

Controlling the Number of HIV Infectives in a Mobile Population

A. Sani ^{a,b}, D. P. Kroese^a

^a Department of Mathematics, University of Queensland, AUSTRALIA,
asani@maths.uq.edu.au, kroese@maths.uq.edu.au.

^b Department of Mathematics, FMIPA, Universitas Haluoleo, Kendari, INDONESIA.

Abstract. The spread of the human immunodeficiency virus (HIV) depends prominently on the migration of people between different regions. An important consequence of this population mobility is that HIV control strategies that are optimal in a regional sense may not be optimal in a national sense. We formulate various mathematical control problems for HIV spread in mobile heterosexual populations, and show how optimal regional control strategies can be obtained that minimize the national spread of HIV. We apply the cross-entropy method to solve these highly multi-modal and non-linear optimization problems. We demonstrate the effectiveness of the method via a range of experiments and illustrate how the form of the optimal control function depends on the mathematical model used for the HIV spread.

Key-words: Epidemic Control, HIV/AIDS, Mobility, Multiple Patches, Cross-Entropy Method.

1 Introduction

The control of infectious diseases such as Acquired Immune Deficiency Syndrome (AIDS), which is caused by HIV, is a challenging public health problem. As the awareness of AIDS increases and the urgency to control it becomes crucial to many nations, more and more local and national governments provide funding to combat the disease. In Indonesia, for example, where a high population mobility among its regions seems to lead to a high risk for the spread of the epidemic [12, 31], many local governments have planned and allocated funds and resources to control the spread of HIV in their regions. However, as the people travel among regions, it is not clear whether an allocation of funds that is optimal for local regions is also optimal for the nation as a whole. The main goal on a country level is to find regional control strategies that minimize the total number of new infectives in the country as a whole, without using more than the allocated budget in each individual region.

In this paper we formulate various mathematical control models for HIV spread, which incorporate complex characteristics such as hetero-sexual transmission and migration among regions. These control models turn out to be highly non-linear and are difficult to solve via standard (convex) optimization procedures. Instead, we solve these problems via the *cross-entropy* (CE) method [24]; to our knowledge this is the first application of the CE method to optimal control. Numerical results demonstrate the efficiency and accuracy of the approach and illustrate how the optimal control functions depend on the models and their parameters.

The literature on optimal control for the spread of HIV is not extensive. One of the reasons is that the mathematical models used to describe the dynamics of HIV spread

tend to be quite complex. For example, the incubation period after infection with HIV is known to be extremely long; the disease can be transmitted in many different ways — via sexual contact, blood transfusions, birthing, and infected syringes — , and the migration of people among subgroups has significant consequences for the dynamics of the epidemic spread [21, 26]. As a consequence, mathematical models involving complex HIV transmission mechanisms usually do not lend themselves to conventional mathematical control techniques such as dynamic programming and convex optimization.

Most studies on epidemic control for HIV therefore focus on influencing the control parameters in the model, e.g., investigating how changing certain control parameters affects the evolution of the number of infected individuals over a certain period of time. The main purpose then is to identify the parameters that have the most significant effect on reducing the number of (new) infectives. For example, in Bernstein [1] three intervention strategies for preventing heterosexual transmission of HIV are considered: (1) reduce the number and rate of change of sex partners, (2) increase condom use, and (3) improve the treatment of sexually transmitted diseases. Both deterministic and stochastic models are considered to simulate the evolution of the AIDS epidemic. The demographic, biological, and behavioral parameters for a severely affected east African city early in the epidemic are estimated and these parameter values are used to compute the spread of HIV under the above intervention strategies. Similar intervention strategies are studied in Hyman [13], for the spread of HIV, and in Kretzschmar et al. [16], for the spread of gonorrhoea and chlamydia.

The approach that we take in this paper is to formulate the HIV intervention problem as a mathematical optimal control problem, that is, finding control functions that optimize some objective function. A review of applications of mathematical control theory to some simple disease models (up to 1976) can be found in Wickwire [30]. For a more recent application see, for example, Clancy and Piunovsky [6], who consider an optimal isolation policy for a deterministic epidemic model by minimizing the cost of isolating infectives from the susceptible population. A similar study is by Greenhalgh [11], who considers an epidemic spreading through a homogeneously mixing population. The control consists here of immunizing susceptible people and isolating and removing infected people from the rest of the population. In Jung et al. [14] a two-strain tuberculosis model is described via a system of ordinary differential equations and two types of treatments are applied as controls, in order to reduce the latent and infectious groups. Standard techniques in optimal control, Hamiltonian and Pontryagin’s Maximum Principles, are applied to derive the optimal solution and the problem is solved numerically for several scenarios. In Kirschner et al. [15] the objective is to minimize the systematic cost of chemotherapy in early HIV treatment. The interaction of the immune system with HIV is described by an ODE model. The control is the percentage effect of chemotherapy on the viral production. The control problem consists here of distributing a limited number of resources over a fixed time horizon, which is a recurring theme in many epidemic intervention studies; see, for example, [3, 5, 23].

The mathematical control problems in the present paper are mostly motivated by the following studies. In Blount et al. [3] an optimal control problem is considered for a discrete-time, deterministic SIS epidemic over a finite time horizon $1, 2, \dots, T$. The objective is to minimize the total cost of the epidemic over the finite time horizon, which accounts for the cost of the new infections produced and the cost of controlling the current infected population. The control consists of a sequence u_1, \dots, u_T of binary values, where

$u_t = 1$ means that a “control” is applied during the t -th time period and $u_t = 0$ means that no control is applied. Two alternative methods, nonlinear programming (NLP) and dynamic programming (DP), are applied to find a best control sequence. It is shown that DP gives qualitatively much better results and is computationally cheaper than NLP. However, the degree of complexity increases exponentially in T , and both approaches are infeasible when either or both the epidemic model and the control function are formulated as (non-linear) *continuous-time* functions. Such problems have been studied in, for example, [3, 5, 23]. In [5, 23], it is shown that it is not generally possible to derive the optimal solution of such complex problem in a closed form and therefore one needs resort to approximation techniques.

Several epidemic models for the spread of HIV in a mobile population were introduced in [25]. There, the infection rates were considered to be fixed for all times. In this paper, we assume that the infection rates for all patches can vary; see also [29]. More specifically, we assume that they are a function of some epidemic control strategy. We consider a continuous-time epidemic problem with continuous-time control functions in multiple populations. The problem is to minimize the number of new infectives over a finite time horizon. The optimal control function is derived via the CE method, as an alternative to NLP and DP [2, 27].

The remainder of this paper is organized as follows. Section 2 describes various models for the evolution of susceptibles and HIV infectives in multiple mobile populations, and formulates the corresponding optimal control problem. Section 3 provides the main CE procedure for solving these highly non-linear control problems. This is clarified with an illustrating example. In Section 4 we present various numerical experiments for the HIV optimal control problem and discuss their outcomes, which are sometimes counter-intuitive. Concluding remarks and directions for future research are given in Section 5.

2 The Models

In this section, we first briefly review various epidemic models for the spread of HIV in mobile heterosexual populations [25], and then formulate the corresponding epidemic control problems, which are the focus of this paper.

2.1 The Epidemic Models

The following models describe the spread of HIV infectives in multiple, sexually active populations, where individuals are allowed to migrate among populations. The mode of transmission is assumed to be via sexual contact only between partners of the opposite sex. This assumption is mainly because heterosexual contact is still the primary mode of HIV infection worldwide. However, one can easily extend the models by including, for example, man-to-man sexual contact or other modes of transmission. Individuals in each population are divided into four groups (compartments): female susceptibles, female infectives, male susceptibles, and male infectives. New susceptibles enter the population at rate B_F (B_M) for females (males). Susceptible individuals leave their groups due to either natural death at rate μ , or because they move to the corresponding infectives group at rate λ_F (λ_M) for females (males). We assume that males are fewer than females (a similar assumption is adopted in [9]) and males choose a female partner from the female subpopulation. Therefore, the rates λ_F and λ_M are assumed to be proportional

to the fraction of infected members of opposite sex, relative to the female subpopulation size. That is,

$$\lambda_F = \beta \frac{I_M(t)}{N_F(t)} \quad \left(\lambda_M = \beta \frac{I_F(t)}{N_F(t)} \right), \quad (1)$$

where β is the infection rate parameter, I_M and I_F are the number of male and female infectives, respectively, and $N_F = S_F + I_F$ is the size of the female subpopulation, with S_F the number of female susceptibles. The rates λ_F and λ_M are called the *forces of infection*, which are the same as the rates of infection per susceptible, defined in [13]. In addition to natural deaths, the infected individuals also leave their groups because they develop AIDS at rate γ . Note that the population in each region does not include the AIDS subpopulation, since AIDS people are assumed to be no longer sexually active due to opportunistic diseases. This is summarized in the scheme in Figure 1.

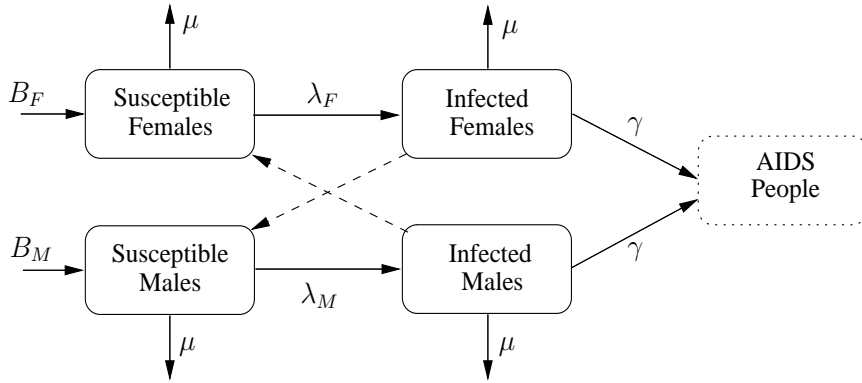


Figure 1: The scheme of the model. Susceptible females (males) are infected by infected males (females) via sexual contact only, indicated by the dashed lines.

Individuals in one population can move to the other populations, where they can spread the disease or be infected by it. Suppose R is the total number of populations, which we will henceforth refer to as *patches*. Let R_i denote patch i and let $v^{(i)}$ stand for the emigration rate from R_i . Suppose, moreover, that an individual leaving R_i migrates to R_j with probability ρ_{ij} . Then, $v_{ij} = \rho_{ij} v^{(i)}$ denotes the migration rate of individuals from R_i to R_j . The mobility of individuals among the patches is illustrated in Figure 2.

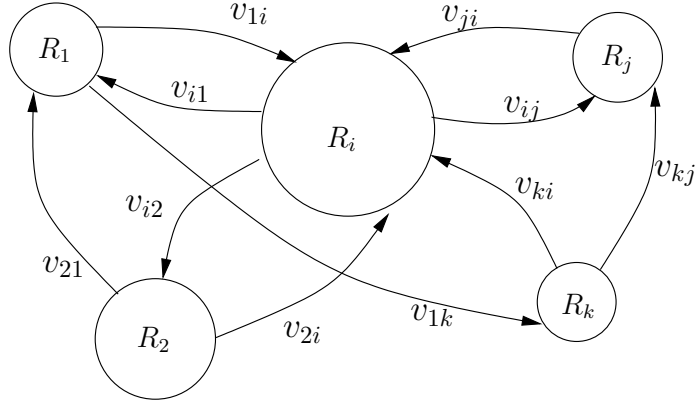


Figure 2: The scheme for the mobility of people among patches. The size of a circle corresponds to the total population size in that patch.

We formulate three different models to describe the dynamics of the system. In the first two models we assume that the individuals do not make an actual move among patches but that there is a force of infection from infected patches to others. We call this *inter-patch infection*. As in [9] and [20], the force of infection is defined as the rate at which a susceptible female (male) in one patch will be infected by an infected male (female) within that patch and from other patches; see (1). When many infectives are present in a patch, the force of infection from this patch to other patches becomes large. We derive two models for this case: one with a constant population size and one with a varying population size. The third model allows individuals to move from one patch to others. Thus, a susceptible female (male) in a patch will be infected by an infected male (female) within that patch only. We consider this case under a varying population size only.

For the case of a constant population size, all individuals that leave the system are replaced (balanced) by an inflow of susceptibles, at a proportion ξ for females and $1 - \xi$ for males. Let $B_F^{(r)}$ and $B_M^{(r)}$ denote the rate at which new susceptible females and males in patch r enter their corresponding compartments, respectively. Then, $B_F^{(r)} = \xi(\mu N^{(r)} + \gamma I^{(r)})$ and $B_M^{(r)} = (1 - \xi)(\mu N^{(r)} + \gamma I^{(r)})$, where $N^{(r)} = N_F^{(r)} + N_M^{(r)}$ is the population size of patch r , with $N_F^{(r)}$ ($N_M^{(r)}$) the size of female (male) subpopulation, and $I^{(r)} = I_F^{(r)} + I_M^{(r)}$. Note that $\gamma I^{(r)}$ represents the rate at which AIDS cases develop. The population sizes $N^{(r)}$, $r = 1, \dots, R$ need not be the same. Let $s_F^{(r)}(t)$, $i_F^{(r)}(t)$, $s_M^{(r)}(t)$, $i_M^{(r)}(t)$ denote the proportion of susceptible females, infected females, susceptible males, and infected males in patch r at time t , respectively, relative to the total population of all patches $N = \sum_{r=1}^R N^{(r)}$. Thus, for example, $s_F^{(r)}(t) = \frac{S_F^{(r)}(t)}{N}$. The scaled inflow rates for susceptible females and males in each patch are given by

$$b_F^{(r)} = \frac{B_F^{(r)}}{N} = \xi(\mu n^{(r)} + \gamma i^{(r)}) \quad \text{and} \quad b_M^{(r)} = \frac{B_M^{(r)}}{N} = (1 - \xi)(\mu n^{(r)} + \gamma i^{(r)}), \quad (2)$$

respectively, where $n^{(r)} = \frac{N^{(r)}}{N}$ and $i^{(r)} = \frac{I^{(r)}}{N} = i_F^{(r)} + i_M^{(r)}$.

For the case of a varying population size, we assume for simplicity that $B_F^{(r)} = B_M^{(r)} =$

B , at a constant rate B , for all patches. In this case the variables $s_F^{(r)}(t), i_F^{(r)}(t), s_M^{(r)}(t), i_M^{(r)}(t)$ denote the proportion of susceptible and infected females, susceptible and infected males in patch r , at time t , relative to a scaling parameter $V = \frac{2RB}{\mu}$, corresponding to the equilibrium population size in the absence of disease. Hence, the scaled inflow rates are the same for susceptible females and males in each patch:

$$b_F^{(r)} = b_M^{(r)} = \frac{\mu}{2R}, \quad (3)$$

where $b_F^{(r)} = \frac{B_F^{(r)}}{V}$ and $b_M^{(r)} = \frac{B_M^{(r)}}{V}$. If the flow rates are not the same for each patch, say $B_F^{(r)} = B_M^{(r)} = B^{(r)}$ with a constant rate $B^{(r)}$ for patch r , one can take, for example, $V = \frac{2B_{tot}}{\mu}$ with $B_{tot} = \sum_{j=1}^R B^{(j)}$ as a scaling parameter. The inflow rates for both susceptible females and males then become $b_F = b_M = \frac{\mu B^{(r)}}{2B_{tot}}$.

Models with Inter-Patch Infection

As mentioned above, we consider two scenarios for this case: the scenario where the size for each population remains constant over time and the case where the population size is varying. The difference concerns the rates at which female (male) susceptibles enter the system. The evolution of susceptibles and infectives among patches for both cases is governed by the following system of differential equations [25]:

$$\begin{aligned} \frac{ds_F^{(r)}}{dt} &= b_F^{(r)} - \sum_{j=1}^R \beta^{(jr)} \frac{i_M^{(j)}}{n_F^{(j)}} s_F^{(r)} - \mu s_F^{(r)}, \\ \frac{di_F^{(r)}}{dt} &= \sum_{j=1}^R \beta^{(jr)} \frac{i_M^{(j)}}{n_F^{(j)}} s_F^{(r)} - (\mu + \gamma) i_F^{(r)}, \\ \frac{ds_M^{(r)}}{dt} &= b_M^{(r)} - \sum_{j=1}^R \beta^{(jr)} \frac{i_F^{(j)}}{n_F^{(j)}} s_M^{(r)} - \mu s_M^{(r)}, \\ \frac{di_M^{(r)}}{dt} &= \sum_{j=1}^R \beta^{(jr)} \frac{i_F^{(j)}}{n_F^{(j)}} s_M^{(r)} - (\mu + \gamma) i_M^{(r)}, \end{aligned} \quad (4)$$

where $n_F^{(r)} = \frac{N_F^{(r)}}{N}$, $n_M^{(r)} = \frac{n_F^{(r)}}{N}$, $b_F^{(r)}$ and $b_M^{(r)}$ are defined in (2) for the case of constant population size and (3) for varying population size. The force of infection in each patch r is described by the term $\sum_{j=1}^R \beta^{(jr)} \frac{i_M^{(j)}}{n_F^{(j)}}$ for susceptible females and $\sum_{j=1}^R \beta^{(jr)} \frac{i_F^{(j)}}{n_F^{(j)}}$ for susceptible males, with $\beta^{(jr)}$ the infection rate parameter from patch j to patch r . Note that the total number of AIDS cases in patch r over a time period $[0, t]$ can be calculated as $\int_0^t \gamma i^{(r)}(u) du$.

Model with Actual Mobility

We consider only the case of varying population size for this model. The formulation of the rate at which a susceptible is infected by an infective in this model is slightly

different from the two previous models with inter-patch infection. In addition, there are two extra terms for each subpopulation, describing the migration rates of individuals among corresponding subpopulations in the patches. Let $n^{(r)} = n_F^{(r)} + n_M^{(r)}$ and $v^{(r)}$ denote the population size and the migration rate of patch r , respectively, relative to a scaling parameter $V = \frac{2RB}{\mu}$. The evolution of the state variables in this model is described by the following system of ODEs:

$$\begin{aligned}
\frac{ds_F^{(r)}}{dt} &= b_F^{(r)} - \beta^{(r)} \frac{i_M^{(r)}}{n_F^{(r)}} s_F^{(r)} - \mu s_F^{(r)} + \sum_{j=1}^R \frac{v_{jr}}{n^{(j)}} s_F^{(j)} - \frac{v^{(r)}}{n^{(r)}} s_F^{(r)}, \\
\frac{di_F^{(r)}}{dt} &= \beta^{(r)} \frac{i_M^{(r)}}{n_F^{(r)}} s_F^{(r)} - (\mu + \gamma) i_F^{(r)} + \sum_{j=1}^R \frac{v_{jr}}{n^{(j)}} i_F^{(j)} - \frac{v^{(r)}}{n^{(r)}} i_F^{(r)}, \\
\frac{ds_M^{(r)}}{dt} &= b_M^{(r)} - \beta^{(r)} \frac{i_F^{(r)}}{n_F^{(r)}} s_M^{(r)} - \mu s_M^{(r)} + \sum_{j=1}^R \frac{v_{jr}}{n^{(j)}} s_M^{(j)} - \frac{v^{(r)}}{n^{(r)}} s_M^{(r)}, \\
\frac{di_M^{(r)}}{dt} &= \beta^{(r)} \frac{i_F^{(r)}}{n_F^{(r)}} s_M^{(r)} - (\mu + \gamma) i_M^{(r)} + \sum_{j=1}^R \frac{v_{jr}}{n^{(j)}} i_M^{(j)} - \frac{v^{(r)}}{n^{(r)}} i_M^{(r)},
\end{aligned} \tag{5}$$

where $\beta^{(r)} \equiv \beta^{(rr)}$ is the parameter of infection rate within patch r and $v_{jr} = \rho_{jr} v^{(j)}$, with ρ_{jr} the probability to visit patch r from patch j . The rates $b_F^{(r)}$ and $b_M^{(r)}$ are defined in (3). We put $v_{rr} = 0$ for all r . Note that all variables (such as $s_F^{(r)}$ and $i_F^{(r)}$) are relative to the parameter V . The dynamic behavior of these models is discussed in [25]. Again, the number of AIDS cases at time t in patch r can be calculated by taking the integral of $\gamma i^{(r)}$.

2.2 Epidemic Optimal Control Formulation

The epidemic control problem is formulated as follows. Suppose each patch r has a fixed budget, say $K^{(r)}$, for $r = 1, \dots, R$, which needs to be spent over a finite time horizon $[0, T]$. Let $u^{(r)} = u^{(r)}(t)$ denote the control function, which is to be interpreted as the amount of budget spent per unit time in patch r . Suppose the parameter of infection rate in patch r , denoted $\beta^{(r)} = \beta(u^{(r)})$, is a decreasing function with respect to the control function $u^{(r)}$. Note that $\beta^{(jr)} = \beta(u^{(j)})$. The function $\beta^{(r)}$ can be, for example, of following form:

$$\beta_1 = \beta(u) = \beta_{\max} - c_1 u, \tag{6}$$

or

$$\beta_2 = \beta(u) = \frac{\beta_{\max}}{1 + c_2 u}, \tag{7}$$

for some nonnegative constants $c_1, c_2 \in \mathbb{R}$ with $c_1 \leq \frac{\beta_{\max}}{u_{\max}}$, where u_{\max} is the largest admissible control value. The function β_2 implies that a relatively small amount of control effort results in a large reduction in the infection rate, but a greater amount of budget is required to reduce the infection rate to its minimum value (i.e., it is costly to completely eliminate an epidemic in the entire population). In this study, we use the form β_2 which is frequently used in epidemic control; see, for example, [4].

The problem we are interested in is to determine the shape of the control function $u^{(r)}$ for each patch r , such that the total number of new infectives in all of the patches

over the time horizon is minimized. Let $\mathbf{u} = \{u^{(r)}, r = 1, \dots, R\}$. Then, our epidemic intervention problems are formulated as follows.

For the models with inter-patch infection

$$\min_{\mathbf{u}} J(\mathbf{u}) = \min_{\mathbf{u}} \int_0^T \sum_{r=1}^R \sum_{j=1}^R \beta^{(jr)} \left[\frac{i_M^{(j)}}{n_F^{(j)}} s_F^{(r)} + \frac{i_F^{(j)}}{n_F^{(j)}} s_M^{(r)} \right] dt, \quad (8)$$

subject to the dynamic constraint (4) and the integral constraints

$$\int_0^T u^{(r)} dt = K^{(r)}, \quad u^{(r)} \geq 0, \quad r = 1, \dots, R. \quad (9)$$

For the model with actual mobility

$$\min_{\mathbf{u}} J(\mathbf{u}) = \min_{\mathbf{u}} \int_0^T \sum_{r=1}^R \beta^{(r)} \left[\frac{i_M^{(r)}}{n_F^{(r)}} s_F^{(r)} + \frac{i_F^{(r)}}{n_F^{(r)}} s_M^{(r)} \right] dt \quad (10)$$

subject to the dynamic constraint (5) and the integral constraints (9).

The terms inside the integral in the objective functions describe the rate at which new infectives are generated. For the case of a constant population size, $J(\mathbf{u})$ is the proportion of new infective cases relative to the total population size N in the R patches. For the case of the varying population size, $J(\mathbf{u})$ is the proportion of new infective cases relative to the scaling parameter $V = \frac{2RB}{\mu}$.

A well-known approach to solving optimal control problems is to first define the Lagrangian of the original, or primal, problem and then consider the adjoint, or dual, problem. This leads in general to a system of coupled differential equations, which sometimes, for simple problems, can be solved explicitly [22]. However, because the constraints in the present problem are highly non-linear and the number of variables is quite high, the above primal-dual approach is difficult to solve analytically [28]. Instead, one could attempt to solve the resulting multipoint boundary-value problem numerically, for example via *indirect* methods such as *multiple shooting* [10, 28].

Other well-known numerical techniques to solve such optimal control problems are NLP and DP. We introduce a simple alternative method by employing a *Cross-Entropy* (CE) technique to solve the problem numerically.

3 Solving Control Problems via CE

In this section we introduce a CE procedure for solving difficult non-linear control problems of the type discussed in the previous section. This is the first reported application of CE to such problems. At the end of this section, we illustrate the approach via a clarifying example.

3.1 CE Method for Optimal Control

The *cross-entropy* (CE) method [24] is a recent Monte Carlo method that has proven to be very successful in solving a wide range of difficult optimization and estimation problems. A gentle tutorial can be found in [8]. The Cross-Entropy (CE) method is an iterative method that involves the following two phases:

1. Generation of a random sample of data (trajectories, vectors, etc.) according to a specified random mechanism.
2. Updating the parameters of the random mechanism, on the basis of the data, in order to produce a “better” sample in the next iteration. This involves a cross-entropy minimization procedure.

Main Procedure

Suppose, we wish to minimize some objective, or performance, function $J(\mathbf{u})$, over a set \mathcal{U} of continuous functions $\mathbf{u} = \{u^{(r)}(t), r = 1, \dots, R, t \in [0, T]\}$, for some finite time T . We can think of \mathcal{U} as a subset of functions in $C^R[0, T]$ (the set of continuous functions from $[0, T]$ to \mathbb{R}^R) that satisfy certain integral and/or dynamic constraints, such as discussed in the previous section. Let us denote the minimum by γ^* , assuming it exists. Thus,

$$\gamma^* = J(\mathbf{u}^*) = \min_{\mathbf{u} \in \mathcal{U}} J(\mathbf{u}). \quad (11)$$

Instead of solving this *functional* optimization program directly, we consider a related *parametric* optimization program, namely

$$\min_{\mathbf{c} \in \mathcal{C}} J(\mathbf{u}_{\mathbf{c}}),$$

where $\mathbf{u}_{\mathbf{c}}$ is a function in \mathcal{U} that is parameterized by some *control vector* $\mathbf{c} \in \mathbb{R}^m$ for some m , and $\mathcal{C} \subset \mathbb{R}^m$ is the collection of such control vectors. Clearly, if the collection $\{\mathbf{u}_{\mathbf{c}}, \mathbf{c} \in \mathcal{C}\}$ is chosen large enough, the solution to the parametric problem should approximate that of the original problem. In particular, let \mathbf{c}^* be the optimal control vector and $\gamma_{\mathbf{c}^*}$ the corresponding optimal value, that is,

$$\gamma_{\mathbf{c}^*} = J(\mathbf{u}_{\mathbf{c}^*}) = \min_{\mathbf{c} \in \mathcal{C}} J(\mathbf{u}_{\mathbf{c}}), \quad (12)$$

then $\gamma^* \approx \gamma_{\mathbf{c}^*}$ and $\mathbf{u}^* \approx \mathbf{u}_{\mathbf{c}^*}$, if the set of parametric control functions can approximate \mathcal{U} well enough. The theoretical convergence properties of the CE method still form an area of active research. Various results proving convergence to the optimal solution can be found in [24], [19] and [7].

A simple way to parameterize the problem is to partition the interval $[0, T]$ into n subintervals $[t_0, t_1], \dots, [t_{n-1}, t_n]$, and let $\mathbf{c} = \{c_i^{(r)}, r = 1, \dots, R; i = 0, \dots, n\}$ be a vector of *control points*, and then define each function $u_{\mathbf{c}}^{(r)}$ via some interpolation of the points $\{(t_i, c_i^{(r)})\}$. The interpolation could, for example, be based on the *finite element method* (FEM) [32], in which case each $u_{\mathbf{c}}^{(r)}$ is of the form

$$u_{\mathbf{c}}^{(r)}(t) = \sum_{i=0}^n c_i^{(r)} w_i(t), \quad (13)$$

where

$$w_i(t) = \begin{cases} \frac{t-t_{i-1}}{t_i-t_{i-1}}, & \text{for } t \in [t_{i-1}, t_i], \\ \frac{t_{i+1}-t}{t_{i+1}-t_i}, & \text{for } t \in [t_i, t_{i+1}], \\ 0, & \text{otherwise.} \end{cases} \quad (14)$$

Remark 3.1 Instead of linear FEM, one can use many different types of *spline* functions, such as the popular cubic *B-Spline*, to represent the control functions, or one can apply the Lagrange interpolating polynomial $L(t)$ of degree $\leq n$, which is given by

$$L(t) = \sum_{j=0}^n l_j(t), \quad \text{where} \quad l_j(t) = a_j \prod_{\substack{i=0 \\ i \neq j}}^n \frac{t - t_i}{t_j - t_i}.$$

For a given control function $\mathbf{u}_{\mathbf{c}}$, the performance value $J(\mathbf{u}_{\mathbf{c}})$ can be evaluated directly by solving numerically the ODE system which describes the dynamic (state) constraints, e.g., via *Runge-Kutta* (RK) techniques. The CE method is employed to find the optimal control vector $\mathbf{c}^* = \{c_i^{(r)*}, r = 1, \dots, R; i = 0, \dots, n\}$. The idea is to generate each control point $c_i^{(r)}$ randomly from a Gaussian distribution with mean $\mu_i^{(r)}$ and standard deviation $\sigma_i^{(r)}$, and to update these parameters via CE to produce better performing control vectors in the next iteration [17]. We summarize the approach in the following algorithm.

Main Algorithm

1. **Initialize:** Choose $\boldsymbol{\mu}_0 = \{\mu_{0i}^{(r)}, r = 1, \dots, R; i = 0, \dots, n\}$ and $\boldsymbol{\sigma}_0 = \{\sigma_{0i}^{(r)}, r = 1, \dots, R; i = 0, \dots, n\}$. Set $k = 1$ (iteration counter).
2. **Draw:** Generate a random sample $\mathbf{C}_1, \dots, \mathbf{C}_M \sim \mathbf{N}(\boldsymbol{\mu}_{k-1}, \boldsymbol{\sigma}_{k-1}^2)$ with $\mathbf{C}_m = \{C_{mi}^{(r)}, r = 1, \dots, R; i = 0, 1, \dots, n\}$. Note that M denotes here the sample size.
3. **Evaluate:** For each control vector \mathbf{C}_m evaluate the objective function $J(\mathbf{u}_{\mathbf{C}_m})$, e.g., by solving the ODE system using RK techniques.
4. **Select:** Find the M_e best performing (=elite) samples, based on the values $\{J(\mathbf{u}_{\mathbf{C}_m})\}$. Let \mathcal{I} be the corresponding set of indices.
5. **Update:** For all $r = 1, \dots, R; i = 0, 1, \dots, n$, let

$$\tilde{\mu}_{ki}^{(r)} = \frac{1}{M_e} \sum_{m \in \mathcal{I}} C_{mi}^{(r)} \quad (15)$$

and

$$\tilde{\sigma}_{ki}^{(r)2} = \frac{1}{M_e} \sum_{m \in \mathcal{I}} \left(C_{mi}^{(r)} - \mu_{ki}^{(r)} \right)^2. \quad (16)$$

6. **Smooth:** For a fixed smoothing parameter $0 < \alpha \leq 1$, let

$$\boldsymbol{\mu}_k = \alpha \tilde{\boldsymbol{\mu}}_k + (1 - \alpha) \boldsymbol{\mu}_{k-1}, \quad \boldsymbol{\sigma}_k = \alpha \tilde{\boldsymbol{\sigma}}_k + (1 - \alpha) \boldsymbol{\sigma}_{k-1} \quad (17)$$

7. **Stop:** Repeat 2–6 until $\max_{i,r} \sigma_{ki}^{(r)} < \varepsilon$. Let L be the final iteration number. Return $\boldsymbol{\mu}_L$ as an estimate of the optimal control parameter \mathbf{c}^* .

Note that the algorithm is completely self-tuning, once the initial parameters, the rarity parameter $\rho = M_e/M$, the smoothing parameter α and the stopping parameter ε are specified. The algorithm is quite flexible under the choice of parameters. Typical values for ρ are 0.01 or 0.1; α usually is chosen in the range 0.6–1. The choice for the initial $\{\mu_{0i}\}$ is quite inconsequential, provided that the $\{\sigma_{0i}\}$ are chosen large enough.

Remark 3.2 If the control problem involves, in addition to dynamic constraints, equality or inequality constraints on the function \mathbf{u} , such as $-1 \leq \mathbf{u}^{(r)} \leq 1$ or $\int_0^T u^{(r)}(t)dt = K$, then, naturally, one would like each random \mathbf{u}_C to satisfy these constraints as well. For the first type of constraints this can be established by *rejecting* samples that do not satisfy the constraints. For equality constraints such as $\int_0^T u^{(r)}(t)dt = K$ such an acceptance-rejection step is not possible. Instead one can choose one of the control points such that the equality constraint holds exactly, where the other control points are chosen adaptively via CE. An alternative approach to deal with constraints on \mathbf{u} is to add a penalty component to the objective function. However, it is often difficult to choose a good penalty function [17].

Remark 3.3 Note that in the above algorithm the choice of the *grid* (or mesh) $\{t_0, \dots, t_n\}$ is assumed to be fixed; only the control points vary. One can easily modify the algorithm to incorporate a grid that gradually refines, in order to obtain a better approximation to \mathbf{u}^* . Starting with a coarse grid initially reduces the simulation effort, and decreasing the mesh size produces a more smooth control function, especially when using linear FEM.

3.2 Example: Optimal Control for the SI Model

The following optimal control example illustrates how the CE technique works and converges to the approximating control function u_{c^*} . It also gives some insight into the numerical results for the more complex models in this paper.

We consider a basic epidemic intervention problem where the evolution of the epidemic is described by a simple continuous-time SI (*susceptible-infective*) model. The objective is to minimize the number of infectives given a limited budget over a fixed time period. This problem is not convex and it is not possible to derive a closed-form solution via the adjoint system (or the standard optimal control theory).

The SI model with a constant population size is defined via the following ODE

$$i'(t) = \beta i(t) (1 - i(t)) - \mu i(t), \quad i(0) = i_0, \quad (18)$$

where $i(t)$ represents the proportion of infectives in the population and μ denotes the death rate of infectives from the population. The proportion of susceptibles is given by $s(t) = 1 - i(t)$. Suppose a fixed amount of budget K is available to control the disease over a fixed time period $[0, T]$ and let the control function $u(t)$ denote the rate at which budget is spent at time t . The infection rate β in (18) is assumed to be a decreasing function of u of the form (7). The problem is to minimize the new infective cases in the given period of time:

$$\min_u J(u) = \min_u \int_0^T \beta(u) i(t) (1 - i(t)) dt, \quad (19)$$

subject to the state constraint (18) and

$$\int_0^T u(t)dt = K, \quad u(t) \geq 0. \quad (20)$$

We assume that initially $i(0) = 0.1$ and we set the other parameters as follows: the time horizon $T = 25$ (years), the death rate $\mu = 0.02$, the available budget $K = 10$, and the maximum infection rate $\beta_{\max} = 0.5$. The infection rate $\beta(u)$ in (19) is defined as

$$\beta(u) = \frac{\beta_{\max}}{1 + u}.$$

Note that if no control is applied ($u = 0$), the infection rate takes its maximum value. When we solve the problem via the CE approach, the approximating control functions for several simulation runs seem to converge to the solution depicted in Figure 3.

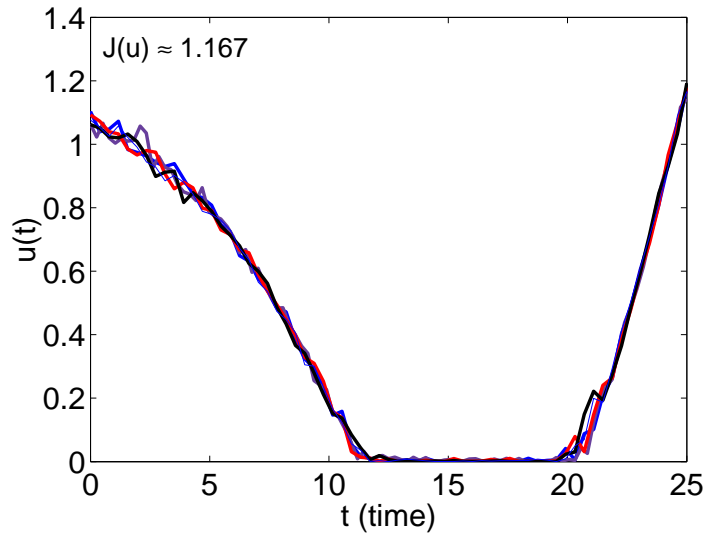


Figure 3: The approximating control function u for the SI control problem, obtained via the CE technique (five different runs).

Figure 4 illustrates how the shape of the final solution depends on the initial proportion of infectives $i(0)$.

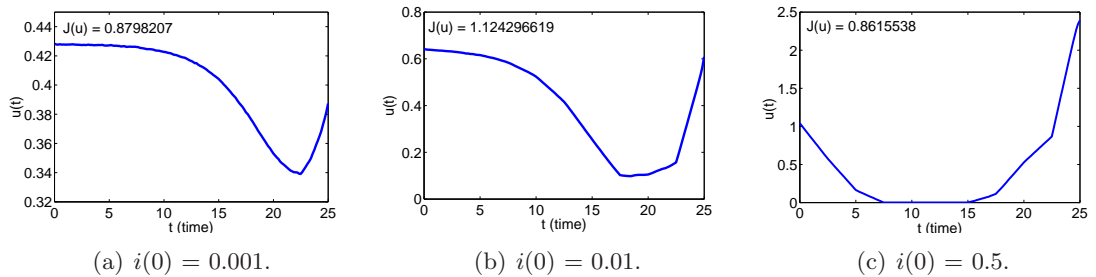


Figure 4: Optimal solutions for three different initial proportions of infectives, $i(0)$.

4 Numerical Experiments and Discussion

In this section, we apply our CE procedure to solve the epidemic control problems introduced in Section 2.2. For clarity, we refer to these epidemic intervention problems as:

- Problem I : the case with inter-patch infection and constant population size,
- Problem II : the case with inter-patch infection and varying population size,
- Problem III: the case with actual mobility.

In the numerical experiments we choose the number of patches $R = 5$ and assume that the initial numbers of infectives and susceptibles are the same for all patches, except for patch 1 where the disease is initially concentrated. In particular, the initial values for the female and male susceptibles are set to $S_F^{(r)}(0) = S_M^{(r)}(0) = 50,000$. The female and male inflow proportions are assumed to be equal, that is, $\xi = 1/2$. We assume that 100 male infectives are initially concentrated in patch 1 only, and no infectives in other patches at time $t = 0$; i.e., $I_M^{(1)}(0) = 100$, $I_F^{(1)}(0) = 0$ and $I_F^{(r)}(0) = I_M^{(r)}(0) = 0$ for $r = 2, \dots, 5$. The available budgets for the patches are $K^{(1)} = 10$, $K^{(2)} = 5$, $K^{(3)} = 2$, $K^{(4)} = 7$, $K^{(5)} = 10$, and the time horizon is fixed at $T = 25$ years. The infection rate is of the form β_2 in (7), with $\beta_{\max}^{(r)} = 0.5$ within each patch and $\beta_{\max}^{(jr)} = 0.3$ from other patches for $j \neq r = 1, \dots, 5$, and $c_2 = 1$ for all patches. Finally, $\mu = 0.02$, $v^{(r)} = 10$ for all r , and $\rho_{ij} = 0.25$ for $i \neq j = 1, \dots, 5$.

4.1 Problem I (Inter-Patch Infection, Constant Population Size)

The shapes of the optimal control functions obtained via CE are depicted in Figure 5. The functions have a “parabolic” shape for all patches (except patch 1 which seems to be “linearly” decreasing) in the interval $[0, 10]$, and are *zero* on $[10, 23]$, followed by a linearly increasing part in $[23, 25]$.

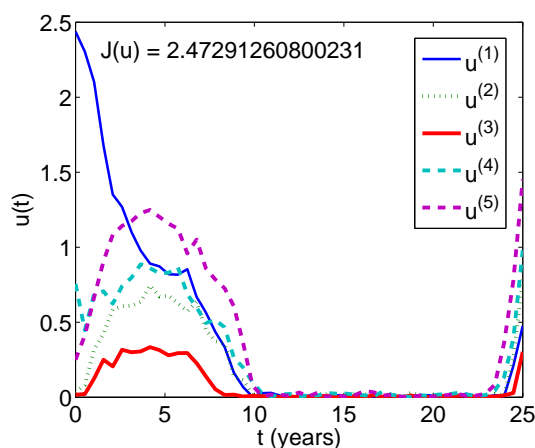


Figure 5: The optimal control functions $u^{(r)}$ of Problem I for patch $r = 1, \dots, 5$ obtained via CE.

Although a little funding should be allocated in the last 2 years, the numerical results via CE suggest that most budget should be concentrated in the first 10 years, but only

a very small portion of the budget is spent in the first one or two years. This result seems reasonable since it might not be worth spending much budget in the first one or two years when no or only a few infectives might be present, except in patch 1.

Comparison with Uniform Strategy

To give some indication of the quality of the CE solution, we compare it with two simple “uniform” strategies, denoted U1 and U2. Given a fixed t_a with $0 < t_a < T$, the U1 scenario means that $u^{(r)}(t) = K^{(r)}/t_a$ for $t \in [0, t_a]$ and $u^{(r)}(t) = 0$ for $t \in (t_a, T]$, and U2 represents $u^{(r)}(t) = 0$ for $t \in [0, T - t_a]$ and $u^{(r)}(t) = K^{(r)}/t_a$ for $t \in [T - t_a, T]$. We vary the variable t_a and solve the corresponding control problem. Figure 6 provides an example how the values of the objective function (8) (the new infective cases relative to the total population size of whole patches) vary for these two scenarios.

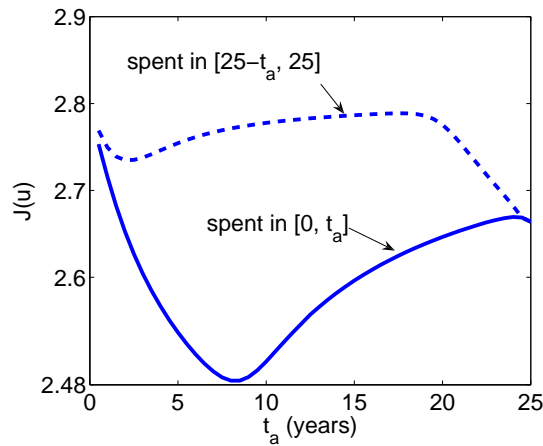


Figure 6: The value of the objective function $J(\mathbf{u})$ for the two uniform strategies, as a function of t_a . The *solid* line corresponds to U1 and the *dashed* line to U2.

We see from Figure 6 that, for this particular problem, policy U1 is always better than U2. The best U1 policy is the one with $t_a \approx 9$ (i.e., spend all budget uniformly and simultaneously in the first 9 years for all patches). However, the best objective value for these scenarios (the bang-bang like structure strategies) is $J(u) > 2.48$ which is greater than that obtained by CE.

Varying Time Horizon T

It is interesting to examine how the shape of the control function changes if one varies the time period T , while keeping other parameters unchanged. Figure 7 shows that the optimal solutions via CE are qualitatively quite different. For $T \leq 10$, all budget is spent almost uniformly in the interval $[0, T]$ and for $T > 10$, the optimal solution has a similar structure to that in Figure 5.

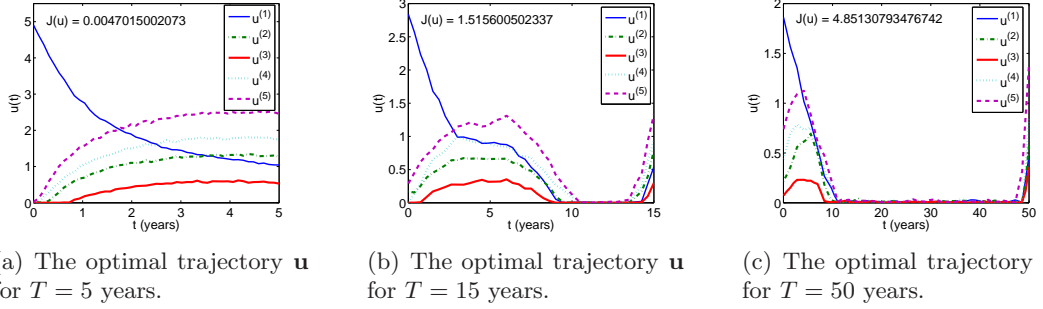


Figure 7: The optimal trajectory \mathbf{u} for Problem I obtained via the CE method as the time horizon T is varied.

Varying Budget $K^{(r)}$

To see how the optimal solution might depend on the amount of available budget for a fixed time horizon $T = 25$, we keep the initial values and other parameters unchanged from the previous situation and vary only $K^{(r)}$. In particular, the available budget is decreased or increased by the same amount for each patch. The changes in the optimal solution are depicted in Figure 8.

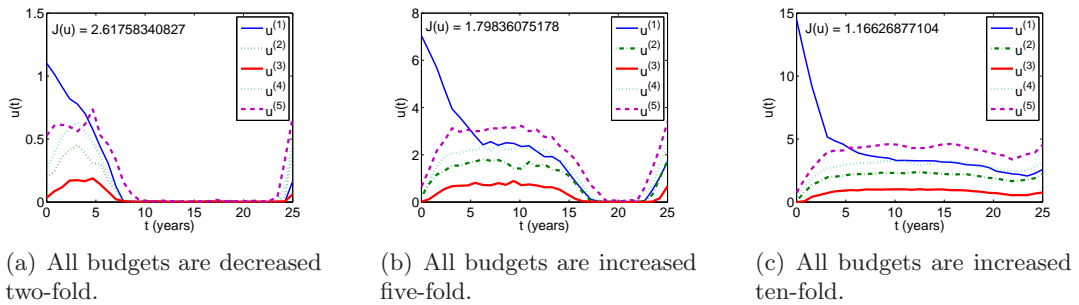


Figure 8: The optimal trajectory \mathbf{u} for Problem I obtained via the CE method as the budgets $K^{(1)} = 10, K^{(2)} = 5, K^{(3)} = 2, K^{(4)} = 7,$ and $K^{(5)} = 10$ are multiplied by a factor $1/2, 5$ and 10 .

Notice that if all budgets are decreased by a factor of 2 or increased by a factor of 5 (see Figure 8(a)-(b)), the shape of $\mathbf{u}_{\mathbf{c}^*}$ qualitatively remains the same as that in Figure 5. However, if all budgets are increased 10-fold, the optimal solution $\mathbf{u}_{\mathbf{c}^*}$ becomes similar to that in Figure 7(a).

To further examine the effect of $K^{(r)}$ on $\mathbf{u}_{\mathbf{c}^*}$, we increase the available budget only in patches 3 and 5. Figure 9 shows that the shapes of the control function for the other patches, $u^{(1)}, u^{(2)},$ and $u^{(4)}$, remain the same as in Figure 5.

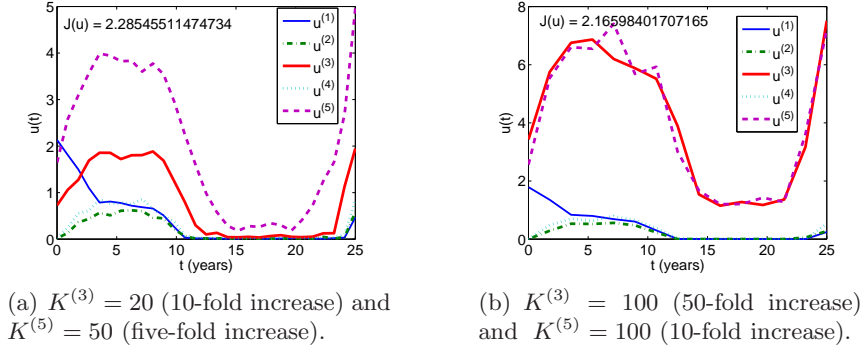


Figure 9: The optimal trajectory \mathbf{u} for Problem I obtained via the CE method as the budgets $K^{(1)} = 10, K^{(2)} = 5, K^{(3)} = 2, K^{(4)} = 7$, and $K^{(5)} = 10$ are varied.

Although the budgets $K^{(3)}$ in Figure 9(a) and $K^{(5)}$ in Figure 9(b) are both increased 10-fold compared to those in Figure 8(c), the numerical results suggest that the shapes of $u^{(3)}$ and $u^{(5)}$ are qualitatively different from those in 8(c), where *all* budgets are increased 10-fold.

Varying Initial Values $I_F^{(r)}(0)$ and $I_M^{(r)}(0)$

To see how the optimal solution depends on the initial values, we change the initial numbers of infectives in the above experiments. The initial numbers of susceptibles are kept the same as those in the previous experiments, i.e., $S_F^{(r)}(0) = S_M^{(r)}(0) = 50,000$. We assume that initially in each patch an equal number of male and female infectives are present. We consider three scenarios: (1) the number of infectives is *small* compared to the number of susceptibles, (2) the number of infectives is *equal* to the number of susceptibles, (3) the number of infectives is *large* compared to the number of susceptibles. In particular, the number of male/female infectives are 100, 50,000 and 450,000, respectively. The results are depicted in Figure 10. Note that the last scenario (Figure 10(c)) shows the optimal solution when the control is introduced in the endemic equilibrium, that is, $S_F^{(r)}(0) = S_M^{(r)}(0) = 50,000, I_F^{(r)}(0) = I_M^{(r)}(0) = 450,000$.

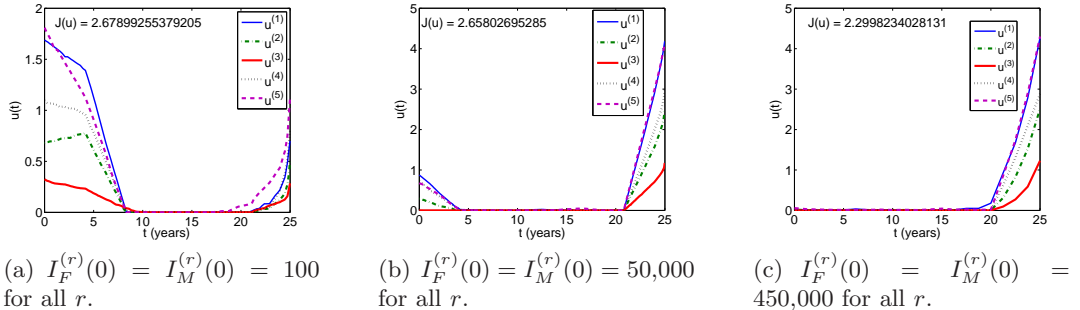


Figure 10: The optimal trajectory \mathbf{u} for Problem I obtained via the CE method for three different scenarios for the number of initial infectives.

We see that now in every patch the control function initially decreases “linearly”, just as was the case for patch 1 in the original scenario of Figure 5, where the infectives are initially concentrated only in patch 1. Interestingly, if the initial number of infectives is large, it is better to concentrate the funding over the last part of the time period.

Changing the value of β_{\max}

Another possible factor influencing the shape of the control function is the value of β_{\max} — the rate of HIV transmission in the absence of any control. This parameter depends on the number of sexual contacts among partners and the probability of infection per sexual contact [9].

Suppose that β_{\max} is decreased by a factor of 10 compared with the previous experiments. Thus, take $\beta_{\max}^{(r)} = 0.05$ within each patch for $r = 1, \dots, 5$, and $\beta_{\max}^{(jr)} = 0.03$ from other patches for $j \neq r = 1, \dots, 5$. These values of β_{\max} are perhaps more realistic in actual HIV cases.

Figure 11 shows that the decrease of β_{\max} , when compared with the original setting of Figure 5, results in a significantly different shape for the controls. In particular, the controls are now spread out over a much larger time period. Note also that the optimal objective function value $J(\mathbf{u}^*)$ is much smaller than in Figure 5.

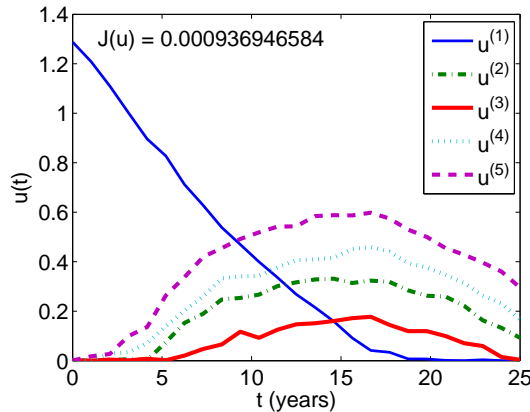


Figure 11: The optimal trajectory \mathbf{u} for Problem I obtained via the CE method, using $\beta_{\max}^{(r)} = 0.05$ within each patch and $\beta_{\max}^{(jr)} = 0.03$ from other patches.

4.2 Problem II (Inter-Patch Infection, Varying Population Size)

For the numerical experiments of Problem II, we use the same initial values and parameters as in Problem I. The flow rates of susceptibles are set to be equal to $B = 1000$ for each patch, which gives the scaling parameter $V = 5 \times 10^5$ (approximately equal to the whole initial population size).

The optimal solution via CE method suggests a much similar result to that in Problem I, namely most portion of the budget should be spent approximately in the first 9.5 years for all patches. Moreover, the control functions in $[0, 9.5]$ have again a “parabolic” shape, except for patch 1 where the infectives are initially concentrated.

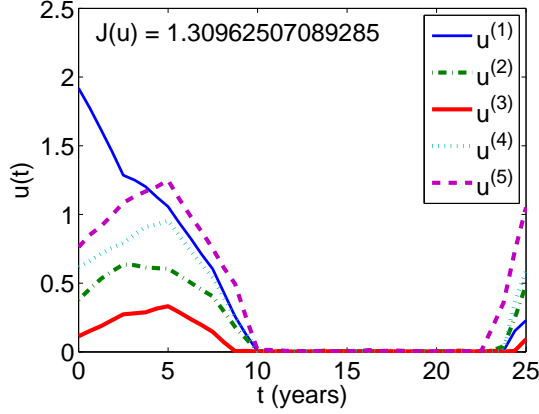


Figure 12: The optimal trajectory \mathbf{u} for Problem II obtained via the CE method.

Thus, the best strategy is to allocate the budget in the first 9.5 years, with “parabolic” control functions for all patches except for patch 1 which has a “linearly” decreasing control function, spend nothing for the next 13 years, and spend the rest of budget in the last 2.5 years in a linearly increasing way, for all patches.

When varying the time horizon T or the budget K in Problem II, the optimal trajectories \mathbf{u} obtained via CE behave much the same as those in Problem I, as shown in Figure 13.

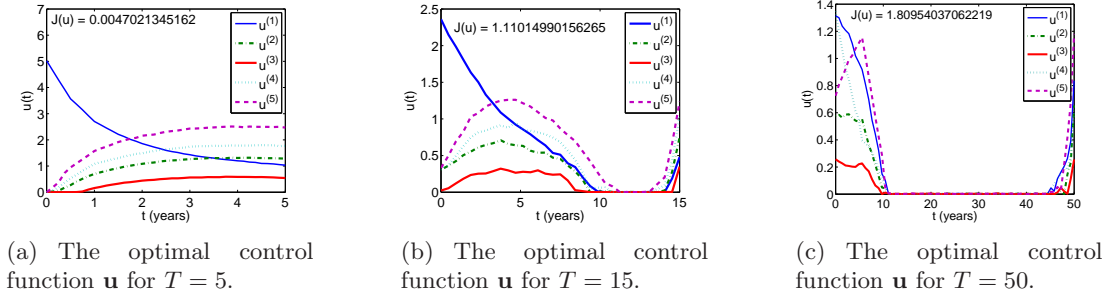


Figure 13: The optimal trajectory \mathbf{u} for Problem II obtained via the CE method, for several different time horizons T .

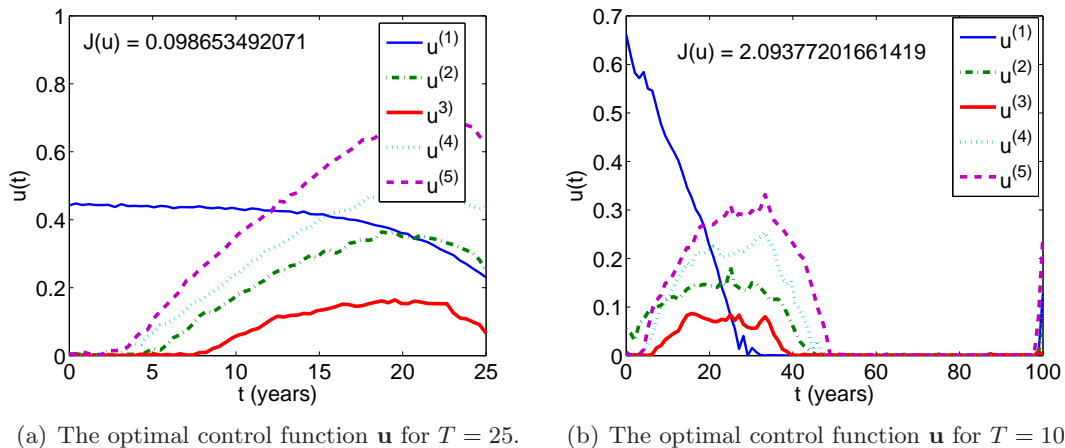
Thus, the optimal strategies suggested via CE for Problem I and Problem II are quantitatively similar. The shapes of the control functions in both models look similar, that is, a parabolic-like function in the first 9.5 years, a zero function for the next 13 years, followed by an linearly increasing function in the last 2.5 years.

Our numerical results suggest that allowing for some reallocation of resources over the time horizon of the problem, rather than allocating resources just once at the beginning (or at the last) of the time horizon, can quantitatively lead to significant decreases in the number of new infectives.

4.3 Problem III (Actual Mobility)

For the numerical experiments of Problem III, we need to specify migration rates of individuals among patches. We set the same migration rate to $v^{(r)} = 10$ per unit time, for all patches, $r = 1, \dots, 5$. We assume that the probability that an individual leaving a patch i visits patch j is equal for all patches, so that $\rho_{ij} = \frac{1}{R-1} = 0.25$, $i \neq j$. Thus, the rate at which individuals leaving a patch i visit patch j is $v_{ij} = \rho_{ij} v^{(i)} = 2.5$. Initial values and other parameters remain the same as specified in the previous simulations for Problems I and II.

Solving the problem via CE gives a result as depicted in Figure 14. Note, however, that the optimal curves are, except for patch 1, significantly lower at the beginning and end of the time period. The optimal controls for Problem III appear to have a different structure to those for Problems I and II. However, if we extend the time scale, say $T = 100$, see Figure 14, we obtain a similar structure as in the previous problems.



(a) The optimal control function \mathbf{u} for $T = 25$. (b) The optimal control function \mathbf{u} for $T = 100$.

Figure 14: The optimal trajectory \mathbf{u} for Problem III obtained via the CE method for 5 patches.

5 Conclusions and Future Research

In this paper, we have formulated various epidemic intervention models for the spread of HIV in multiple populations and introduced a new CE technique to solve these problems. The numerical results indicate that the shapes of the control functions for the different models are qualitatively similar. If the budget is limited and the time horizon large, the control functions for the patches that are initially free of infectives exhibit a mutual structure: the first part of the control function is concave, starting and ending at zero, the second part is zero (no control), and the third part is linearly increasing. However, the control shapes depend also importantly on the time horizon, available budget, and the initial number of infectives. The numerical experiments suggest that the controls for the different patches are highly synchronized. Moreover, they indicate that the optimal trajectories qualitatively have similar form. Considering other control parameters such as isolating infectives, HIV/AIDS campaign programming, contact tracing, etc., in a more complex model with an age structure, risk groups and levels of infectivity could also be a

future study. Another possible direction for future research is to formulate the epidemic intervention as a Continuous-Time Markov Decision Process and compare the results to those in this study. Finally, since the application of the CE technique to optimal control problems is relatively new, more analytical investigations in this area would be welcome.

Acknowledgments

The first author is grateful to ADS-AUSAID for funding his PhD scholarship. The work of the second author is supported by the Australian Research Council (Discovery Grant DP0558957).

References

- [1] R. S. Berstein, D. C. Sokal, S. T. Seitz, B. Auvert, J. Stover, and W. Naamara. Simulating the control of a heterosexual HIV epidemic in a severely affected East African city. *Interfaces*, 28(3):101–126, 1998.
- [2] D. P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, Belmont, Massachusetts, 2nd edition, 2002.
- [3] S. Blount, A. Galambosi, and S. Yakowitz. Nonlinear and dynamic programming for epidemic intervention. *Appl. Math. Comput.*, 86(2-3):123–136, 1997.
- [4] S. Blount and S. Yakowitz. A computational method for the study of stochastic epidemics. *Math. Comput. Modelling*, 32(1-2):139–154, 2000.
- [5] M. L. Brandeau, G. S. Zaric, and A. Ritcher. Optimal resource allocation for epidemic control among multiple independent populations: Beyond cost effectiveness analysis. *Journal of Health Economics*, 22(4):575–598, 2003.
- [6] D. Clancy and A. B. Piunovsky. An explicit optimal isolation policy for a deterministic epidemic model. *Appl. Math. Comput.*, 163(3):1109–1121, 2005.
- [7] A. Costa and J. Owen and D. P. Kroese. Convergence properties of the cross-entropy method for discrete optimization *Operations Research Letters*, 5:573–580, 2007.
- [8] P-T. De Boer, D.P Kroese, S. Mannor, and R.Y. Rubinstein. A tutorial on the cross-entropy method. *Annals of Operations Research*, 134:19–67, 2005.
- [9] K. Dietz. On the transmission dynamics of HIV. *Math. Biosci.*, 90:397–414, 1988.
- [10] G. Fraser-Andrews. A Multiple-shooting technique for optimal control. *Journal of Optimization Theory and Applications*, 102 (2):299–313, 1999.
- [11] David Greenhalgh. Stochastic processes in epidemic modelling and simulation. In *Stochastic processes: modelling and simulation*, volume 21 of *Handbook of Statist.*, pages 285–335. North-Holland, Amsterdam, 2003.
- [12] G. Hugo. Indonesia, internal and international population mobility: implications for the spread of HIV/AIDS. Technical report, University of Adelaide, Australia, November 2001.

- [13] J. M. Hyman, J. Li, and E. A. Stanley. Modeling the impact of random screening and contact tracing in reducing the spread of HIV. *Math. Biosci.*, 181(1):17–54, 2003.
- [14] E. Jung, S. Lenhart, and Z. Feng. Optimal control of treatments in a two-strain tuberculosis model. *Discrete and Continuous Dynamical Systems-Series B*, 2(4):473–482, 2002.
- [15] D. Kirschner, S. Lenhart, and S. Serbin. Optimal control of the chemotherapy HIV. *Journal of Mathematical Biology*, 35:775–792, 1997.
- [16] M. Kretzschmar, Y.T.H.P. van Duynhoven, and A.J. Severijnen. Modeling prevention strategies for gonorrhea and chlamydia using stochastic network simulations. *Am. J. Epidem.*, 144:306–317, 1996.
- [17] D.P. Kroese, S. Porotsky, and R.Y. Rubinstein. The cross-entropy method for continuous multi-extremal optimization. *Methodology and Computing in Applied Probability.*, 8:383–407, 2006.
- [18] P. Lory. Enlarging the domain of convergence for multiple shooting by the homotopy method. *Numerische Mathematik.*, 35:231–240, 1980.
- [19] L. Margolin. On the convergence of the cross-entropy method. *Annals of Operations Research.*, 134: 201–214, 2005.
- [20] R. M. May, R. M. Anderson, and A. R. McLean. Possible demographic consequences of HIV/AIDS epidemics. II. Assuming HIV infection does not necessarily lead to AIDS. In *Mathematical Approaches to Problems in Resource Management and Epidemiology (Ithaca, NY, 1987)*, volume 81 of *Lecture Notes in Biomath.*, pages 220–248. Springer, Berlin, 1989.
- [21] C. J. Mode and C. K. Sleeman. *Stochastic Processes in Epidemiology, HIV/AIDS, Other Infectious Diseases and Computers*. World Scientific, 2000.
- [22] D. R. Pinch. *Optimal Control and the Calculus of Variations*. Oxford Science Publications, 2002.
- [23] A. Richter, M.L. Brandeau, and D.K Owens. An analysis of optimal resource allocation for HIV prevention among injection drug users and nonusers. *Medical Decision Making*, 19:167–179, 1999.
- [24] R.Y. Rubinstein and D.P. Kroese. *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte- Carlo Simulation, and Machine Learning*. Springer-Verlag, 2004.
- [25] A. Sani, D.P. Kroese, and P.K. Pollett. Stochastic models for the spread of HIV in a mobile heterosexual population. *Math. Biosci.*, 208:98–124, 2007.
- [26] R. B. Schinazi. On the role of social cluster in the transmission of infectious diseases. *Theoret. Population Biol.*, 61:163–169, 2002.
- [27] M. Sniedovich. *Dynamic Programming*. Marcel Dekker, Inc., 1992.

- [28] O. von Stryk and R. Bulirsch. Direct and indirect methods for trajectory optimization. *Ann. Oper. Res.*, 37:357–373, 1992.
- [29] H. J. Whitaker and C. P. Farrington. Infection with varying contact rates: Application to varicella. *Biometrics*, 60(3):615–623, 2004.
- [30] K. Wickwire. Mathematical models for the control of pests and infectious diseases: a survey. *Theoret. Population Biology*, 11(2):182–238, 1977.
- [31] S. Woods, editor. *Report on the Global AIDS Epidemic: 4th Global Report*. Joint United Nations Programme on HIV/AIDS (UNAIDS), 2004.
- [32] O. C. Zienkiewicz and R. L. Taylor. *The Finite Element Method*. London; Boston: Butterworth-Heinemann, 5th edition, 2000.