isolation policy. An example of a cost running cost function is given by the following parameter values:

$$
\begin{aligned}
L_1 &= 1.0, \\
M_1 &= 0.0, \\
N_1 &= 500.0, \\
L_2 &= 5.0, \\
M_2 &= 0.0, \\
N_2 &= 555.0.
\end{aligned}
\tag{56}
$$

The results of Monte Carlo simulations are plotted in Figures 18 - 21.
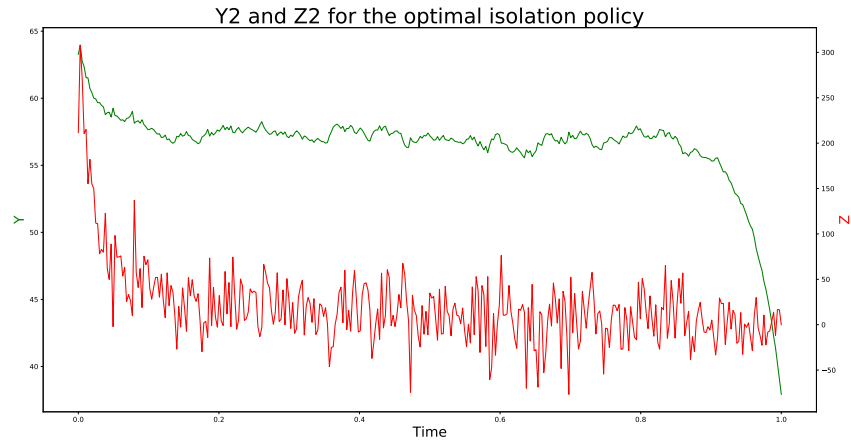
Figure 4: Plot of $Y_{2,t}$ (primary axis) and $Z_{2,t}$ (secondary axis) for equation (33).

A low cost running cost function is given by the following set of coefficients:

$$
\begin{aligned}
L_1 &= 0.1, \\
M_1 &= 0.0, \\
N_1 &= 50.0, \\
L_2 &= 3.0, \\
M_2 &= 0.0, \\
N_2 &= 35.0.
\end{aligned}
\tag{57}
$$

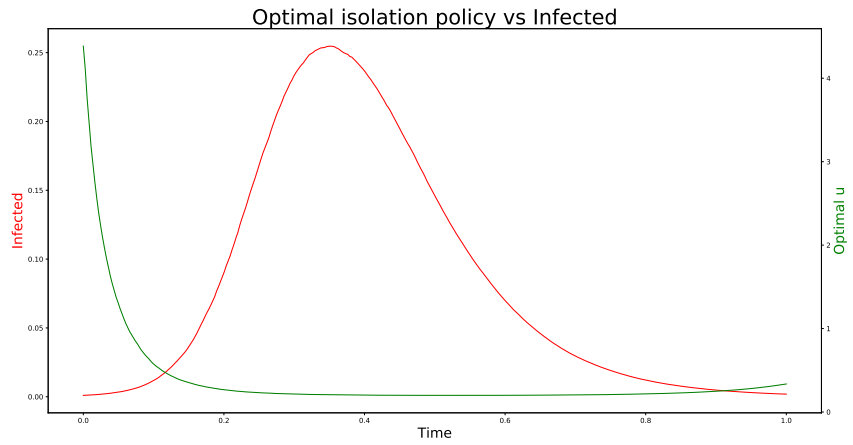The results of Monte Carlo simulations are plotted in Figures 22 - 25.
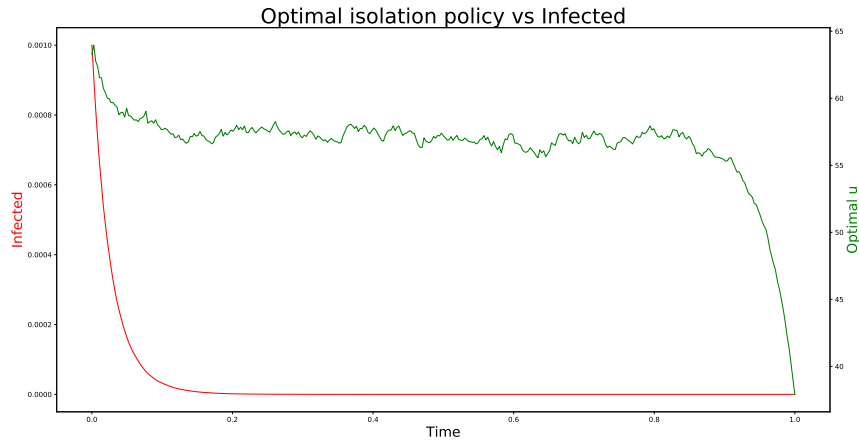
Figure 5: Plot of the Infected fraction (primary axis) versus the optimal isolation policy (secondary axis) under high cost running cost function.

References

[1] Anderson, R. M., and May, R. M.: Population biology of infectious diseases: Part I, *Nature*, **280**, 361 - 366 (1971).

[2] Bellman, R.: *Dynamic Programming*, Princeton University Press (1957).

[3] Bolzoni, L., Bonacini, E., Soresina, C., and Groppi, M.: Time-optimal control strategies in SIR epidemic models, *Mathematical Biosciences*, **292**, 86 - 96 (2017).

[4] Brauer, F., Castillo-Chavez, C., and Feng, Z.: *Mathematical Models in Epidemiology*, Springer (2019).

[5] Britton, T.: Stochastic epidemic models: A survey, *Mathematical Biosciences*, **210**, 24 - 35 (2010).

[6] Bartlett, M. S.: Deterministic and stochastic models for recurrent epidemics, *Proc. Third Berkeley Symposium on Math. Statistics and Probability*, **4**, 81 -109 (1956).

[7] Bouchard B., Touzi N.: *Discrete Time Approximation and Monte-Carlo Simulation of Backward Stochastic Differential Equations*, Stochastic Processes and their Applications, Vol 111, 2, 175-206 (2004).

[8] Doyle, A. C.: *A Scandal in Bohemia*.

[9] El Karoui, N., Peng, S. and Quenez, M. C.: Backward stochastic differential equations in finance, *Math. Finance*, **7**, 1 - 71 (1997).
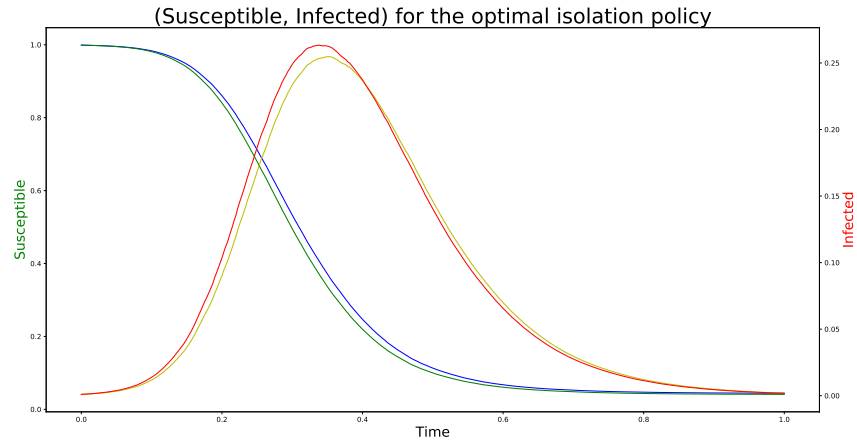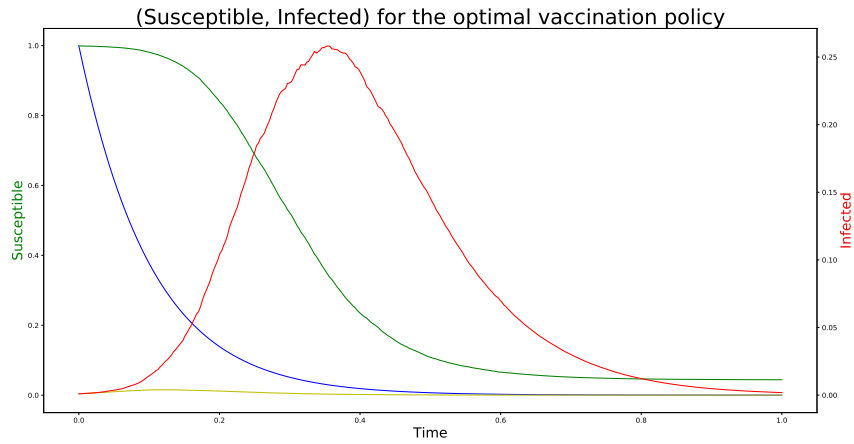
Figure 6: Raw (Susceptible, Infected) versus optimal (Susceptible, Infected) under the isolation policy assuming low cost running cost function. The read and green lines are the average values of the raw Susceptible (primary axis) and Infected (secondary axis) fractions, respectively, while the yellow and blue lines represent the averages of the corresponding optimal values.

[10] Flemming, W. H., Soner, H. M.: *Controlled Markov processes and viscosity solutions*, Springer-Verlag (1992).

[11] Gray, A., Greenhalgh, D., Hu, L., Mao, X., and Pan, J.: A stochastic differential equation SIS epidemic model, *SIAM J. on Appl. Math.*, **71**, 876 - 902 (2011).

[12] Hairer, E., Lubich, C., and Wanner, G.: Geometric numerical integration illustrated by the StörmerVerlet method, *Acta Numerica*, 399 450 (2003).

[13] Hansen, E., and Day, T.: Optimal control of epidemics with limited resources, *J. Math. Biol.*, **62**, 423 - 451 (2011).

[14] Han, J., Jentzen, A., and E, W.: Solving high-dimensional partial differential equations using deep learning, *Proc. Nat. Acad. Sci.*, **115**, 8505 - 8510 (2018).

[15] Kermack, W. O., and McKendrick, A. G.: A contribution to the mathematical theory of epidemics, *Proc. Roy. Soc. Lond.* **A115**, 700 - 721 (1927).

[16] Leimkuhler, B., and Reich, S.: *Simulating Hamiltonian Dynamics*, Cambridge University Press (2004).

[17] Lesniewski, A., and Lesniewski, N.: Options on infectious diseases, `https://arxiv.org/abs/2003.07992` (2020).

[18] Lesniewski, A., and Richter, A.: Managing counterparty credit risk via backward stochastic differential equations, `https://arxiv.org/abs/1608.03237` (2016).
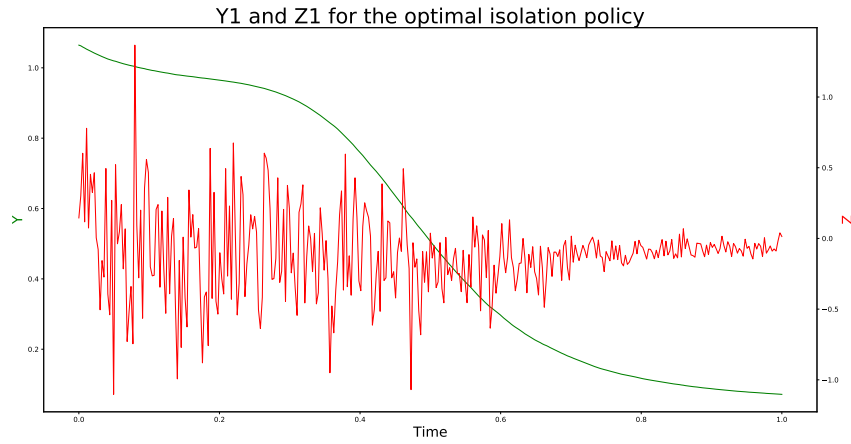
Figure 7: Plot of $Y_{1,t}$ (primary axis) and $Z_{1,t}$ (secondary axis) for equation (33).

[19] Longstaff, F. A., and Schwartz, E. S.: *Valuing American Options by Simulation: A Simple Least-Squares Approach*, Rev. Fin. Stud., 14 (1), 113 - 147 (2001).

[20] Pham, H.: *Continuous-time Stochastic Control and Optimization with Financial Applications*, Springer (2009).

[21] Pontryagin, L. S., Boltyanskii, V. G., Gamkrelidze, R.V., and Mishchenko, E. F.:*The Mathematical Theory of Optimal Processes*, Interscience (1962).

[22] Yong, J., and Zhou, X. U.: *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer (1999).
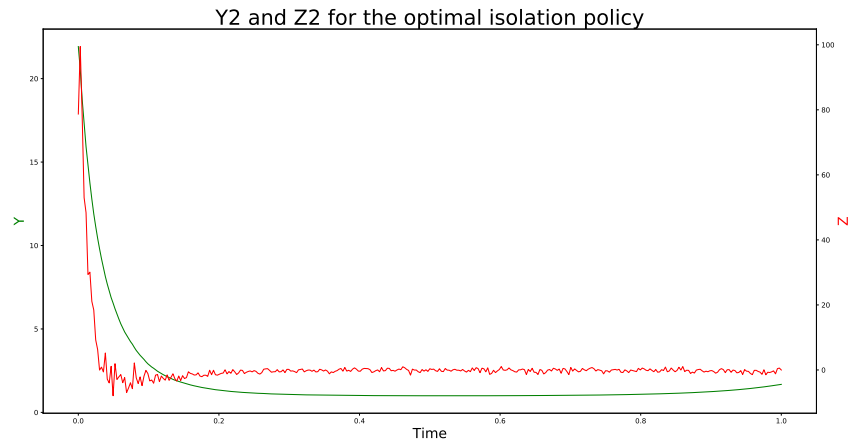
Figure 8: Plot of $Y_{2,t}$ (primary axis) and $Z_{2,t}$ (secondary axis) for equation (33).

Figure 9: Plot of the Infected fraction (primary axis) versus the optimal isolation policy (secondary axis) under low cost running cost function.

Figure 10: Raw (Susceptible, Infected) versus optimal (Susceptible, Infected) under the vaccination policy assuming high cost running cost function. The read and green lines are the average values of the raw Susceptible (primary axis) and Infected (secondary axis) fractions, respectively, while the yellow and blue lines represent the averages of the corresponding optimal values.
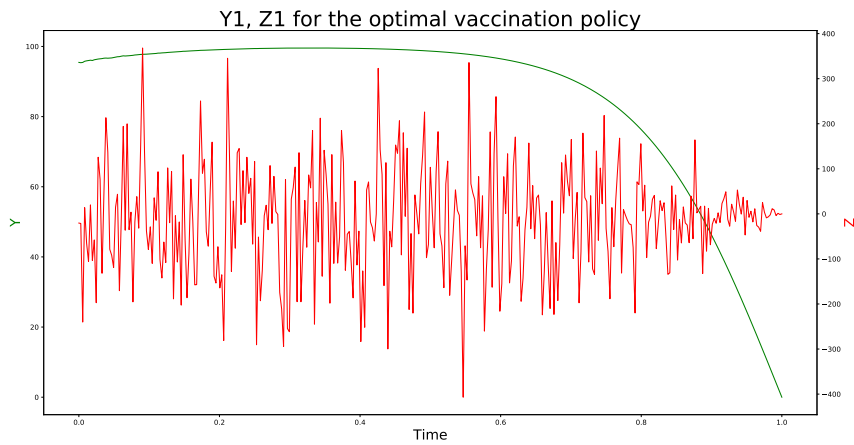
Figure 11: Plot of $Y_{1,t}$ (primary axis) and $Z_{1,t}$ (secondary axis) for equation (34).
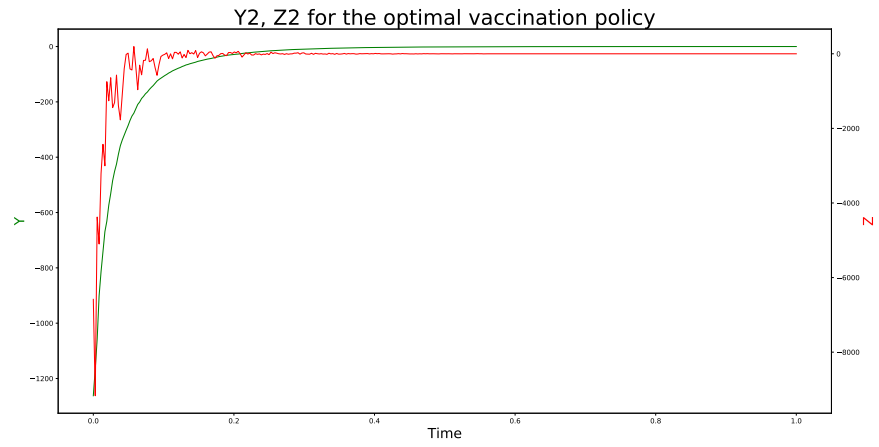
Figure 12: Plot of $Y_{2,t}$ (primary axis) and $Z_{2,t}$ (secondary axis) for equation (34).
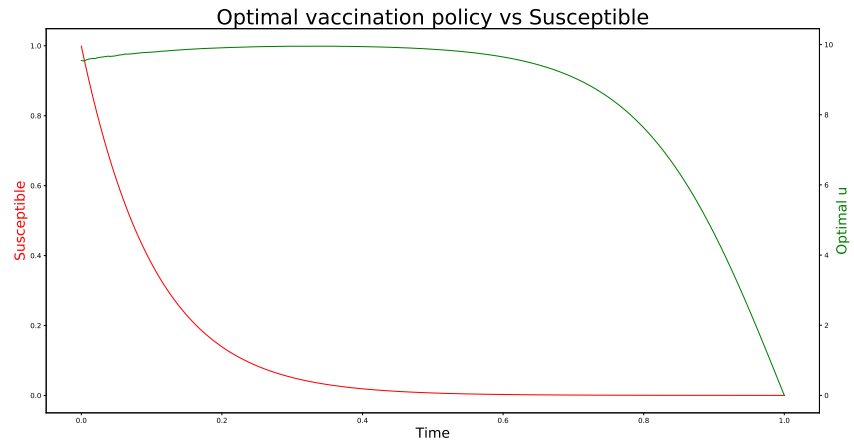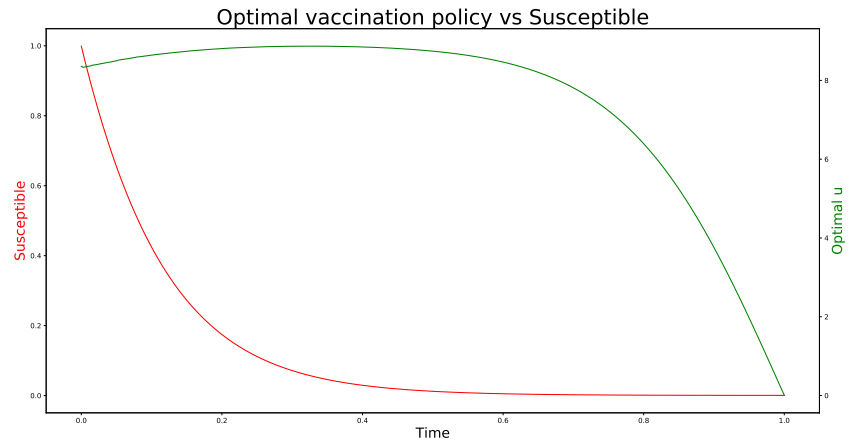

Figure 13: Plot of the Infected fraction (primary axis) versus the optimal vaccination policy (secondary axis) under high cost running cost function.
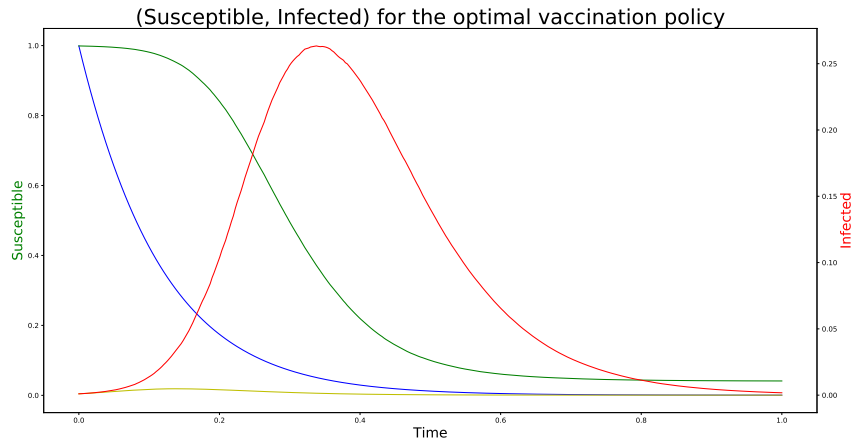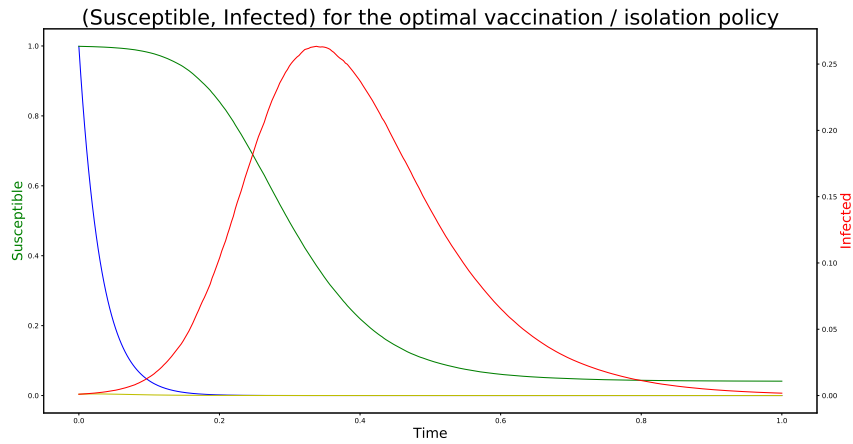
Figure 14: Raw (Susceptible, Infected) versus optimal (Susceptible, Infected) under the vaccination policy assuming low cost running cost function. The read and green lines are the average values of the raw Susceptible (primary axis) and Infected (secondary axis) fractions, respectively, while the yellow and blue lines represent the averages of the corresponding optimal values.
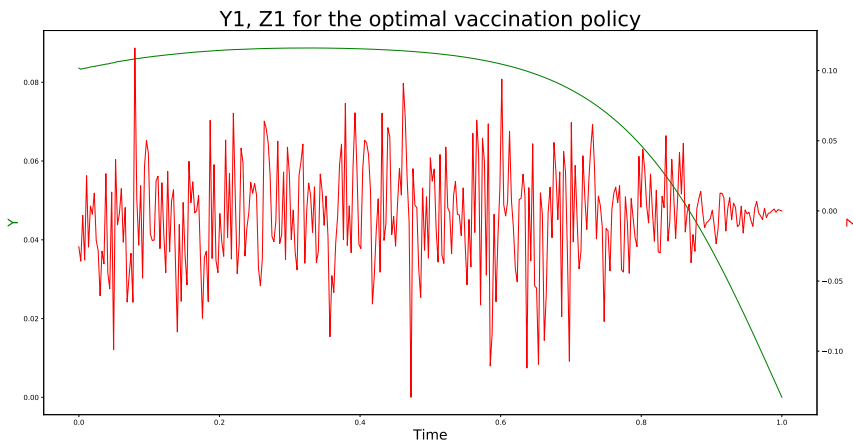

Figure 15: Plot of $Y_{1,t}$ (primary axis) and $Z_{1,t}$ (secondary axis) for equation (34).
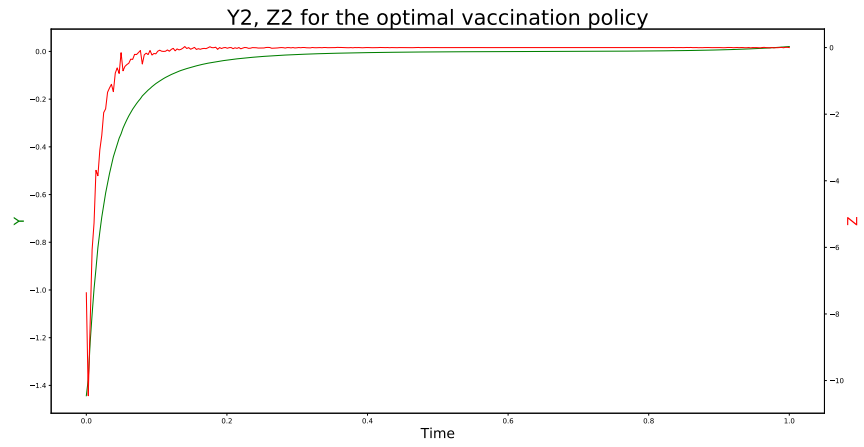
Figure 16: Plot of $Y_{2,t}$ (primary axis) and $Z_{2,t}$ (secondary axis) for equation (34).

Figure 17: Plot of the Infected fraction (primary axis) versus the optimal vaccination policy (secondary axis) under low cost running cost function.

Figure 18: Raw (Susceptible, Infected) versus optimal (Susceptible, Infected) under the vacci-nation / isolation policy assuming high cost running cost function. The read and green lines are the average values of the raw Susceptible (primary axis) and Infected (secondary axis) fractions, respectively, while the yellow and blue lines represent the averages of the corresponding optimal values.
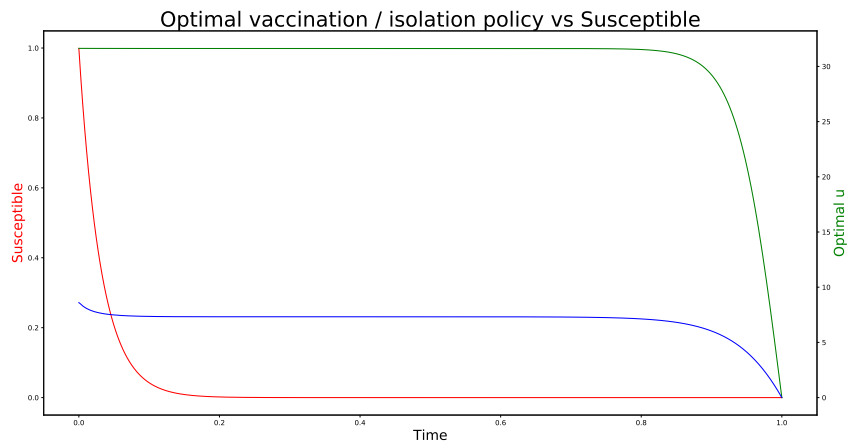
Figure 19: Optimal isolation (green line) and vaccination (blue line) policies (secondary axis) versus the Susceptible (red line) fraction (primary axis) for the high cost running cost function.
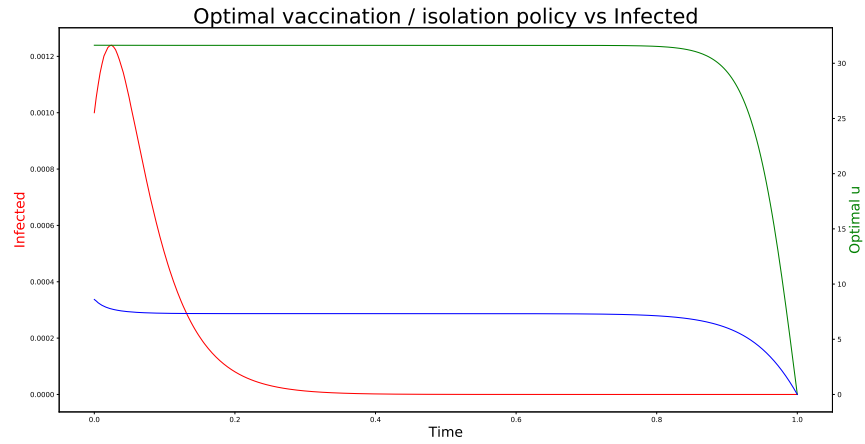
Figure 20: Optimal isolation (green line) and vaccination (blue line) policies (secondary axis) versus the Infected (red line) fraction (primary axis) for the high cost running cost function.
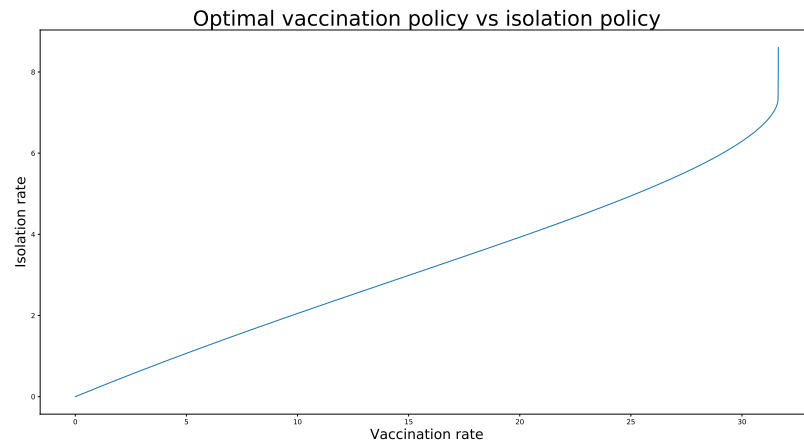

Figure 21: Plot of the optimal isolation policy versus optimal vaccination policy under high cost running cost function.
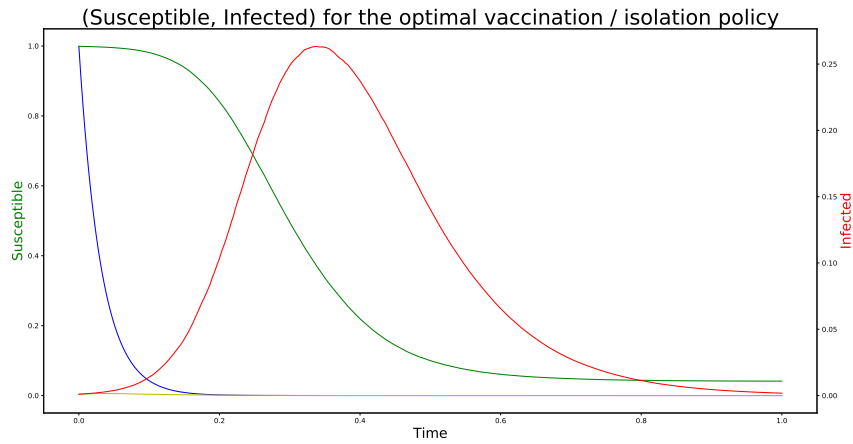
Figure 22: Raw (Susceptible, Infected) versus optimal (Susceptible, Infected) under the vacci-
nation / isolation policy assuming low cost running cost function. The read and green lines are
the average values of the raw Susceptible (primary axis) and Infected (secondary axis) fractions,
respectively, while the yellow and blue lines represent the averages of the corresponding optimal
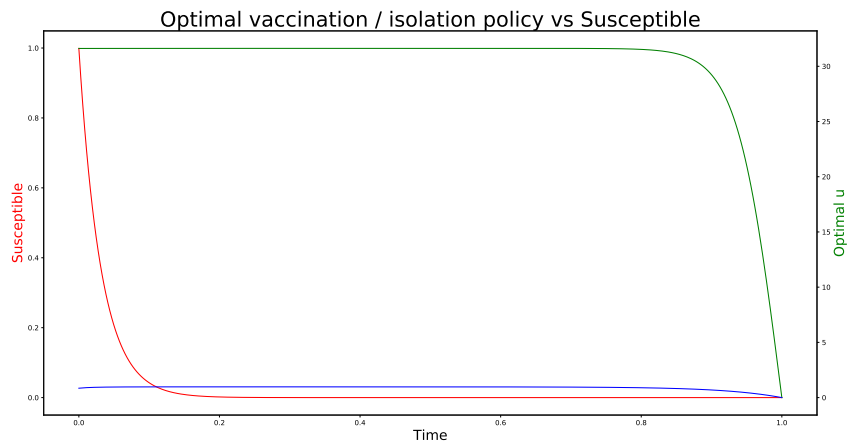values.

Figure 23: Optimal isolation (green line) and vaccination (blue line) policies (secondary axis)
versus the Susceptible (red line) fraction (primary axis) for the low cost running cost function.
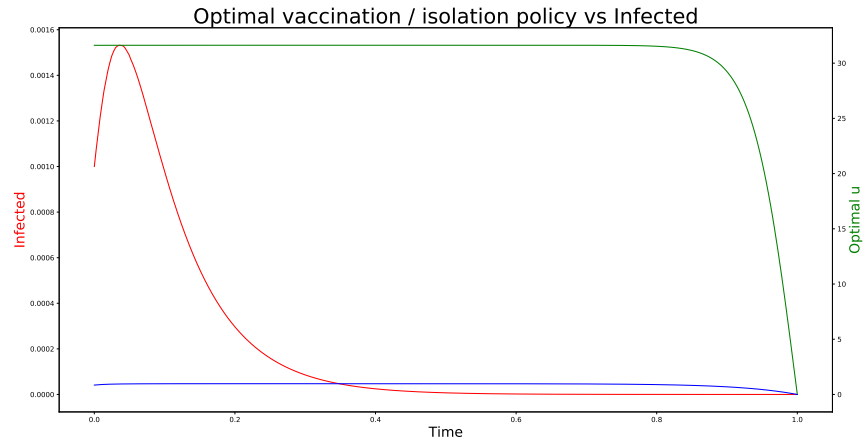
Figure 24: Optimal isolation (green line) and vaccination (blue line) policies (secondary axis) versus the Infected (red line) fraction (primary axis) for the low cost running cost function.
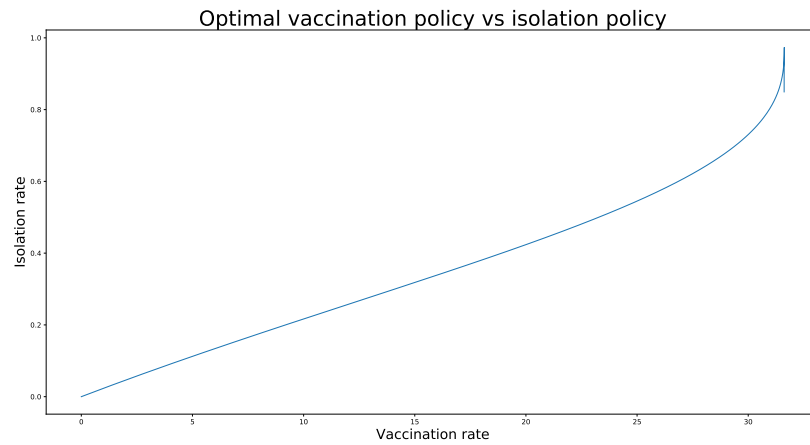

Figure 25: Plot of the optimal isolation policy versus optimal vaccination policy under low cost running cost function.