https://doi.org/10.1142/9789811263033_0009

Chapter 9

Solving Singular Control Problems in Mathematical Biology Using PASA

Summer Atkins[*,§], **Mahya Aghaee**[†,¶], **Maia Martcheva**[‡,‖], **and William Hager**[‡,**]

[*] *Department of Mathematics, Louisiana State University, Baton Rouge, LA, USA*
[†] *Laboratory for Systems Medicine, College of Medicine, University of Florida Gainesville, Florida, USA*
[‡] *Department of Mathematics, University of Florida Gainesville, Florida, USA*
[§] *srnatkins@lsu.edu*
[¶] *mahya.aghaee@medicine.ufl.edu*
[‖] *maia@ufl.edu*
[**] *hager@ufl.edu*

In this chapter, we demonstrate how to use a nonlinear polyhedral constrained optimization solver called the Polyhedral Active Set Algorithm (PASA) for solving a general singular control problem. We present a method for discretizing a general optimal control problem involving the use of the gradient of the Lagrangian for computing the gradient of the cost functional so that PASA can be applied. When a numerical solution contains artifacts that resemble "chattering," a phenomenon where the control oscillates wildly along the singular region, we recommend a method of regularizing the singular control problem by adding a term to the cost functional that measures a scalar multiple of the total variation of the control, where the scalar is viewed as a tuning parameter. We then demonstrate PASA's performance on three singular control problems that give rise to different applications of mathematical biology. We also provide some exposition on the heuristics that we use in determining an appropriate size for the tuning parameter.

Keywords: Singular control, Total variation, Bounded variation Regularization, Pontryagin's Minimum Principle, Switching function, Fishery problem, Plant problem, SIR problem

9.1 Introduction

Optimal control theory is a tool that is used in mathematical biology for observing how a dynamical system behaves when employing one or many variables that can be controlled outside of that system. Mathematical biologists apply optimal control theory to disease models of immunologic and epidemic types,[1–4] to management decisions in harvesting,[5,6] and to resource allocation models.[7,8] In practice, mathematical biologists tend to construct optimal control problems with quadratic dependence on the control. Problems of this structure are well-behaved in the sense that there are established methods of proving existence and uniqueness of an optimal control.[9–11] In addition, for numerically solving problems of this form, many employ the forward–backward sweep method, a numerical procedure presented in Lenhart and Workman's book[12] that involves combinations of the forward application and the backward application of a fourth-order Runge–Kutta method. We direct the reader to Refs. 13–16 which are excellent surveys of other numerical procedures, such as gradient methods, quasi-Newton methods, shooting methods, and collocation methods, that are used within the optimization community for solving optimal control problems.

Control problems in biology tend to depend quadratically with respect to the control due to the construction of the objective or cost functional, which is the functional that is being optimized with respect to the control variables. The construction of the objective functional is an essential component to optimal control theory because it measures our criteria for determining what control strategy is deemed "best." In mathematical biology, the costs are frequently nonlinear and depend on the states and the controls, and a cost term for a particular control may be the sum of a bilinear term in that control and one for the states and a quadratic term in the control. Frequently, the quadratic term has a lower coefficient.

With regard to the principle of parsimony, it is difficult to justify the use of a quadratic term for representing the cost of administering a control. A linear term would be a more realistic representation of the cost of applying a control. For control problems that are linear in the control, it is possible to obtain a solution that is piecewise constant where the constant values correspond to the bounds of the control. An optimal control of this structure, which is often called a "bang-bang" control, can be readily interpreted and implemented. These characteristics compel many (see Refs. 5,6,8 and 17–20), to use control problems with linear dependence on the control for biological models.

There are however evident setbacks to using optimal control problems in which the control appears linearly. For one thing, these problems are much more difficult to solve analytically due to the potential existence of a singular subarc. As demonstrated in Refs. 5,6 and 18, procedures for obtaining an explicit formula for the singular case involve taking one or many time derivatives of the switching function as a means to gain a system of equations. Often the first-order and second-order necessary conditions for optimality, which are respectively named Pontryagin's Maximum Principle[10] and the Generalized Legendre–Clebsch Condition/Kelley's Condition,[21–24] are checked. Naturally, methods for explicitly solving singular control problems increase in difficulty when multiple state variables and multiple control variables are involved.

Additionally, numerically solving singular control problems are inevitably problematic. If the parameters of an optimal control problem that is linear in the control are set to where all optimal control variables are bang-bang, then the forward–backward sweep method[12] can successfully run. However, if the presence of a singular control is a possibility, then the forward–backward sweep method is not advisable. In Foroozandeh *et al.*,[25] Foroozandeh, De Pinho, Shamsi test four numerical solvers, including the Imperial College London Optimal Control Software (ICLOCS)[26] and the Gauss Pseudospectral Optimization Software (GPOPS),[27] by solving a singular optimal control problem for Autonomous Underwater Vehicles (AUV). They find that three of these methods have difficulty in detecting the structure of the optimal control and in accurately computing the switching points without *a priori* information. Switching points are the corresponding points in time when an optimal control switches from singular to non-singular and vice versa. When solving for the AUV problem, both ICLOCS and GPOPS obtain a control that exhibits oscillations within the singular region, causing both methods to be unable to provide direct information about the switching points. Foroozandeh *et al.* conclude that only the mixed binary nonlinear programming method (MBNLP)[28] is successful in accurately approximating the optimal control to the AUV problem and its switching points.

One of the predominate issues associated with singular control problems is the concept of "chattering," which is also known as the "Fuller Phenomenon" (see Ref. 24). As mentioned in Zelikin and Borisov's book,[24] an optimal control is said to be *chattering* if the control oscillates infinitely many times between the bounds of the control over a finite region. It is thought that such an event occurs in singular control problems when the optimal control has a singular subarc that cannot be directly joined to a non-singular (bang) region without promoting oscillations. In MacDanell and Powers,[21] MacDanell and Powers present some necessary conditions for joining singular and nonsingular subarcs. Zelikin and Borisov[24] mention other theorems that can be used to verify when chattering is present. It is possible that the discretization of an optimal control problem causes a numerical method to generate numerical artifacts that resemble chattering even though the optimal control does not chatter. However, it is difficult to determine when a numerical solution is exhibiting many oscillations due to chattering or due to numerical artifacts.

Regardless of the situation, it is evident that a wildly oscillating optimal control is an unrealistic procedure to implement. A way to bypass this issue is to solve for a regularized version of the optimal control problem. In Ref. 29, Yang *et al.* present methods of regularizing optimal control problems by adding a penalty term to the cost functional of the original problem. The penalization terms suggested in Yang *et al.*[29] are: a weighted parameter times the L^2 norm of the control, a weighted parameter times the L^2 norm of the derivative of the control, and a weighted parameter times the L^2 norm of the second derivative of the control. In Ding and Lenhart,[6] Ding and Lenhart employ a penalty term to a harvesting optimal control problem to avoid a potential chattering result found in the control h. The penalty term that was applied for this problem consisted of a weighted parameter times $\|\nabla h\|^2_{L^2}$. The penalty term that is used in Ding and Lenhart[6] adds convexity properties to the control problem which can be beneficial for verifying existence and uniqueness of an optimal control. Additionally, $\|\nabla h\|_{L^2}^2$ can be discretized to where it can serve as a crude estimate for the total variation of the control. However, Lenhart and Ding[6] need to use variational inequalities to solve for their problem, and incorporating such a penalty term restricts their set of admissible controls into a functional space that requires its controls to be differentiable.

The penalty term that we believe to show the most promise has been recently suggested in Capognigro, Ghezzi, Piccoli, and Trelat's work.[30] In Ref. 30, recommend a method of regularizing chattering in optimal control problems by adding a penalty term that represents a tuning parameter times the total variation of the control. Total variation regularization or bounded variation regularization influences the numerical solution in a way that reduces the number of oscillations. In Caponigro *et al.*,[30] this penalty is applied to the Fuller problem, which is the classical example that introduced the concept of chattering, and they obtain a "quasi" optimal solution to the Fuller problem that does not chatter. Caponigro *et al.* also prove that the optimal value of the penalized problem converges to the associated optimal value of the original problem as the penalty weight parameter p tends to zero.

In this chapter, we demonstrate how to use a nonlinear polyhedral constrained optimization solver called PASA,[1] developed by Hager and Zhang,[31] to solve a general singular control problem that is being regularized by use of a total variation term.[30] We recommend PASA because it is user friendly to those who are not as acquainted with optimization techniques for optimal control problems, and it is freely accessible to use on MATLAB for Linux and Unix operating systems. According to Hager and Zhang,[31] PASA consists of two phases, with the first phase being the gradient projection algorithm and the second phase being any algorithm that is used for solving linearly constrained optimization problems. The gradient projection algorithm is an optimization solver commonly used for bounded constrained optimization problems. When applying gradient descent to a bounded constrained optimization problem, it is possible to obtain

[1]To access PASA software that can be used on MATLAB for Linux and Unix operating systems, download the SuiteOPT Version 1.0.0 software given on https://people.clas.ufl.edu/hager/software. For future reference, any updates to the software will be uploaded to this link, and to access older versions of the software, use the same link and then select "Software Archive."

an iterate that lies outside of the feasible set due to the negative direction of the gradient. Projected gradient method, takes this issue into consideration by adding additional steps involving projecting points outside the feasible set onto the feasible set. For more information on projected gradient methods we direct the reader to Refs. 31–36. Using PASA to solve optimal control problems involves converting an optimal control problem into a discretized optimization problem.

In this chapter, the discretization for the general singular control problem involves using explicit Euler's method for the state equations, and left-rectangular integral approximation for the original cost functional. Additionally, we need to ensure that the discretized and regularized cost functional is differentiable, which requires performing a decomposition of the absolute value terms associated with the total variation penalty. We use the gradient of the Lagrangian function of the discretized optimal control problem for computing the gradient to the discretized cost functional. Conveniently, the process of ensuring that the gradient of the Lagrangian is equal to the gradient of the discretized objective functional yields a discretization procedure for the adjoint equations.

Further, we demonstrate PASA's performance on three singular control problems that are being regularized via bounded variation. The order of these examples increases in difficulty based upon the number of state variables and the number of control variables. Additionally, each example gives rise to different applications of mathematical biology. Explicit formulas for the singular case and for the switching points are given for the first two examples, which allow us to compare PASA's numerical results with the exact solution. For each problem, we illustrate the discretization process and then present numerical results that were obtained when solving for both the unpenalized and penalized problem. We also provide some exposition on the heuristics that we use in determining an appropriate size for the tuning parameter ρ.

The first example is a fishery harvesting problem that was originally presented in Clark's[37] and later restated in Lenhart and Workman's.[12] The fishery problem consists of one state variable and one control variable, where the state variable represents the fish population and the control variable represents harvesting effort. In Lenhart and Workman,[12] an explicit formula for the singular case is obtained by using Pontryagin's Maximum Principle[10]; however, forward–backward sweep is unable to solve for the problem whenever parameters are set to ensure existence of a singular subarc.

The second example is from King and Roughgarden's work,[8] and the optimal control problem is a resource allocation model for studying an annual plant's allocation procedure in distributing photosynthate. The control problem consists of two state variables and one control variable. One variable measures the weight of the components of a plant that correspond to vegetative growth, while the other variable measures the weight of the components of a plant that correspond to reproductive growth. The control variable used in this problem represents the fraction of photosynthate being reserved for vegetative growth. King and Roughgarden use Pontryagin's Maximum Principle[10] and find conditions based upon the parameters of the problem for determining when the optimal control would be bang–bang or concatenations of singular and bang control. They verify that their explicit formula for the singular subarc satisfies the generalized Legendre–Clebsch Condition and the strengthened Legendre–Clebsch Condition.[21–24] Additionally, King and Roughgarden use one of MacDanell and Power's Junction theorems[21] to show the optimal control to the problem satisfies the necessary conditions for joining singular and non-singular subarcs. When using PASA to solve for this plant problem, we only need to use the regularization term for a degenerate case of the problem.

The final example, which is from Ledzewicz, Aghaee, and Schättler's,[17] is an optimal control problem where the three state variables correspond to an SIR model with demography. An SIR model is a compartmental model that is used for modeling the spread of an infectious disease in a population, where the population is divided into the following three classes: (1) S is the class of individuals who are susceptible to the disease; (2) I is the class of infected individuals who are assumed to be infectious; and (3) R is the class of individuals who have recovered from the disease and are considered immune to the disease. For books covering mathematical models for epidemiology, we recommend Brauer and Castillo-Chavez's[38] and Martcheva's.[39] The optimal control problem used in Ledzewicz *et al.*[17] consists of two control variables where one control represents vaccination while the other represents treatment. Ledzewicz *et al.* numerically solve this problem with parameters set to where a singular subarc is present in the optimal vaccination strategy while the optimal treatment strategy obtained appears to be bang-bang. When using PASA, we regularize only the vaccination control via bounded variation since the treatment control contains no oscillations. In the Appendix section, we provide the MATLAB code that was used for solving the last example to illustrate how to write up these optimal control problems.

9.2 Discretization of the Regularized Control Problem via Bounded Variation

The following problem is the optimal control problem of interest:

$$\begin{aligned} \min_{u\in\mathcal{A}} \quad & J(\boldsymbol{u}) = \int_0^T g(\boldsymbol{x}(t),\boldsymbol{u}(t))\mathrm{d}t \\ \text{sub. to} \quad & \dot{x}_i(t) = f_i(\boldsymbol{x}(t),\boldsymbol{u}(t)) \quad \text{for all } i=1,\ldots,n, \\ & x_i(0) = x_{i,0}, \quad \text{for all } i=1,\ldots,n, \\ & \xi_j \le u_j(t) \le \omega_j \quad \text{for all } t\in[0,T] \quad \text{and} \quad \text{for all } j=1,\ldots,m, \end{aligned} \tag{9.1}$$

where functions $f_1,\ldots,f_n$, and g are assumed to be continuously differentiable in all arguments. In addition, we assume that if any of the m control variables u_i appear in functions $f_1,\ldots,f_n$ and g, then u_i appears linearly. The class of admissible controls is the following:

$$\mathcal{A} = \{\boldsymbol{u}\in(L^1(0,T))^m | \boldsymbol{\xi}\le\boldsymbol{u}(t)\le\boldsymbol{\omega} \quad \text{for all } t\in[0,T]\}.$$

We assume that the conditions for the Filippov–Cesari Existence Theorem[9] hold for problem (9.1). State vector $\boldsymbol{x}\in\mathbb{R}^n$ consists of n state variables that satisfy the state equations and the initial conditions given in problem (9.1).

The common procedure for solving problem (9.1) is employing Pontryagin's Maximum Principle[10] to generate the first-order necessary conditions for optimality. We first define the Hamiltonian function to problem (9.1), $H(\boldsymbol{x},\boldsymbol{u},\boldsymbol{\lambda})$, as

$$\begin{aligned} H(\boldsymbol{x},\boldsymbol{u},\boldsymbol{\lambda}) &= g(\boldsymbol{x},\boldsymbol{u}) + \boldsymbol{\lambda}^T\boldsymbol{f}(\boldsymbol{x},\boldsymbol{u}) \\ &= g(\boldsymbol{x},\boldsymbol{u}) + \sum_{i=0}^{n}\lambda_i(t)f_i(\boldsymbol{x},\boldsymbol{u}), \end{aligned}$$

where $\boldsymbol{\lambda} \in \mathbb{R}^n$ is the adjoint vector and the superscript T means transpose. We have from Pontryagin's Maximum Principle[10] that if $\boldsymbol{u}^*$ is the optimal control to problem (9.1) with corresponding trajectory $\boldsymbol{x}^*$, then there exists a non-zero adjoint vector $\boldsymbol{\lambda}^*$ that is a solution to the following adjoint system:

$$\dot{\lambda}_\ell(t) = -\frac{\partial H(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\lambda})}{\partial x_\ell}, \quad \text{for all } \ell = 1, \ldots, n,$$
$$\lambda_\ell(T) = 0, \quad \text{for all } \ell = 1, \ldots, n,$$

and satisfies

$$H(\boldsymbol{u}^*, \boldsymbol{x}^*, \boldsymbol{\lambda}^*) = \min_{\boldsymbol{u} \in \mathcal{A}} H(\boldsymbol{u}, \boldsymbol{x}^*, \boldsymbol{\lambda}^*).$$

Using our definition of the Hamiltonian function, the adjoint equations are

$$\dot{\lambda}_\ell(t) = -\frac{\partial g(\boldsymbol{x}, \boldsymbol{u})}{\partial x_\ell} - \sum_{i=1}^{n} \lambda_i(t) \frac{\partial f_i(\boldsymbol{x}, \boldsymbol{u})}{\partial x_\ell} \quad \text{for all } \ell = 1, \ldots, n, \tag{9.2}$$

with transversality conditions being

$$\lambda_\ell(T) = 0, \quad \text{for all } \ell = 1, \ldots, n. \tag{9.3}$$

Based on assumptions of functions g and $f_1, \ldots, f_n$ in problem (9.1), the Hamiltonian is linear in the control.

For demonstration purposes, we assume that every component of control vector $\boldsymbol{u}$ in problem (9.1) needs to be regularized via bounded variation.[30] This means that when numerically solving problem (9.1) without this regularization term, we obtain oscillatory numerical artifacts or other unwarranted numerical artifacts in each control. To regularize problem (9.1) via bounded variation, we introduce a tuning vector $\boldsymbol{\rho}$ where $0 \leq \rho_j < 1$ for all $j = 1, \ldots, m$, and we present Conway's[40] definition of the total variation function V of a real or complex valued function u that is defined on the interval $[a, b]$. Let $\mathbb{K}$ be the field of complex numbers or the field of real numbers. Given $u : [a, b] \to \mathbb{K}$, the total variation of u on $[a, b]$ is defined to be

$$V(u) = \sup_{\mathcal{P}} \sum_{k=0}^{N_P - 1} |u(t_{k+1}) - u(t_k)|,$$

where $\mathcal{P} = \{P = \{t_0, t_1, \ldots, t_{N_P}\} : \ a \leq t_0 \leq t_1 \leq \cdots \leq t_{N_p} \leq b\}$. Note that if we assume that function u is real-valued and piecewise constant on $[a, b]$, then the total variation of u on $[a, b]$ is the sum of the absolute value of the jumps in u.

The regularization of problem (9.1) via bounded variation is

$$\begin{aligned}
\min_{\boldsymbol{u} \in \mathcal{A}} \quad & J_{\boldsymbol{\rho}}(\boldsymbol{u}) = \textstyle\int_0^T g(\boldsymbol{x}(t)\boldsymbol{u}(t))\mathrm{d}t + \sum\limits_{j=1}^{m} \rho_j V(u_j) \\
\text{sub.to} \quad & \dot{x}_i(t) = f_i(\boldsymbol{x}(t), \boldsymbol{u}(t)), \quad \text{for all } i = 1, \ldots, n, \\
& x_i(0) = x_{i,0}, \quad \text{for all } i = 1, \ldots, n, \\
& \xi_j \leq u_j(t) \leq \omega_j \ \text{ for all } t \in [0, T] \ \text{ and } \ \text{for all } j = 1, \ldots, m,
\end{aligned} \tag{9.4}$$

where $V(u_j)$ is the total variation of u_j on interval $[0, T]$ and $0 \leq \rho_j < 1$ is the bounded variation penalty parameter associated with control variable u_j for all $j = 1, \ldots, m$. For this problem, we assume that all control variables need to be penalized. However, if (based upon observations of the numerical solutions for the unregularized problem (9.1)) we notice that one of the control variables u_ℓ exhibits no numerical artifacts or unusual oscillations, then we recommend solving problem (9.4) with the corresponding tuning parameter ρ_ℓ set to being zero. We can construct the Hamiltonian function that corresponds to problem (9.4), and the Hamiltonian gives the same adjoint equations (9.2) and transversality conditions (9.3).

For using PASA to numerically solve the regularized problem, we need to discretize problem (9.4) and discretize the adjoint equations (9.2). We present the method of discretizing the regularized problem only, but we emphasize that we can use PASA to solve the discretized unpenalized problem by numerically solving the discretized regularized problem when the vector $\boldsymbol{\rho}$ is set to being the zero vector. We begin by partitioning time interval $[0, T]$, by using $N+1$ equally spaced nodes. For all $i = 1, \ldots, n$ and $k = 0, \ldots, N$, we denote $x_{i,k} = x_i(t_k)$, and we emphasize that component $x_{i,0}$ is the initial value given in problem (9.4). So for each $i = 1, \ldots, n$, we have that $\boldsymbol{x}_i \in \mathbb{R}^{N+1}$ with $x_{i,0}$ being the initial value given. We assume that for all $j = 1, \ldots, m$, control variable u_j is constant over each mesh interval. For all $j = 1, \ldots, m$ and $k = 0, \ldots, N-2$, we denote $u_{j,k} = u_j(t)$ for all $t_k \leq t < t_{k+1}$, and for all $j = 1, \ldots, m$, we denote

$u_{j,N-1} = u_j(t)$ for all $t_{N-1} \leq t \leq t_N = T$. So for each $j = 1, \ldots, m$ we have that $\boldsymbol{u}_j \in \mathbb{R}^N$. The assumption of each control variable u_j being constant on each mesh interval allows us to express the total variation of u_j on $[0, T]$ in terms of the particular partition of $[0, T]$ that is used for the discretization:

$$V(u_j) = \sum_{k=0}^{N-2} |u_{j,k+1} - u_{j,k}|.$$

For simplicity of discussion, we use left-rectangular integral approximation to discretize the integral used in objective functional $J_{\boldsymbol{\rho}}$, and we use forward Euler's method to discretize the state equations given in problem (9.4). In general, we recommend using an explicit scheme for discretizing the state equations if the dynamics of the system have only initial conditions involved. We then have the following:

$$\begin{aligned}
\min \quad & J_{\boldsymbol{\rho}}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m) \\
& \qquad = \sum_{k=0}^{N-1} h g(\boldsymbol{x}_{\cdot,k}, \boldsymbol{u}_{\cdot,k}) + \sum_{j=1}^{m} \rho_j \sum_{k=0}^{N-2} |u_{j,k+1} - u_{j,k}| \\
& x_{i,k+1} = x_{i,k} + h f_i(\boldsymbol{x}_{\cdot,k}, \boldsymbol{u}_{\cdot,k}) \quad \text{for all } i = 1, \ldots, n \text{ and} \\
& \qquad k = 0, \ldots, N-1, \\
& \qquad \xi_j \leq u_{j,k} \leq \omega_j, \quad \text{for all } j = 1, \ldots, m \text{ and } k = 0, \ldots, N-1,
\end{aligned} \tag{9.5}$$

where $h = \frac{T}{N}$ is the mesh size, $\boldsymbol{x}_{\cdot,k} = [x_{1,k}, x_{2,k}, \ldots, x_{n,k}]$, and $\boldsymbol{u}_{\cdot,k} = [u_{1,k}, u_{2,k}, \ldots, u_{m,k}]$ for all $k = 0, \ldots, N-1$.

Since PASA consists of a phase that uses a projected gradient method, we need the objective function in problem (9.5) to be differentiable, which is not the case due to the absolute value terms that correspond to the discretization of the total variation function. We need to perform a decomposition of each absolute value term so that J_ρ can be differentiable. For each $j = 1, \ldots, m$, we introduce two $N-1$ dimensional vectors $\boldsymbol{\zeta}_j$ and $\boldsymbol{\iota}_j$ whose entries are non-negative, and every entry of $\boldsymbol{\zeta}_j$ and $\boldsymbol{\iota}_j$ is defined as

$$|u_{j,k+1} - u_{j,k}| = \zeta_{j,k} + \iota_{j,k} \quad \text{for all } k = 0, \ldots, N-2.$$

Another way of viewing $\boldsymbol{\zeta}_j$ and $\boldsymbol{\iota}_j$, is that each component $\zeta_{j,k}$ and $\iota_{j,k}$ will be defined based upon the following conditions:

$$\begin{aligned}
&\text{Condition 1 : If } u_{j,k+1} - u_{j,k} > 0, \text{ then} \\
&\qquad \zeta_{j,k} = u_{j,k+1} - u_{j,k} \text{ and } \iota_{j,k} = 0, \\
&\text{Condition 2 : If } u_{j,k+1} - u_{j,k} \le 0, \text{ then} \\
&\qquad \zeta_{j,k} = 0 \quad \text{and} \quad \iota_{j,k} = -(u_{j,k+1} - u_{j,k}).
\end{aligned}$$

Employing this decomposition to problem (9.5) yields

$$\begin{aligned}
\min \quad & J_{\boldsymbol{\rho}}(\boldsymbol{u}_1, \boldsymbol{\zeta}_1, \ldots, \boldsymbol{u}_m, \boldsymbol{\zeta}_m, \boldsymbol{\iota}_m) \\
& \quad = \sum_{k=0}^{N-1} hg(\boldsymbol{x}_{\cdot,k}, \boldsymbol{u}_{\cdot,k}) + \sum_{j=1}^{m} \left(\rho_j \sum_{k=0}^{N-2} (\zeta_{j,k} + \iota_{j,k}) \right) \\
& x_{i,k+1} = x_{i,k} + hf_i(\boldsymbol{x}_{\cdot,k}, \boldsymbol{u}_{\cdot,k}), \quad \text{for all } i = 1, \ldots, n \text{ and} \\
& \qquad\qquad k = 0, \ldots, N-1, \\
& \xi_j \le u_{j,k} \le \omega_j, \quad \text{for all } j = 1, \ldots, m \text{ and } k = 0, \ldots, N-1, \\
& u_{j,k+1} - u_{j,k} = \zeta_{j,k} - \iota_{j,k}, \quad \text{for all } j = 1, \ldots, m \text{ and} \\
& \qquad\qquad k = 0, \ldots, N-2, \\
& 0 \le \boldsymbol{\zeta}_j, \quad 0 \le \boldsymbol{\iota}_j \quad \text{for all } j = 1, \ldots, m.
\end{aligned} \tag{9.6}$$

For the above problem, we are now minimizing $J_{\boldsymbol{\rho}}$ with respect to vectors $\boldsymbol{u}_1, \boldsymbol{\zeta}_1, \boldsymbol{\iota}_1, \ldots, \boldsymbol{u}_m, \boldsymbol{\zeta}_m$, and $\boldsymbol{\iota}_m$. The constraints associated with each $\boldsymbol{\zeta}_j$ and $\boldsymbol{\iota}_j$, are constraints that PASA can interpret. For all $j = 1, \ldots, m$, the equality constraints associated with each $\boldsymbol{\zeta}_j$ and $\boldsymbol{\iota}_j$ in problem (9.6) are linear and can be expressed as

$$\left[\boldsymbol{A}_j \middle| -\boldsymbol{I}_{N-1} \middle| \boldsymbol{I}_{N-1} \right] \begin{bmatrix} \boldsymbol{u}_j \\ \hline \boldsymbol{\zeta}_j \\ \hline \boldsymbol{\iota}_j \end{bmatrix} = \boldsymbol{0},$$

where $\boldsymbol{I}_{N-1}$ is the identity matrix of dimension $N-1$, $\mathbf{0}$ is an $N-1$ dimensional all zeros vector, and $\boldsymbol{A_j}$ is an $N-1$ by N matrix defined as

$$\boldsymbol{A_j} = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -1 & 1 \end{bmatrix}. \tag{9.7}$$

Moreover, the equality constraints of the decomposition vectors given in problem (9.6) can be written as

$$\left[\boldsymbol{A_1}\middle|-\boldsymbol{I}_{N-1}\middle|\boldsymbol{I}_{N-1}\middle|\cdots\middle|\boldsymbol{A_j}\middle|-\boldsymbol{I}_{N-1}\middle|\boldsymbol{I}_{N-1}\middle|\cdots\middle|\boldsymbol{A_m}\middle|-\boldsymbol{I}_{N-1}\middle|\boldsymbol{I}_{N-1}\right] \begin{bmatrix} \boldsymbol{u}_1 \\ \hline \boldsymbol{\zeta}_1 \\ \hline \boldsymbol{\iota}_1 \\ \hline \vdots \\ \hline \boldsymbol{u}_j \\ \hline \boldsymbol{\zeta}_j \\ \hline \boldsymbol{\iota}_j \\ \hline \vdots \\ \hline \boldsymbol{u}_m \\ \hline \boldsymbol{\zeta}_m \\ \hline \boldsymbol{\iota}_m \end{bmatrix} = \mathbf{0},$$

where $\boldsymbol{A_j}$ is defined as (9.7) for all $j = 1, \ldots, m$ and $\mathbf{0} \in \mathbb{R}^{N-1}$.

From problem (9.6), we wish to find the gradient of $J_{\boldsymbol{\rho}}$ with respect to $\boldsymbol{u}_1, \boldsymbol{\zeta}_1, \boldsymbol{\iota}_1, \ldots, \boldsymbol{u}_m, \boldsymbol{\zeta}_m$, and $\boldsymbol{\iota}_m$. We compute the gradient of the Lagrangian of problem (9.6) with respect to $\boldsymbol{u}_1, \boldsymbol{\zeta}_1, \boldsymbol{\iota}_1, \ldots, \boldsymbol{u}_m, \boldsymbol{\zeta}_m$, and $\boldsymbol{\iota}_m$ to find $\nabla J_{\boldsymbol{\rho}}$. This is necessary because based upon problem (9.6), state variables $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ can be viewed as functions of $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$. So when computing the gradient of $J_{\boldsymbol{\rho}}$ with respect to $\boldsymbol{u}_1, \boldsymbol{\zeta}_1, \boldsymbol{\iota}_1, \ldots, \boldsymbol{u}_m, \boldsymbol{\zeta}_m$, and $\boldsymbol{\iota}_m$, we should consider that vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ depend on the controls.

The discretized problem can be put in the following general form for which the technique of using the Lagrangian to compute the gradient of a cost functional can be used:

$$\begin{aligned} \min \quad & G(\boldsymbol{x}, \boldsymbol{u}) \\ & \boldsymbol{F}(\boldsymbol{x}, \boldsymbol{u}) = \mathbf{0}, \end{aligned} \tag{9.8}$$

where $\boldsymbol{u} \in \mathbb{R}^m$, $\boldsymbol{x} \in \mathbb{R}^n$, $G : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, and $\boldsymbol{F} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ are differentiable; moreover, it is assumed in problem (9.8) that we can uniquely solve for $\boldsymbol{x}$ in term of $\boldsymbol{u}$. Based on these assumptions, we can rewrite problem (9.8) as

$$\begin{aligned} \min \quad & \mathcal{J}(\boldsymbol{u}) = G(\boldsymbol{x}(\boldsymbol{u}), \boldsymbol{u}) \\ & \boldsymbol{F}(\boldsymbol{x}(\boldsymbol{u}), \boldsymbol{u}) = \mathbf{0}, \end{aligned} \tag{9.9}$$

where $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{u})$ denotes the unique solution of $\boldsymbol{F}(\boldsymbol{x}, \boldsymbol{u}) = \mathbf{0}$ for a given $\boldsymbol{u} \in \mathbb{R}^m$. The following result can be deduced from the implicit function theorem and the chain rule (see Ref. [41, Remark 3.2], Ref. [42]):

Theorem 9.1. *If the Jacobian $\nabla_{\boldsymbol{x}} \boldsymbol{F}(\boldsymbol{x}, \boldsymbol{u})$ is invertible for each $\boldsymbol{u} \in \mathbb{R}^m$ and $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{u})$, then for each $\boldsymbol{u} \in \mathbb{R}^m$ the gradient of $\mathcal{J}$ in problem* (9.9) *is*

$$\nabla_{\boldsymbol{u}} \mathcal{J}(\boldsymbol{u}) = \nabla_{\boldsymbol{u}} \mathcal{L}(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\lambda})_{|\boldsymbol{x}=\boldsymbol{x}(\boldsymbol{u})}, \tag{9.10}$$

where $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ is the Lagrangian of problem (9.9),

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\lambda}) = G(\boldsymbol{x}, \boldsymbol{u}) + \boldsymbol{\lambda}^T \boldsymbol{F}(\boldsymbol{x}, \boldsymbol{u}),$$

and $\boldsymbol{\lambda}$ is chosen such that

$$\nabla_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\lambda}) = \nabla_{\boldsymbol{x}} G(\boldsymbol{x}, \boldsymbol{u}) + \boldsymbol{\lambda}^T \nabla_{\boldsymbol{x}} \boldsymbol{F}(\boldsymbol{x}, \boldsymbol{u}) = \mathbf{0}. \tag{9.11}$$

Before computing the Lagrangian to problem (9.6), we rewrite the state equations accordingly:

$$\begin{aligned} & -x_{i,k+1} + x_{i,k} + hf_i(\boldsymbol{x}_{\cdot,k}, \boldsymbol{u}_{\cdot,k}) = 0, \\ & \quad \text{for all } i = 1, \ldots, n \quad \text{and} \quad k = 0, \ldots, N-1. \end{aligned}$$

The Lagrangian to problem (9.6) is then

$$\begin{aligned} & \mathcal{L}(\boldsymbol{u}_1, \boldsymbol{\zeta}_1, \ldots, \boldsymbol{u}_m, \boldsymbol{\zeta}_m, \boldsymbol{\iota}_m) \\ & \quad = \sum_{k=0}^{N-1} hg(\boldsymbol{x}_{\cdot,k}, \boldsymbol{u}_{\cdot,k}) + \sum_{j=1}^{m} \rho_j \sum_{k=0}^{N-2} (\zeta_{j,k} + \iota_{j,k}) \\ & \qquad + \sum_{i=1}^{n} \sum_{k=0}^{N-1} \lambda_{i,k}(-x_{i,k+1} + x_{i,k} + hf_i(\boldsymbol{x}_{\cdot,k}, \boldsymbol{u}_{\cdot,k})), \end{aligned} \tag{9.12}$$

where $\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_n \in \mathbb{R}^{N-1}$ are the Lagrange multiplier vectors. For all $j = 1, \ldots, m$ and for all $k = 0, \ldots, N-1$, we compute the partial derivative of $\mathcal{L}$ with respect to $u_{j,k}$ and obtain

$$\frac{\partial}{\partial u_{j,k}}\mathcal{L} = h\frac{\partial}{\partial u_{j,k}}g(\boldsymbol{x}_{\boldsymbol{\cdot},k}, \boldsymbol{u}_{\boldsymbol{\cdot},k}) + \sum_{i=1}^{n} \lambda_{i,k} \left(h\frac{\partial}{\partial u_{j,k}} f_i(\boldsymbol{x}_{\boldsymbol{\cdot},k}, \boldsymbol{u}_{\boldsymbol{\cdot},k}) \right).$$

For all $j = 1, \ldots, m$, we have that

$$\boldsymbol{\nabla}_{\boldsymbol{u_j}} \boldsymbol{\mathcal{L}} = h \begin{bmatrix} \dfrac{\partial g(\boldsymbol{x}_{\boldsymbol{\cdot},0}, \boldsymbol{u}_{\boldsymbol{\cdot},0})}{\partial u_{j,0}} + \sum_{i=1}^{n} \lambda_{i,0} \left(\frac{\partial f_i(\boldsymbol{x}_{\boldsymbol{\cdot},0}, \boldsymbol{u}_{\boldsymbol{\cdot},0})}{\partial u_{j,0}} \right) \\ \vdots \\ \dfrac{\partial g(\boldsymbol{x}_{\boldsymbol{\cdot},k}, \boldsymbol{u}_{\boldsymbol{\cdot},k})}{\partial u_{j,k}} + \sum_{i=1}^{n} \lambda_{i,k} \left(\frac{\partial f_i(\boldsymbol{x}_{\boldsymbol{\cdot},k}, \boldsymbol{u}_{\boldsymbol{\cdot},k})}{\partial u_{j,k}} \right) \\ \vdots \\ \dfrac{\partial g(\boldsymbol{x}_{\boldsymbol{\cdot},N-1}, \boldsymbol{u}_{\boldsymbol{\cdot},N-1})}{\partial u_{j,N-1}} + \sum_{i=1}^{n} \lambda_{i,N-1} \left(\frac{\partial f_i(\boldsymbol{x}_{\boldsymbol{\cdot},N-1}, \boldsymbol{u}_{\boldsymbol{\cdot},N-1})}{\partial u_{j,N-1}} \right) \end{bmatrix} \tag{9.13}$$

where $\boldsymbol{\nabla}_{\boldsymbol{u_j}} \boldsymbol{\mathcal{L}} \in \mathbb{R}^N$. For all $j = 1, \ldots, m$ and for all $k = 0, \ldots, N-2$, we compute the partial derivative of $\mathcal{L}$ with respect to $\zeta_{j,k}$ and the partial derivative of $\mathcal{L}$ with respect to $\iota_{j,k}$, as follows:

$$\frac{\partial}{\partial \zeta_{j,k}}\mathcal{L} = \rho_j,$$
$$\frac{\partial}{\partial \iota_{j,k}}\mathcal{L} = \rho_j.$$

We can then say for all $j = 1, \ldots, m$ that

$$\boldsymbol{\nabla}_{\boldsymbol{\zeta_j}} \boldsymbol{\mathcal{L}} = \begin{bmatrix} \rho_j \\ \vdots \\ \rho_j \end{bmatrix}, \tag{9.14}$$

and

$$\boldsymbol{\nabla}_{\boldsymbol{\iota_j}}\boldsymbol{\mathcal{L}} = \begin{bmatrix} \rho_j \\ \vdots \\ \rho_j \end{bmatrix} \tag{9.15}$$

where $\boldsymbol{\nabla}_{\boldsymbol{\zeta_j}}\boldsymbol{\mathcal{L}} \in \mathbb{R}^{N-2}$ and $\boldsymbol{\nabla}_{\boldsymbol{\iota_j}}\boldsymbol{\mathcal{L}} \in \mathbb{R}^{N-2}$. By Theorem 9.1, provided that Lagrange multiplier vectors $\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_n$ satisfy theorem condition (9.11), we have that

$$\boldsymbol{\nabla J} = \boldsymbol{\nabla \mathcal{L}} = \left[\begin{array}{c} \boldsymbol{\nabla}_{\boldsymbol{u_1}}\boldsymbol{\mathcal{L}} \\ \boldsymbol{\nabla}_{\boldsymbol{\zeta_1}}\boldsymbol{\mathcal{L}} \\ \boldsymbol{\nabla}_{\boldsymbol{\iota_1}}\boldsymbol{\mathcal{L}} \\ \hline \vdots \\ \hline \boldsymbol{\nabla}_{\boldsymbol{u_k}}\boldsymbol{\mathcal{L}} \\ \boldsymbol{\nabla}_{\boldsymbol{\zeta_k}}\boldsymbol{\mathcal{L}} \\ \boldsymbol{\nabla}_{\boldsymbol{\iota_k}}\boldsymbol{\mathcal{L}} \\ \hline \vdots \\ \hline \boldsymbol{\nabla}_{\boldsymbol{u_m}}\boldsymbol{\mathcal{L}} \\ \boldsymbol{\nabla}_{\boldsymbol{\zeta_m}}\boldsymbol{\mathcal{L}} \\ \boldsymbol{\nabla}_{\boldsymbol{\iota_m}}\boldsymbol{\mathcal{L}} \end{array}\right] \tag{9.16}$$

where entries of $\boldsymbol{\nabla \mathcal{L}}$ are defined in equations (9.13)–(9.15). Conveniently, our method of finding vectors $\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_n$ that satisfy condition (9.11) produces a discretization of the adjoint equations (9.2). For all $i = 1, \ldots, n$ and for $k = 1, \ldots, N-1$, we compute the partial derivative of $\mathcal{L}$ with respect to $x_{i,k}$, as follows:

$$\frac{\partial \mathcal{L}}{\partial x_{i,k}} = h\frac{\partial g(\boldsymbol{x}_{\cdot,k}, \boldsymbol{u}_{\cdot,k})}{\partial x_{i,k}} - \lambda_{i,k-1} + \lambda_{i,k} + \sum_{\ell=1}^{n} \lambda_{\ell,k}\left(h\frac{\partial f_\ell(\boldsymbol{x}_{\cdot,k}, \boldsymbol{u}_{\cdot,k})}{\partial x_{i,k}}\right).$$

Additionally, we have that

$$\frac{\partial \mathcal{L}}{\partial x_{i,N}} = -\lambda_{i,N-1}.$$

We did not take the partial derivative of $\mathcal{L}$ with respect to $x_{i,0}$ since $x_{i,0}$ is a known value for all $i = 1, \ldots, n$.

To satisfy theorem condition (9.11), we set $\frac{\partial}{\partial x_{i,k}}\mathcal{L}$ equal to zero for all $i = 1, \ldots, n$ and $k = 1, \ldots, N$, and solve for $\lambda_{i,k-1}$. We obtain the following for all $i = 1, \ldots, n$:

$$\lambda_{i,k-1} = h\frac{\partial g(\boldsymbol{x}_{\cdot,k}, \boldsymbol{u}_{\cdot,k})}{\partial x_{i,k}} + \lambda_{i,k} + \sum_{\ell=1}^{n} \lambda_{\ell,k}\left(h\frac{\partial f_\ell(\boldsymbol{x}_{\cdot,k}, \boldsymbol{u}_{\cdot,k})}{\partial x_{i,k}}\right) \quad (9.17)$$

$$\text{for } k = 1, \ldots, N-1, \quad \text{and}$$

$$\lambda_{i,N-1} = 0. \quad (9.18)$$

The above equations not only allow us to use the gradient of the Lagrangian of problem (9.5) to find the gradient of $J_{\boldsymbol{\rho}}$, but also give us a discretization for adjoint equations (9.2). Additionally, for all $i = 1, \ldots, n$, equation (9.18) is analogous to the transversality condition given in (9.3).

9.3 Example 1: The Fishery Problem

In this section, we focus on a basic resource model on harvesting, which was first presented in Clark's.[37] We present the fishery problem as stated in Lenhart and Workman [12, Example 17.4], where a logistic growth function is utilized within the fishery model.

$$\begin{aligned} \max_{u\in\mathcal{A}} \quad & J(u) = \textstyle\int_0^T (pqx(t) - c)u(t)\,\mathrm{d}t \\ \text{sub.to} \quad & x'(t) = x(t)(1 - x(t)) - qu(t)x(t), \\ & x(0) = x_0 > 0, \\ & \qquad 0 \le u(t) \le M \end{aligned} \quad (9.19)$$

In problem (9.19), $u(t)$ is the control variable that measures the effort put into harvesting fish at time t, and $x(t)$ measures the total population of fish at time t. We are assuming that there is a maximum harvesting rate, M. Parameter q represents the "catchability" of a fish and parameter c represents the cost of harvesting one unit of fish. Parameter p is the selling price for one unit of harvested fish.

The objective functional J is constructed to represent the total profit, revenue less cost, of harvesting fish over time interval $[0, T]$. The set of admissible controls, $\mathcal{A}$, for problem (9.19) is

$$\mathcal{A} = \{u \in L^1(0,T) : \quad \text{for all } t \in [0,T],\ u(t) \in A,\ A = [0, M]\}.$$

Existence of an optimal control for problem (9.19) follows from the Filippov–Cesari Existence Theorem.[9]

According to Lenhart and Workman,[12] the standard forward–backward sweep method does not converge if parameters in problem (9.19) are set to where the optimal control u^* contains a singular region. In this section, we present an explicit formula for the singular case, which is obtained via Pontryagin's Maximum Principle,[10] and we show that the singular case satisfies the generalized Legendre–Clebsch Condition.[21–24] Additionally, we present a set of assumptions on the parameters to the problem in order to obtain an optimal harvesting strategy that begins singular and switches to the maximum harvesting rate. For this scenario, an explicit formula for the switching point is obtained. With parameters set to meet those particular assumptions, we use the explicit solution to problem (9.19) to test PASA's accuracy in solving for the regularized problem for varying values of the tuning parameter. We also discuss how to discretize for both the fishery problem and the regularized version of the problem in which a bounded variation regularization term is applied. And finally, we present some empirical evidence for convergence between the numerical solution obtained by PASA for the regularized fishery problem and the exact solution to the fishery problem.

9.3.1 *Explicitly solving the fishery problem*

Lenhart and Workman demonstrate a method of using Pontryagin's Maximum Principle[10,12] and properties of the switching function to solve for the singular case to problem (9.19). Lenhart and Workman also discuss conditions for existence of a singular case. Since we use a numerical solver that is used for solving minimization problems, we provide an analytical solution to the minimization problem that is equivalent to problem (9.19). The equivalent minimization problem is obtained by negating the objective functional $J(u)$ in problem (9.19):

$$\begin{aligned} \min_u \quad & J(u) = \textstyle\int_0^T -(pqx(t) - c)u(t)\mathrm{d}t \\ \text{sub. to} \quad & x'(t) = x(t)(1 - x(t)) - qu(t)x(t), \\ & x(0) = x_0 > 0, \\ & \qquad 0 \le u(t) \le M. \end{aligned} \tag{9.20}$$

Our procedure for solving problem (9.20) is analogous to what is used in Lenhart and Workman.[12] We use Pontryagin's Maximum Principle[10] to solve problem (9.20). The Hamiltonian for the above problem is

$$H(x, u, \lambda) = (c - pqx)u + \lambda(x - x^2 - qux), \tag{9.21}$$

where λ is the adjoint variable. By taking the partial derivative of the Hamiltonian with respect to state variable x, we obtain the adjoint equation associated with adjoint variable λ, which is

$$\lambda'(t) = -\frac{\partial H}{\partial x} = pqu - \lambda + 2\lambda x + q\lambda u \tag{9.22}$$

with the transversality condition being

$$\lambda(T) = 0. \tag{9.23}$$

We also use the Hamiltonian given in (9.21) to compute the switching function corresponding to problem (9.20)

$$\psi(t) = \frac{\partial H}{\partial u} = c - pqx(t) - q\lambda(t)x(t). \tag{9.24}$$

Based on Pontryagin's Maximum Principle,[10] if there exists an optimal pair (u^*, x^*) for problem (9.20), then there exists λ^*, satisfying adjoint equation (9.22) and terminal condition (9.23), where $H(x^*, u^*, \lambda^*) \le H(x^*, u, \lambda^*)$ for all admissible controls u. Additionally, if u^* is the optimal control, then u^* must have the following form:

$$u^*(t) = \begin{cases} 0 & \text{whenever } \psi(t) > 0, \\ \text{singular} & \text{whenever } \psi(t) = 0, \\ M & \text{whenever } \psi(t) < 0. \end{cases} \tag{9.25}$$

To solve for the singular case we suppose $\psi(t) \equiv 0$ on some subinterval $I \subset [0,T]$. Assuming that parameter $c > 0$, then we have from the switching function given in (9.24) that both x and $pq + q\lambda$ are nonzero on interval I. By setting the ψ equal to zero and solving for λ, we have that

$$\lambda^*(t) = \frac{c - pqx(t)}{qx(t)} \tag{9.26}$$

on the interval I. We differentiate equation (9.26) and use the state equation given in problem (9.20) to obtain the following:

$$\begin{aligned} \lambda'(t) &= -\frac{c}{qx^2}(x(1-x) - qux) \\ &= -\frac{c}{qx} + \frac{c}{q} + \frac{cu}{x}. \end{aligned} \tag{9.27}$$

We rewrite equation (9.22) by using expression (9.26), and we get

$$\lambda'(t) = -\frac{c}{qx} + p + \frac{2c}{q} - 2px + \frac{cu}{x}. \tag{9.28}$$

Equating expressions (9.27) and (9.28) gives a solution for x on the singular interval which is

$$x^*(t) = \frac{c + pq}{2pq}.$$

Since x^* is constant on I, we have that $x'(t) = 0$ on I. We use the singular solution for x^* and set the state equation found in problem (9.20) equal to zero to solve for u^*. On the interval I, we have that

$$\begin{aligned} 0 = x'(t) &= x^*(1 - x^*) - qux^* \\ u^*(t) &= \frac{1 - x^*}{q} \\ u^*(t) &= \frac{pq - c}{2pq^2}. \end{aligned}$$

Additionally, x^* being constant on the singular region and expression (9.26) would imply that λ^* is also constant on the singular region with constant value being as follows:

$$\lambda^*(t) = \frac{p(c - pq)}{c + pq}.$$

By substituting in the constant solutions for u^*, x^*, and λ^* into adjoint equation (9.22), we have that the right-hand side of the adjoint equation is zero, as desired. To conclude, we find that if a singular region, I, exists, then u^*, x^*, and λ^* are all constant on I, where

$$u^* = \frac{pq - c}{2pq^2}, \quad x^* = \frac{c + pq}{2pq}, \quad \lambda^* = \frac{p(c - pq)}{c + pq}. \tag{9.29}$$

We wish to show that the singular case solution satisfies the second-order necessary condition of optimality, which is referred to as the generalized Legendre–Clebsch Condition [21,22] or Kelley's condition.[23,24,43] Before showing that the singular cases given in (9.29) satisfy the Legendre–Clebsch Condition and/or Kelley's condition, we present some parameter assumptions on problem (9.20).

Assumption 9.1. Parameters $p, c, M, q > 0$ are set to satisfy $0 < pq - c < 2pq^2M$.

Assumption 9.2. Initial value x_0 is set to equal $\frac{c+pq}{2pq}$, which is the constant value that is associated to the state solution corresponding to singular u^*.

Note that Assumption 9.1 implies that the singular case solution u^* satisfies the bounded constraints that are assumed on the control. Assumption 9.2 dynamically forces the problem to yield an optimal control that begins singular. The generalized Legendre–Clebsch Condition involves finding what is called the *order* of a singular arc, which is defined as being the integer q such that $(\frac{d^{2q}}{dt^{2q}}\frac{\partial H}{\partial u})$ is the lowest order total derivative of the partial derivative of the Hamiltonian with respect to u, in which control u appears explicitly. We use the state equations given in problem (9.20) and the adjoint equations given in (9.22) to find the first and second time derivative of the switching function (9.24):

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\psi = pq(x^2 - x) - q\lambda x^2, \\
&\frac{\mathrm{d}^2}{\mathrm{d}t^2}\psi = pq(2x-1)(x - x^2) - q\lambda x^2 + [pq^2 x + q^2\lambda x^2 - 3pq^2x^2]u.
\end{aligned} \tag{9.30}$$

From above, we have that the order of the singular arc is $q = 1$. We need to show that if u^* is an optimal singular control on some interval of order q, then it is necessary that

$$(-1)^q \frac{\partial}{\partial u}\left[\frac{\mathrm{d}^2}{\mathrm{d}t^2}\left(\frac{\partial H}{\partial u}\right)\right]\bigg|_{x=x^*,\lambda=\lambda^*} \geq 0. \tag{9.31}$$

We take the partial derivative of (9.30) with respect to u and evaluate at the corresponding singular case solutions for x and λ given in (9.29):

$$\frac{\partial}{\partial u}\left[\frac{\mathrm{d}^2}{\mathrm{d}t^2}\left(\frac{\partial H}{\partial u}\right)\right]\bigg|_{x=\frac{c+pq}{2pq},\lambda=\frac{p(c-pq)}{2pq^2}}$$
$$= \frac{q(c+pq)}{2} - \frac{3(c+pq)^2}{4p} - \frac{(pq-c)(c+pq)^2}{8p^2q^2}.$$

Note that the term $\frac{(pq-c)(c+pq)^2}{8p^2q^2}$ is positive by Assumption 9.1. Using algebra, we combine the first two terms in the above equation to obtain

$$\frac{\partial}{\partial u}\left[\frac{\mathrm{d}^2}{\mathrm{d}t^2}\left(\frac{\partial H}{\partial u}\right)\right]\bigg|_{x=\frac{c+pq}{2pq},\lambda=\frac{p(c-pq)}{2pq^2}}$$
$$= -\frac{3c^2}{4p} - cq - \frac{pq^2}{4} - \frac{(pq-c)(c+pq)^2}{8p^2q^2} < 0.$$

By multiplying the above inequality by $(-1)^q$ where $q = 1$, we then have the second-order necessary condition of optimality (9.31) being satisfied.

Note that only Assumption 9.1 is needed in proving that the singular case satisfies the Legendre–Clebsch condition; however, we can use both assumptions to obtain a control that begins singular and switches to the maximum harvesting rate, where an explicit formula for the switching point can be obtained. This particular scenario for problem (9.20) makes it an excellent candidate problem to use for testing PASA's accuracy in solving the regularized variant of this problem. We first verify Assumptions 9.1 and 9.2 imply that u^* is not singular on the entire time interval $[0, T]$. By using the transversality condition, $\lambda(T) = 0$, we recognize that the singular case solution for λ given in (9.26) is 0 if and only if parameters are set to either satisfy $p = 0$ or $c - pq = 0$, and Assumption 9.1 ensures that both cases are not possible. Let $0 < t^* < T$ be the time when u^* switches from being singular to non-singular, and let $I = [0, t^*)$ be the interval corresponding to when the optimal control u^* is singular.

By looking at the objective functional for problem (9.20), intuition tells us that $u^* \neq 0$ on $[t^*, T]$, and we can prove this by using proof by contradiction. Assume that

$$u^*(t) = \begin{cases} \dfrac{pq - c}{2pq^2} & 0 \leq t < t^* \\ 0 & t^* \leq t \leq T \end{cases}$$

is the optimal control to problem (9.20). Consider the following admissible control $\hat{v}$ where $\hat{v}$ is singular over the entire interval, i.e.,

$$\hat{v}(t) = \frac{pq - c}{2pq^2} \quad \text{for all } t \in [0, T].$$

By assumption of u^* being optimal for problem (9.20), we have $J(\hat{v}) \geq J(u^*)$. Since $\hat{v} = u^*$ on the interval $[0, t^*]$ and $u^* \equiv 0$ on interval $[t^*, T]$, we obtain the following:

$$J(\hat{v}) = J(u^*) - \int_{t^*}^{T} ((pqx_{\hat{v}}(t) - c)\hat{v}(t))\mathrm{d}t,$$

where $x_{\hat{v}}(t)$ is the corresponding solution to the state equation given in problem (9.20). A contradiction is obtained if we can show that $pqx_{\hat{v}} - c > 0$ for all $t \in [t^*, T]$ because the following implies $J(\hat{v}) \leq J(u^*)$. Since $\hat{v}$ is singular on the interval $[t^*, T]$, we have that $x_{\hat{v}}$ is the corresponding singular case solution given in (9.29). We then have that

$$\begin{aligned} pqx_{\hat{v}} - c &= pq\left(\frac{c + pq}{2pq}\right) - c \\ &= \frac{1}{2}(pq - c) \\ &> 0, \end{aligned}$$

where the above inequality holds by Assumption 9.1. Therefore, we have our contradiction.

It then follows by Pontryagin's Maximum Principle that the optimal harvesting policy for problem (9.20) with parameters set to satisfy Assumptions 9.1 and 9.2 is a control that begins singular and switches once to the maximal harvesting effort, meaning $u^* \equiv M$ on the interval $I = [t^*, T]$. We can find an explicit expression for t^*. Assume that $u^*(t) = M$ for all $t \in [t^*, T]$, then the adjoint equation (9.22) becomes

$$\lambda'(t) = pqM - \lambda + 2\lambda x + q\lambda M. \tag{9.32}$$

Now, λ is continuous at t^*. Hence,

$$\lambda(t^*) = \lim_{t \to t^{*-}} \lambda(t) = \frac{p(c - pq)}{c + pc},$$

which is the solution for equation (9.22) when u is singular. We can then solve for equation (9.32) to where λ must satisfy the terminal condition, $\lambda(T) = 0$, and the condition that

$$\lambda(t^*) = \frac{p(c - pq)}{c + pc}.$$

This condition allows us to obtain an explicit solution for t^*. When solving differential equation (9.32), we use a standard method that is used in solving linear differential equations. We rewrite (9.32) as

$$\lambda'(t) + \lambda(t)(\alpha - 2x(t)) = pqM, \tag{9.33}$$

where $\alpha = 1 - qM$. We let $v(t)$ be the appropriate integrating factor, which is defined as

$$v(t) = \mathrm{e}^{\int_{t^*}^{t} (\alpha - 2x(\tau))\mathrm{d}\tau}. \tag{9.34}$$

Multiplying $v(t)$ to both sides of Eqs. (9.33), integrating over the interval (t^*, t), and rearranging terms yields

$$\lambda(t) = \frac{1}{v(t)} \left[v(t^*)\lambda(t^*) + pqM \int_{t^*}^{t} v(\tau)\mathrm{d}\tau \right]. \tag{9.35}$$

Now in order to evaluate $v(t)$ and $\int_{t^*}^{t} v(\tau)\mathrm{d}\tau$, we need an explicit solution for $x(t)$ over the interval $[t^*, T]$.

Given that $u^*(t) = M$ for all $t \in [t^*, T]$, the state equation given in problem (9.20) over the specified interval becomes

$$\frac{\mathrm{d}x}{\mathrm{d}t} = x(1-x) - qMx, \tag{9.36}$$

which is separable. To solve the above equation, we separate variables x and t

$$\frac{\mathrm{d}x}{x(\alpha - x)} = \mathrm{d}t,$$

where $\alpha = 1 - qM$. We perform a partial fraction decomposition on the left-hand side of the above equation, integrate both sides, and exponentiate to obtain the following:

$$\frac{x}{|\alpha - x|} = K\mathrm{e}^{\alpha t}, \tag{9.37}$$

where K is some constant. We have that state variable x is continuous at switching point t^*. By continuity and Assumption 9.2 we have

$$x^*(t^*) = \lim_{t \to t^{*-}} x^*(t) = x_0 = \frac{c + pq}{2pq},$$

which is the state solution value associated with the singular case. We can use the above equality to solve for constant value K found in Eq. (9.37). Evaluating Eq. (9.37) at $t = t^*$ yields the following:

$$K = \frac{x_0 \mathrm{e}^{-\alpha t^*}}{|\alpha - x_0|} = \frac{(c + pq)\mathrm{e}^{-\alpha t^*}}{\gamma}, \tag{9.38}$$

where $\gamma = |2\alpha pq - c - pq|$. We obtain an explicit solution of Eq. (9.36), by rewriting Eq. (9.37) as

$$\frac{x}{\alpha - x} = \hat{K}\mathrm{e}^{\alpha t} \quad \text{where } \hat{K} = \begin{cases} K & \alpha - x \geq 0, \\ -K & \alpha - x < 0 \end{cases}.$$

Solving for the above equation yields

$$x(t) = \frac{\alpha \hat{K} \mathrm{e}^{\alpha t}}{1 + \hat{K}\mathrm{e}^{\alpha t}} \quad \text{for } t \in [t^*, T]. \tag{9.39}$$

Note that the value of $\hat{K}$ seems to depend on the sign of $\alpha - x(t)$. However, we can use continuity of x and the structure of the solutions for x on $[t^*, T]$ to show that the value of $\hat{K}$ only depends on the sign of $\alpha - x_0$. But first, we prove the following:

Proposition 9.1. *Assumptions* 9.1 *and* 9.2 *imply* $\alpha - x_0 < 0$ *where* $\alpha = 1 - qM$.

Proof. By Assumption 9.1, parameters $c, p, q, M > 0$ are chosen to satisfy $0 < pq - c$ and $pq - c < 2pq^2M$. Dividing both sides of the second inequality by $2pq$ yields

$$\frac{pq - c}{2pq} < qM.$$

Negating the inequality and adding one to both sides yields

$$\alpha < 1 + \frac{c - pq}{2pq} = \frac{c + pq}{2pq},$$

and the right-hand side of the inequality is x_0, by Assumption 9.2. Therefore, we have $\alpha - x_0 < 0$. □

Recall that by Assumption 9.2 and Eqs. (9.29) and (9.39), we have the following solution for the state variable:

$$x(t) = \begin{cases} x_0 = \dfrac{c + pq}{2pq} & 0 \leq t \leq t^* \\ \dfrac{\alpha \hat{K} \mathrm{e}^{\alpha t}}{1 + \hat{K}\mathrm{e}^{\alpha t}} & t^* \leq t \leq T \end{cases}, \tag{9.40}$$

where the sign of $\hat{K}$ is determined by the sign $\alpha - x(t)$. Since x is continuous on the entire time interval and since x is constant on the singular region, we have that $x(t^*) = x_0$. Hence, at $t = t^*$, $\hat{K}$ is determined by $\alpha - x(t^*) = \alpha - x_0$. Differentiating $x(t)$ along the non-singular region yields

$$x'(t) = \frac{\alpha^2 \hat{K} \mathrm{e}^{\alpha t}}{(1 + \hat{K}\mathrm{e}^{\alpha t})^2},$$

which is either strictly positive or strictly negative based upon the sign of $\hat{K}$. Since $\hat{K}$ is negative at $t = t^*$, there is some open interval containing t^* such that the function x is a non-increasing function. Note also

$$\lim_{t\to\infty} \frac{\alpha \hat{K} \mathrm{e}^{\alpha t}}{1+\hat{K}\mathrm{e}^{\alpha t}} = \alpha,$$

so this function has a horizontal asymptote being $x_{\text{hor}} = \alpha$ on the tx-plane. This implies that $x(t)$ given in (9.40) remains above the horizontal asymptote on the interval $[t^*, T]$ even though the function is non-increasing on (t^*, T). In conclusion, the sign of $\alpha - x_0$ determines the structure of $x(t)$ on the non-singular region. Using Proposition 9.1 and Eq. (9.40) we then conclude $x(t)$ is of the following form:

$$x(t) = \begin{cases} x_0 = \dfrac{c+pq}{2pq} & 0 \le t \le t^* \\ \dfrac{\alpha K \mathrm{e}^{\alpha t}}{-1+K\mathrm{e}^{\alpha t}} & t^* \le t \le T \end{cases},$$

where K is defined on Eq. (9.38). We now use the explicit solution for variable $x(t)$ to evaluate $v(t)$ given in (9.34) as follows:

$$v(t) = \exp\left(\alpha(t-t^*) - 2\int_{t^*}^{t} \frac{\alpha K \mathrm{e}^{\alpha\tau}}{-1+K\mathrm{e}^{\alpha\tau}}\mathrm{d}\tau\right).$$

We use a u-substitution to evaluate the integral term in v. After applying some logarithm rules, we obtain the following:

$$v(t) = \left(\frac{-1+K\mathrm{e}^{\alpha t^*}}{-1+K\mathrm{e}^{\alpha t}}\right)^2 \mathrm{e}^{\alpha(t-t^*)}. \tag{9.41}$$

To evaluate $\int_{t^*}^{t} v(\tau)\mathrm{d}\tau$, we need to use a u-substitution method. Let $\sigma(\tau) = -1+K\mathrm{e}^{\alpha\tau}$, then $\int_{t^*}^{t} v(\tau)\mathrm{d}\tau$ with v given in (9.41) becomes

$$\begin{aligned}\int_{t^*}^{t} v(\tau)\mathrm{d}\tau &= \int_{t^*}^{t} \left(\frac{-1+K\mathrm{e}^{\alpha t^*}}{-1+K\mathrm{e}^{\alpha\tau}}\right)^2 \mathrm{e}^{\alpha(\tau-t^*)}\mathrm{d}\tau \\ &= \int_{\sigma(t^*)}^{\sigma(t)} \frac{(-1+K\mathrm{e}^{\alpha t^*})^2 \mathrm{e}^{-\alpha t^*}}{\alpha K \sigma^2}\mathrm{d}\sigma \\ &= \frac{(-1+K\mathrm{e}^{\alpha t^*})^2}{\alpha K}\mathrm{e}^{-\alpha t^*}\left[\frac{1}{-1+K\mathrm{e}^{\alpha t^*}} - \frac{1}{-1+K\mathrm{e}^{\alpha t}}\right]\end{aligned}$$

$$= \frac{e^{-\alpha t^*}(-1+Ke^{\alpha t^*})}{\alpha K}\left[1 - \frac{-1+Ke^{\alpha t^*}}{-1+Ke^{\alpha t}}\right]$$

$$= \frac{e^{-\alpha t^*}(-1+Ke^{\alpha t^*})}{\alpha(-1+Ke^{\alpha t})}\left[e^{\alpha t} - e^{\alpha t^*}\right]$$

$$\int_{t^*}^{t} v(\tau)\mathrm{d}\tau = \frac{(-1+Ke^{\alpha t^*})}{\alpha(-1+Ke^{\alpha t})}\left[e^{\alpha(t-t^*)} - 1\right].$$

Now we use equation (9.35) to evaluate $\lambda(t)$ as follows:

$$\lambda(t) = \left(\frac{-1+Ke^{\alpha t}}{-1+Ke^{\alpha t^*}}\right)^2 e^{\alpha(t^*-t)} f(t). \tag{9.42}$$

where

$$f(t) = \frac{p(c-pq)}{c+pq} + \frac{pqM}{\alpha}\left(\frac{-1+Ke^{\alpha t^*}}{-1+Ke^{\alpha t}}\right)\left(e^{\alpha(t-t^*)} - 1\right)$$

To find t^*, we set $\lambda(T) = 0$ and solve for t^*. Note that the term $-1 + Ke^{\alpha t}$ used in Eq. (9.42) must be non-zero for all $t \in [t^*, T]$, otherwise the state variable solution given in Eq. (9.39) is not defined for all $t \in [t^*, T]$. Using $-1+Ke^{\alpha t}$ for all $t \in [t^*, T]$ with $t^* < T$ allows us to conclude that $\lambda(T) = 0$ if and only if $f(T) = 0$. Consequently, for finding t^* we set $f(T) = 0$ and solve for t^*.

$$0 = f(T)$$

$$0 = \frac{c-pq}{c+pq} + \frac{qM}{\alpha}\left(\frac{-1+Ke^{\alpha t^*}}{-1+Ke^{\alpha T}}\right)\left(e^{\alpha(T-t^*)} - 1\right)$$

We multiply both sides of the above equation by $\alpha(c+pq)(-1+Ke^{\alpha T})$ to obtain

$$0 = \alpha(c-pq)(-1+Ke^{\alpha T}) + qM(c+pq)(-1+Ke^{\alpha t^*})\left(e^{\alpha(T-t^*)} - 1\right).$$

Substituting the value for K given in (9.38) into the above equation and multiplying everything by γ yields

$$0 = \alpha(c-pq)\left(-\gamma + (c+pq)e^{\alpha(T-t^*)}\right)$$
$$+\ qM(c+pq)[-\gamma + (c+pq)]\left(e^{\alpha(T-t^*)} - 1\right).$$

We rearrange terms from the above equation to isolate expression $\mathrm{e}^{\alpha(T-t^*)}$

$$\mathrm{e}^{\alpha(T-t^*)} = \frac{\gamma\alpha(c-pq) + qM(c+pq)(-\gamma+(c+pq))}{(c+pq)[\alpha(c-pq)+qM(-\gamma+(c+pq))]}.$$

We take the natural logarithm of both sides of the above equation and rearrange terms to find that

$$t^* = T - \frac{1}{\alpha}\ln\left[\frac{\gamma\alpha(c-pq) + qM(c+pq)(-\gamma+(c+pq))}{(c+pq)[\alpha(c-pq)+qM(-\gamma+(c+pq))]}\right],$$

where $\alpha = 1 - qM$ and $\gamma = |pq - c - 2pq^2M|$. We simplify t^* more by substituting in $\alpha = 1 - qM$ into the above expression:

$$t^* = T - \frac{1}{1-qM}\ln\left[\frac{-\gamma(pq-c) - 2\gamma cqM + qM(c+pq)^2}{(c+pq)[(c-pq)+2pq^2M - \gamma qM]}\right]. \quad (9.43)$$

To summarize, Assumptions 9.1 and 9.2 imply that the optimal control u^* to problem (9.20) must begin singular and switch once to the non-singular case where $u^* \equiv M$ on $[t^*, T]$ with t^* given in (9.43). Additionally, the solutions for u^*, x^*, and λ^* are the following:

$$u^*(t) = \begin{cases} \dfrac{pq-c}{2pq^2} & 0 \le t < t^*, \\ M & t^* \le t \le T, \end{cases} \quad (9.44)$$

$$x^*(t) = \begin{cases} \dfrac{c+pq}{2pq} & 0 \le t \le t^*, \\ \dfrac{\alpha K\mathrm{e}^{\alpha t}}{-1+K\mathrm{e}^{\alpha t}} & t^* \le t \le T, \end{cases} \quad (9.45)$$

and

$$\lambda^*(t) = \begin{cases} \dfrac{p(c-pq)}{c+pq} & 0 \le t \le t^*, \\ \left(\dfrac{-1+K\mathrm{e}^{\alpha t}}{-1+K\mathrm{e}^{\alpha t^*}}\right)^2 \mathrm{e}^{\alpha(t^*-t)} f(t) & t^* \le t \le T, \end{cases} \quad (9.46)$$

where

$$f(t) = \frac{p(c-pq)}{c+pq} + \frac{pqM}{\alpha}\left(\frac{-1+K\mathrm{e}^{\alpha t^*}}{-1+K\mathrm{e}^{\alpha t}}\right)\left(\mathrm{e}^{\alpha(t-t^*)} - 1\right),$$

$$\alpha = 1 - qM,$$

$$K = \frac{x_0 e^{-\alpha t^*}}{|\alpha - x_0|} = \frac{(c + pq)e^{-\alpha t^*}}{\gamma},$$

and

$$\gamma = |2\alpha pq - c - pq|.$$

9.3.2 *Discretization of the fishery problem*

For numerically solving problem (9.20), we first discretize and then optimize. We use the polyhedral active set algorithm (PASA), which was developed by Hager and Zhang,[31] to find an optimal solution to the discretized problem. Additionally, we need to discretize the adjoint equation associated with problem (9.20), which is

$$\lambda'(t) = pqu - \lambda + 2\lambda x + q\lambda u, \tag{9.47}$$

with the transversality condition being

$$\lambda(T) = 0.$$

For discretizing problem (9.20), we assume that control u is constant over each mesh interval. We partition time interval $[0, T]$, by using $N + 1$ equally spaced nodes, $0 = t_0 < t_1 < \cdots < t_N = T$. For all $k = 0, 1, \ldots, N$, we assume that the $x_k = x(t_k)$. For the control, we denote $u_k = u(t)$ for all $t_k \le t < t_{k+1}$ when $k = 0, \ldots, N-2$ and $u_{N-1} = u(t)$ for all $t_{N-1} \le t \le t_N$. So we have $\boldsymbol{x} \in \mathbb{R}^{N+1}$ while $\boldsymbol{u} \in \mathbb{R}^N$. We use a left-rectangular integral approximation for objective function J in (9.20), and we use forward Euler's method to approximate the state equation in (9.20). The discretization of problem (9.20) is then

$$\begin{aligned}
\min \quad J(\boldsymbol{u}) &= \sum_{k=0}^{N-1} h(c - pqx_k)u_k \\
x_{k+1} &= x_k + h(1 - x_k - qu_k)x_k \quad \text{for all } 0 \le k \le N-1, \\
&\quad x_0 > 0, \\
&\quad 0 \le u_k \le M \quad \text{for all } 0 \le k \le N-1,
\end{aligned} \tag{9.48}$$

where $h = T/N$ is the mesh size and the first component of state vector, x_0, is set to being the initial condition associated with the state equation given in problem (9.20).

Since PASA uses the gradient projection algorithm for one of its phases, we need to compute the gradient of the cost functional for problem (9.48). We use Theorem 9.1 to find $\nabla_{\boldsymbol{u}} J$, which requires finding the Lagrangian to problem (9.48) and its gradient. Additionally, we need to construct a Lagrange multiplier vector $\boldsymbol{\lambda}$ that satisfies Eq. (9.11). Consequently, the Lagrange multiplier vector that satisfies Eq. (9.11) produces the numerical scheme that is used for discretizing adjoint equation (9.47) and produces the transversality condition (9.23). To compute the Lagrangian to problem (9.48), we first need to arrange the discretized state equations accordingly

$$-x_{k+1} + x_k + h(x_k - x_k^2 - qu_k x_k) = 0 \quad \text{for all } k = 0, 1, \ldots, N-1. \tag{9.49}$$

The Lagrangian to problem (9.48) is

$$\begin{aligned}\mathcal{L}(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\lambda}) = &\sum_{k=0}^{N-1} (h(c - pqx_k)u_k) \\ &+ \sum_{k=0}^{N-1} \lambda_k(-x_{k+1} + x_k + h(x_k - x_k^2 - qu_k x_k)),\end{aligned}$$

where $\boldsymbol{\lambda} \in \mathbb{R}^{N-1}$ is the Lagrange multiplier vector. Note that we need not worry about the inequality constraints associated with the bounds of the control when computing the Lagrangian to problem (9.48) because these bounds are not being entered into the cost functional J. By taking the partial derivative of $\mathcal{L}$ with respect to u_k, we obtain

$$\frac{\partial \mathcal{L}}{\partial u_k} = h(c - pqx_k - q\lambda_k x_k) \quad \text{for } k = 0, \ldots, N-1.$$

By Theorem 9.1, we have that

$$\nabla_{\boldsymbol{u}} J = \nabla_{\boldsymbol{u}} \mathcal{L} = \begin{bmatrix} h(c - pqx_0 - q\lambda_0 x_0) \\ \vdots \\ h(c - pqx_k - q\lambda_k x_k) \\ \vdots \\ h(c - pqx_{N-1} - q\lambda_{N-1} x_{N-1}) \end{bmatrix}, \tag{9.50}$$

provided that Eq. (9.11) is satisfied.

To satisfy Eq. (9.11), we take the partial derivative of $\mathcal{L}$ with respect to x_k for all $k = 1, \ldots, N$, and note that we do not take the partial derivative of $\mathcal{L}$ with respect to x_0 because x_0 is a known value. Taking the partial derivative of $\mathcal{L}$ with respect to the state vector components yield the following expressions:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial x_k} &= h(-pqu_k + \lambda_k(1 - 2x_k - qu_k)) \\ &\quad + \lambda_k - \lambda_{k-1} \quad \text{for } k = 1, \ldots, N-1, \end{aligned} \tag{9.51}$$

$$\frac{\partial \mathcal{L}}{x_N} = -\lambda_{N-1} \tag{9.52}$$

To align with Eq. (9.11), we set expressions (9.51) and (9.52) equal to zero and solve for λ_{k-1} for all $k = 1, \ldots, N$:

$$\begin{aligned}\lambda_{k-1} &= \lambda_k + h(-pqu_k + \lambda_k(1 - 2x_k - qu_k)) \\ &\qquad \text{for } k = 1, \ldots, N-1,\end{aligned} \tag{9.53}$$

$$\lambda_{N-1} = 0. \tag{9.54}$$

Expressions (9.53) and (9.54) serve as the discretization for the costate equation (9.47) and the transversality condition (9.23).

We use PASA to solve for the regularized version of problem (9.20) where the penalty applied to the problem is a bounded variation penalty, as suggested in Capognigro *et al.*,.[30] The regularized version of problem (9.20) is as follows:

$$\begin{aligned}\min \quad J_\rho(u) &= \int_0^T (c - pqx)u\,dx + \rho V(u) \\ x'(t) &= (1 - x)x - qux,\ x(0) = x_0 > 0, \\ 0 &\leq u(t) \leq M,\end{aligned} \tag{9.55}$$

where $0 \leq \rho < 1$ is a tuning parameter and $V(u)$ measures the total variation of u which is

$$V(u) = \sup_{\mathcal{P}} \sum_{i=0}^{n_P - 1} |u(t_{i+1}) - u(t_i)| \tag{9.56}$$

where $\mathcal{P} = \{P = \{t_0, t_1, \ldots, t_{n_P}\} :\ \text{P is a partition of } [0, T]\}$. Note that if we numerically solve problem (9.55) with $\rho = 0$, then the problem is not being regularized via bounded variation. We use the same procedure as before to discretize problem (9.55). Assuming control u to be constant over each mesh interval allows us to express the total variation of u as being the sum of the absolute value of the jumps of u. For a sufficiently small mesh size h, we have

$$V(u) = \sum_{k=0}^{N-1} |u(t_{k+1}) - u(t_k)|.$$

The discretized version of problem (9.55) is

$$\begin{aligned}
\min \quad J_\rho(\boldsymbol{u}) &= \sum_{k=0}^{N-1} (h(c - pqx_k)u_k) + \rho \sum_{k=0}^{N-1} |u_{k+1} - u_k| \\
x_{k+1} &= x_k + h(1 - x_k - qu_k)x_k \quad \text{for all } 0 \le k \le N-1, \\
&x_0 > 0, \\
&0 \le u_k \le M \quad \text{for all } 0 \le k \le N-1.
\end{aligned} \tag{9.57}$$

Because PASA involves a gradient scheme, we should be concerned about the absolute value terms that are used in problem (9.57)'s cost function. We suggest a decomposition of each absolute value term in J_ρ to ensure that J_ρ is differentiable. We introduce two $N-1$ vectors $\boldsymbol{\zeta}$ and $\boldsymbol{\iota}$ whose entries are non-negative. Each entry of $\boldsymbol{\zeta}$ and $\boldsymbol{\iota}$ is defined as

$$|u_{k+1} - u_k| = \zeta_k + \iota_k, \quad \text{for all } k = 0, \ldots, N-2.$$

An equivalent way of expressing the above equation is to assign values to the components of $\boldsymbol{\zeta}$ and $\boldsymbol{\iota}$ based upon the following conditions:

Condition 1: If $u_{k+1} - u_k > 0,$ then $\zeta_k = u_{k+1} - u_k$ and $\iota_k = 0;$

Condition 2: If $u_{k+1} - u_k \le 0,$ then $\zeta_k = 0$ and $\iota_k = -(u_{k+1} - u_k).$

This decomposition converts problem (9.57) into the following:

$$\begin{aligned}
\min \quad J_\rho(\boldsymbol{u}, \boldsymbol{\zeta}, \boldsymbol{\iota}) &= \sum_{k=0}^{N-1} (h(c - pqx_k)u_k) + \rho \sum_{k=0}^{N-2} (\zeta_k + \iota_k) \\
&x_{k+1} = x_k + h(1 - x_k - qu_k)x_k \quad \text{for all } k = 0, \ldots, N-1, \\
&x_0 > 0, \\
&0 \le u_k \le M \quad \text{for all } k = 0, \ldots, N-1, \\
&u_{k+1} - u_k = \zeta_k - \iota_k \quad \text{for all } k = 0, \ldots, N-2, \\
&\zeta_k \ge 0 \text{ and } \iota_k \ge 0 \quad \text{for all } 1, \ldots, N-1.
\end{aligned} \tag{9.58}$$

For problem (9.58), we are minimizing the regularized objective function with respect to vectors $\boldsymbol{u}, \boldsymbol{\zeta}$, and $\boldsymbol{\iota}$. The constraints associated with $\boldsymbol{\zeta}$ and $\boldsymbol{\iota}$, are constraints that PASA can interpret. The equality constraints associated with $\boldsymbol{\zeta}$ and $\boldsymbol{\iota}$ can be written accordingly as

$$\left[\, \boldsymbol{A} \middle| -\boldsymbol{I}_{N-1} \middle| \boldsymbol{I}_{N-1} \,\right] \left[\begin{array}{c} \boldsymbol{u} \\ \hline \boldsymbol{\zeta} \\ \hline \boldsymbol{\iota} \end{array}\right] = \boldsymbol{0}, \tag{9.59}$$

where $\boldsymbol{I}_{N-1}$ is the identity matrix with dimension $N-1$, $\boldsymbol{0}$ is the $N-1$ dimensional all zeros vector, and $\boldsymbol{A}$ is the $N-1 \times N$ dimensional sparse matrix given in (9.7). Since problem (9.58) is optimizing J_ρ with respect to $\boldsymbol{u}$, $\boldsymbol{\zeta}$, and $\boldsymbol{\iota}$, the Lagrangian of problem (9.58) is

$$\begin{aligned} \mathcal{L}(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\zeta}, \boldsymbol{\iota}, \boldsymbol{\lambda}) = {} & \sum_{k=0}^{N-1} (h(c - pqx_k)u_k) + \rho \sum_{k=0}^{N-2} (\zeta_k + \iota_k) \\ & + \sum_{k=0}^{N-1} \lambda_k(-x_{k+1} + x_k + h(x_k - x_k^2 - qu_kx_k)). \end{aligned}$$

If we use Theorem (9.1) to find the gradient of J_ρ, we have the following:

$$\nabla_{\boldsymbol{u},\boldsymbol{\zeta},\boldsymbol{\iota}} J_\rho = \nabla_{\boldsymbol{u},\boldsymbol{\zeta},\boldsymbol{\iota}} \mathcal{L} = \left[\begin{array}{c} \nabla_{\boldsymbol{u}} J \\ \hline \rho \\ \vdots \\ \rho \\ \hline \rho \\ \vdots \\ \rho \end{array}\right], \tag{9.60}$$

where $\nabla_{\boldsymbol{u}} J$ is defined in Eq. (9.50), provided that the theorem's condition (9.11) is satisfied. As before, we satisfy condition (9.11) by taking the partial derivative of $\mathcal{L}$ with respect to each x_k for $k = 1, \ldots, N$ and set each partial derivative equal to 0 and solve for the components of $\boldsymbol{\lambda}$. Performing this procedure yields Eqs. (9.53) and (9.54), and they serve as our discretization procedure of the adjoint equation (9.22) and our generalization of the transversality condition (9.23).

9.3.2.1 *Summary of the discretization*

We convert the regularized harvesting problem (9.55) into problem (9.58) by performing the following steps.

(1) *Discretization of the state equation*: We use forward Euler's method to discretize the state equation which is given in Eq. (9.49).
(2) *Discretization of the objective functional and the decomposition of absolute value terms*: We use a left-rectangular integral approximation for discretizing the integral shown in (9.55). We also assume u as being piecewise constant over each mesh interval, which allows us to convert the total variation term given in (9.56) into the finite series of absolute value terms that is used in problem (9.57). Because we use a gradient scheme for solving the discretized problem, we need to decompose the absolute value terms by introducing two $N-1$ vectors $\boldsymbol{\zeta}$ and $\boldsymbol{\iota}$ that satisfy the constraints that are included in problem (9.58).
(3) *Finding the gradient of the regularized objective functional*: By Theorem 9.1, we use the gradient of the Lagrangian to problem (9.58) to find $\nabla_{[\boldsymbol{u},\boldsymbol{\zeta},\boldsymbol{\iota}]}J_\rho$. The formula for $\nabla_{[\boldsymbol{u},\boldsymbol{\zeta},\boldsymbol{\iota}]}J_\rho$ is given in Eq. (9.60), where as $\nabla_{\boldsymbol{u}}J$ is given in equation (9.50).
(4) *Discretization of the adjoint variable*: In order to apply Theorem 9.1 for computing the gradient of the penalized objective function, adjoint vector $\boldsymbol{\lambda}$ needs to satisfy condition (9.11). We set each partial derivative of the Lagrangian $\mathcal{L}$ with respect to x_k equal to 0 and solve for λ_{k-1}. This results in discretizing adjoint equation (9.22) and the transversality condition $\lambda(T) = 0$. The discretized equations are given in Eqs. (9.53) and (9.54).

9.3.3 *Numerical results of the fishery problem*

We wish to numerically solve problem (9.20) with the following parameters defined in Table 9.1. Note that these parameter values satisfy Assumptions 9.1 and 9.2, so the optimal harvesting strategy is of the form given in (9.44). Based on Table 9.1 and Eq. (9.43), $t^* \approx 9.5392$, which is between values 0 and $T = 10$. Additionally, from these parameter values the singular control solution is $u^* = 0.1875$ which is between the bounds 0 and $M = 1$, as desired. After using Table 9.1 to evaluate t^* from (9.43), u^* from (9.44), x^* from (9.45),

Table 9.1. Parameter settings for the fishery problem.

Parameter	Description	Value
T	Terminal time	10
p	Selling price of one unit of fish	2
q	"Catchability" of the fish	2
c	Cost of harvesting one unit of fish	1
M	Maximum harvest effort	1
x_0	Initial population size of fish	$\frac{c+pq}{2pq} = 0.625$

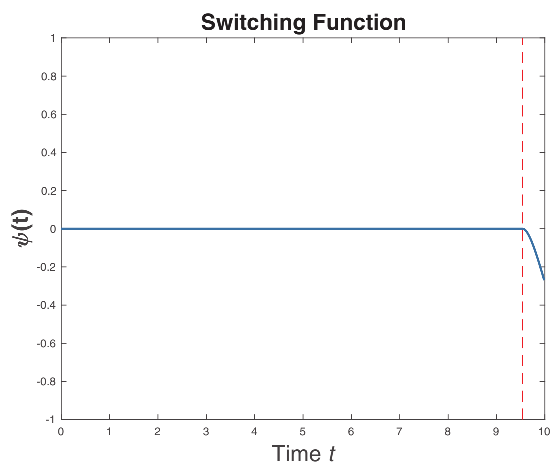
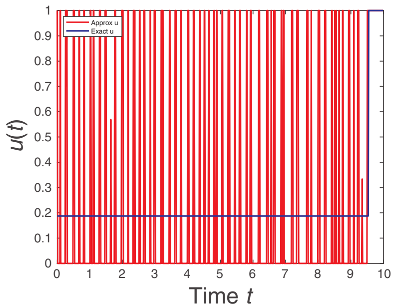


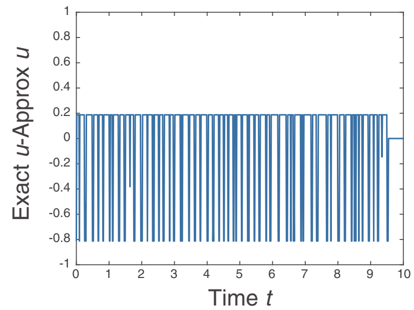
Fig. 9.1. Plot of the Switching Function. The red dotted line is the vertical line $t = t^* \approx 9.5392$.

and λ^* from (9.46), we substitute these solutions into the switching function (9.24) to see if u^* satisfies (9.25). In Figure 9.1, we plot the switching function ψ that we obtained. Regard that the switching function is zero when u^* is singular and becomes negative at the interval $[t^*, T]$, which is when $u^* = M$. So we have that the harvesting policy u^* for (9.44) satisfies (9.25). This is equivalent to saying that u^* satisfies Pontryagin's Maximum Principle,[10] which is the first-order necessary condition for optimality.

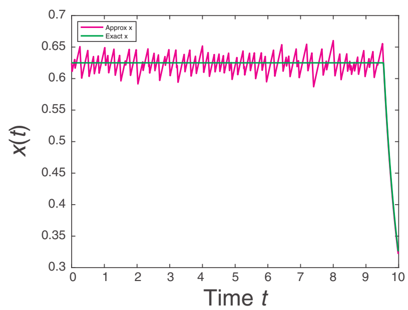
We use PASA to solve problem (9.20) with parameter settings given in Table 9.1. Our initial guess for control u is $u(t) = 0$ for all $t \in [0, T]$, and the stopping tolerance is set to 10^{-10}. We partition



(a) Unpenalized control $\hat{u}$ (red) vs optimal control u^* (blue)



(b) $u^* - \hat{u}$ (cyan).



(c) Fish population solution $\hat{x}$ (pink) corresponding to $\hat{u}$ vs exact state solution x^* (green)

Fig. 9.2. Unpenalized Results for the fishery problem: Time interval $[0, T]$ is partitioned to have $N = 750$ mesh intervals and PASA stopping tolerance is tol$=10^{-10}$.

$[0, T]$ to where there are $N = 750$ mesh intervals, and we use the discretization method that is presented in Section 9.3.2. We first observe PASA's approximation of the optimal control for problem (9.20) when no bounded variation regularization is being applied, which we denote as being $\hat{u}$. In Figure 9.2(a), a plot of $\hat{u}$ is shown in red and is compared to the exact solution, u^*, which is shown in blue. In Figure 9.2(a), $\hat{u}$ has oscillations that resemble chattering on the singular region. Additionally, we reran experiments with a tighter stopping tolerance and a finer partition of $[0, T]$, and PASA still obtained an oscillatory solution. It is likely that the discretization of problem (9.20) causes PASA to generate the oscillatory artifacts. We computed a left-rectangular integral approximation of the total profit of the harvested fish over time interval $[0, T]$, (i.e., the cost functional for the equivalent maximization problem (9.19)) when employing harvesting policy, $\hat{u}$, and optimal harvesting policy, u^*, and found that the $-J(\hat{u}) \approx 3.074$ and $-J(u^*) \approx 3.0575$ where J is the cost functional used in (9.20).

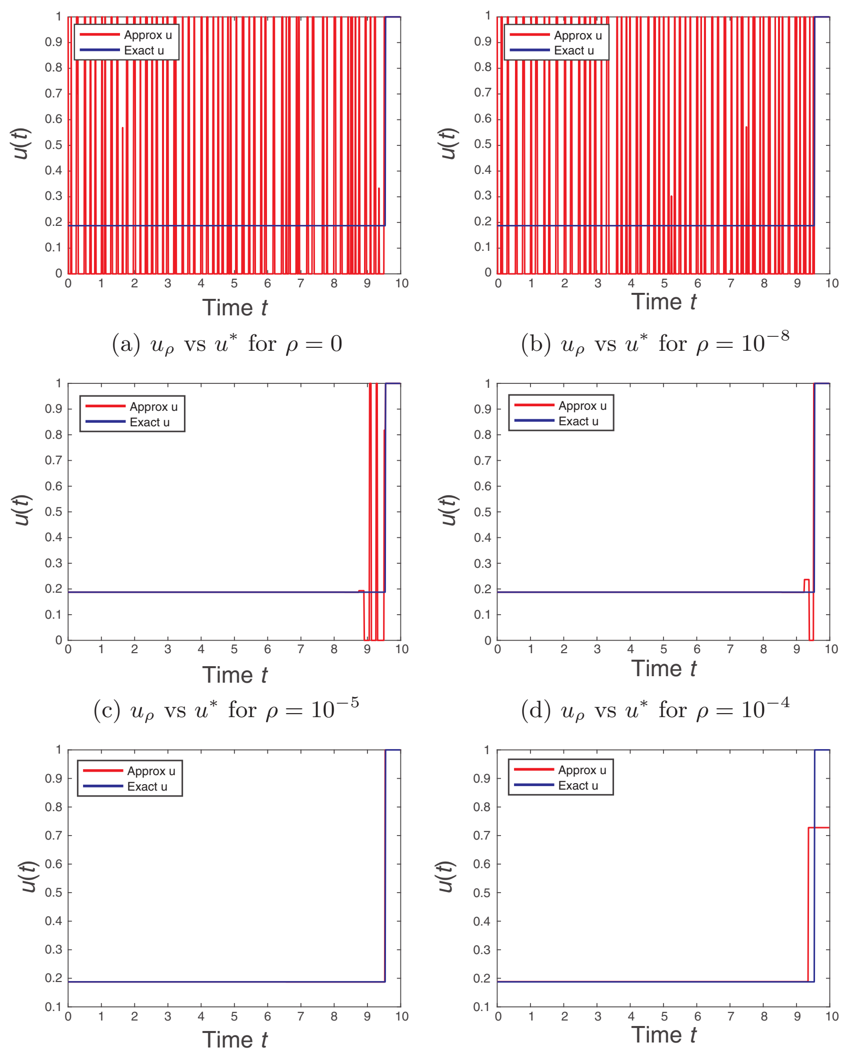
The major concern associated with the approximate solution obtained in Figure 9.2(a) is that it is an unrealistic harvesting strategy to use. We wish to regularize problem (9.20) by adding a bounded variation term that reduces the number of oscillations. Before getting into the results, we first discuss our methods for determining which penalty parameter values give the best solution. We have some advantages for choosing an appropriate tuning parameter value due to knowing the analytic solution. However, for problems in which we do not know what u^* looks like we recommend looking at the plots of PASA's approximation of the regularized control u_ρ. Firstly, we would rule out a penalty parameter value if the plot of u_ρ contained any unusual jumps and/or oscillations. If such a solution occurred, we suggest that the penalty parameter value is too small and is generating a solution that is similar to the unpenalized solution. Secondly, we would rule out a ρ value if the corresponding penalized solution u_ρ does not closely align with Pontryagin's Maximum Principle.[10] In this example, Pontryagin's Maximum Principle implies that u^* is a piecewise constant function where $u^*(t) \in \{0, \frac{pq-c}{2pq^2}, M\}$ for all $t \in [0, T]$. So if the penalized solution u_ρ ever takes on values that were not $0, M$, or $\frac{pq-c}{2pq^2}$, then we find value ρ to be suspect. Thirdly, one might find it useful to look at the plots of the switching function that are associated with the penalized solution u_ρ. This third suggestion might be the most useful if one were penalizing an optimal control problem where the singular case solution cannot be found explicitly. The sign of the switching function can help one verify whether or not u_ρ aligns with Pontryagin's Maximum Principle.

Since we have the advantage of comparing the regularized solution u_ρ with the exact solution u^*, we can determine which penalty parameter value is appropriate by comparing the plots between u^* and u_ρ as well as the plots of $u^* - u_\rho$. Additionally, we look at the instance in time when the penalized solution u_ρ switched from singular to non-singular to see if it is close to when the switching point should occur, i.e., when $t^* \approx 9.5392$.

When calculating the associated switching point for u_ρ, we look at the first instance when $u_\rho = M$. Due to the discretization of problem (9.55), each approximated switching point is some mesh point value of $[0, T]$. We also use the L_1 norm difference between u^* and u_ρ for determining which penalty parameter value gives the closest approximation to u^*. We observe the L_∞ norm difference between u^* and u_ρ as well; however, this computation may not be helpful since approximated switching points are restricted to mesh points of the partitioned time interval.

Using parameter settings given in Table 9.1, we use PASA to solve problem (9.55) for varying values of penalty parameter $\rho \in \{10^{-9}, 10^{-8}, 10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$. Our initial guess for problem (9.55) is $u(t) = 0$ for all $t \in [0, T]$. We partition $[0, T]$ to where there are $N = 750$ mesh intervals, and we use the discretization method described in Section 9.3.2.1. Additionally, the stopping tolerance is set to 10^{-10}. Results corresponding to the penalized solutions that PASA obtained are found in Figures 9.3 and 9.4 and Table 9.2. Note that from looking at the plots of the penalized controls of Figure 9.2 alone, we find that $\rho = 10^{-2}$ is the most appropriate penalty parameter value. We believe that even if we do not know what the explicit formula for the singular case was for problem (9.55), the plots of the regularized controls should lead us to choose $\rho = 10^{-2}$ as being the appropriate tuning parameter because no oscillations are occurring and the end behavior of the penalized control matches the non-oscillating region from the unpenalized control. This yields some evidence that we could potentially use PASA for solving penalized control problems without *a priori* information of the optimal control.

We continue to explain how we ruled out all cases except for $\rho = 10^{-2}$. In Figure 9.3(b), we rule out tuning parameter value $\rho = 10^{-8}$, since its corresponding solution u_ρ is oscillating in the same manner as the unpenalized solution, whose plot is given in Figure 9.3(a). We do not provide plots of penalized solution u_ρ that corresponded to when $\rho = 10^{-9}, 10^{-7}, 10^{-6}$ because their plots possessed many oscillations similar to Figure 9.3(b). Additionally, one could look at



(a) u_ρ vs u^* for $\rho = 0$

(b) u_ρ vs u^* for $\rho = 10^{-8}$

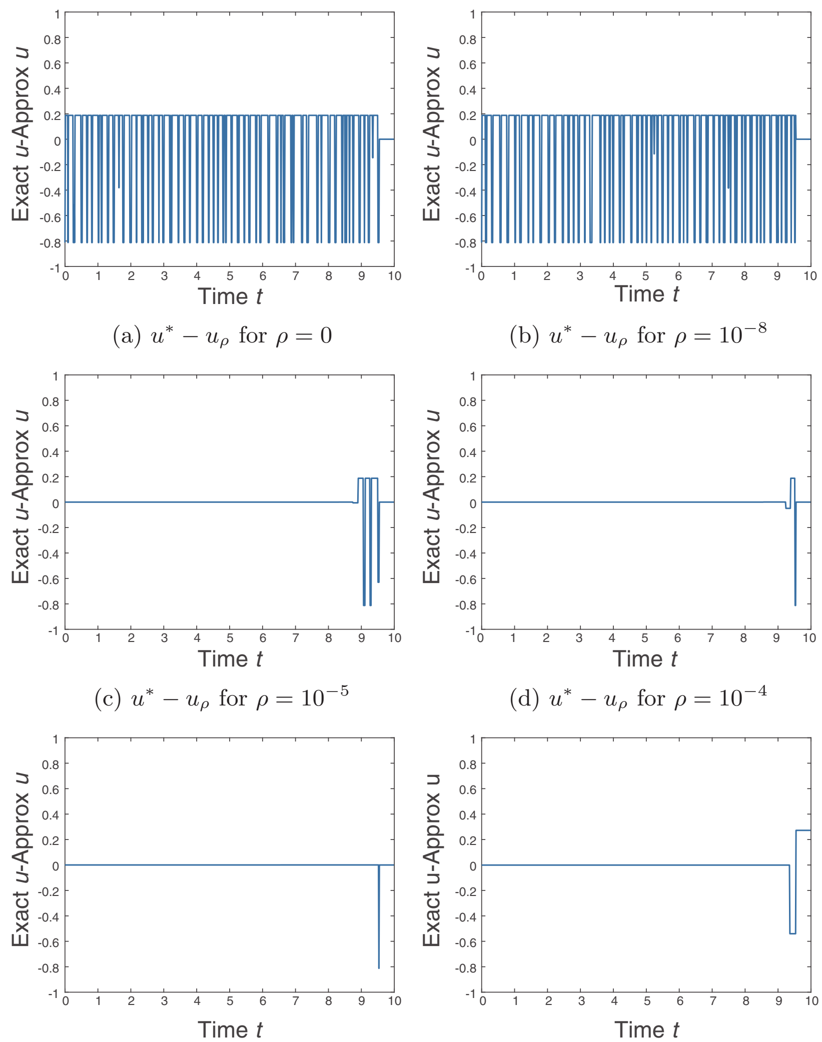
(c) u_ρ vs u^* for $\rho = 10^{-5}$

(d) u_ρ vs u^* for $\rho = 10^{-4}$

(e) u_ρ vs u^* for $\rho = 10^{-2}$

(f) u_ρ vs u^* for $\rho = 10^{-1}$

Fig. 9.3. Plots of the regularized control u_ρ (red) vs the optimal control u^* (blue) for varying values of the tuning parameter ρ. Time interval $[0, T]$ is partitioned to have $N = 750$ mesh intervals. PASA stopping tolerance is tol=10^{-10}.



(a) $u^* - u_\rho$ for $\rho = 0$

(b) $u^* - u_\rho$ for $\rho = 10^{-8}$

(c) $u^* - u_\rho$ for $\rho = 10^{-5}$

(d) $u^* - u_\rho$ for $\rho = 10^{-4}$

(e) $u^* - u_\rho$ for $\rho = 10^{-2}$

(f) $u^* - u_\rho$ for $\rho = 10^{-1}$

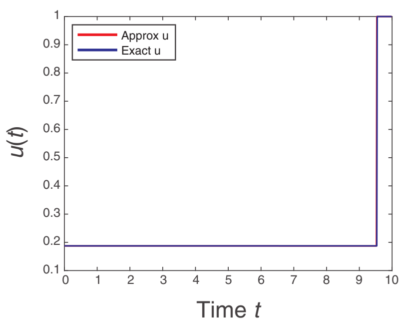
Fig. 9.4. Plots of the optimal control minus the regularized control $u^* - u_\rho$ (cyan) for varying values of the tuning parameter ρ. Time interval $[0, T]$ is partitioned to have $N = 750$ mesh intervals. PASA stopping tolerance is tol=10^{-10}.

Table 9.2. Varying tuning parameter table for the fishery problem.

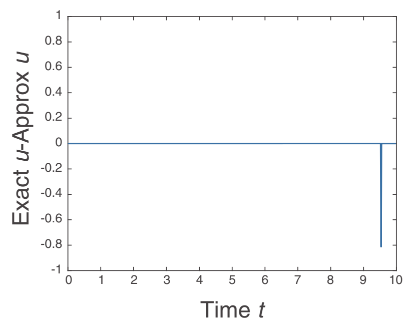
Parameters	ρ	$\|u^* - u_\rho\|_{L^1}$	$\|u^* - u_\rho\|_\infty$	Switch	Runtime
$N = 750$	0	2.89548286	0.8125	0^*	8.82
$\text{tol} = 10^{-10}$	10^{-9}	2.89262350	0.8125	0^*	14.44
$T = 10$	10^{-8}	2.89708395	0.8125	0^*	22.11
$M = 1$	10^{-7}	2.89799306	0.8125	0^*	28.45
$p = 2$	10^{-6}	2.88774180	0.8125	0.04^*	34.88
$q = 2$	10^{-5}	0.19705717	0.8125	9.0667	6.58
$c = 1$	10^{-4}	0.05427589	0.8125	9.5200	1.25
$x_0 = 0.625$	10^{-3}	0.01628685	0.8125	**9.5333**	0.68
$h = 0.0133$	10^{-2}	**0.01282480**	0.8125	**9.5333**	**0.38**
	10^{-1}	0.22780044	0.5402	9.9867	0.43

Note: The total variation of our exact solution is $v(u^*) = 0.8125$, so if our approximated solution u_ρ switched from the singular case to m at a time different from u^*, then $\|u^* - u_\rho\|_\infty = v(u^*)$. the switch column is indicating the first instance in time t when $u_\rho(t) = m$. the $*$-values on the switch column indicate that the numerical solution is oscillating. The Runtime column indicates the time (in seconds) it takes for PASA to solve the problem.

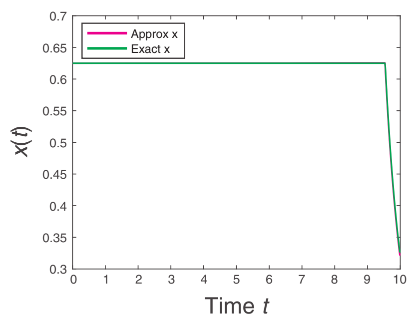
the $\|u^* - u_\rho\|_{L^1}$ column in Table 9.2 to see that the penalized solution for u_ρ for $\rho \in \{10^{-9}, 10^{-8}, 10^{-7}, 10^{-6}\}$ are negligible in comparison to the unpenalized solution. In Figures 9.3(c) and 9.4(c), we start to see some improvements in the singular region of u_ρ when $\rho = 10^{-5}$. However, due to the oscillations appearing in time interval $(8.5, 9.5)$, we dismiss $\rho = 10^{-5}$. We observe in Figure 9.3(d) that increasing the penalty parameter ρ to 10^{-4} significantly reduces the number of oscillations along the singular region. However, around time interval $[9.24, 9.37]$, $u_\rho(t) \approx 0.2366$, which is a value different from the bounds of the control and the singular case solution. We rule out penalty $\rho = 10^{-4}$ for this reason. We do not provide a figure of the solution we obtained when penalty parameter value $\rho = 10^{-3}$, but we dismiss this penalty parameter value due to similar reasoning that was used when $\rho = 10^{-4}$. We also rule out tuning parameter $\rho = 10^{-1}$ because as illustrated in Figure 9.3(f), u_ρ is constant around $[9.36, 10]$ with constant value being approximately 0.7277. For this case, we say that $\rho = 10^{-1}$ is over-penalizing the control along the non-singular region. Based on Table 9.2 and Figures 9.3(e) and 9.4(e), we find that penalty parameter $\rho = 10^{-2}$ is the most appropriate penalty parameter. For $\rho = 10^{-2}$, we have that u_ρ switches to the non-singular case at the



(a) u_ρ (red) vs u^* (blue) with $\rho = 10^{-2}$



(b) $u^* - u_\rho$ (cyan) with $\rho = 10^{-2}$



(c) Approximate State x_ρ (pink) vs Exact State x^* (green) with $\rho = 10^{-2}$

Fig. 9.5. Results for when $\rho = 10^{-2}$: Time interval $[0, T]$ is partitioned to have $N = 750$ mesh intervals and PASA stopping tolerance is tol$=10^{-10}$.

node that is closest to the actual switching point. In addition, u_ρ have the lowest L^1 norm error in Table 9.2 when the penalty parameter is set to $\rho = 10^{-2}$. In Figure 9.5(c), we have a plot of the state solution corresponding to u_ρ when $\rho = 10^{-2}$. Observe that the plot of the approximated fish population corresponding to u_ρ lines up almost perfectly with the true solution.

After finding an appropriate tuning parameter value for problem (9.55), we then study the numerical convergence rate between PASA's solution to the discretized regularized problem with tuning parameter $\rho = 10^{-2}$ and the exact solution to problem (9.20). We observe the L^1 norm error between the exact solution u^* and the regularized solution u_h for varying mesh sizes $h \in \{0.2, 0.1, 0.05, 0.025, 0.0125, 0.00625, 0.003125\}$. In Table 9.3, we have

Table 9.3. Convergence analysis between the penalized solution and the exact solution to fishery problem.

Parameters	h	err_h	$\frac{\text{err}_h}{\text{err}_{h/2}}$	$\log_2\left(\frac{\text{err}_h}{\text{err}_{h/2}}\right)$
$\text{tol} = 10^{-10}$	0.2	0.28091381		
$\rho = 10^{-2}$	0.1	0.12111065	2.31948057	1.21380176
	0.05	0.00089762	134.92448804	7.07600840
	0.025	0.02736103	0.03280644	−4.92987718
	0.0125	0.01089531	2.51126802	1.32841601
	0.00625	0.00527063	2.06717301	1.04765914
	0.003125	0.00165509	3.18450288	1.67106818

Note: Here $err_h = \|u_h - u^*\|_{L^1}$, where u_h is the penalized solution for Fishery problem when the mesh size is h.

err_h representing the $\|u^* - u_h\|_{L^1}$. Note that for almost all mesh size values the L^1 norm error between u_h and u^* is relatively close to the mesh size value. However, in the case when $h = 0.05$, we have that the L^1 norm error is significantly less than the mesh size. This is because the mesh size influences the discretization of time interval $[0, T]$ to where one mesh point is exceedingly close to the true switching point value. We look at the last column of Table 9.3 to determine what the rate of convergence between the penalized solution and the exact solution is. All but two entries in the last column of Table 9.3 are values that are slightly greater than one, and the entries that are not close to one involved using the L^1 norm error for when $h = 0.05$. Based on the values of the last column of Table 9.3, we find the convergence rate as being slightly better than a linear rate. Additionally, form Table 9.3, we take the natural logarithm of values found in columns 2 and 3, and use least squares method to indicate if there is a linear relationship between $\ln(h)$ and $\ln(\text{err}_h)$. Since we view data point $(\ln(0.05), \ln(0.00089762))$ as being an outlier, we also perform least squares between $\ln(h)$ and $\ln(\text{err}_h)$ without the outlier point. In Figure 9.6, the blue line does a better job with fitting to the data than the red line. In Table 9.4, we see that the goodness of fit associated with the blue line is very strong. We look at the slope associated with the blue line to determine the rate of convergence. Since the slope of the blue line is approximately 1.200738, we have further indication that the rate of convergence between the penalized solution and the exact solution is better than linear.

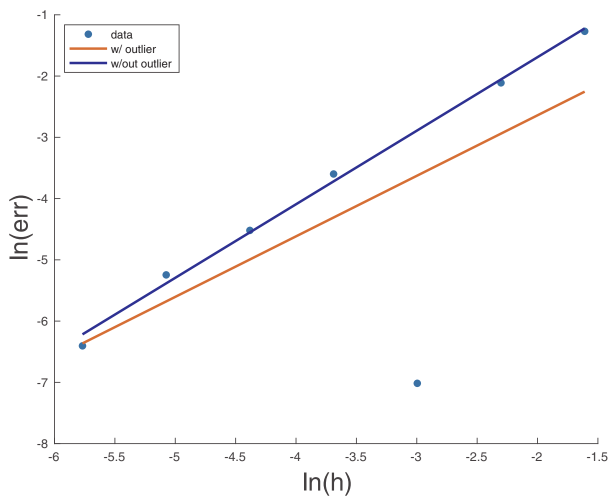


Fig. 9.6. Linear fit of data points $(\ln(h), \ln(\text{err}_h))$ from Table 9.3. Red line is the linear fit of the data points including outlier point $(\ln(0.05), \ln(0.00089762))$. Blue line is the linear fit of data points excluding the outlier point.

Table 9.4. Linear fit table.

$\ln(\text{err}_h) = m(\ln(h)) + b$	**Slope** m	$\boldsymbol{y}$**-intercept** b	r^2
With outlier	0.988064	-0.6645189	0.4826212
Without outlier	1.200738	0.7100375	0.9959617

9.4 Example 2: The Plant Problem

In this section, we implement a biological optimal control problem from David King and Jonathan Roughgarden[8] that can be solved analytically. In Ref. 8, King and Roughgarden use optimal control theory to study the allocation strategies that annual plants possess when distributing photosynthate to components of the plant that pertain to *vegetative growth* and to components that pertain to *reproductive growth*. The control variable $u(t)$ involved represents the fraction of photosynthate being reserved for vegetative growth. The remaining photosynthate, $1 - u(t)$, is assigned to aid in the reproductive growth processes. The annual plants that are being modeled are assumed to live in an environment where consecutive seasons are identical. Since $u(t)$ represents a fraction, the control variable is assumed to be bounded in the obvious way:

$$0 \leq u(t) \leq 1 \quad \text{for all } t \in [0, T], \tag{9.61}$$

where T represents the maximum season length. The optimal control problem has two state variables where one state x_1 represents the weight of the vegetative part of a plant, while the second state x_2 represents the weight of the reproductive part of a plant. The construction of the state equations is a continuous version of a discrete time model of reproduction for annual plants that was developed by Cohen[44] and is given by

$$x_1'(t) = u(t)x_1, \qquad x_1(0) > 0, \tag{9.62}$$

$$x_2'(t) = (1 - u(t))x_2, \quad x_2(0) \geq 0. \tag{9.63}$$

For the construction of an objective functional, King and Roughgarden[8] aim to find the strategy that maximizes total reproduction and assert that such a strategy is "expected in organisms if natural selection maximizes total reproduction." Their construction of a cost functional that measures the total reproduction produced by a plant is based on Cohen's discrete model in[45] which uses the expectation of the logarithm of seed yield to express the long-term rate of population increases observed in annual plants. The objective functional is chosen to have natural logarithm of x_2, which helps this problem to have singular arcs. The optimal control problem is then

$$\max_{u \in \mathcal{A}} J(u) = \int_0^T \ln(x_2(t)) \mathrm{d}t \tag{9.64}$$

subject to the state equations (9.62)–(9.63) and control bounds (9.61). The set of the admissible controls is defined $\mathcal{A} = \{u \in L^1(0, T) : \text{for all } t \in [0, T], \ u(t) \in A, \ A = [0, 1]\}$. Existence of an optimal control follows from the Filippov–Cesari Existence Theorem.[9] An explicit solution is obtained for problem (9.64) that satisfies Pontryagin's Maximum Principle,[10] the Generalized Legendre Clebsch Condition or Kelley's Condition,[22–24,43] the Strengthened Generalized Legendre Clebsch Condition, and McDanell and Powers's Junction Theorem.[21] The structure of the optimal allocation strategy u^* depends on the maximum season length T, the initial weight of the vegetative part of the plant $x_1(0)$, and the initial weight of the reproductive part of the plant $x_2(0)$.

In this section, we provide a summary of the solution to the equivalent minimization problem as follows:

$$\begin{array}{ll} \min & -J(u) = -\int_0^T \ln(x_2(t))\mathrm{d}t \\ \text{s.t.} & \dot{x}_1(t) = u(t)x_1, \\ & \dot{x}_2(t) = (1-u(t))x_1, \\ & x_1(0) = x_{1,0} > 0, \quad x_2(0) = x_{2,0} \geq 0, \\ & 0 \leq u(t) \leq 1. \end{array} \tag{9.65}$$

Additionally we demonstrate how to discretize the penalized version of problem (9.65)

$$\begin{array}{ll} \min & J_\rho(u) = -\int_0^T \ln(x_2(t))\mathrm{d}t + \rho V(u) \\ \text{s.t.} & \dot{x}_1(t) = u(t)x_1, \\ & \dot{x}_2(t) = (1-u(t))x_1, \\ & x_1(0) = x_{1,0} > 0, \quad x_2(0) = x_{2,0} \geq 0, \\ & 0 \leq u(t) \leq 1, \end{array} \tag{9.66}$$

where $0 \leq \rho < 1$ is the bounded variation tuning parameter and $V(u)$ measures the total variation of control u, as defined in Eq. (9.56). We then use PASA to numerically solve problems (9.65) and (9.66) when parameters $x_{1,0}$, $x_{2,0}$, and T are set in such a way where u^* contains a singular subarc.

9.4.1 *Solving for the singular case to the plant problem*

Before we present a summary of the solution to problem (9.65), we discuss some terminology that is used in King and Roughgarden's work[8] to describe the events when $u^*(t) = 0$ and when $u^*(t) = 1$. We say that u^* is *purely reproductive* at time t when $u^*(t) = 0$. We say that u^* is *purely vegetative* at time t when $u^*(t) = 1$. Our method for obtaining a solution to the minimization problem (9.65) parallels with what is used in King and Roughgarden.[8] Our method for solving problem (9.65) involves Pontryagin's Maximum Principle.[10] We compute the Hamiltonian to problem (9.65) as

$$H = -\ln x_2 + x_1(\lambda_1 - \lambda_2)u + x_1\lambda_2.$$

We take the partial derivatives of the Hamiltonian with respect to x_1 and x_2 to construct differential equations for costate variables λ_1 and λ_2 as follows:

$$\lambda_1'(t) = -\frac{\partial H}{\partial x_1} = (\lambda_2 - \lambda_1)u - \lambda_2, \tag{9.67}$$

$$\lambda_2'(t) = -\frac{\partial H}{\partial x_2} = \frac{1}{x_2}, \tag{9.68}$$

and use the transversality conditions to obtain boundary conditions for the costate equations

$$\lambda_1(T) = \lambda_2(T) = 0. \tag{9.69}$$

We construct the switching function $\psi(t)$ by taking the partial derivative of the Hamiltonian with respect to control u as:

$$\psi(t) = x_1(\lambda_1 - \lambda_2) \tag{9.70}$$

Since $x_1(t) > 0$ for all t, the sign of ψ depends on the sign of $\lambda_1 - \lambda_2$. By Pontryagin's Maximum Principle, the optimal allocation strategy $u^*(t)$ is as follows:

$$u^*(t) = \begin{cases} 0 & \text{whenever } \lambda_1(t) > \lambda_2(t), \\ \text{singular} & \text{whenever } \lambda_1(t) = \lambda_2(t), \\ 1 & \text{whenever } \lambda_1(t) < \lambda_2(t). \end{cases} \tag{9.71}$$

To solve for the singular case, we assume that there is an interval $I \subset [0, T]$ where $\lambda_1(t) = \lambda_1(t)$ for all $t \in I$. This implies that $\lambda_1'(t) \equiv \lambda_2'(t)$ on I. From adjoint equations (9.67)–(9.68), we have that the following holds on I:

$$(\lambda_2(t) - \lambda_1(t))u(t) - \lambda_2(t) = \frac{1}{x_2(t)},$$

which yields the following on interval I:

$$\lambda_2(t) = -\frac{1}{x_2(t)}. \tag{9.72}$$

It follows from Eqs. (9.68) and (9.72) that

$$\dot{\lambda}_2 = -\lambda_2$$

for all $t \in I$. Solving for the above differential equation yields

$$\lambda_2(t) = A\mathrm{e}^{-t},$$

where $A = -\frac{1}{x_2(t_1)}\mathrm{e}^{t_1}$ and t_1 is the time when the singular subarc begins.

We can construct a differential equation for $x_2(t)$ on interval I by differentiating Eq. (9.72) with respect to time and rearranging the terms as follows:

$$\dot{x}_2 = x_2. \tag{9.73}$$

Equations (9.63) and (9.73) yield

$$(1 - u(t))x_1 = x_2,$$

which implies that

$$\frac{x_2(t)}{x_1(t)} = 1 - u(t) \tag{9.74}$$

on interval I. We differentiate (9.74) to get

$$-\dot{u}(t) = \frac{x_1\dot{x}_2 - \dot{x}_1 x_2}{x_1^2}.$$

We replace $\dot{x}_1$ and $\dot{x}_2$ in the above equation with state Eq. (9.62) and equation (9.73), respectively, and obtain the following:

$$-\dot{u}(t) = \frac{x_2}{x_1}(1 - u(t)). \tag{9.75}$$

Substituting (9.74) into (9.75) yields

$$\frac{\dot{u}(t)}{(1 - u(t))^2} = -1.$$

We solve for the separable equation above and obtain the following solution:

$$u(t) = 1 - \frac{1}{C - t}, \tag{9.76}$$

on interval I, where C is some constant. Observe that from Eq. (9.76), (9.74) becomes

$$\frac{x_2(t)}{x_1(t)} = \frac{1}{C - t}, \tag{9.77}$$

on interval I.

In order to find constant term C we need to find the time, t_2, when the singular subarc ends. As mentioned in,[8] we can be certain that u^* is not singular at time T because the singular case solution for adjoint variable $\lambda_2(t)$ given in equation (9.72) would not satisfy the transversality condition (9.69). Now from the adjoint equations (9.67) and (9.68) and the transversality conditions (9.69), we have that $\dot{\lambda}_1(t) \to 0$ and $\dot{\lambda}_2(t) \to \frac{1}{x_2(T)}$ as $t \to T$. The derivatives of λ_1 and λ_2 near T and the transversality conditions imply that $\lambda_1(t) > \lambda_2(t)$ for some interval, namely $(t_2, T]$, and so from (9.71), we have that $u^*(t) \equiv 0$ on $(t_2, T]$. So, an optimal control that possesses a singular region switches to being purely reproductive on t_2 and remains purely reproductive on the interval $(t_2, T]$.

In order to find t_2, we must first solve for the state equations and the associated adjoint equations on time interval $[t_2, T]$. From the state equations (9.62)–(9.63), $u(t) \equiv 0$ on $[t_2, T]$ implies the following:

$$\begin{aligned} x_1(t) &= x_1(t_2), \\ x_2(t) &= x_1(t_2)(t - t_2) + x_2(t_2), \end{aligned} \tag{9.78}$$

where

$$x_1(t_2) = (C - t_2)x_2(t_2), \quad \text{and} \quad x_2(t_2) = x_2(t_1)\mathrm{e}^{t_2 - t_1}.$$

Differentiating $x_2(t)$ given in Eq. (9.78) yields

$$\dot{x}_2(t) = x_1(t) = x_1(t_2). \tag{9.79}$$

We use the differential found from the above separable equation and the transversality condition (9.69) to solve for the adjoint equation (9.68).

$$\lambda_2(T) - \lambda_2(t) = \int_{x_2(t)}^{x_2(T)} \frac{1}{x_2} \frac{\mathrm{d}x_2}{x_1(t_2)}$$

$$\lambda_2(t) = \frac{1}{x_1(t_2)} \ln\left(\frac{x_2(t)}{x_2(T)}\right) \text{ for } t \in [t_2, T]. \tag{9.80}$$

We use Eq. (9.80) and adjoint equation (9.67) to obtain the following differential equation for λ_1:

$$\dot{\lambda}_1(t) = -\frac{1}{x_1(t_2)} \ln\left(\frac{x_2(t)}{x_2(T)}\right). \tag{9.81}$$

Again, we use the differential obtained from separable equation (9.79) and the terminal condition for λ_1 (9.69) to solve the above differential equation as follows:

$$\lambda_1(T) - \lambda_1(t) = -\frac{1}{(x_1(t_2))^2} \int_{x_2(t)}^{x_2(T)} (\ln x_2 - \ln x_2(T))\mathrm{d}x_2$$

$$-\lambda_1(t) = -\frac{1}{(x_1(t_2))^2} \left[x_2(t) - x_2(T) + x_2(t) \ln\left(\frac{x_2(T)}{x_2(t)}\right)\right].$$

By negating the above equality and applying logarithmic rules, we obtain a solution for $\lambda_1(t)$:

$$\lambda_1(t) = -\frac{1}{(x_1(t_2))^2} \left[x_2(t) \ln\left(\frac{x_2(t)}{x_2(T)}\right) + x_2(T) - x_2(t)\right]$$
$$\text{for } t \in [t_2, T]. \tag{9.82}$$

After finding the solutions for the state and adjoint equations on interval $(t_2, T]$, we are ready to solve for t_2. Since $\lambda_1(t)$ and $\lambda_2(t)$ are continuous on $[0, T]$ and $\lambda_1(t) \equiv \lambda_2(t)$ on the singular case, we have $\lambda_1(t_2) = \lambda_2(t_2)$ We equate Eqs. (9.82) and (9.80) at $t = t_2$ and solve for t_2 as follows:

$$-\frac{1}{(x_1(t_2))^2} \left[x_2(t_2) \ln\left(\frac{x_2(t_2)}{x_2(T)}\right) + x_2(T) - x_2(t_2)\right]$$
$$= -\frac{1}{x_1(t_2)} \ln\left(\frac{x_2(t_2)}{x_2(T)}\right).$$

We multiply the above equation by $x_1(t_2)$ and rearrange the terms to obtain the following equation:

$$x_2(T) - x_2(t_2) = (x_1(t_2) + x_2(t_2)) \ln\left(\frac{x_2(T)}{x_2(t_2)}\right). \tag{9.83}$$

We use (9.78) to evaluate $x_2(T)$, which yields $x_2(T) = x_1(t_2)$ $(T - t_2) + x_2(t_2)$, and apply it to Eq. (9.83) as follows:

$$T - t_2 = \left[1 + \frac{x_2(t_2)}{x_1(t_2)}\right] \left[\ln\left(\frac{x_1(t_2)}{x_2(t_2)}(T - t_2) + 1\right)\right]. \tag{9.84}$$

We also have $\dot{\lambda}_1(t_2) = \dot{\lambda}_2(t_2)$, so using (9.72), (9.81), and $x_2(T) = x_1(t_2)(T - t_2) + x_2(t_2)$, we find that

$$\frac{x_1(t_2)}{x_2(t_2)} = \ln\left(\frac{x_1(t_2)}{x_2(t_2)}(T - t_2) + 1\right). \tag{9.85}$$

We rewrite Eqs. (9.84) and (9.85) into terms of y and z, where $y = T - t_2$ and $z = \frac{x_2(t_2)}{x_1(t_2)}$, to obtain a nonlinear system of equations that can be solved numerically through a nonlinear solver. Consequently, we have the following:

$$y = T - t_2 \approx 2.79328213, \quad \text{and} \tag{9.86}$$

$$z = \frac{x_2(t_2)}{x_1(t_2)} \approx 0.55763674. \tag{9.87}$$

Using $x_2(T) = x_1(t_2)(T - t_2) + x_2(t_2)$ and $x_1(t) = x_1(t_2)$ for $t \in [t_2, T]$ yields

$$\frac{x_2(T)}{x_1(T)} = (T - t_2) + \frac{x_2(t_2)}{x_1(t_2)}.$$

From Eqs. (9.86) and (9.87), we then have

$$\frac{x_2(T)}{x_1(T)} \approx 3.35091887. \tag{9.88}$$

Additionally, from Eqs. (9.84) and (9.85) we have

$$T - t_2 = \left[1 + \frac{x_2(t_2)}{x_1(t_2)}\right] \frac{x_1(t_2)}{x_2(t_2)} = 1 + \frac{x_1(t_2)}{x_2(t_2)},$$

which implies that

$$\frac{x_2(t_2)}{x_1(t_2)} = \frac{1}{(T-1)-t_2}.$$

Finally, we can evaluate Eq. (9.77) at $t = t_2$ and use the above equation to solve for constant C, and we have that $C = T - 1$.

9.4.1.1 *Summary of singular case solution to the plant problem*

In Section 9.4.1, we find that if the optimal control u^* to problem (9.65) contains a singular subarc where u^* becomes singular at time t_1, then u^* is the following on the interval $[t_1, T]$:

$$u^*(t) = \begin{cases} 1 - \dfrac{1}{T-1-t} & \text{when } t \in [t_1, t_2), \\ 0, & \text{when } t \in [t_2, T], \end{cases} \tag{9.89}$$

where $t_2 \approx T - 2.79328213$. The corresponding state solutions x_1 and x_2 are as follows:

$$x_1(t) = \begin{cases} (T-1-t)x_2(t) & \text{when } t \in [t_1, t_2], \\ x_1(t_2) & \text{when } t \in [t_2, T], \end{cases} \tag{9.90}$$

$$x_2(t) = \begin{cases} x_2(t_1)\mathrm{e}^{t-t_1} & \text{when } t \in [t_1, t_2], \\ x_1(t_2)(t-t_2) + x_2(t_2) & \text{when } t \in [t_2, T], \end{cases} \tag{9.91}$$

where $x_1(t_2) = (T-1-t_2)x_2(t_2)$ and $x_2(t_2) = x_2(t_1)\mathrm{e}^{t_2-t_1}$. The corresponding solution to the adjoint variables λ_1 and λ_2 are as follows:

$$\lambda_1(t) = \begin{cases} \lambda_2(t) & \text{when } t \in [t_1, t_2], \\ -\dfrac{1}{(x_1(t_2))^2}\left[x_2(t)\ln\left(\dfrac{x_2(t)}{x_2(T)}\right)\right. & \\ \qquad \left. + x_2(T) - x_2(t)\right] & \text{when } t \in [t_2, T], \end{cases} \tag{9.92}$$

$$\lambda_2(t) = \begin{cases} -\dfrac{1}{x_2(t_1)}\mathrm{e}^{t_1-t} & \text{when } t \in [t_1, t_2], \\ \dfrac{1}{x_1(t_2)}\ln\left(\dfrac{x_2(t)}{x_2(T)}\right) & \text{when } t \in [t_2, T]. \end{cases} \tag{9.93}$$

Additionally, King and Roughgarden's work[8] uses the generalized Legendre–Clebsch Condition,[21–23] also referred to as Kelley's condition,[24,43] to show that the second order necessary condition for optimality is satisfied. The generalized Legendre–Clebsch Condition involves finding what is called the *order* of a singular arc, which is defined as the integer q being such that $\left(\frac{\mathrm{d}^{2q}}{\mathrm{d}t^{2q}}\frac{\partial H}{\partial u}\right)$ is the lowest order total derivative of the partial derivative of the Hamiltonian with respect to u, in which control u appears explicitly. By using Eqs. (9.62), (9.63), (9.67), (9.68), and (9.70), we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial H}{\partial u} = \frac{\mathrm{d}}{\mathrm{d}t}\psi = -x_1\lambda_2 - \frac{x_1}{x_2},$$
$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\partial H}{\partial u} = \frac{\mathrm{d}^2}{\mathrm{d}t^2}\psi = \left(-x_1\lambda_2 - \frac{x_1}{x_2} - \frac{x_1^2}{x_2^2}\right)u - \frac{x_1}{x_2} + \frac{x_1^2}{x_2^2}.$$

So for problem (9.65), the order of singular subarc u^* is 1. By General Legendre Clebsch Condition, if u^* is an optimal singular control on some interval $[t_1, t_2)$ of order q, then it is necessary that

$$(-1)^q\frac{\partial}{\partial u}\left[\frac{\mathrm{d}^{2q}}{\mathrm{d}t^{2q}}\left(\frac{\partial H}{\partial u}\right)\right] \geq 0$$

if the extremum is a minimum, and the inequality is reversed if the extremum is a maximum.[21] Additionally, if the above inequality is strict, then the strengthened Legendre–Clebsch Condition holds. On singular region $[t_1, t_2)$, we have that

$$(-1)^1\frac{\partial}{\partial u}\left[\frac{\mathrm{d}^2}{\mathrm{d}t^2}\left(\frac{\partial H}{\partial u}\right)\right] = \frac{x_1^2}{x_2^2} > 0,$$

so the singular arc is of order 1 and satisfies both the generalized Legendre–Clebsch Condition and the strengthened Legendre–Clebsch Condition.

What we have yet to discuss is conditions for problem (9.65) that determine when the optimal control u^* is bang-bang or concatenations of bang and singular controls. In King and Roughgarden,[8] King and Roughgarden construct a series of conditions for determining the structure of the optimal control solution u^* to problem (9.65). Most of the conditions are based upon the value of terminal time T and the ratio of the reproductive to vegetative weight at time $t = 0$. We provide a summary of these conditions (see Ref. 8 for details on how these conditions are constructed):

(1) The optimal allocation strategy cannot be purely reproductive on the entire time interval ($u^*(t)$ cannot be zero over the entire time interval $[0, T]$) unless $T \leq 3.35091887$ and $\frac{x_2(0)}{x_1(0)} \leq 3.35091887 - T$.

(2) If $T > 3.35091887$ and $0 \leq \frac{x_2(0)}{x_1(0)} < 0.55763674e^{T-2.79328213}$, the optimal allocation, u^*, contains a singular subarc.

   (a) If $\frac{x_2(0)}{x_1(0)} = \frac{1}{T-1}$, then the optimal allocation strategy u^* begins with a singular subarc and switches to being purely reproductive at time $t_2 = T - 2.79328213$.

   (b) If $0 \leq \frac{x_2(0)}{x_1(0)} < \frac{1}{T-1}$, then the optimal control u^* contributes to reproductive growth before and after the singular subarc occurs.

   (c) If $\frac{1}{T-1} < \frac{x_2(0)}{x_1(0)} < 0.55763674e^{T-2.79328213}$, then u^* begins contributing to purely vegetative growth before the singular subarc occurs. At time $t_2 = T - 2.79328213$, u^* switches from being singular to being purely reproductive.

(3) If $\frac{x_2(0)}{x_1(0)} \geq 0.55763674e^{T-2.79328213}$, then the optimal control must be bang-bang, where u^* begins with purely vegetative growth and switches once to being purely reproductive.

The value $\frac{1}{T-1}$ corresponds to evaluating the right-hand side of equation (9.77) at $t = 0$. Additionally, some of the numerical values shown in the above conditions correspond to Eqs. (9.86), (9.87), and (9.88).

We are only interested in finding the exact solutions to problem (9.65) in the cases where the optimal control contains a singular subarc, i.e., Cases 2a, 2b, and 2c. For Case 2a, we use Eqs. (9.89)–(9.93) with $t_1 = 0$ to construct the exact solution which is as follows:

Case 2a: If $T > 3.35091887$, $0 \leq \frac{x_2(0)}{x_1(0)} < 0.55763674e^{T-2.79328213}$, and $\frac{x_2(0)}{x_1(0)} = \frac{1}{T-1}$, then

$$u^*(t) = \begin{cases} 1 - \dfrac{1}{T - t - 1} & 0 \leq t \leq t_2, \\ 0 & t_2 < t \leq T, \end{cases} \tag{9.94}$$

$$x_1(t) = \begin{cases} x_2(t)(T-t-1) & 0 \le t \le t_2, \\ x_1(t_2) & t_2 \le t \le T, \end{cases}$$

$$x_2(t) = \begin{cases} x_2(0)e^t & 0 \le t \le t_2, \\ x_1(t_2)(t-t_2) + x_2(t_2) & t_2 \le t \le T, \end{cases}$$

$$\lambda_1(t) = \begin{cases} \lambda_2(t) - \dfrac{1}{(x_1(t_2))^2} & 0 \le t \le t_2, \\ \qquad \times \left[x_2(t)\ln\left(\dfrac{x_2(t)}{x_2(T)}\right) + x_2(T) - x_2(t)\right] & t_2 \le t \le T, \end{cases}$$

$$\lambda_2(t) = \begin{cases} -\dfrac{1}{x_2(0)}e^{-t} & 0 \le t \le t_2, \\ \dfrac{1}{x_1(t_2)}\ln\left(\dfrac{x_2(t)}{x_2(T)}\right) & t_2 \le t \le T, \end{cases}$$

where $t_2 = T - 2.79328213$, $x_1(t_2) = (T-1-t_2)x_2(t_2)$, and $x_2(t_2) = x_2(0)e^{t_2}$.

For Case 2b, we solve the state equations (9.62) and (9.63) on the interval $[0, t_1]$ with u set to being 0, and then we use continuity of the state variables to get an explicit solution for t_1. Then we solve the adjoint equations (9.67) and (9.68) on $[0, t_1]$ where we use Eqs. (9.92)–(9.93) and continuity of λ_1 and λ_2 at t_1 to gain the following boundary conditions: $\lambda_1(t_1) = \lambda_2(t_1) = -\frac{1}{x_2(t_1)}$.

The exact solution for Case 2b is as follows:

Case 2b: If $T > 3.35091887$, $0 \le \frac{x_2(0)}{x_1(0)} < 0.55763674e^{T-2.79328213}$, and $\frac{x_2(0)}{x_1(0)} < \frac{1}{T-1}$, then

$$u^*(t) = \begin{cases} 0 & 0 \le t < t_1, \\ 1 - \dfrac{1}{T-t-1} & t_1 \le t \le t_2, \\ 0 & t_2 < t \le T, \end{cases}$$

$$x_1(t) = \begin{cases} x_1(0) & 0 \le t \le t_1, \\ (T - t - 1)x_2(t) & t_1 \le t \le t_2, \\ x_1(t_2) & t_2 \le t \le T, \end{cases}$$

$$x_2(t) = \begin{cases} x_1(0)t + x_2(0) & 0 \le t \le t_1, \\ x_2(t_1)\mathrm{e}^{t-t_1} & t_1 \le t \le t_2, \\ x_1(t_2)(t - t_2) + x_2(t_2) & t_2 \le t \le T, \end{cases}$$

$$\lambda_1(t) = \begin{cases} \dfrac{x_2(t)}{(x_1(0))^2} \ln\left(\dfrac{x_2(t_1)}{x_2(t)}\right) + \dfrac{x_2(t) - x_2(t_1)}{(x_1(0))^2} & \\ \quad + \dfrac{x_2(t) - x_2(t_1)}{x_2(t_1)x_1(0)} - \dfrac{1}{x_2(t_1)} & 0 \le t \le t_1, \\ \lambda_2(t) - \dfrac{x_2(t)}{(x_1(t_2))^2} & t_1 \le t \le t_2, \\ \quad \times \left(\ln\left(\dfrac{x_2(t)}{x_2(T)}\right) - 1\right) - \dfrac{x_2(T)}{(x_1(t_2))^2} & t_2 \le t \le T, \end{cases}$$

$$\lambda_2(t) = \begin{cases} \dfrac{1}{x_1(0)} \ln\left(\dfrac{x_2(t)}{x_2(t_1)}\right) - \dfrac{1}{x_2(t_1)} & 0 \le t \le t_1, \\ -\dfrac{1}{x_2(t_1)}\mathrm{e}^{t_1-t} & t_1 \le t \le t_2, \\ \dfrac{1}{x_1(t_2)} \ln\left(\dfrac{x_2(t)}{x_2(T)}\right) & t_2 \le t \le T, \end{cases}$$

where $t_2 \approx T - 2.7932821329007607$, $x_1(t_2) = x_2(t_2)(T - t_2 - 1)$, and $x_2(t_1) = x_1(0)t_1 + x_2(0)$. Also, switch t_1 is the solution to the following equation:

$$\frac{x_2(t_1)}{x_1(t_1)} = \frac{1}{T - t_1 - 1}.$$

The solution to the above equation is the following expression:

$$t_{1,\pm} = \frac{1}{2}\left[T - 1 - \frac{x_2(0)}{x_1(0)} \pm \sqrt{T^2 + 2T\left(\frac{x_2(0)}{x_1(0)} - 1\right) - 3 + \left(\frac{x_2(0)}{x_1(0)}\right)^2}\,\right],$$

and we choose t_1 to be the solution that is between the values 0 and t_2.

For Case 2c, we solve the state equations (9.62) and (9.63) on the interval $[0, t_1]$ with u set to being 1, and then use continuity of the state variables to obtain an equation that can be used to solve for t_1. Then we solve the adjoint equations (9.67) and (9.68) on $[0, t_1]$, where we used Eqs. (9.92)–(9.93) and continuity of λ_1 and λ_2 at t_1 to gain the following boundary conditions: $\lambda_1(t_1) = \lambda_2(t_1) = -\frac{1}{x_2(t_1)}$.

The exact solution for Case 2c is as follows:

Case 2c: If $T > 3.35091887$ and $\frac{1}{T-1} < \frac{x_2(0)}{x_1(0)} \leq 0.55763673\mathrm{e}^{T-2.79328213}$, then

$$u^*(t) = \begin{cases} 1 & 0 \leq t < t_1, \\ 1 - \dfrac{1}{T-t-1} & t_1 \leq t \leq t_2, \\ 0 & t_2 < t \leq T, \end{cases}$$

$$x_1(t) = \begin{cases} x_1(0)\mathrm{e}^{t} & 0 \leq t \leq t_1, \\ x_2(t)(T-t-1) & t_1 \leq t \leq t_2, \\ x_1(t_2) & t_2 \leq t \leq T, \end{cases}$$

$$x_2(t) = \begin{cases} x_2(0) & 0 \leq t \leq t_1, \\ x_2(t_1)\mathrm{e}^{t-t_1} & t_1 \leq t \leq t_2, \\ x_1(t_2)(t-t_2) + x_2(t_2) & t_2 \leq t \leq T, \end{cases}$$

$$\lambda_1(t) = \begin{cases} -\dfrac{1}{x_2(0)}\mathrm{e}^{t_1-t} & 0 \leq t \leq t_1, \\ \lambda_2(t) - \dfrac{x_2(t)}{(x_1(t_2))^2} & t_1 \leq t \leq t_2, \\ \quad \times \left(\ln\left(\dfrac{x_2(t)}{x_2(T)}\right) - 1\right) - \dfrac{x_2(T)}{(x_1(t_2)^2} & t_2 \leq t \leq T, \end{cases}$$

$$\lambda_2(t) = \begin{cases} \dfrac{1}{x_2(0)}(t - t_1) - \dfrac{1}{x_2(t_1)} & 0 \leq t \leq t_1, \\ -\dfrac{1}{x_2(0)}\mathrm{e}^{t_1 - t} & t_1 \leq t \leq t_2, \\ \dfrac{1}{x_1(t_2)} \ln\left(\dfrac{x_2(t)}{x_2(T)}\right) & t_2 \leq t \leq T, \end{cases}$$

where $t_2 = T - 2.79328213$, $x_1(t_1) = x_1(0)\mathrm{e}^{t_1}$, $x_2(t_1) = x_2(0)$, $x_1(t_2) = x_2(t_2)(T - t_2 - 1)$, and $x_2(t_2) = x_2(t_1)\mathrm{e}^{t_2 - t_1}$. Also, switch t_1 can be obtained by using a nonlinear solver for the following nonlinear equation:

$$\frac{x_2(0)}{x_1(0)}\mathrm{e}^{-}t_1 = \frac{1}{T - 1 - t_1}.$$

9.4.2 *Discretization of the plant problem*

We discretize the following regularized problem:

$$\begin{aligned} \min \quad & J_\rho(u) = -\textstyle\int_0^T \ln(x_2(t))\mathrm{d}t + \rho V(u) \\ \text{s.t.} \quad & \dot{x}_1(t) = u(t)x_1, \\ & \dot{x}_2(t) = (1 - u(t))x_1, \\ & x_1(0) = x_{1,0} > 0, \quad x_2(0) = x_{2,0} \geq 0, \\ & \qquad 0 \leq u(t) \leq 1, \end{aligned} \tag{9.95}$$

where $0 \leq \rho < 1$ is the bounded variation penalty parameter and $V(u)$ measures the total variation of control u, as defined in Eq. (9.56). We also discretize the following adjoint equations:

$$\lambda_1'(t) = -\frac{\partial H}{\partial x_1} = (\lambda_2 - \lambda_1)u - \lambda_2, \tag{9.96}$$

$$\lambda_2'(t) = -\frac{\partial H}{\partial x_2} = \frac{1}{x_2}, \tag{9.97}$$

with $\lambda_1(T) = \lambda_2(T) = 0$ being the transversality conditions.

We assume the control u is constant over each mesh interval. We partition time interval $[0, T]$, by using $N+1$ equally spaced nodes, $0 = t_0 < t_1 < \cdots < t_N = T$. For all $k = 0, 1, \ldots, N$, we assume that $x_{1,k} = x_1(t_k)$ and $x_{2,k} = x_2(t_k)$. For the control, we denote $u_k = u(t)$ for all $t_k \leq t < t_{k+1}$ when $k = 0, \ldots N-2$ and $u_{N-1} = u(t)$ for all $t_{N-1} \leq t \leq t_N$. So we have $\boldsymbol{x}_1 \in \mathbb{R}^{N+1}$, $\boldsymbol{x}_2 \in \mathbb{R}^{N+1}$, while $u \in \mathbb{R}^N$. We use a left-rectangular integral approximation for objective functional J_ρ in problem (9.95). The discretization of problem (9.95) is then

$$
\begin{aligned}
\min \quad & J_\rho(\boldsymbol{u}) = \sum_{k=0}^{N-1} (-h \ln (x_{2,k})) + \rho \sum_{k=0}^{N-2} |u_{k+1} - u_k| \\
& x_{1,k+1} = x_{1,k} + h(u_k x_{1,k}) \quad \text{for all } k = 0, \ldots, N-1, \\
& x_{2,k+1} = x_{2,k} + h((1-u_k)x_{1,k}) \quad \text{for all } k = 0, \ldots, N-1, \\
& \qquad x_{1,0} > 0, \quad x_{2,0} \geq 0 \\
& 0 \leq u_k \leq 1 \quad \text{for all } k = 0, \ldots, N-1,
\end{aligned}
$$

where $h = \frac{T}{N}$ is the mesh size and the first component of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ is set to being the initial condition associated with the state equations. Since PASA uses a gradient scheme for one of its phases, we need the cost functional to be differentiable. We perform a decomposition on each absolute value term in J_ρ to ensure that J_ρ is differentiable. We introduce two $N-1$ vectors $\boldsymbol{\zeta}$ and $\boldsymbol{\iota}$ whose entries are nonnegative. Each entry of $\boldsymbol{\zeta}$ and $\boldsymbol{\iota}$ is defined as

$$
|u_{k+1} - u_k| = \zeta_k + \iota_k \quad \text{for all } k = 0, \ldots, N-2.
$$

With this decomposition in mind, the discretized penalized problem is the following:

$$
\begin{aligned}
\min \quad & J_\rho(\boldsymbol{u}, \boldsymbol{\zeta}, \boldsymbol{\iota}) = \sum_{k=0}^{N-1} (-h \ln (x_{2,k})) + \rho \sum_{k=0}^{N-2} (\zeta_k + \iota_k) \\
& x_{1,k+1} = x_{1,k} + h(u_k x_{1,k}) \quad \text{for all } k = 0, \ldots, N-1, \\
& x_{2,k+1} = x_{2,k} + h(1-u_k)x_{1,k} \quad \text{for all } k = 0, \ldots, N-1, \\
& \qquad x_{1,0} > 0 \quad x_{2,0} \geq 0, \\
& 0 \leq u_k \leq \quad \text{for all } k = 0, \ldots, N-1, \\
& u_{k+1} - u_k = \zeta_k - \iota_k \quad \text{for all } k = 0, \ldots, N-2, \\
& \zeta_k \geq 0, \quad \iota_k \geq 0 \quad \text{for all } k = 0, \ldots, N-1.
\end{aligned} \tag{9.98}
$$

Note that for problem (9.98), we are minimizing the penalized objective functional with respect to three vectors, $\boldsymbol{u}, \boldsymbol{\zeta}$, and $\boldsymbol{\iota}$. The equality constraints associated with $\boldsymbol{\zeta}$ and $\boldsymbol{\iota}$ are linear constraints that PASA can interpret. The equality constraints associated with $\boldsymbol{\zeta}$ and $\boldsymbol{\iota}$ can be written like so:

$$\left[\boldsymbol{A} \middle| -\boldsymbol{I}_{N-1} \middle| \boldsymbol{I}_{N-1}\right] \begin{bmatrix} \boldsymbol{u} \\ \hline \boldsymbol{\zeta} \\ \hline \boldsymbol{\iota} \end{bmatrix} = \mathbf{0},$$

where $\boldsymbol{I}_{N-1}$ is the identity matrix with dimension $N-1$, $\mathbf{0}$ is the $N-1$ dimensional all zeros vector, and $\boldsymbol{A}$ is an $N-1 \times N$ sparse matrix defined on Eq. (9.7).

For finding the gradient of J_ρ in problem (9.98), we use Theorem 9.1. In order to compute the Lagrangian of problem (9.98), we rewrite the discretized state equations accordingly:

$$-x_{1,k+1} + x_{1,k} + h(u_k x_{1,k}) = 0, \quad \text{for all } k = 0, \dots, N,$$
$$-x_{2,k+1} + x_{2,k} + h(1-u_k)x_{1,k} = 0, \quad \text{for all } k = 0, \dots, N.$$

The Lagrangian of problem (9.98) is then

$$\begin{aligned} \mathcal{L}(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{u}, \boldsymbol{\zeta}, \boldsymbol{\iota}) = & \sum_{k=0}^{N-1} (-h \ln(x_{2,k})) + \rho \sum_{k=0}^{N-2} (\zeta_k + \iota_k) \\ & + \sum_{k=0}^{N-1} \lambda_{1,k}(-x_{1,k+1} + x_{1,k} + h(u_k x_{1,k})) \\ & + \sum_{k=0}^{N-1} \lambda_{2,k}(-x_{2,k+1} + x_{2,k} + h((1-u_k)x_{1,k})), \end{aligned}$$

where $\boldsymbol{\lambda_1}, \boldsymbol{\lambda_2} \in \mathbb{R}^{N-1}$ are the Lagrangian multiplier vectors. We take the partial derivative of $\mathcal{L}$ with respect to u_k and obtain

$$\frac{\partial \mathcal{L}}{\partial u_k} = h(x_{1,k}(\lambda_{1,k} - \lambda_{2,k})) \quad \text{for all } k = 0, \dots, N-1.$$

By using Theorem 9.1, we obtain the following:

$$\nabla_{\boldsymbol{u},\boldsymbol{\zeta},\boldsymbol{\iota}} J_\rho = \nabla_{\boldsymbol{u},\boldsymbol{\zeta},\boldsymbol{\iota}} \mathcal{L} = \left[\begin{array}{c} h(x_{1,0}(\lambda_{1,0} - \lambda_{2,0})) \\ \vdots \\ h(x_{1,k}(\lambda_{1,k} - \lambda_{2,k}) \\ \vdots \\ h(x_{1,N-1})(\lambda_{1,N-1} - \lambda_{2,N-1}) \\ \hline \rho \\ \vdots \\ \rho \\ \hline \rho \\ \vdots \\ \rho \end{array}\right], \qquad (9.99)$$

provided that condition (9.11) in Theorem 9.1 is satisfied. To satisfy (9.11), we take the partial derivative of the Lagrangian $\mathcal{L}$ with respect to $x_{1,k}$ and $x_{2,k}$ for all $k = 1, \ldots, N$, and note that we are not taking the partial derivative of $\mathcal{L}$ with respect to $x_{1,0}$ and $x_{2,0}$ because they are known values. Taking the partial derivative of $\mathcal{L}$ with respect to each state vector components yields the following expressions:

$$\frac{\partial \mathcal{L}}{\partial x_{1,k}} = h(\lambda_{1,k} u_k + \lambda_{2,k}(1 - u_k)) + \lambda_{1,k} - \lambda_{1,k-1} \quad \text{for } k = 1, \ldots, N-1 \qquad (9.100)$$

$$\frac{\partial \mathcal{L}}{\partial x_{1,N}} = -\lambda_{1,N-1} \qquad (9.101)$$

$$\frac{\partial \mathcal{L}}{\partial x_{2,k}} = -h\left(\frac{1}{x_{2,k}}\right) + \lambda_{2,k} - \lambda_{2,k-1} \quad \text{for } k = 1, \ldots, N-1 \qquad (9.102)$$

$$\frac{\partial \mathcal{L}}{\partial x_{2,N}} = -\lambda_{2,N-1}. \qquad (9.103)$$

For satisfying condition (9.11) from Theorem 9.1, we set the above equations equal to zeros and perform the following steps: solve for

$\lambda_{1,k-1}$ in Eq. (9.100); solve for $\lambda_{1,N-1}$ in Eq. (9.101); solve for $\lambda_{2,k-1}$ in Eq. (9.102); and solve for $\lambda_{2,N-1}$ in equation (9.103). Consequently, we generate a discretization for the adjoint equations (9.67) and (9.68) that also produces the transversality conditions (9.69) as follows:

$$\lambda_{1,k-1} = \lambda_{1,k} + h(\lambda_{1,k}u_k + \lambda_{2,k}(1-u_k)) \quad \text{for } k = 1, \ldots, N-1, \tag{9.104}$$

$$\lambda_{1,N-1} = 0, \tag{9.105}$$

$$\lambda_{2,k-1} = \lambda_{2,k} - h\left(\frac{1}{x_{2,k}}\right) \quad \text{for } k = 1, \ldots, N-1 \quad \text{and} \tag{9.106}$$

$$\lambda_{2,N-1} = 0. \tag{9.107}$$

9.4.3 *Numerical results of the plant problem*

We use PASA to numerically solve problem (9.65) with stopping tolerance set to being 10^{-10} and with our initial guess for the control being $u(t) = 0$ over the entire time interval $[0, T]$. We partition the time interval to where there are $N = 750$ mesh intervals with mesh size being $h = 0.00667$, and the discretization process is as described in Section 9.4.2 with $\rho = 0$. We are interested in solving the problem when the parameters are set so that a singular case occurs. King and Roughgarden[8] mention that from a biological standpoint, the initial weight of the reproductive part of a plant, $x_2(0)$, is always zero since germination involves vegetative growth only. Consequently, the only realistic situation where the exact solution to problem (9.65) contains a singular subarc is if the exact solution is of Case 2b with $x_{2,0} = 0$. However, we still would like to observe how PASA performs for all possible cases when the exact solution to problem (9.65) contains a singular subarc. Table 9.5 describes the parameter settings that we used for solving problem (9.65).

When parameters are set to the values shown in column three of Table 9.5, the exact solution to problem (9.65) will be of Case 2a where control u^* begins singular and switches to the purely reproductive case at $t_2 \approx 2.2067$. When parameters are set to the values shown in column four of Table 9.5, the exact solution to problem (9.65) is of Case 2b where control u^* begins purely reproductive,

Table 9.5. Numerical values of parameters used in computations for the plant problem.

Parameter	Description	Case 2a	Case 2b	Case 2c
T	Terminal time	5	5	5
$x_{1,0}$	Initial value for vegetative weight	4	1	1
$x_{2,0}$	Initial value for reproductive weight	1	10^{-4}	2

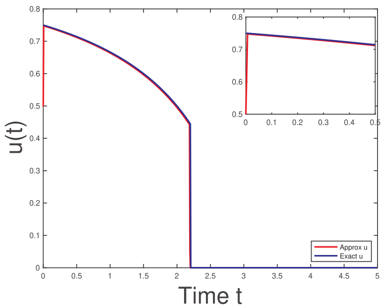
Table 9.6. Results from unpenalized solution to the plant problem for each case.

Case	$\lVert u^* - \hat{u}\rVert_{L^1}$	$\lVert u^* - \hat{u}\rVert_{L^\infty}$	t_1^*	$\hat{t}_1$	t_2^*	$\hat{t}_2$	$-J(u^*)$	$-J(\hat{u})$	Runtime
2a	0.017557	0.444444	NA	NA	2.2067	2.2	11.7874	11.7875	46.25
2b	0.017519	0.444444	0.2678	0.2667	2.2067	2.2	3.6009	3.6010	32.97
2c	0.013588	0.444444	1.5778	1.5733	2.2067	2.2	8.6130	8.6130	7.48

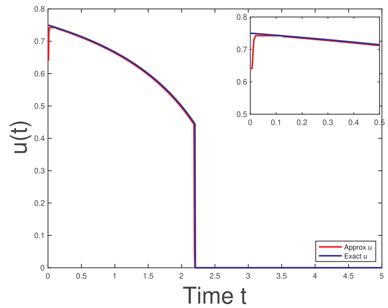
Note: The starred notation corresponds to the exact solution, while the hat notation corresponds to the approximated solution. Note $-J(u)$ is the left rectangular integral approximation of $\int_0^T \ln(x(t))dt$. We have that $\lVert u^* - \hat{u}\rVert_{L^\infty(0,T)} = 0.444444$ for each case. This is due to the approximated switch from singular to purely reproduction, $\hat{t_2}$, being off by one node. The Runtime column gives the time (in seconds) it takes for PASA to solve the problem.

switches to the singular case solution at $t_1 \approx 0.2678$, and switches back to being purely reproductive at $t_2 \approx 2.2067$. We want to emphasize that in Case 2b parameter settings, $x_{2,0} = 10^{-4}$ because we want this initial value to be a close approximation of zero without having any issues in computing the integral approximation for the objective functional $J(u) = \int_0^T (-\ln x_2(t))\mathrm{d}t$. When parameters are set to the values shown in the last column of Table 9.5, the exact solution to problem (9.65) is of Case 2c where control u^* begins purely vegetative, switches to the singular case solution at $t_1 \approx 1.5778$, and switches back to being purely reproductive at $t_2 \approx 2.2067$.

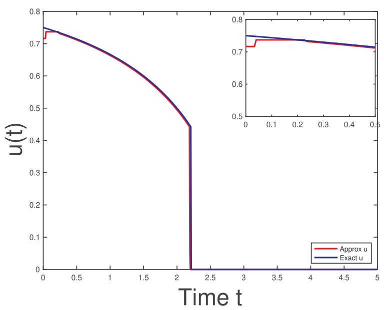
For each case, we first observe the unregularized solutions that PASA obtained when solving for problem (9.65) to see if it is even necessary to penalize this problem. When numerically solving the unpenalized problem, we obtain solutions that are not oscillatory for all three cases. Figures and descriptions corresponding to the unregularized solutions are given in Appendix Section 9.A.1 and in Table 9.6. We find PASA's unpenalized solution to be a sufficient approximation to the true solution to problem (9.65) in Cases 2b



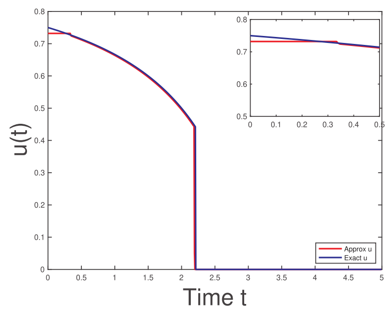
(a) u_ρ vs u^* for $\rho = 0$



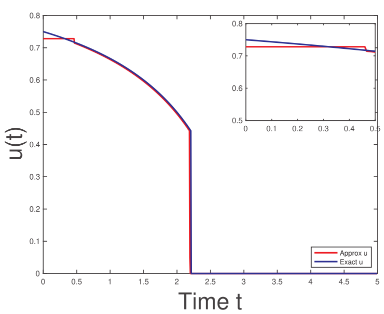
(b) u_ρ vs u^* for $\rho = 10^{-6}$



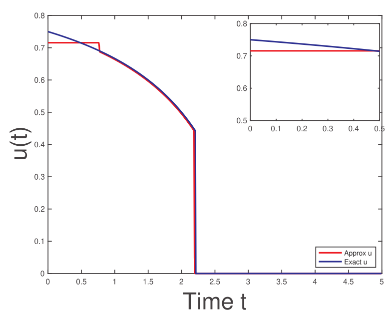
(c) u_ρ vs u^* for $\rho = 10^{-5}$



(d) u_ρ vs u^* for $\rho = 10^{-4}$



(e) u_ρ vs u^* for $\rho = 10^{-3}$



(f) u_ρ vs u^* for $\rho = 10^{-2}$

Fig. 9.7. Varying penalty to plant problem (Case 2a): Plots of regularized control u_ρ (red) vs optimal control u^* (blue) for tuning parameter $\rho \in \{0, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}\}$. The top right corner of each figure is a zoomed-in plot of the same figure over the time interval $[0, 0.5]$.

and 2c. However, in Figure 9.7(a) the unpenalized approximation for Case 2a, $\hat{u}$, appears to have an unusual dip at $t = 0$. Based upon explicit solution (9.94) and the parameters setting given in Table 9.5, $u^*(0) = 0.75$, but for the approximated, unpenalized solution we have $\hat{u}(0) \approx 0.4980$. We use PASA to solve for Problem (9.65) for Case 2a with $N = 1,000$ and $tol = 10^{-12}$ to see if such changes would provide any improvements for the initial value. The unregularized solution that PASA obtained when $N = 1,000$ and $tol = 10^{-12}$ still possesses a dip at $t = 0$ where $u(0) \approx 0.4985$. The reason behind this initial jump is because Case 2a is a degenerate case. When looking at the conditions between becoming a solution of Case 2a, Case 2b, and Case 2c, one notices that they all share a condition that pertains to how the ratio of initial values relates with the fraction $\frac{1}{T-1}$. When having parameters set to where $\frac{x_{2,0}}{x_{1,0}} = \frac{1}{T-1}$, the discretization of problem (9.65) can cause any numerical solver to converge to a solution that does not begin singular.

Although there is no oscillations in Figure 9.7(a), we add a bounded variation penalty term to the objective functional of problem (9.65) to see if the approximated penalized control, u_ρ, does not dip as significantly as the unpenalized solution did at $t = 0$. As before, we partition time interval $[0, T]$ to where there are $N = 750$ mesh intervals and use the discretization method described in Section 9.4.2. We use PASA to solve the regularized problem (9.66) for varying values of the tuning parameter $\rho \in \{0, 10^{-9}, 10^{-8}, 10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$, with initial guess for our control being $u(t) = 0$ over the entire time interval. We have parameters set to being the third column in Table 9.5, with stopping tolerance set to being 10^{-10}. In Table 9.7, we record the L^1 and L^∞ norm errors between the exact solution u^* and the penalized solution that PASA obtained, u_ρ, the approximated value of $u_\rho(0)$, the approximated switching point t_2, and the runtime that was needed to solve for penalized problem (9.66). In column four of Table 9.7, we are finding $\|u^* - u_\rho\|_{L^\infty(0,2)}$ because otherwise the whole column would be 0.444444, which is what we observed in Table 9.6. Observe from Table 9.7, the penalty parameter starts to influence the behavior of u_ρ when $\rho \geq 10^{-6}$. However, it is worth noting that for penalty parameter values $\rho = 10^{-9}, 10^{-8}$, 10^{-7}, PASA obtained a solution comparable to the unpenalized solution at a faster rate. Based upon

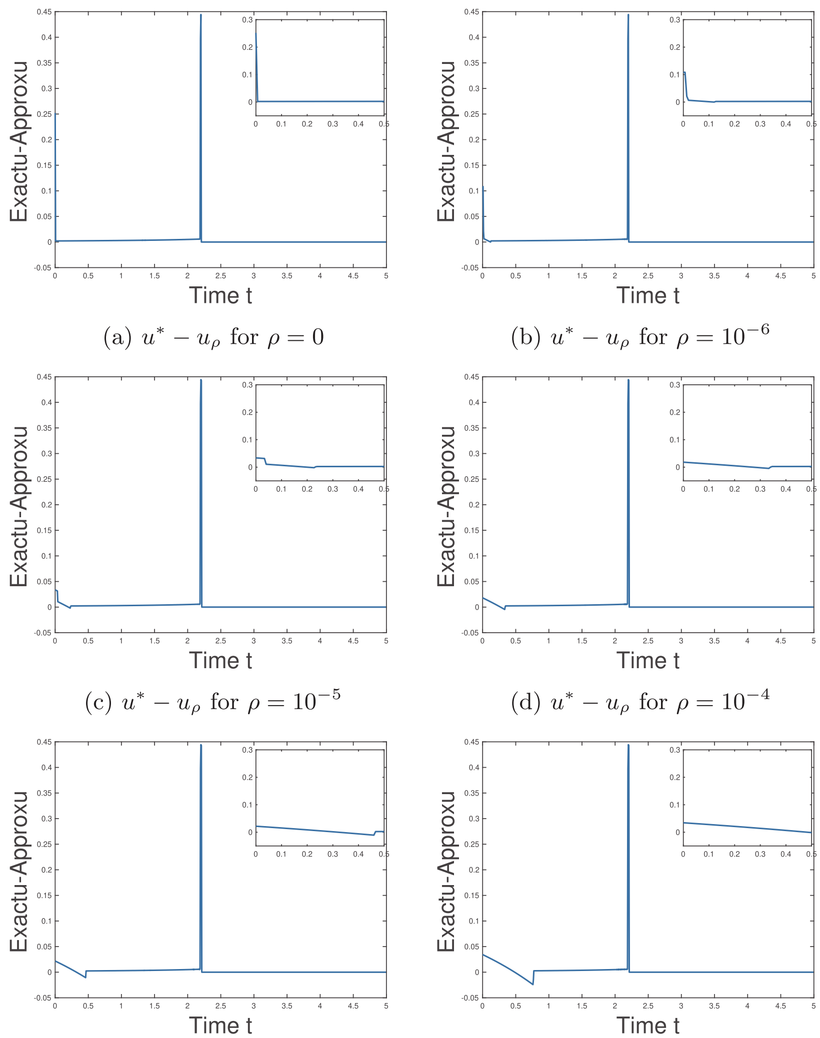
Table 9.7. Varying penalty parameter for plant problem.

ρ	$\|u_\rho - u^*\|_{L^1(0,T)}$	$\|u_\rho - u^*\|_{L^\infty(0,2)}$	$u_\rho(0)$	Switch t_2	Runtime
0	0.01755676	0.25200195	0.4980	2.2	46.25
10^{-9}	0.01755676	0.25153629	0.4985	2.2	45.81
10^{-8}	0.01755672	0.24953915	0.5005	2.2	36.36
10^{-7}	0.01755623	0.23038155	0.5196	2.2	32.01
10^{-6}	**0.01755173**	0.10869044	0.6413	**2.2067**	27.54
10^{-5}	0.01760049	0.03367945	0.7163	2.2667	32.82
10^{-4}	0.01785072	**0.01809606**	**0.7319**	2.2667	27.60
10^{-3}	0.01928274	0.02181486	0.7282	2.3667	26.27
10^{-2}	0.02628408	0.03439004	0.7156	2.6067	23.42
10^{-1}	0.06203874	0.07234748	0.6777	2.3667	**9.96**

Note: Case 2a. $N = 750$, tol $= 10^{-10}$, and $h = 0.00667$.

Table 9.7, u_ρ was closest to u^* with respect to the L^1 norm when $\rho = 10^{-6}$, and u_ρ switched to being non-singular at the node that was closest to the true switching point. The penalized solution u_ρ that was obtained when $\rho = 10^{-4}$ had an initial value that was the closest to the true solution's initial value $u^*(0) = 0.75$; however, this improvement seems to be at the expense of increasing the L^1 norm error between u_ρ and u^*.

In Figures 9.7 and 9.8, we provide a chart of figures that pertain to u_ρ and a chart of figures that pertain to $u^* - u_\rho$ for $\rho = 0, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}$, and 10^{-2}. Note that in Figures 9.7 and 9.8, the penalized solutions do not begin singular immediately; however, in comparison to the unpenalized solution, they give better approximations to $u^*(0)$. In Figure 9.7(d), the penalized solution associated with penalty parameter $\rho = 10^{-4}$ begins constant with constant value approximately being 0.7319 and switches to the singular case solution approximately at $t = 0.32$. Notice in Figures 9.7(e) and 9.7(f) that these penalized solutions also begin constant at values close to 0.7, but remain constant for a longer time period. Note also in the sixth column of Table 9.7 that for values $\rho \geq 10^{-5}$, u_ρ starts to overestimate the point at which the solution should switch from singular to non-singular. In Figure 9.7(b), we find that the unpenalized solution u_ρ associated with penalty parameter value $\rho = 10^{-6}$ did the best job in improving the approximated initial value without deviating too much from the exact solution. In Figure 9.9, we have the trajectories



(a) $u^* - u_\rho$ for $\rho = 0$

(b) $u^* - u_\rho$ for $\rho = 10^{-6}$

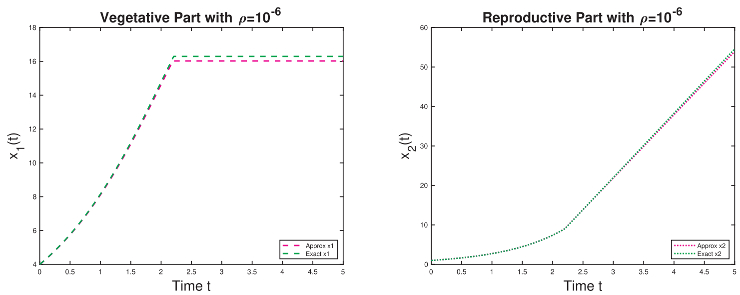
(c) $u^* - u_\rho$ for $\rho = 10^{-5}$

(d) $u^* - u_\rho$ for $\rho = 10^{-4}$

(e) $u^* - u_\rho$ for $\rho = 10^{-3}$

(f) $u^* - u_\rho$ for $\rho = 10^{-2}$

Fig. 9.8. Varying penalty to plant problem (Case 2a): Plots of $u^* - u_\rho$ (cyan) for tuning parameter $\rho \in \{0, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}\}$. The top right corner of each figure is a zoomed-in plot of the same figure over the time interval $[0, 0.5]$.



(a) Exact State x_1 (green) vs $x_{1,\rho}$ with $\rho = 10^{-6}$ (pink)

(b) Exact State x_2 (green) vs $x_{2,\rho}$ with $\rho = 10^{-6}$ (pink)

Fig. 9.9. Corresponding trajectories of x_1 and x_2 for penalized plant problem with $\rho = 10^{-6}$.

of x_1 and x_2 corresponding to u_ρ when $\rho = 10^{-6}$. The corresponding state solutions to the penalized control are comparable to the exact state solutions.

9.5 Example 3: The SIR Problem

Our next example is from Ledzewicz, Aghaee, and Schättler's article.[17] This example demonstrates how PASA can be applied to problems in the absence of a true solution. In Ref. 17, Ledzewicz *et al.* use an SIR model with demography to study the 2013–2016 outbreak of Ebola in West Africa. In an SIR model, the total population N is divided into the following three compartments: the susceptible class (S), the infectious class (I), and the recovered class (R). The dynamics in consideration are based on a modification of a system of equations that was first presented in Brauer and Castillo-Chavez's book[38]:

$$\dot{S} = \gamma N - \nu S - \beta \frac{IS}{N} + \rho R, \quad S(0) = S_0, \tag{9.108}$$

$$\dot{I} = \beta \frac{IS}{N} - (\nu + \mu) I - \alpha I, \quad I(0) = I_0, \tag{9.109}$$

$$\dot{R} = -\nu R - \rho R + \alpha I, \quad R(0) = R_0. \tag{9.110}$$

The Greek letters in the equations above represent constant parameters. In Eq. (9.108), the coefficient γ is the birth rate per unit time, and it is assumed that the population of newborns are immediately classified as being susceptible to the disease. Standard incidence is assumed in the model with transmission rate being β, and it is assumed that infected members recover from Ebola at rate α. Parameter ν is the natural death rate while parameter μ is the death rate due to infection. We emphasize here that for this model ρ is not the bounded variation tuning parameter, but instead ρ is the parameter that measures the rate at which a recovered individual becomes again susceptible to Ebola.

In Ref. 17, Ledzewicz, *et al.* construct an optimal control problem to gain understanding on how treatment and a theoretical vaccination of Ebola should be applied to best contribute to limiting the spread of the disease. For the state equations, they incorporate controls, u and v, into Eqs. (9.108)–(9.110), where u represents the vaccination rate and v represents the treatment rate. The modified dynamics are as follows:

$$\dot{S} = \gamma N - \nu S - \beta \frac{IS}{N} + \rho R - \kappa S u, \quad S(0) = S_0, \tag{9.111}$$

$$\dot{I} = \beta \frac{IS}{N} - (\nu + \mu) I - \alpha I - \eta I v, \quad I(0) = I_0, \tag{9.112}$$

$$\dot{R} = -\nu R - \rho R + \kappa S u + \alpha I + \eta I v, \quad R(0) = R_0, \tag{9.113}$$

where κ and η are denoted as being the efficacy of vaccination and treatment, respectively. For constructing of an objective functional, Ledzewicz *et al.*[17] have the following goal in mind: *"Given initial population sizes of all three classes, S_0, I_0, and R_0, find the best strategy in terms of the combined efforts of vaccination and treatment that minimizes the number of infectious persons while at the same time also taking into account the cost of vaccination and treatment."* The objective functional in consideration for minimization is as follows:

$$J(u, v) = \int_0^T (aI(t) + bu(t) + cv(t))\mathrm{d}t \tag{9.114}$$

The objective functional J is intended to represent the weighted average of the number of infectious persons and costs of vaccination and treatment. The optimal control problem is as follows:

$$
\begin{aligned}
\min_{(u,v)\in\mathcal{A}} \quad & J(u,v) = \int_0^T (aI(t) + bu(t) + cv(t))\mathrm{d}t \\
\text{s.t.} \quad & \dot{S} = \gamma N - \nu S - \beta \tfrac{IS}{N} + \rho R - \kappa S u, \\
& \dot{I} = \beta \tfrac{IS}{N} - (\nu + \mu) I - \alpha I - \eta I v, \\
& \dot{R} = -\nu R - \rho R + \kappa S u + \alpha I + \eta I v, \\
& S(0) = S_0, \quad I(0) = I_0, \quad R(0) = R_0,
\end{aligned}
\tag{9.115}
$$

where $\mathcal{A}$ is the set of admissible controls, which is assumed as being the set of all Lebesgue measurable functions $u : [0,T] \to [0, u_{\max}]$ and $v : [0,T] \to [0, v_{\max}]$, where $u_{\max}$ is the maximum vaccination rate and $v_{\max}$ is the maximum treatment rate. The assumptions on $\mathcal{A}$ ensure existence of an optimal solution to problem (9.115), which follows from Fleming and Rishel's .[11]

For problem (9.115), Ledzewicz *et al.*[17] used Pontryagin's Maximum Principle[10] to find the first-order necessary conditions for optimality. By rewriting the state equations into a multi-input control-affine system of vector form and using Lie derivatives they were able to compute the switching functions for controls u and v as well as multiple derivatives of the switching function, allowing the Legendre–Clebsch Condition to be checked. Their methods of using Lie brackets are particularly useful in determining existence of a singular control when observing problems with many states and control variables, and we direct the reader to their article[17] and to Schättler and Ledzewicz's book [46] for more details on their procedures.

In Ref. 17, Ledzewicz *et al.* use the Tomlab algorithm PROPT,[47] a collocation solver, to obtain an approximate solution to problem (9.115) with the numerical values of parameters set to where a singular subarc is present in u^*, while v^* contains no singular subarc. In this section, we discuss how to discretize the penalized Ebola problem given in problem (9.116), where the penalty term involved is based on the total variation of the vaccination control u.[30] Additionally, in Appendix Section 9.A.2, we provide the MATLAB file, demoOC.m, to demonstrate how to use PASA to solve problem (9.116). We use PASA to solve both the unpenalized and penalized problems with parameters set to being what was used in Ledzewicz *et al.*[17] and remark about PASA's performance on this problem.

9.5.1 *Discretization of the SIR problem*

We discretize the following regularized problem:

$$\begin{aligned}
\min_{(u,v)\in\mathcal{A}} \quad & J_p(u,v) = \textstyle\int_0^T (aI(t) + bu(t) + cv(t))\mathrm{d}t + pV(u) \\
\text{s.t.} \quad & \dot{S} = \gamma N - \nu S - \beta \tfrac{IS}{N} + \rho R - \kappa S u, \\
& \dot{I} = \beta \tfrac{IS}{N} - (\nu+\mu)I - \alpha I - \eta I v, \\
& \dot{R} = -\nu R - \rho R + \kappa S u + \alpha I + \eta I v, \\
& S(0) = S_0, \quad I(0) = I_0, \quad R(0) = R_0,
\end{aligned} \tag{9.116}$$

where $0 \le p < 1$ is the bounded variation tuning parameter and function V measures the total variation of the control which is defined in Eq. (9.56). Since we are numerically solving the problem with parameters set to where control v^* has no singular subarcs, it is likely that v need not be penalized. Additionally, we discretize the following adjoint equations:

$$\begin{aligned}
\dot{\lambda}_S = {} & \lambda_S \left(-\gamma + \nu + \beta \frac{I}{N} - \beta \frac{IS}{N^2} + \kappa u \right) \\
& + \lambda_I \left(\beta \frac{IS}{N^2} - \beta \frac{I}{N} \right) - \lambda_R \kappa u,
\end{aligned} \tag{9.117}$$

$$\begin{aligned}
\dot{\lambda}_I = {} & -a + \lambda_S \left(-\gamma + \beta \frac{S}{N} - \beta \frac{IS}{N^2} \right) \\
& + \lambda_I \left(-\beta \frac{S}{N} + \beta \frac{IS}{N^2} + (\nu + \mu + \alpha) + \eta v \right) - \lambda_R (\alpha + \eta v),
\end{aligned} \tag{9.118}$$

$$\dot{\lambda}_R = -\lambda_S \left(\gamma + \beta \frac{IS}{N^2} + \rho \right) + \lambda_I \left(\beta \frac{IS}{N^2} \right) + \lambda_R (\nu + \rho), \tag{9.119}$$

with transversality conditions being

$$\lambda_S(T) = 0, \quad \lambda_I(T) = 0, \quad \lambda_R(T) = 0. \tag{9.120}$$

First we assume that controls u and v are constant over each mesh interval. We partition time interval $[0, T]$, by using $n+1$ equally spaced nodes, $0 = t_0 < t_1 < \cdots < t_n = T$. For all $k = 0, 1, \ldots, n$, we assume that $S_k = S(t_k)$, $I_k = I(t_k)$, and $R_k = R(t_k)$. For the controls we denote $u_k = u(t)$ and $v_k = v(t)$ for all $t_k \le t < t_{k+1}$ when $k = 0, \ldots, n-2$. Additionally, we denote $u_{n-1} = u(t)$ and $v_{n-1} = v(t)$ for all $t_{n-1} \le t \le t_n$. We then have $\boldsymbol{S} \in \mathbb{R}^{n+1}, \boldsymbol{I} \in \mathbb{R}^{n+1}$, and $\boldsymbol{R} \in \mathbb{R}^{n+1}$, while $\boldsymbol{u} \in \mathbb{R}^n$ and $\boldsymbol{v} \in \mathbb{R}^n$. Furthermore, $\boldsymbol{N} \in \mathbb{R}^{n+1}$ where $N_k = N(t_k) = S_k + I_k + R_k$ for all $k = 0, 1, \ldots, n$. We use a left-rectangular integral approximation for approximating the objective functional $J_p(u, v)$ in problem (9.116), and we use forward Euler's method for discretizing the state equations. Since PASA uses a gradient scheme for one of its phases, we need the cost function to be differentiable. We suggest a decomposition of each absolute value term arising from the total variation term in J_ρ to ensure that J_ρ is differentiable. We introduce two $n-1$ vectors, $\boldsymbol{\zeta}$ and $\boldsymbol{\iota}$, whose entries are nonnegative. Each entry of $\boldsymbol{\zeta}$ and $\boldsymbol{\iota}$ is defined as

$$|u_{k+1} - u_k| = \zeta_k + \iota_k \quad \text{for all } k = 0, \ldots, n-2.$$

The discretization of problem (9.116) then becomes

$$\begin{aligned}
\min \quad J_p(\boldsymbol{u}, \boldsymbol{\zeta}, \boldsymbol{\iota}, \boldsymbol{v}) &= \sum_{k=0}^{n-1} h(aI_k + bu_k + cv_k) + p \sum_{k=0}^{n-1} (\zeta_k + \iota_k) \\
S_{k+1} &= S_k + h\left(\gamma N_k - \nu S_k - \beta \tfrac{I_k S_k}{N_k} + \rho R_k - \kappa S_k u_k\right) \\
&\quad \text{for } k = 0, \ldots, n-1, \\
I_{k+1} &= I_k + h\left(\beta \tfrac{I_k S_k}{N_k} - (\nu + \mu + \alpha) I_k - \eta I_k v_k\right) \\
&\quad \text{for } k = 0, \ldots, n-1, \\
R_{k+1} &= R_k + h(-\nu R_k - \rho R_k + \kappa S_k u_k + \alpha I_k + \eta I_k v_k) \\
0 \le u_k &\le u_{\max} \quad \text{for } k = 0, \ldots, n-1, \\
0 \le v_k &\le v_{\max} \quad \text{for } k = 0, \ldots, n-1, \\
u_{k+1} - u_k &= \zeta_k - \iota_k; \quad \text{for } k = 0, \ldots, n-2, \\
\zeta_k \ge 0, \iota_k &\ge 0 \quad \text{for } k = 0, \ldots, n-1,
\end{aligned} \tag{9.121}$$

where $h = \frac{T}{n}$ is the mesh size and the first components of S, I, and R are set to be the initial conditions associated with the state equations. Note that, for the above problem, we are minimizing the penalized objective function with respect to vectors $\boldsymbol{u}, \boldsymbol{\zeta}, \boldsymbol{\iota}$, and $\boldsymbol{v}$. The equality constraints associated with $\boldsymbol{\zeta}$ and ι are linear constraints that PASA can interpret. The equality constraints associated with $\boldsymbol{\zeta}$ and $\boldsymbol{\iota}$ are written as

$$\left[\,\boldsymbol{A}\middle|-\boldsymbol{I}_{n-1}\middle|\boldsymbol{I}_{n-1}\middle|\boldsymbol{O}_{n-1,n}\,\right]\begin{bmatrix}\boldsymbol{u}\\ \hline \boldsymbol{\zeta}\\ \hline \boldsymbol{\iota}\\ \hline \boldsymbol{v}\end{bmatrix}=\boldsymbol{0}, \tag{9.122}$$

where $\boldsymbol{I}_{n-1}$ is the identity matrix with dimension $n-1$, $\boldsymbol{O}_{n-1,n}$ is an $n-1\times n$ all zeros matrix, $\boldsymbol{0}$ is the $n-1$ all zeros vector, and $\boldsymbol{A}$ is an $n-1\times n$ dimensional sparse matrix give in Eq. (9.7).

For finding the gradient of J_p in problem (9.121), we use Theorem 9.1. The Lagrangian of problem (9.121) is as follows

$$\begin{aligned}\mathcal{L}=&\sum_{k=0}^{n-1}h(aI_k+bu_k+cv_k)+p\sum_{k=0}^{n-1}(\zeta_k-\iota_k)\\&+\sum_{k=0}^{n-1}\lambda_{S,k}\left(-S_{k+1}+S_k+h\left(\gamma N_k-\nu S_k\right.\right.\\&\qquad\left.\left.-\beta\frac{I_kS_k}{N_k}+\rho R_k-\kappa S_ku_k\right)\right)\\&+\sum_{k=0}^{n-1}\lambda_{I,k}\left(-I_{k+1}+I_k+h\left(\beta\frac{I_kS_k}{N_k}-(\nu+\mu+\alpha)I_k-\eta I_kv_k\right)\right)\\&+\sum_{k=0}^{n-1}\lambda_{R,k}\left(-R_{k+1}+R_k+h\left(-\nu R_k-\rho R_k\right.\right.\\&\qquad\left.\left.+\kappa S_ku_k+\alpha I_k+\eta I_kv_k\right)\right),\end{aligned} \tag{9.123}$$

where $\boldsymbol{\lambda}_S\in\mathbb{R}^{n-1}$, $\boldsymbol{\lambda}_I\in\mathbb{R}^{n-1}$, $\boldsymbol{\lambda}_R\in\mathbb{R}^{n-1}$ are the Lagrangian multiplier vectors. As suggested in Theorem 9.1, we use the gradient of the Lagrangian with respect to vectors $\boldsymbol{u}$, $\boldsymbol{\zeta}$, $\boldsymbol{\iota}$, and $\boldsymbol{v}$ to find $\boldsymbol{\nabla}_{\boldsymbol{u},\boldsymbol{\zeta},\boldsymbol{\iota},\boldsymbol{v}}\boldsymbol{J_p}\in\mathbb{R}^{2(n-1)+2(n-2)}$. This is necessary because based upon the state equations (9.111)–(9.113) of problem (9.116), S, I, and R can be viewed as functions of u and v. So when computing $\boldsymbol{\nabla}_{\boldsymbol{u},\boldsymbol{\zeta},\boldsymbol{\iota},\boldsymbol{v}}\boldsymbol{J_p}\in\mathbb{R}^{2(n-1)+2(n-2)}$, we should consider that $\boldsymbol{S}$, $\boldsymbol{I}$, and $\boldsymbol{R}$ depend on $\boldsymbol{u}$ and $\boldsymbol{v}$. We find the partial derivative of $\mathcal{L}$ with respect to u_k, ζ_k, ι_k, and v_k, and obtain the following:

$$\frac{\partial \mathcal{L}}{\partial u_k} = hb + h\kappa S_k(\lambda_{R,k} - \lambda_{S,k}) \quad \text{for all } k = 0, \ldots, n-1$$
$$\frac{\partial \mathcal{L}}{\partial \zeta_k} = p \quad \text{for all } k = 0, \ldots, n-2$$
$$\frac{\partial \mathcal{L}}{\partial \iota_k} = p \quad \text{for all } k = 0, \ldots, n-2$$
$$\frac{\partial \mathcal{L}}{\partial v_k} = hc + h\eta I_k(\lambda_{R,k} - \lambda_{I,k}) \quad \text{for all } k = 0, \ldots, n-1.$$

By Theorem 9.1, we compute $\boldsymbol{\nabla_{u,\zeta,\iota,v} J_p} \in \mathbb{R}^{2(n-1)+2(n-2)}$ as follows

$$\boldsymbol{\nabla_{u,\zeta,\iota,v} J_p} = \boldsymbol{\nabla_{u,\zeta,\iota,v} \mathcal{L}} = \left[\begin{array}{c} hb + h\kappa S_0(\lambda_{R,0} - \lambda_{S,0}) \\ \vdots \\ hb + h\kappa S_k(\lambda_{R,k} - \lambda_{S,k}) \\ \vdots \\ hb + h\kappa S_{n-1}(\lambda_{R,n-1} - \lambda_{S,n-1}) \\ \hline p \\ \vdots \\ p \\ \hline p \\ \vdots \\ p \\ \hline hc + h\eta I_k(\lambda_{R,0} - \lambda_{I,0}) \\ \vdots \\ hc + h\eta I_k(\lambda_{R,k} - \lambda_{I,k}) \\ \vdots \\ hc + h\eta I_{n-1}(\lambda_{R,n-1} - \lambda_{I,n-1}) \end{array}\right], \tag{9.124}$$

provided that the Lagrange multiplier vectors satisfy condition (9.11) in Theorem 9.1. To satisfy condition (9.11), we take the partial derivative of the Lagrangian $\mathcal{L}$ with respect to S_k, I_k, and R_k for all $k = 1, \dots, n$. Note that we are not finding the partial derivative of $\mathcal{L}$ with respect to S_0, I_0, and R_0 since these entries are the known values. Taking the partial derivative of the Lagrangian, given in Eq. (9.123), with respect to S_k, I_k, and R_k yields the following:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial S_k} &= -\lambda_{S,k-1} + \lambda_{S,k} + h\lambda_{S,k}\left(\gamma - \nu - \beta\frac{I_k}{N_k} + \beta\frac{I_kS_k}{N_k^2} - \kappa u_k\right) \\
&\quad + h\lambda_{I,k}\left(\beta\frac{I_k}{N_k} - \beta\frac{I_kS_k}{N_k^2}\right) + h\lambda_{R,k}(\kappa u_k), \qquad (9.125) \\
&\quad \text{for } k = 1, \dots, n-1 \quad \text{and} \\
\frac{\partial \mathcal{L}}{\partial S_n} &= -\lambda_{S,n-1}, \qquad (9.126) \\
\frac{\partial \mathcal{L}}{\partial I_k} &= ha - \lambda_{I,k-1} + \lambda_{I,k} \\
&\quad + h\lambda_{I,k}\left(\beta\frac{S_k}{N_k} - \beta\frac{I_kS_k}{N_k^2} - (\nu + \mu + \alpha) - \eta v_k\right) \\
&\quad + h\lambda_{S,k}\left(\gamma - \beta\frac{S_k}{N_k} + \beta\frac{I_kS_k}{N_k^2}\right) + h\lambda_{R,k}(\alpha + \eta v_k), \qquad (9.127) \\
&\quad \text{for } k = 1, \dots, n-1 \quad \text{and} \\
\frac{\partial \mathcal{L}}{\partial I_n} &= -\lambda_{I,n-1}, \qquad (9.128) \\
\frac{\partial \mathcal{L}}{\partial R_k} &= -\lambda_{R,k-1} + \lambda_{R,k} - h\lambda_{R,k}(\nu + \rho) \\
&\quad + h\lambda_{S,k}\left(\gamma + \rho + \beta\frac{I_kS_k}{N_k^2}\right) - h\lambda_{I,k}\left(\beta\frac{I_kS_k}{N_k^2}\right), \qquad (9.129) \\
&\quad \text{for } k = 1, \dots, n-1 \text{ and} \\
\frac{\partial \mathcal{L}}{\partial R_n} &= -\lambda_{R,n-1}. \qquad (9.130)
\end{aligned}$$

For finding multiplier vectors $\boldsymbol{\lambda}_S$, $\boldsymbol{\lambda}_I$, $\boldsymbol{\lambda}_R$ that satisfy Theorem condition (9.11), we set the above equations (9.125)–(9.130) equal to zero and do the following: solve for $\lambda_{S,k-1}$ in Eq. (9.125), solve for

$\lambda_{S,n-1}$ in Eq. (9.126), solve for $\lambda_{I,k-1}$ in Eq. (9.127), solve for $\lambda_{I,n-1}$ in Eq. (9.128), solve for $\lambda_{R,k-1}$ in Eqs. (9.129), and solve for $\lambda_{R,n-1}$ in Eq. (9.130). After performing the following steps and rearranging terms, we generate a discretization of Eqs. (9.117)–(9.119) that satisfy the transversality conditions (9.120):

$$\begin{aligned}\lambda_{S,k-1} &= \lambda_{S,k} + h\lambda_{S,k}\left(\gamma - \nu - \beta\frac{I_k}{N_k} + \beta\frac{I_k S_k}{N_k^2} - \kappa u_k\right)\\ &\quad + h\lambda_{I,k}\left(\beta\frac{I_k}{N_k} - \beta\frac{I_k S_k}{N_k^2}\right) + h\lambda_{R,k}(\kappa u_k), \qquad (9.131)\\ &\quad \text{for } k = 1,\ldots,n-1, \text{ and}\\ \lambda_{S,n-1} &= 0, \qquad (9.132)\\ \lambda_{I,k-1} &= \lambda_{I,k} + ha + h\lambda_{S,k}\left(\gamma - \beta\frac{S_k}{N_k} + \beta\frac{I_k S_k}{N_k^2}\right)\\ &\quad + h\lambda_{I,k}\left(\beta\frac{S_k}{N_k} - \beta\frac{I_k S_k}{N_k^2}\right) + h\lambda_{R,k}(\alpha + \eta v_k), \qquad (9.133)\\ &\quad \text{for } k = 1,\ldots,n-1, \text{ and}\\ \lambda_{I,n-1} &= 0, \qquad (9.134)\\ \lambda_{R,k-1} &= \lambda_{R,k} + h\lambda_{S,k}\left(\gamma + \rho + \beta\frac{I_k S_k}{N_k^2}\right) - h\lambda_{I,k}\left(\beta\frac{I_k S_k}{N_k^2}\right)\\ &\quad - h\lambda_{R,k}(\nu + \rho), \quad \text{for } k = 1,\ldots,n-1, \text{ and} \qquad (9.135)\\ \lambda_{R,n-1} &= 0. \qquad (9.136)\end{aligned}$$

9.5.2 *Numerical results of the SIR problem*

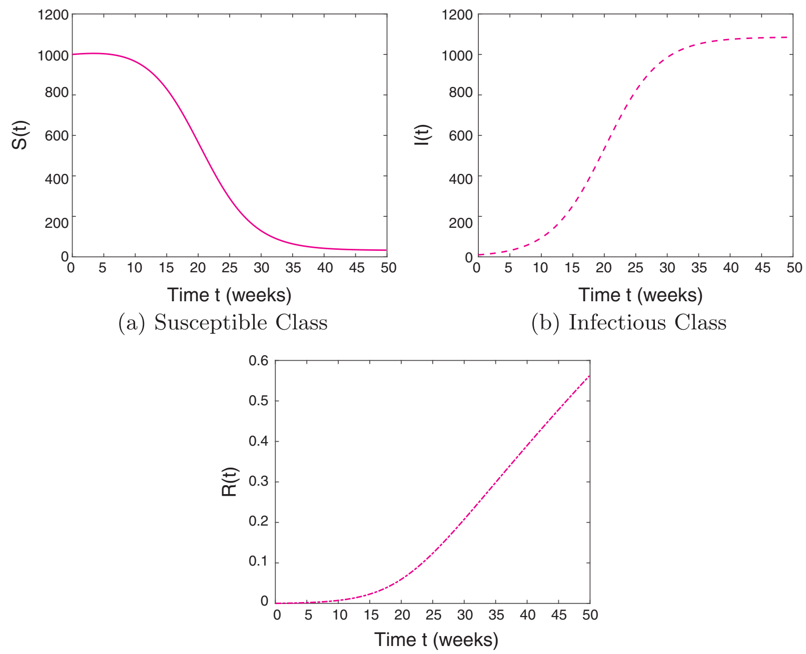
We first use PASA to numerically solve the unregularized problem given in (9.115) with stopping tolerance set to being 10^{-8}. Our initial guess for the vaccination and treatment control is $u(t) = 0$ and $v(t) = 0$ over the entire time interval, respectively. The numerical parameters, which are given in Table 9.8, are set to the values that are used in Ledzewicz *et al.*[17] because we want to see if our numerical results are comparable to theirs. Now, Ledzewicz *et al.*[17] are investigating optimal vaccine and treatment strategies for a theoretical

Table 9.8. Numerical values of parameters used in computations.

Parameter	Numerical value	Parameter description
γ	0.00683	Birth rate of the population
ν	0.00188	Natural death rate of the population
β	0.2426	Rate of infectiousness of the disease
μ	0.005	Disease induced death rate
α	0.00002	Rate at which disease is overcome
ρ	0.007	Resensitization rate
κ	0.3	Effectiveness of vaccination
η	0.1	Effectiveness of treatment
a	5	Weight of I used in cost functional
b	50	Weight for cost of vaccination
c	300	Weight for cost of treatment
T	50	Time horizon (weeks)
$u_{\max}$	1	Maximum vaccination rate
$v_{\max}$	1	Maximum treatment rate
S_0	1000	Initial condition for S
I_0	10	Initial condition for I
R_0	0	Initial condition for R

vaccine and treatment for Ebola, and one can infer from their parameter settings that they are not concerned with realistic limitations in widely developing and administering these theoretical controls. For example, based on these parameter values, a vaccination strategy being $u(t) = u_{\max} = 1$ implies that we are vaccinating approximately 97% of the susceptible population immediately at time t. When considering realistic limitations in developing and administering vaccines, one might want to consider setting $u_{\max}$ to a value that is less than one.

We partition the time interval to where there are $n = 750$ mesh intervals with mesh size being $h = 0.0667$, and use the discretization methods presented in Section 9.5.1 with penalty parameter set to being $p = 0$. The results corresponding to the approximated solution that PASA obtained for problem (9.115) are given in Figure 9.11. Observe in Figure 9.11(a) that the numerical solution associated with the vaccination control, $\hat{u}$, begins with a full dose segment (i.e., $\hat{u} = 1$), followed by an interval of many oscillations, and ends with the control being turned off. The many oscillations approximately start at time $t_{u,1} = 12.933$ and end at time $t_{u,2} = 36.533$. We looked



(a) Susceptible Class

(b) Infectious Class

(c) Recovered Class

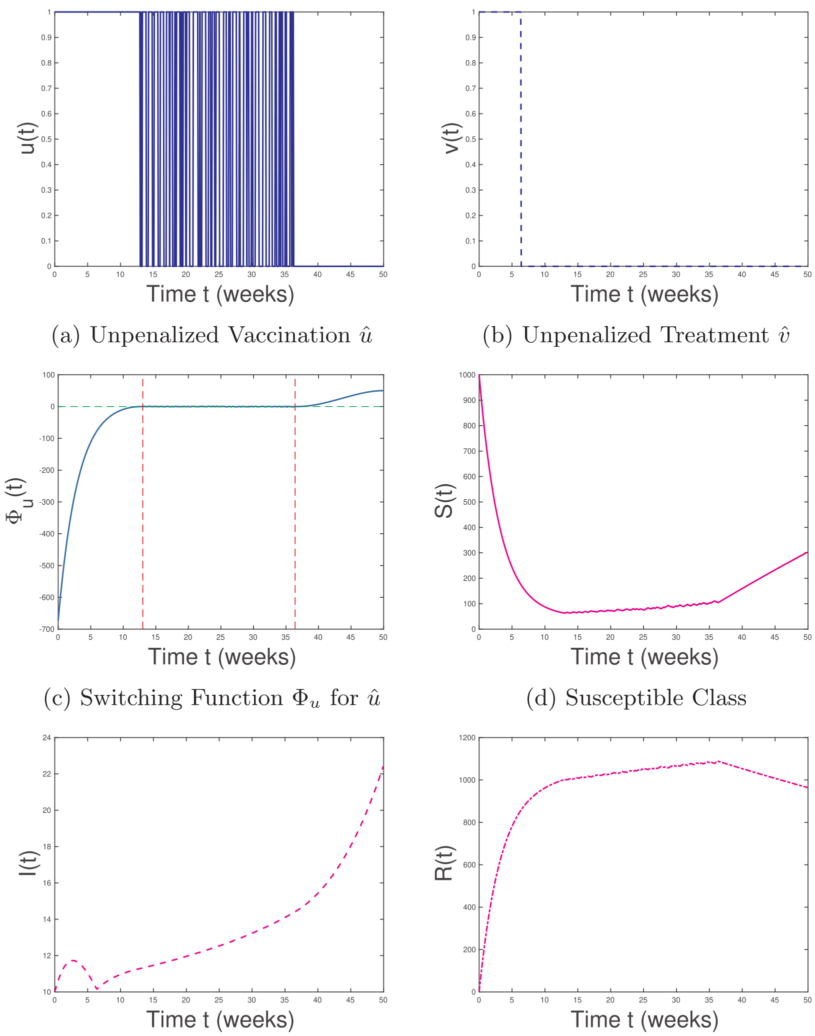
Fig. 9.10. Spread of disease without vaccination and treatment.

at the corresponding switching function to $\hat{u}$, which is $\Phi_u(t) = b + (\lambda_R(t) - \lambda_S(t))\kappa S(t)$ (see Ref. 17 for construction of switching function) to verify that the oscillating region corresponds to a singular subarc. It is necessary that the optimal control u^* is as follows:

$$u^*(t) = \begin{cases} 0 & \text{if } \Phi_u(t) > 0, \\ u_{\max} & \text{if } \Phi_u(t) < 0, \end{cases} \tag{9.137}$$

and $u^*(t)$ is singular if Φ_u is zero over an open interval of time. A plot of the switching function Φ_u is given in Figure 9.11(c).

Although we observe in Figures 9.11(c)–9.11(e) that the dynamics of S, I, and R corresponding to the approximated controls $\hat{u}$ and $\hat{v}$ give significantly favorable results in comparison to the dynamics associated with no application of vaccination or treatment, which is given in Figure 9.10, we recognize that the wild oscillations found in $\hat{u}$ makes it an unrealistic strategy to implement. In order to obtain a



(a) Unpenalized Vaccination $\hat{u}$ (b) Unpenalized Treatment $\hat{v}$

(c) Switching Function Φ_u for $\hat{u}$ (d) Susceptible Class

(e) Infectious Class (f) Recovered Class

Fig. 9.11. Spread of disease with unpenalized vaccination and treatment.

more realistic vaccination strategy, we penalize control u via a total variation penalty term.[30] We are not penalizing control v, since control v is a piecewise constant function that can be easily interpreted.

We use PASA to solve for penalized problem (9.116) with stopping tolerance set to 10^{-8} and with parameter value settings given in Table 9.8. Our initial guess for controls are $u(t) = 0$ and $v(t) = 0$ over the entire time interval. We partitioned $[0, T]$ to $n = 750$ mesh intervals of size $h = 0.0667$, and the discretization procedure is given in Section 9.5.1. Additionally in Section 9.A.2, we have a MATLAB file given that can be used for solving the problem. Note that this MATLAB file is associated with the most up to date version of PASA (SuiteOPT Version 2.0.0), but our experiments presented here are using an older version of PASA (SuiteOPT Version 1.0.0). We numerically solve for problem (9.116) for varying values of penalty parameter $p \in \{10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$. If the approximated control, u_p, corresponding to penalty parameter value p is no longer oscillating and if the sign of the approximated switching, Φ_{u_p}, aligns with Eq. (9.137), then we consider p as being a suitable penalty parameter value to use.

In Table 9.9, we record an approximation of the unpenalized cost functional, given in Eq. (9.114), being evaluated at the penalized solution u_p. The values corresponding to $J(u_p)$ are remarkably similar to the approximated objective value that corresponds to the unpenalized solution $\hat{u}$. We also record PASA's runtime performance (in

Table 9.9. Varying penalty results.

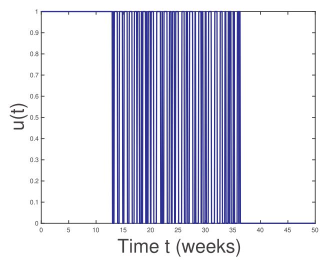
	p	$J(u_p)$	$t_{u,1}$	$t_{u,2}$	Runtime
$N = 750$	0	6572.955	12.933	36.533	16.77
$h = 0.0667$	10^{-5}	6572.956	13.067	36.400	17.07
$\text{tol} = 10^{-8}$	10^{-4}	6572.948	13.133	36.400	20.58
	10^{-3}	6573.018	12.933	36.533	19.12
	10^{-2}	6573.101	12.467	36.000	56.58
	10^{-1}	6575.429	12.333	37.067	10.47

Note: $J(u_p)$ is the approximated unpenalized cost functional value, $t_{u,1}$ and $t_{u,2}$ are the approximated switches corresponding to each solution. The Runtime column corresponds to the time (in seconds) it took to run PASA for each problem.

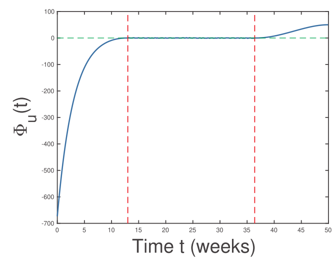
seconds) on solving problem (9.116), and we record the approximated switches of u_p, $t_{u,1}$, and $t_{u,2}$. When numerically solving problem (9.116), we observe oscillations in all cases except for when $p = 10^{-1}$. In Figure 9.12, we provide the plots of the penalized solution u_p and the plots of the corresponding switching function when the penalty parameter was set to being 10^{-3}, 10^{-2}, and 10^{-1}. Figures 9.12(a), 9.12(b), and 9.12(c) illustrate a trend of how increasing the tuning parameter p influences the penalized solution to where many oscillations no longer occur. We could increase the penalty parameter values larger than 10^{-1}; however, we did not do so because the approximated switching function associated u_p when $p = 10^{-1}$, given in Figure 9.12, gave indication that the structure of u_p aligns with Eq. (9.137). Meaning we have a non-oscillatory vaccination strategy that is not only more realistic to implement, but also satisfies the first-order necessary conditions of optimality.[10] In addition, we do not increase the tuning parameter to values larger than 10^{-1} due to the possibility of over-penalization. In terms of runtime performances, we have that PASA performed the fastest when the penalty parameter was set to being 10^{-1}.

We recognize in Figure 9.12 that the singular region associated with u_p when the bounded variation tuning parameter is set to $p = 10^{-1}$ is a rather crude approximation. However, if we want a smoother presentation of the singular region, we recommend either increasing the number of mesh intervals used for the discretization and/or tightening the stopping tolerance. In Figure 9.13, we provide a solution that PASA obtained in solving for penalized problem (9.116) with $p = 10^{-1}$ when time interval $[0, T]$ was partitioned to have 2000 mesh intervals with mesh size being $h = 0.025$. The approximated singular subarc that is given in Figure 9.13 is comparable to the approximation that Ledezewicz *et al.*[17] obtained.

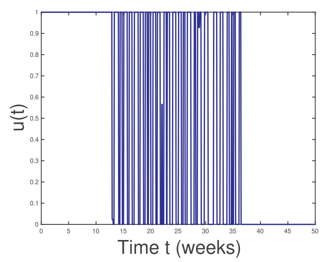
Although these results are theoretical due to no vaccination for Ebola, optimal control theory is used to see how vaccination and treatment can influence the spread of the disease. We observe in Figure 9.10 that without treatment and vaccination, the susceptible class starts to decline due to many individuals becoming infectious, and few individuals recover from the disease. When incorporating the vaccination strategy and treatment strategy that is respectively,



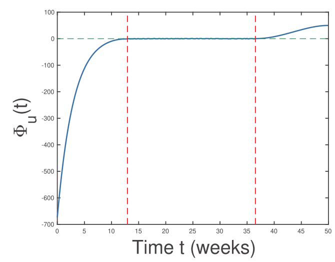
(a) u_p with $p = 0$



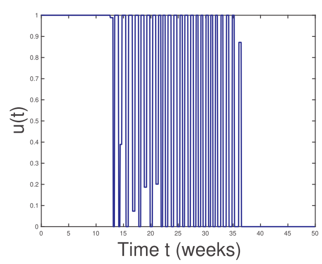
(b) Φ_{u_p} with $p = 0$



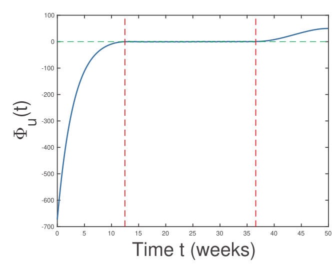
(c) u_p with $p = 10^{-3}$



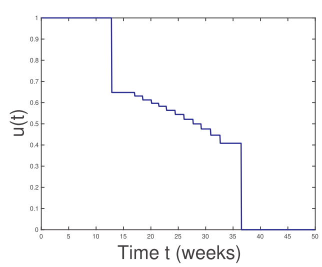
(d) Φ_{u_p} with $p = 10^{-3}$



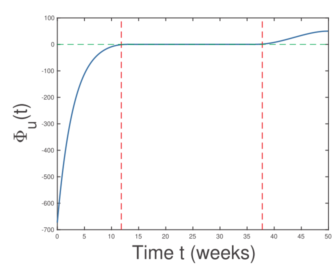
(e) u_p with $p = 10^{-2}$



(f) Φ_{u_p} with $p = 10^{-2}$

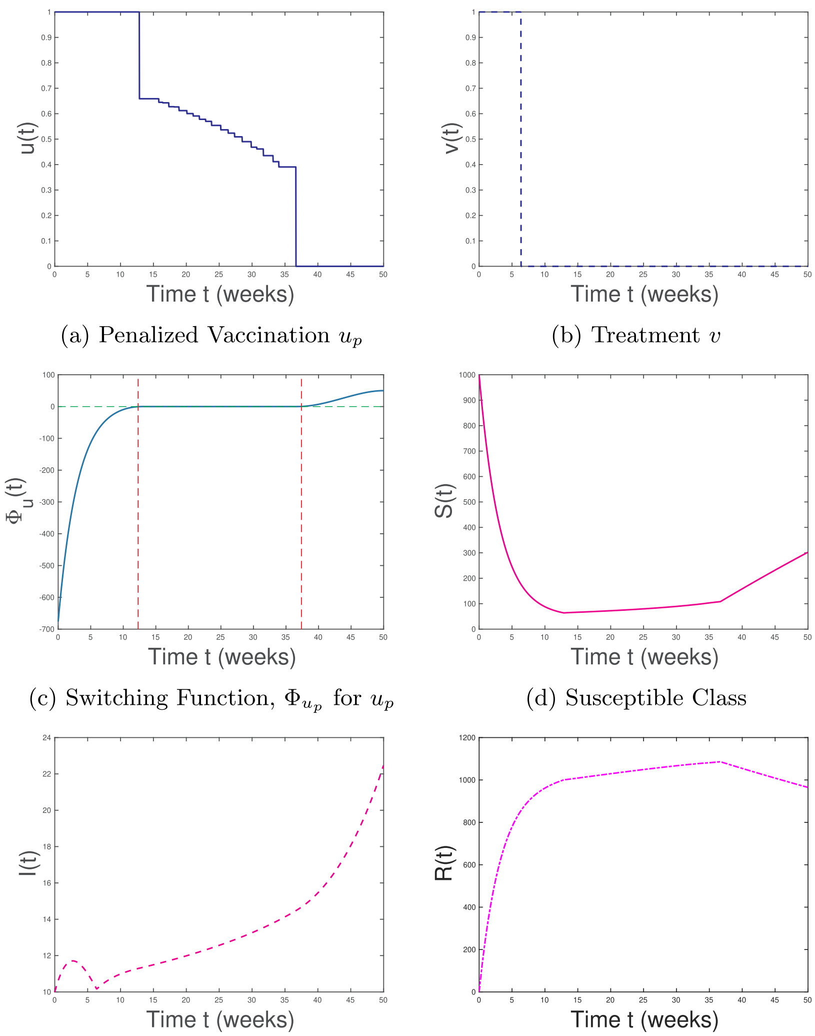


(g) u_p with $p = 10^{-1}$



(h) Φ_{u_p} with $p = 10^{-1}$

Fig. 9.12. Plots of the regularized vaccination strategy u_p and its corresponding switching function ϕ_u for varying tuning parameter values for SIR Problem.



(a) Penalized Vaccination u_p

(b) Treatment v

(c) Switching Function, Φ_{u_p} for u_p

(d) Susceptible Class

(e) Infectious Class

(f) Recovered Class

Fig. 9.13. Spread of the disease with regularized vaccination and unregularized treatment. Tuning parameter: $p = 10^{-1}$, mesh intervals: $n = 2,000$, and Error Tolerance: 10^{-8}.

given in Figures 9.13(a) and 9.13(b), we observe in Figures 9.13(a)–9.13(b) a large portion of the susceptible class transfer to the recovered class through vaccination. The positive influence that treatment has on the infectious class occurs approximately in the first 6 weeks when the treatment strategy is set to treat all infectious individuals. We believe that the weighted cost of treatment parameter value c is influences the optimal treatment strategy to switch to incorporating no treatment at $t_{v,1} \approx 6.4$ weeks. Reducing the weighted cost of treatment parameter would alter the optimal treatment strategy to extending the region(s) when $v^* = v_{\max}$.

9.6 Concluding Remarks

We demonstrate how Hager and Zhang's[31] Polyhedral Active Set Algorithm (PASA) can be implemented for numerically solving optimal control problems that are linear in the control. If problems of this form possess a singular subarc, we recommend regularizing such problems by using a penalty term based on the total variation of the control as suggested in Caponigro *et al.*.[30] Total variation regularization allows for PASA to yield approximations that do not oscillate wildly over the singular region. We provide a discretization method for a general optimal control problem that uses Euler's method for the state variables and uses the gradient of the Lagrangian of the discretized optimal control problem to approximate the gradient of the cost functional that is being used. Additionally, the use of the Lagrangian consequently aids in the discretization of the adjoint equations.

We then present in detail three optimal control problems that apply to biological models. For the first two problems, an explicit solution satisfying the first-order necessary conditions of optimality[10] is found. In the third problem, we can verify existence of a singular subarc for one of the control variables used. For all three examples, we use PASA to numerically solve each problem with parameters set to where the existence of a singular subarc is possible. For the example associated with the plant problem that was presented in King and Roughgarden's,[8] three cases are investigated where each case determines the beginning behavior of the optimal allocation strategy. PASA solves the unpenalized plant problem for all three cases, and in all three cases no oscillations are found. However, in the degenerate case, the case corresponding to the optimal control beginning singular, the unpenalized solution has an unusual dip at $t = 0$. For this case, we regularize the problem via total variation and obtain a penalized solution that serves to be a better approximation of the true solution. When using PASA to solve the unpenalized fishery problem[12,37] and to solve the unpenalized SIR problem,[17] we obtain numerical solutions that possessed oscillations along the singular region. We then apply a total variation regularization for varying values of the penalty parameter. We find that the oscillatory solution can be controlled by increasing the penalty parameter size.

In these examples, we find that observing the plots of the approximated penalized solution and of the corresponding switching function are effective heuristics for determining whether or not the penalty parameter size is over or under-penalizing the solution. For the fishery problem, we use additional information such as the approximated switching points and the L^1 norm difference between the true solution and the penalized solution to help determine what penalty parameter value is the most appropriate choice in comparison to all of the other penalty parameter values being tested. For the SIR problem, we find that the penalized solution that PASA obtains is comparable to the approximated solution that is presented in Ledzewicz *et al.*,[17] which was obtained via a collocation method called PROPT.[47]

Overall, the use of PASA and the use of total variation regularization make the process of numerically solving singular control problems more tractable. With the aid of PASA and total variation penalization, mathematical biologists can consider constructing an optimal control problem with a cost functional that possesses a more accurate representation of the cost in implementing the control, which aligns more with the principle of parsimony than when using an optimal control problem with quadratic dependence. Solving an optimal control problem with a more accurate representation of the cost functional ensures that the optimal control associated with the problem truly meets the criteria that are being considered for determining which control strategy is the best relative to the state dynamics being studied. Additionally, using an optimal control problem where the control variables appear linearly can potentially yield optimal control solutions with regions that are both easy to interpret and implement. Incorporating a more practical, uncomplicated, and accurate optimal control into a dynamical system that is intended to model some biological phenomenon can provide more compelling and elucidating results that are applicable to practice and worthwhile to study.

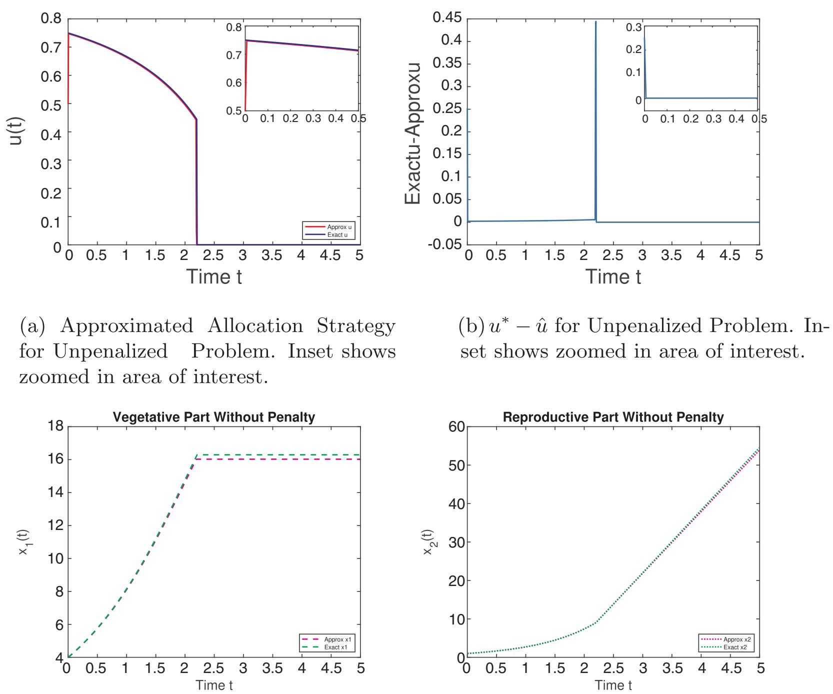
Acknowledgments

Support by the National Science Foundation (NSF) under Grants 1522629 and 1819002, and by the Office of Naval Research under Grants N00014-15-1-2048 and N00014-18-1-2100 is gratefully acknowledged. S. Atkins was supported in part by the University of Florida's Graduate Student Fellowship. S. Atkins also received student travel support for the CMPD5 conference from NSF CMPD5 Grant 1917506, The Center for Applied Mathematics (CAM), and University of Florida's CLAS. M. Martcheva was supported in part by NSF CMPD5 Grant 1917506 and from the University of Florida's Department of Mathematics. We also are grateful to Suzanne Lenhart for her helpful comments and suggestions of problems to explore.

Appendix 9

9.A.1 *Numerical Results Associated with the Unpenalized Plant Problem*

The unpenalized solutions that PASA obtained when parameters were set to be of Cases 2a, 2b and 2c are shown in Figures 9.A.14, 9.A.15, and 9.A.16, respectively; and numerical results associated with each case are provided in Table 9.6. As you can see in Figures 9.A.14(a), 9.A.15(a), and 9.A.16(a), the unpenalized control found for each case does not present any oscillatory artifacts. We see in Figures 9.A.15(b) and 9.A.16(b) that other than the discrepancies arising from the switching points, the unpenalized solutions for Case 2b and Case 2c only differ slightly from their respective exact solution along the singular region. The state solutions that correspond to the unpenalized solution for Cases 2a, 2b, and 2c, which are respectively shown in Figures 9.A.16(a)–9.A.16(d), Figures 9.A.15(c)–9.A.15(d) and Figures 9.A.16(c)–9.A.16(d), compare relatively well with the state solutions that correspond to u^*. In Figures 9.A.14(c), 9.A.15(c), and 9.A.16(c), we see that in each case the approximate solutions



(a) Approximated Allocation Strategy for Unpenalized Problem. Inset shows zoomed in area of interest.

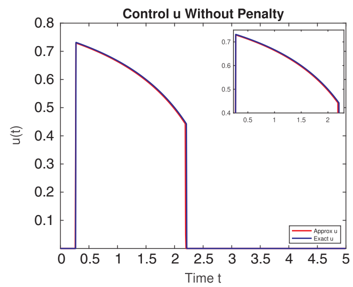
(b) $u^* - \hat{u}$ for Unpenalized Problem. Inset shows zoomed in area of interest.

(c) Vegetative Part of Plant from Unpenalized $\hat{u}$

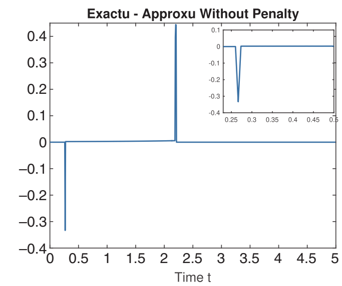
(d) Reproductive Part of Plant from Unpenalized $\hat{u}$

Fig. 9.A.14. Results from solving unpenalized plant problem (Case 2a).

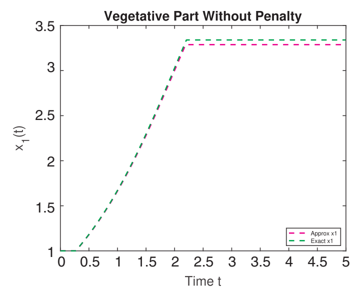
associated with the vegetative parts of the plant do give slightly underestimated values along the interval $[1.5, T]$. This could be due to the small discrepancies in $\hat{u}$ along the singular region. Moreover, if we look at the analytic solution for x_1 on $[t_2, T]$ given in (9.90), we can see how errors between u^* and $\hat{u}$ and how switching prematurely to the non-singular case could contribute errors in x_1 along the non-singular region $[t_2, T]$. Now in Figures 9.A.14, 9.A.15, and 9.A.16 we see that solutions associated with the reproductive parts of plants give slightly underestimated values on the interval $[t_2, T]$. From the explicit solution of state variable x_2 given in Eq. (9.91), we see that state variable x_2 is a line on time interval $[t_2, T]$ with slope being $x_1(t_2)$. So for each case, the approximated solution for x_2 presents underestimated values along time interval $[t_2, T]$ because



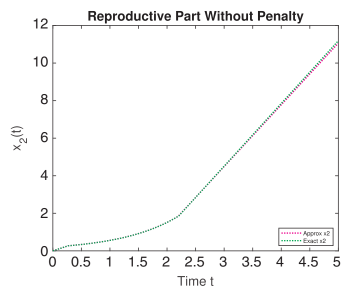
(a) Approximated Allocation Strategy for Unpenalized Problem. Inset shows zoomed in area of interest.



(b) $u^* - \hat{u}$ for Unpenalized Problem. Inset shows zoomed in area of interest.



(c) Vegetative Part of Plant from Unpenalized $\hat{u}$



(d) Reproductive Part of Plant from Unpenalized $\hat{u}$

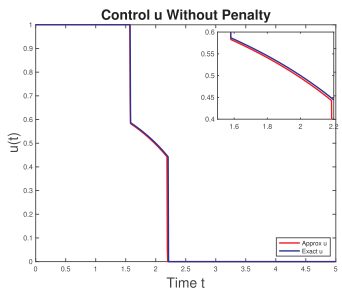
Fig. 9.A.15. Results from solving unpenalized plant problem (Case 2b).

the approximation of state variable x_1 yields underestimated values near t_2. Despite these differences, we find PASA's unpenalized solution to be a sufficient approximation to the true solution to problem (9.65) in Cases 2b and 2c.

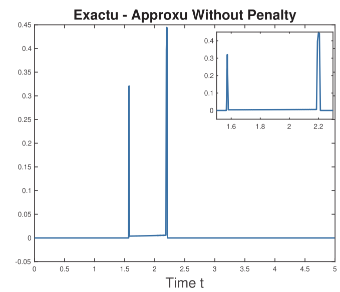
9.A.2 *MATLAB code for the SIR problem*

The following is the MATLAB file called demoOC.m. It can be used in solving problem (9.116). This file is included when downloading the PASA package, and it is located in the following directory:

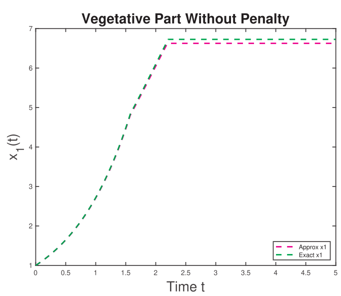
SuiteOPT/PASA/MATLAB. This demo file is associated with the most up-to-date version of PASA, which is SuiteOPT Version 2.0.0. To access PASA software which can be used on MATLAB for



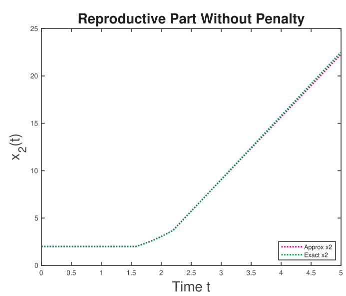
(a) Approximated Allocation Strategy for Unpenalized Problem. Inset shows zoomed in area of interest.



(b) $u^* - \hat{u}$ for Unpenalized Problem. Inset shows zoomed in area of interest.



(c) Vegetative Part of Plant From Unpenalized $\hat{u}$



(d) Reproductive Part of Plant From Unpenalized $\hat{u}$

Fig. 9.A.16. Results from solving unpenalized plant problem (Case 2c).

Linux and Unix operating systems, download the SuiteOPT Version 2.0.0 software given on https://people.clas.ufl.edu/hager/software. For future reference, any updates to the software will be uploaded to this link, and to access older versions of the software, use the same link and then select “Software Archive.” We recommend the reader to read MATLAB file readme.m which is located in the same directory as demoOC.m. The readme.m file goes into detail about the inputs that are used for running PASA. Note that we have state variables initialized as n-vectors rather than $n+1$-vectors because from the discretization of problem (9.116) given in Section 9.5.1, the $(n+1)$th mesh point of S, I, and R is not present in the following: the approximation of the penalized objective functional, the approximation of gradient of the objective functional, and the discretized adjoint equations. We only used the $(n+1)$th mesh points to generalize the transversality conditions for the adjoint equations. We still are discretizing $[0, T]$ with n mesh intervals, which is why the mesh size is set to being $h = T/n$.

```
% This test problem is an optimal control problem that arises in the
% treatment of the ebola virus (see *). An important feature of the problem is
% that it is singular. If the dynamics were discretized and the cost
% was optimized, then the solution oscillates wildly in the singular
% region, so the optimal control in that region would not be determined.
% The oscillations are removed using a penalty term based on the total
% variation of the optimal control. The dynamics are discretized using
% Euler's method with a constant control on each mesh interval. If u_i
% is the control on the i-th interval, then we add the constraint
%
% u_{i+1} - u_i = iota_i - zeta_i
%
% where zeta_i and iota_i >= 0. The penalty term that we add to the objective
% function is p * sum_{i = 1}^n iota_i + zeta_i, which is p times the total
% variation in the control. Besides the control u, which can be singular,
% there is another control v which is bang/bang. The optimization in this
% problem is over x = [u, zeta, iota, v] where u and v are both constrained
% to the interval [0, 1], while the only constraint on zeta and iota is the
% nonnegativity constraint. If there are n intervals in the mesh, then the
```

```
% dimension of x is 4*n - 2 since u and v both have n components while
% zeta and iota have n-1 components. For pasa, the constraints should be
% written in the form bl <= A*x <= bu and lo <= x <= hi. Thus a row of the A
% matrix corresponds to the constraint:
%
%           0 <= u_i - u_{i+1} + iota_i - zeta_i <= 0, 1 <= i <= n-1
%
% The state variable in the control problem is the triple (S, I, R)
% corresponding to the susceptible, infected, and recovered individuals.
% The control u is vaccination rate while v is referred to as treatment rate.
% The controls are linear in both the dynamics and the cost, which leads
% to the singularity of the solution.
%
%   * Optimal Control for a SIR Epidemiological Model with Time-varying
%   Populations by Urszula Ledzewicz, Mahya Aghaee, and Heinz Schaettler,
%   2016 IEEE Conference on Control Applications (CCA), pp. 1268-1273,
%   DOI: 10.1109/CCA.2016.7587981

function demoOC
    % ------ Initialize constant parameters ------ %
    global gamma beta nu rho kappa mu alpha eta T n a b c p
    clf ;
    gamma = 0.00683 ;  %birth rate of the population
    nu = 0.00188 ;     %natural death rate of the population
    beta = 0.2426 ;    %rate of infectiousness of the disease
    rho = 0.007 ;      %resensitization rate
    kappa = 0.3 ;      %effectiveness of vaccination
    mu = 0.005 ;       %disease induced death rate
```

```
    alpha = 0.00002 ;  %rate at which disease is overcome
    eta = 0.1 ;        %effectiveness of treatment
    T = 50 ;           %time horizon (weeks)
    n = 500 ;          % Dimension of u and v, the number of mesh intervals
    a = 5 ;            % Constant in cost function
    b = 50 ;           % Constant in cost function
    c = 300 ;          % Constant in cost function
    p = 1e-1 ;         % penalty parameter in cost function
    umax = 1 ;         % maximum vaccination rate
    vmax = 1;          % maximum treatment rate
    h = T/n ;          % Step size
    S = zeros (n, 1) ;
    I = zeros (n, 1) ;
    R = zeros (n, 1) ;

%% --- Store problem description in a structure which we call pasadata --- %
    % ------------- Setup sparse matrix A ------------- %
    A1 = spdiags([ones(n-1,1) -ones(n-1,1)], [0,1], n-1, n) ;
    A2 = speye(n-1) ;
    A = sparse (n-1, 4*n-2) ;
    % A1 corresponds to the u_i - u_{i+1} term in the contraint while
    % A2 is used for the zeta and iota terms in the constraint
    A(:, 1:3*n-2) = [A1 -A2 A2] ;

    % put A in the structure
    pasadata.A = A ;

    % If the constraint bl <= A*x <= bu is present, then pasa uses the
    % dimensions of A to determine the number of linear constraints
    % and the number of components in x. Thus if the constraint matrix lies
    % in the upper left nrow by ncol submatrix of a larger matrix Afull,
```

```
    % set pasadata.A = Afull(1:nrow, 1:ncol). In the example above, the
    % statement A = sparse (n-1, 4*n-2) specified the dimensions of A

    % -------- store the bounds for A*x -------- %
    pasadata.bl = zeros (n-1, 1) ;
    pasadata.bu = zeros (n-1, 1) ;

    % --------- store the bounds for x --------- %
    pasadata.lo = zeros(4*n-2,1) ;
    pasadata.hi = ...
              [umax*ones(n,1); inf*ones(n-1,1); inf*ones(n-1,1); vmax*ones(n,1)] ;

    % The codes to evaluate the cost function and its gradient appear below.
    % Store the name of the codes in the pasadata structure.
    pasadata.grad  = @grad ;  % objective gradient
    pasadata.value = @cost ;  % objective value

%% ------------------- User defined parameter values for pasa ---------------- %%
    % Type "pasa readme" for discussion of the parameters.
    % By default, there is no printing of statistics.
    pasadata.pasa.PrintStat = 1 ;    % print statistics for used routines
    pasadata.pasa.grad_tol = 1.e-8 ; % PASA stopping tolerance (1.e-6 default)

    % ----------------- Call pasa to determine optimal x ----------------- %
    [x, stats] = pasa (pasadata) ;

    % Since pasadata.pasa.PrintStat = 1, the statistics are displayed at
    % the end of the run. Since the stats structure was included as an
    % output, the corresponding numerical entries can be found in the
```

```
    % structures stats.pasa and stats.cg

    % ------------------------ Plot the states ------------------------ %
    u = x (1:n) ;           % extract the control u from the returned solution x
    v = x (3*n-1:4*n-2) ;   % extract the control v from the returned solution x
    [S, I, R] = state (u, v) ; % the state associated with the optimal controls

    t = linspace (0, T-h, n) ;
    subplot(3,2,1)
    plot (t, S,'linewidth',2) ;
    xlabel('Time')
    ylabel('S')
    subplot(3,2,2)
    plot (t, I, 'linewidth',2) ;
    xlabel('Time')
    ylabel('I')
    subplot(3,2,3)
    plot (t, R, 'linewidth',2) ;
    xlabel('Time')
    ylabel('R')

    % ------------------- Plot the controls u and v ------------------- %
    subplot(3,2,5);
    plot (t, u,'linewidth',2) ;
    xlabel('Time')
    ylabel({'Control u','(Vaccination)'})
    subplot(3,2,6);
    plot (t, v, 'linewidth',2) ;
    xlabel('Time')
    ylabel({'Control v','(Treatment)'})

%% -------------------- User defined functions for pasa -------------------- %%
    % ----- Objective function ----- %
    function J = cost(x)
        h = T/n ;
        u = x(1:n) ;
```

```
        zeta = x(n+1:2*n-1) ;
        iota = x(2*n:3*n-2) ;
        v = x(3*n-1: 4*n-2) ;
        [S, I, R] = state (u, v) ;
        J = h*(a*sum(I) + b*sum(u) + c*sum(v)) ;
        J = J + p*sum(zeta + iota) ;
    end

    % ---- Gradient of objective function ---- %
    function g = grad(x)
        h = T/n ;
        u = x(1:n) ;
        v = x(3*n-1: 4*n-2) ;

        % Compute state and costate
        [S, I, R] = state (u, v);
        [lS, lI, lR] = costate (S, I, R, u, v) ;

        % Update gradient values
        for i=1:n
            Fu(i) = h*(b - lS(i)*kappa*S(i) + lR(i)*kappa*S(i)) ;
            Fv(i) = h*(c - lI(i)*eta*I(i) + lR(i)*eta*I(i)) ;
        end

        % Store gradient values in array g to return
        g = [Fu, p*ones(1,n-1), p*ones(1,n-1), Fv] ;
    end

    % ---- State ---- %
    function [S, I, R] = state (u, v)
        h = T/n;
        S(1) = 1000 ;
        I(1) = 10 ;
        R(1) = 0 ;

        for i = 1:n-1
            N = S(i) + I(i) + R(i) ;
            S(i + 1) = S(i) + h*(gamma*N - nu*S(i) - (beta*S(i)*I(i))/N ...
                + rho*R(i) - kappa*S(i)*u(i)) ;
```

```
                I(i + 1) = I(i) + h*(beta*S(i)*I(i)/N - (nu + mu + alpha)*I(i)...
                    - eta*I(i)*v(i)) ;
                R(i + 1) = R(i) + h*( - nu*R(i) - rho*R(i)...
                    +  kappa*S(i)*u(i) + alpha*I(i) + eta*I(i)*v(i)) ;
            end
        end

        % ----- Costate ----- %
        function [lS, lI, lR] = costate (S, I, R, u, v)
            h = T/n ;
            lS(n) = 0 ;
            lI(n) = 0 ;
            lR(n) = 0 ;

            for i=n:-1:2
                N = S(i) + I(i) + R(i) ;
                lS(i-1) = lS(i)+ h*lS(i)*(gamma - nu - beta*(I(i)/N) ...
                                + beta*(S(i)*I(i)/(N^2)) - kappa*u(i))...
                                + h*lI(i)*(beta*(I(i)/N) - beta*(S(i)*I(i)/(N^2)))...
                                + h*lR(i)*(kappa*u(i)) ;
                lI(i-1) = lI(i)+ h*a +h*lS(i)*(gamma - (beta*S(i))/N ...
                                + (beta*S(i)*I(i))/N^2)...
                                + h*lI(i)*(( beta*S(i))/N - (beta*S(i)*I(i))/(N^2) ...
                                - (nu + mu + alpha) - eta*v(i))...
                                + h*lR(i)*(alpha + eta*v(i)) ;
                lR(i-1) = lR(i)+ h*lS(i)*(gamma + (beta*S(i)*I(i))/(N^2) + rho)...
                                + h* lI(i)*(- (beta*S(i)*I(i))/(N^2) ) ...
                                + h* lR(i)*(- nu - rho) ;
            end
        end
    end
```

References

1. Joshi, H. R. (2002). Optimal control of an HIV immunology model, *Optim. Control Appl. Methods* **23**, pp. 199–213.
2. Kirschner, D., Lenhart, S., and Serbin, S. (1997). Optimal control of the chemotherapy of HIV, *J. Math. Biol.* **35**, pp. 775–792.
3. Jung, E., Lenhart, S., and Feng, Z. (2005). Optimal control of treatments in a two strain tuberculosis model, *Discrete Contin. Dyn. Syst.* **2**, pp. 473–482.
4. Ögren, P. and Martin, C. (1997). Vaccination strategies for epidemics in highly mobile populations, *Appl. Math. Comput.* **127**, pp. 315–322.
5. Neubert, M. (2003). Marine reserves and optimal harvesting, *Ecol. Lett.* **6**, pp. 843–849, doi: 10.1046/j.1461-0248.2003.00493.x.
6. Ding, W. and Lenhart, S. (2009). Optimal harvesting of a spatially explicit fishery model, *Natural Resource Modeling* **22**, pp. 173–211.
7. Gaff, H., Joshi, H. R., and Lenhart, S. (2007). Optimal harvesting during an invasion of a sublethal plant pathogen, *Environ. Develop. Econ.* **12**, 4, pp. 673–686.
8. King, D. and Roughgarden, J. (1982). Graded allocation between vegetative and reproductive growth for annual plants in growing seasons of random length, *Theoret. Popul. Biol.* **22**, 1, pp. 1–16, doi: 10.1016/0040-5809(82)90032-6.
9. Seierstad, A. and Sydæter, K. (1987). *Optimal Control Theory with Economic Applications*, Vol. 24 of Advanced Textbooks in Economics North-Holland Publishing Co., Amsterdam.
10. Pontryagin, L., Boltyanskii, V., Gamkrelidze, R., and Mishchenko, E. (1964). *The Mathematical Theory of Optimal Processes*, Macmillan, New York.
11. Fleming, W. H. and Rishel, R. (1975). *Deterministic and Stochastic Optimal Control*, Springer-Verlag, New York, NY.
12. Lenhart, S. and Workman, J. T. (2007). *Optimal Control Applied to Biological Models*, Chapman and Hall/CRC Press, Boca Raton, FL.
13. Miele, A. (1975). Recent advances in gradient algorithms for optimal control problems, *J. Optim. Theory Appl.* **17**, pp. 361–430.
14. Edge, E. and Powers, W. (1976). Function-space quasi-newton algorithms for optimal control problems with bounded controls and singular arcs, *J. Optim. Theory Appl.* **20**, 4, pp. 455–479.
15. Betts, J. T. (1998). Survey of numerical methods for trajectory optimization, *J. Guid. Control Dyn.* **21**, pp. 193–207.
16. Rao, A. (2010). A survey of numerical methods for optimal control, *Adv. Astronaut. Sci.* **135**, p. 32.

17. Ledzewicz, U., Aghaee, M., and Schättler, H. (2016). Optimal control for a SIR epidemiological model with time-varying populations, *IEEE Conference on Control Applications (CCA)*, pp. 1268–1273, doi: 10.1109/CCA.2016.7587981.
18. Ledzewicz, U. and Schättler, H. (2011). On optimal singular controls for a general SIR-model with vaccination and treatment, in *Proceedings of the 8th AIMS Conference*, pp. 981–990.
19. Shen, M., Xiao, Y., Rong, L., and Meyers, L. A. (2019). Conflict and accord of optimal treatment strategies for HIV infection within and between hosts, *Math. Biosci.* **309**, pp. 107–117.
20. Joshi, H. R., Herrera, G. E., Lenhart, S., and Neubert, M. (2009). Optimal dynamic harvest of a mobile renewable resource, *Nat. Resource Model.* **22**, 2, pp. 322–343.
21. McDanell, J. and Powers, W. (1971). Necessary conditions for joining optimal singular and nonsingular subarcs, *SIAM J. Control* **9**, pp. 161–173.
22. Robbins, H. (1967). A generalized Legendre-Clebsch condition for the singular cases of optimal control, *IBM J. Res. Dev.* **11**, 4, pp. 361–372.
23. Lewis, R. (1980). Definitions of order and junction conditions in singular optimal control problems, *SIAM J. Control Optim.* **18**, pp. 21–32.
24. Zelikin, M. I. and Borisov, V. (1994). *Theory of Chattering Control: with Applications to Astronautics, Robotics, Economics, and Engineering*, Birkhäuser, Boston, MA.
25. Foroozandeh, Z., De Pinho, M. D. R., and Shamsi, M. (2019). On numerical methods for singular optimal control problems: An application to an AUV problem, *Discrete Contin. Dyn. Syst. — B* **24**, 5, pp. 2219–2235, doi: 10.3934/dcdsb.2019092.
26. Van Wyk, E. J., Falugi, P., and Kerrigan, E. C. (2010). ICLOCS2: Imperial college London optimal control software, http://www.ee.ic.ac.uk/ICLOCS/. Accessed 15 September, 2020.
27. Patterson, M. A. and Rao, A. (2014). GPOPS-II: A MATLAB software for solving multiple-phase optimal control probles using hp-adaptive qaussian quadrature collocation methods and spares non-linear programming, *ACM Trans. Math. Softw.* **41**, 1, pp. 1–37, doi: 10.1145/2558904.
28. Foroozandeh, Z., De Pinho, M. D. R., and Shamsi, M. (2017). A mixed binary non-linear programming approach for the numerical solution of a family of singular optimal control problems, *Int. J. Control* pp. 1–35, doi: 10.1080/00207179.2017.1399216.
29. Yang, X.-q., Teo, K. L., and Caccetta, L. (2001). *Optimization Methods and Applications*, Vol. 52 of Applied Optimization. Springer, New York, NY.

30. Caponigro, M., Ghezzi, R., Piccoli, B., and Trelat, E. (2018). Regularization of chattering phenomena via bounded variation controls, *IEEE Trans. Automatic Control* **63**, 7, pp. 2046–2060, doi: 10.1109/TAC.2018.2810540, arXiv:1303.5796.
31. Hager, W. W. and Zhang, H. (2016). An active set algorithm for nonlinear optimization with polyhedral constraints, *Sci. China Math.* **59**, 8, pp. 1525–1542, doi: 10.1007/s11425-016-0300-6, arXiv:1606.01992.
32. Bertsekas, D. P. (1976). On the Goldstein-Levitin-Polyak gradient projection method, *IEEE Trans. Automat. Control* **21**, pp. 174–184.
33. Calamai, P. and Moré, J. J. (1987). Projected gradient for linearly constrained problems, *Math Program* **39**, pp. 93–116.
34. Goldstein, A. A. (1964). Convex programming in Hilbert space, *Bull. Amer. Math. Soc.* **70**, pp. 709–710.
35. Levitin, E. S. and Polyak, B. T. (1966). Constrained minimization problems, *USSR Comput. Math. Math. Phys.* **6**, pp. 1–50.
36. McCormick, G. P. and Tapia, R. A. (1972). The gradient projection method under mild differentiability conditions, *SIAM J. Control* **10**, pp. 93–98.
37. Clark, C. W. (1990). *Mathematical Bioeconomics: The Optimal Management of Renewable Resources*, Wiley, New York.
38. Brauer, F. and Castillo-Chávez, C. (2001). *Mathematical Models in Population Biology and Epidemiology*, Springer, New York, NY.
39. Martcheva, M. (2015). *An Introduction to Mathematical Epidemiology*, Springer, New York, NY.
40. Conway, J. (1978). *Functions of One Complex Variable*, Springer-Verlag, New York.
41. Hager, W. W. (2000). Runge-Kutta methods in optimal control and the transformed adjoint system, *Numer. Math.* **87**, pp. 247–282.
42. Hager, W. W. and Rostamian, R. (1987). Optimal coatings, bang-bang controls, and gradient techniques, *Optim. Control Appl. Meth.* **8**, pp. 1–20.
43. Kelley, H., Kopp, R., and Moyer, H. (1967). Singular extremals, *Topics Optim.: Math. Sci. and Eng.* **31**, pp. 63–101.
44. Cohen, D. (1971). Maximizing final yield when growth is limited by time or by limiting resources, *J. Theor. Biol.* pp. 299–307, doi: 10.1016/0022-5193(71)90068-3.
45. Cohen, D. (1967). Optimizing reproduction in a randomly varying environment when a correlation may exist between the conditions at the time a choice has to be made and the subsequent outcome,

J. Theor. Biol., **16**, 1, pp. 1–14, doi: https://doi.org/10.1016/0022-5193(67)90050-1, http://www.sciencedirect.com/science/article/pii/0022519367900501.

46. Schättler, H. and Ledzewicz, U. (2012). *Geometric Optimal Control*, Springer, New York, NY.
47. Rutquist, P. and Edvall, M. (2010). *PROPT: Matlab Optimal Control Software*, Tomlab Optimization, Inc., Pullman, WA.