A NEW COMPUTATIONAL METHOD FOR OPTIMIZING NONLINEAR IMPULSIVE SYSTEMS

Qun Lin, Ryan Loxton, Kok Lay Teo, and Yong Hong Wu

Department of Mathematics and Statistics
Curtin University of Technology, Perth, Australia
Corresponding author email: R.Loxton@curtin.edu.au

Abstract. In this paper, we consider a system that evolves by switching between several subsystems of ordinary differential equations. The switching mechanism in this system induces an instantaneous change in the system's state, which can be controlled through a set of decision parameters. We develop a new computational method, based on nonlinear programming, for optimizing the system parameters and the subsystem switching times. We then successfully apply this method to two interesting examples.

Keywords. Nonlinear optimal control, Impulsive system, Gradient-based optimization, Nonlinear programming, Time-scaling transformation.

AMS (MOS) subject classification: 49J15, 49N25, 90C30.

1 Introduction

Many real-life systems operate by switching between different subsystems or modes. Such systems are called *switched systems*. An example of a switched system is a switched-capacitor DC-DC power converter, which operates by periodically changing its circuit topology [6,14]. Other examples of switched systems include robots [2], locomotives [7,8], hybrid power generators [19], and biochemical reactors [3,4].

In some switched systems, changing mode causes an instantaneous change in the system's state—a so-called *state jump*. For example, changing the circuit topology of a switched-capacitor DC-DC power converter causes a sudden voltage drop in the power converter's capacitors. Switched systems of this type, in which subsystem switches are accompanied by state jumps, are called *impulsive systems*.

In this paper, we consider an impulsive system whose subsystems are described by nonlinear ordinary differential equations. The state jumps in this system can be controlled through a set of system parameters. Our goal is to choose values for these system parameters and the subsystem switching times to minimize a given cost function.

This type of dynamic optimization problem is a major computational challenge. There are two reasons for this:

(i) The governing impulsive system is difficult to integrate numerically because its switching times are variable.

(ii) The partial derivatives of the cost function with respect to the switching times are not well-defined.

In this paper, we overcome these difficulties by introducing an *equivalent* optimization problem that is governed by an impulsive system with *fixed* switching times. This new problem is much easier to solve than the original, and its solution can be used to obtain the optimal parameters and optimal switching times for the original impulsive system.

This approach is inspired by the so-called *time-scaling transformation*, which was first introduced in [10] to compute optimal switching times for bang-bang control problems. Under this transformation, the switching times are mapped to fixed points in a new time horizon. The time-scaling transformation has already been applied to a variety of optimization and optimal control problems involving switched and impulsive systems (see [11-13,17,20,21]). Typically, the transformation works by introducing a new time variable s and relating s to the original time variable t through the dynamic equations

$$\frac{dt(s)}{ds} = v(s), \tag{1}$$

$$t(0) = 0, \tag{2}$$

where v is a non-negative piecewise constant function. In this paper, we describe a new way of applying the time-scaling transformation that does not use equations (1)-(2).

The other main contribution of this paper is a novel method for computing the gradient of the cost function. This method can be used in conjunction with a standard nonlinear programming algorithm (see [15,16]) to optimize the switching times and system parameters. Unlike the methods in [11,12,20], which involve integrating a costate system backwards in time, our new method involves integrating an auxiliary system *forward* in time. Thus, since both the state and auxiliary systems are integrated in the same direction, our new method is very convenient to implement.

2　Problem Statement

Consider the following impulsive system with $m \geq 2$ subsystems:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}^i\big(\boldsymbol{x}(t), \boldsymbol{\zeta}\big), \qquad t \in (\tau_{i-1}, \tau_i), \qquad i = 1, \ldots, m, \tag{3}$$

and

$$\boldsymbol{x}(\tau_i) = \boldsymbol{x}(\tau_i^+) = \begin{cases} \boldsymbol{x}^0, & \text{if } i = 0, & \text{(4a)} \\ \boldsymbol{h}^i\big(\boldsymbol{x}(\tau_i^-), \boldsymbol{\zeta}\big), & \text{if } i = 1, \ldots, m, & \text{(4b)} \end{cases}$$

where $\tau_0 \triangleq 0$, $\tau_m \triangleq T$, and $T > 0$ is a given final time; τ_i, $i = 1, \ldots, m-1$, are the subsystem switching times; $\boldsymbol{x}(t) \in \mathbb{R}^n$ is the system's state at time t; $\boldsymbol{x}^0 \in \mathbb{R}^n$ is the system's given initial state; $\boldsymbol{\zeta} \in \mathbb{R}^r$ is a vector of system parameters; and $\boldsymbol{f}^i : \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^n$, $i = 1, \ldots, m$, and $\boldsymbol{h}^i : \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^n$, $i = 1, \ldots, m$, are given functions.

The subsystem switching times are decision variables chosen by the system designer. They must satisfy the following constraints:

$$\tau_i - \tau_{i-1} \geq \Delta_i, \qquad i = 1, \ldots, m, \tag{5}$$

where $\Delta_i > 0$ is the minimum duration of subsystem i. Clearly,

$$\Delta_1 + \cdots + \Delta_m \leq T.$$

Let $\mathcal{T}$ denote the set of all $\boldsymbol{\tau} \in \mathbb{R}^{m-1}$ satisfying (5).

The system parameters are also chosen by the system designer. They are subject to the following constraints:

$$a_j \leq \zeta_j \leq b_j, \qquad j = 1, \ldots, r, \tag{6}$$

where a_j and b_j are given real numbers such that $a_j < b_j$. Let $\mathcal{W}$ denote the set of all $\boldsymbol{\zeta} \in \mathbb{R}^r$ satisfying (6).

We make the following assumptions.

Assumption 1. The functions $\boldsymbol{f}^i$ and $\boldsymbol{h}^i$, $i = 1, \ldots, m$, are continuously differentiable.

Assumption 2. There exists a real number $L_1 > 0$ such that

$$\|\boldsymbol{f}^i(\boldsymbol{v}, \boldsymbol{w})\| \leq L_1(1 + \|\boldsymbol{v}\|), \qquad (\boldsymbol{v}, \boldsymbol{w}) \in \mathbb{R}^n \times \mathcal{W}, \qquad i = 1, \ldots, m,$$

where $\| \cdot \|$ denotes the Euclidean norm.

Assumptions 1-2 ensure that the impulsive system (3)-(4) has a unique solution $\boldsymbol{x}(\cdot | \boldsymbol{\tau}, \boldsymbol{\zeta})$ corresponding to each pair $(\boldsymbol{\tau}, \boldsymbol{\zeta}) \in \mathcal{T} \times \mathcal{W}$ (see Theorem 3.1.6 of [1]).

We assume that the system's operating cost depends on the state immediately before and after each switch. Accordingly, we define a *cost function* as follows:

$$J(\boldsymbol{\tau}, \boldsymbol{\zeta}) \triangleq \sum_{i=1}^{m} \Psi_i\big(\boldsymbol{\zeta}, \boldsymbol{x}(\tau_i^- | \boldsymbol{\tau}, \boldsymbol{\zeta}), \boldsymbol{x}(\tau_i^+ | \boldsymbol{\tau}, \boldsymbol{\zeta})\big), \qquad (\boldsymbol{\tau}, \boldsymbol{\zeta}) \in \mathcal{T} \times \mathcal{W}, \tag{7}$$

where $\Psi_i : \mathbb{R}^r \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, are given continuously differentiable functions. Note that we can easily incorporate an integral term into (7) by introducing a dummy state variable. For example, consider the following integral term:

$$\sum_{i=1}^{m} \int_{\tau_{i-1}}^{\tau_i} \mathcal{L}_i\big(\boldsymbol{x}(t | \boldsymbol{\tau}, \boldsymbol{\zeta}), \boldsymbol{\zeta}\big) dt.$$

It is clear that this term can be replaced by $v(\tau_m^-)$, where v satisfies the dynamics

$$\dot{v}(t) = \mathcal{L}_i\big(\boldsymbol{x}(t), \boldsymbol{\zeta}\big), \qquad t \in (\tau_{i-1}, \tau_i), \qquad i = 1, \ldots, m,$$

and

$$v(\tau_i^+) = \begin{cases} 0, & i = 0, \\ v(\tau_i^-), & i = 1, \ldots, m. \end{cases}$$

We want to choose the switching times $\tau_1, \ldots, \tau_{m-1}$ and the system parameters $\zeta_1, \ldots, \zeta_r$ to minimize the cost function (7) subject to the constraints (5) and (6). We state this problem formally below.

Problem P.   *Find a pair* $(\boldsymbol{\tau}^*, \boldsymbol{\zeta}^*) \in \mathcal{T} \times \mathcal{W}$ *such that*

$$J(\boldsymbol{\tau}^*, \boldsymbol{\zeta}^*) = \inf_{(\boldsymbol{\tau}, \boldsymbol{\zeta}) \in \mathcal{T} \times \mathcal{W}} J(\boldsymbol{\tau}, \boldsymbol{\zeta}).$$

3   An Equivalent Problem

The decision variables in Problem P are the subsystem switching times $\tau_1, \ldots, \tau_{m-1}$ and the system parameters $\zeta_1, \ldots, \zeta_r$. As we mentioned in the introduction, the partial derivatives of the cost function with respect to the switching times are not well-defined (see [10]). Hence, nonlinear programming techniques, which use the cost function's partial derivatives to compute search directions, cannot solve Problem P directly. In this section, we transform Problem P into an equivalent problem that is easier to solve.

3.1   Problem Statement

Let

$$\Theta \triangleq \big\{ \boldsymbol{\theta} \in \mathbb{R}^m : \theta_i \geq \Delta_i, \, i = 1, \ldots, m; \, \theta_1 + \cdots + \theta_m = T \big\}.$$

Consider the following impulsive system:

$$\dot{\boldsymbol{y}}(s) = \theta_i \boldsymbol{f}^i\big(\boldsymbol{y}(s), \boldsymbol{\zeta}\big), \qquad s \in (i-1, i), \qquad i = 1, \ldots, m, \qquad (8)$$

and

$$\boldsymbol{y}(i) = \boldsymbol{y}(i^+) = \begin{cases} \boldsymbol{x}^0, & \text{if } i = 0, & (9a) \\ \boldsymbol{h}^i\big(\boldsymbol{y}(i^-), \boldsymbol{\zeta}\big), & \text{if } i = 1, \ldots, m, & (9b) \end{cases}$$

where $(\boldsymbol{\theta}, \boldsymbol{\zeta}) \in \Theta \times \mathcal{W}$. Let $\boldsymbol{y}(\cdot|\boldsymbol{\theta}, \boldsymbol{\zeta})$ denote the solution of (8)-(9) corresponding to the pair $(\boldsymbol{\theta}, \boldsymbol{\zeta}) \in \Theta \times \mathcal{W}$. Note that subsystem switches in (8)-(9) occur at the fixed times $s = 1, \ldots, m-1$.

Define a new cost function $\tilde{J}$ as follows:

$$\tilde{J}(\boldsymbol{\theta}, \boldsymbol{\zeta}) \triangleq \sum_{i=1}^m \Psi_i\big(\boldsymbol{\zeta}, \boldsymbol{y}(i^-|\boldsymbol{\theta}, \boldsymbol{\zeta}), \boldsymbol{y}(i^+|\boldsymbol{\theta}, \boldsymbol{\zeta})\big), \qquad (\boldsymbol{\theta}, \boldsymbol{\zeta}) \in \Theta \times \mathcal{W}.$$

Consider the following optimization problem.

Problem $\tilde{\mathrm{P}}$.   *Find a pair $(\boldsymbol{\theta}^*, \boldsymbol{\zeta}^*) \in \Theta \times \mathcal{W}$ such that*

$$\tilde{J}(\boldsymbol{\theta}^*, \boldsymbol{\zeta}^*) = \inf_{(\boldsymbol{\theta}, \boldsymbol{\zeta}) \in \Theta \times \mathcal{W}} \tilde{J}(\boldsymbol{\theta}, \boldsymbol{\zeta}).$$

In the next subsection, we will show that Problem $\tilde{\mathrm{P}}$ is equivalent to Problem P. This means that a solution of Problem P can be obtained from a solution of Problem $\tilde{\mathrm{P}}$, and vice versa.

3.2   Equivalence of Problems P and $\tilde{\mathrm{P}}$

For each $\boldsymbol{\theta} \in \Theta$, define a corresponding function $\mu(\cdot|\boldsymbol{\theta}) : [0, m] \to \mathbb{R}$ as follows:

$$\mu(s|\boldsymbol{\theta}) \triangleq \begin{cases} \sum_{k=1}^{\lfloor s \rfloor} \theta_k + \theta_{\lfloor s \rfloor + 1}(s - \lfloor s \rfloor), & \text{if } s \in [0, m), \\ T, & \text{if } s = m, \end{cases}$$

where $\lfloor \cdot \rfloor$ denotes the floor function. It is easy to see that $\mu(\cdot|\boldsymbol{\theta})$ is continuous, non-negative, and strictly increasing. Furthermore,

$$\mu(i|\boldsymbol{\theta}) - \mu(i - 1|\boldsymbol{\theta}) = \theta_i \geq \Delta_i, \qquad i = 1, \dots, m. \tag{10}$$

Now, for each $\boldsymbol{\theta} \in \Theta$, define

$$\tilde{\boldsymbol{\tau}}(\boldsymbol{\theta}) \triangleq \left[\mu(1|\boldsymbol{\theta}), \, \mu(2|\boldsymbol{\theta}), \, \dots, \, \mu(m - 1|\boldsymbol{\theta}) \right]^T \in \mathbb{R}^{m-1}.$$

It follows from (10) that $\tilde{\boldsymbol{\tau}}(\boldsymbol{\theta}) \in \mathcal{T}$. Thus, the components of $\tilde{\boldsymbol{\tau}}(\boldsymbol{\theta})$ can be used as the switching times in (3)-(4). Our first result links the solution of (3)-(4) with the solution of (8)-(9).

Theorem 1. *For each pair $(\boldsymbol{\theta}, \boldsymbol{\zeta}) \in \Theta \times \mathcal{W}$,*

$$\boldsymbol{y}(s|\boldsymbol{\theta}, \boldsymbol{\zeta}) = \boldsymbol{x}(\mu(s|\boldsymbol{\theta})|\tilde{\boldsymbol{\tau}}(\boldsymbol{\theta}), \boldsymbol{\zeta}), \qquad s \in [0, m].$$

Proof. Let $(\boldsymbol{\theta}, \boldsymbol{\zeta}) \in \Theta \times \mathcal{W}$ be arbitrary but fixed. For simplicity, we write μ instead of $\mu(\cdot|\boldsymbol{\theta})$, $\boldsymbol{x}$ instead of $\boldsymbol{x}(\cdot|\tilde{\boldsymbol{\tau}}(\boldsymbol{\theta}), \boldsymbol{\zeta})$, and $\boldsymbol{y}$ instead of $\boldsymbol{y}(\cdot|\boldsymbol{\theta}, \boldsymbol{\zeta})$. This notation will not cause confusion because both $\boldsymbol{\theta}$ and $\boldsymbol{\zeta}$ are fixed.

Clearly,

$$\dot{\mu}(s) = \theta_i, \qquad s \in (i - 1, i), \qquad i = 1, \dots, m. \tag{11}$$

Furthermore, since μ is strictly increasing,

$$\mu(i - 1) < \mu(s) < \mu(i), \qquad s \in (i - 1, i), \qquad i = 1, \dots, m. \tag{12}$$

Define

$$\tilde{\boldsymbol{x}}(s) \triangleq \boldsymbol{x}(\mu(s)|\tilde{\boldsymbol{\tau}}(\boldsymbol{\theta}), \boldsymbol{\zeta}), \qquad s \in [0, m].$$

By virtue of (3), (11), and (12), we have

$$\dot{\tilde{\boldsymbol{x}}}(s) = \dot{\mu}(s)\boldsymbol{f}^i\big(\tilde{\boldsymbol{x}}(s),\boldsymbol{\zeta}\big) = \theta_i\boldsymbol{f}^i\big(\tilde{\boldsymbol{x}}(s),\boldsymbol{\zeta}\big), \quad s \in (i-1,i), \quad i = 1,\ldots,m. \quad (13)$$

Since μ is continuous and strictly increasing,

$$\tilde{\boldsymbol{x}}(i^-) = \lim_{s\to i^-} \boldsymbol{x}(\mu(s)) = \lim_{t\to\mu(i)^-} \boldsymbol{x}(t) = \boldsymbol{x}(\mu(i)^-), \qquad i = 1,\ldots,m. \quad (14)$$

Similarly,

$$\tilde{\boldsymbol{x}}(i^+) = \boldsymbol{x}(\mu(i)^+), \qquad i = 1,\ldots,m. \quad (15)$$

It follows from (4b), (14), and (15) that

$$\tilde{\boldsymbol{x}}(i^+) = \boldsymbol{x}(\mu(i)^+) = \boldsymbol{h}^i\big(\boldsymbol{x}(\mu(i)^-),\boldsymbol{\zeta}\big) = \boldsymbol{h}^i\big(\tilde{\boldsymbol{x}}(i^-),\boldsymbol{\zeta}\big), \quad i = 1,\ldots,m. \quad (16)$$

Furthermore,

$$\tilde{\boldsymbol{x}}(0) = \boldsymbol{x}(0) = \boldsymbol{x}^0. \quad (17)$$

Equations (13), (16), and (17) show that $\tilde{\boldsymbol{x}}$ is the unique solution of (8)-(9). Thus,

$$\boldsymbol{y}(s) = \tilde{\boldsymbol{x}}(s) = \boldsymbol{x}(\mu(s)), \qquad s \in [0,m],$$

as required. □

We now use Theorem 1 to prove the following important result.

Theorem 2. *For each pair* $(\boldsymbol{\theta},\boldsymbol{\zeta}) \in \Theta \times \mathcal{W}$,

$$\tilde{J}(\boldsymbol{\theta},\boldsymbol{\zeta}) = J(\tilde{\boldsymbol{\tau}}(\boldsymbol{\theta}),\boldsymbol{\zeta}).$$

Proof. Let $(\boldsymbol{\theta},\boldsymbol{\zeta}) \in \Theta \times \mathcal{W}$ be arbitrary but fixed. Furthermore, let μ, $\boldsymbol{x}$, and $\boldsymbol{y}$ be as defined in the proof of Theorem 1.

Recall from Theorem 1 that

$$\boldsymbol{y}(s) = \boldsymbol{x}(\mu(s)), \qquad s \in [0,m].$$

Thus, since μ is continuous and increasing, for each $i = 1,\ldots,m$,

$$\boldsymbol{y}(i^-) = \lim_{s\to i^-} \boldsymbol{y}(s) = \lim_{s\to i^-} \boldsymbol{x}(\mu(s)) = \lim_{t\to\mu(i)^-} \boldsymbol{x}(t) = \boldsymbol{x}(\mu(i)^-). \quad (18)$$

Similarly,

$$\boldsymbol{y}(i^+) = \boldsymbol{x}(\mu(i)^+), \qquad i = 1,\ldots,m. \quad (19)$$

From (18) and (19) we obtain

$$\begin{aligned}
\tilde{J}(\boldsymbol{\theta},\boldsymbol{\zeta}) &= \sum_{i=1}^m \Psi_i\big(\boldsymbol{\zeta},\boldsymbol{y}(i^-),\boldsymbol{y}(i^+)\big) \\
&= \sum_{i=1}^m \Psi_i\big(\boldsymbol{\zeta},\boldsymbol{x}(\mu(i)^-),\boldsymbol{x}(\mu(i)^+)\big) \\
&= J(\tilde{\boldsymbol{\tau}}(\boldsymbol{\theta}),\boldsymbol{\zeta}),
\end{aligned}$$

which completes the proof. □

We now prove our main result: that Problem $\tilde{P}$ is equivalent to Problem P.

Theorem 3. *Let* $(\boldsymbol{\theta}^*, \boldsymbol{\zeta}^*) \in \Theta \times \mathcal{W}$. *Then* $(\boldsymbol{\theta}^*, \boldsymbol{\zeta}^*)$ *is optimal for Problem* $\tilde{P}$ *if and only if* $(\tilde{\boldsymbol{\tau}}(\boldsymbol{\theta}^*), \boldsymbol{\zeta}^*)$ *is optimal for Problem P.*

Proof. Suppose that $(\boldsymbol{\theta}^*, \boldsymbol{\zeta}^*) \in \Theta \times \mathcal{W}$ is optimal for Problem $\tilde{P}$. Then $\tilde{\boldsymbol{\tau}}(\boldsymbol{\theta}^*) \in \mathcal{T}$ and thus $(\tilde{\boldsymbol{\tau}}(\boldsymbol{\theta}^*), \boldsymbol{\zeta}^*)$ is feasible for Problem P. Let $(\boldsymbol{\tau}, \boldsymbol{\zeta}) \in \mathcal{T} \times \mathcal{W}$ be arbitrary but fixed. Define

$$\theta_i \triangleq \tau_i - \tau_{i-1}, \qquad i = 1, \ldots, m.$$

Then

$$\theta_i \geq \Delta_i, \qquad i = 1, \ldots, m,$$

and

$$\sum_{i=1}^{m} \theta_i = \sum_{i=1}^{m} (\tau_i - \tau_{i-1}) = \tau_m - \tau_0 = T.$$

Thus, $\boldsymbol{\theta} \in \Theta$. Furthermore,

$$\mu(i|\boldsymbol{\theta}) = \sum_{k=1}^{i} \theta_k = \sum_{k=1}^{i} (\tau_k - \tau_{k-1}) = \tau_i, \qquad i = 1, \ldots, m-1.$$

This shows that $\tilde{\boldsymbol{\tau}}(\boldsymbol{\theta}) = \boldsymbol{\tau}$. Now, by Theorem 2,

$$J(\tilde{\boldsymbol{\tau}}(\boldsymbol{\theta}^*), \boldsymbol{\zeta}^*) = \tilde{J}(\boldsymbol{\theta}^*, \boldsymbol{\zeta}^*) \leq \tilde{J}(\boldsymbol{\theta}, \boldsymbol{\zeta}) = J(\tilde{\boldsymbol{\tau}}(\boldsymbol{\theta}), \boldsymbol{\zeta}) = J(\boldsymbol{\tau}, \boldsymbol{\zeta}).$$

Since $(\boldsymbol{\tau}, \boldsymbol{\zeta})$ was chosen arbitrarily, this inequality shows that $(\tilde{\boldsymbol{\tau}}(\boldsymbol{\theta}^*), \boldsymbol{\zeta}^*)$ is an optimal solution for Problem P.

Conversely, suppose $(\tilde{\boldsymbol{\tau}}(\boldsymbol{\theta}^*), \boldsymbol{\zeta}^*) \in \mathcal{T} \times \mathcal{W}$ is optimal for Problem P and let $(\boldsymbol{\theta}, \boldsymbol{\zeta}) \in \Theta \times \mathcal{W}$ be arbitrary but fixed. Then $(\tilde{\boldsymbol{\tau}}(\boldsymbol{\theta}), \boldsymbol{\zeta}) \in \mathcal{T} \times \mathcal{W}$. Hence, by Theorem 2,

$$\tilde{J}(\boldsymbol{\theta}^*, \boldsymbol{\zeta}^*) = J(\tilde{\boldsymbol{\tau}}(\boldsymbol{\theta}^*), \boldsymbol{\zeta}^*) \leq J(\tilde{\boldsymbol{\tau}}(\boldsymbol{\theta}), \boldsymbol{\zeta}) = \tilde{J}(\boldsymbol{\theta}, \boldsymbol{\zeta}).$$

This shows that $(\boldsymbol{\theta}^*, \boldsymbol{\zeta}^*)$ is optimal for Problem $\tilde{P}$. $\square$

4   Gradient Computation for Problem $\tilde{P}$

In this section, we develop an algorithm for computing the gradient of $\tilde{J}$. This algorithm can be combined with any nonlinear programming method—for example, a conjugate gradient method (see [15,16])—to solve Problem $\tilde{P}$.

Let

$$\delta_{k,i} \triangleq \begin{cases} 1, & \text{if } k = i, \\ 0, & \text{otherwise,} \end{cases}$$

and

$$\hat{\delta}_{k,i} \triangleq \begin{cases} 1, & \text{if } k \leq i, \\ 0, & \text{otherwise.} \end{cases}$$

For each $k = 1, \ldots, m$, define the following auxiliary system:

$$\dot{\boldsymbol{\psi}}^k(s) = \hat{\delta}_{k,i}\theta_i \frac{\partial \boldsymbol{f}^i\big(\boldsymbol{y}(s|\boldsymbol{\theta},\boldsymbol{\zeta}),\boldsymbol{\zeta}\big)}{\partial \boldsymbol{x}}\boldsymbol{\psi}^k(s) \tag{20}$$
$$+ \delta_{k,i}\boldsymbol{f}^i\big(\boldsymbol{y}(s|\boldsymbol{\theta},\boldsymbol{\zeta}),\boldsymbol{\zeta}\big), \quad s \in (i-1,i), \quad i = 1, \ldots, m,$$

and

$$\boldsymbol{\psi}^k(i) = \boldsymbol{\psi}^k(i^+) = \begin{cases} \boldsymbol{0}, & \text{if } i = 0, \tag{21a} \\[2mm] \dfrac{\partial \boldsymbol{h}^i\big(\boldsymbol{y}(i^-|\boldsymbol{\theta},\boldsymbol{\zeta}),\boldsymbol{\zeta}\big)}{\partial \boldsymbol{x}}\boldsymbol{\psi}^k(i^-), & \text{if } i = 1, \ldots, m, \quad \text{(21b)} \end{cases}$$

where $(\boldsymbol{\theta},\boldsymbol{\zeta}) \in \Theta \times \mathcal{W}$. Let $\boldsymbol{\psi}^k(\cdot|\boldsymbol{\theta},\boldsymbol{\zeta})$ denote the solution of (20)-(21).

For each $j = 1, \ldots, r$, define another auxiliary system as follows:

$$\dot{\boldsymbol{\phi}}^j(s) = \theta_i \frac{\partial \boldsymbol{f}^i\big(\boldsymbol{y}(s|\boldsymbol{\theta},\boldsymbol{\zeta}),\boldsymbol{\zeta}\big)}{\partial \boldsymbol{x}}\boldsymbol{\phi}^j(s) \tag{22}$$
$$+ \theta_i \frac{\partial \boldsymbol{f}^i\big(\boldsymbol{y}(s|\boldsymbol{\theta},\boldsymbol{\zeta}),\boldsymbol{\zeta}\big)}{\partial \zeta_j}, \quad s \in (i-1,i), \quad i = 1, \ldots, m,$$

and

$$\boldsymbol{\phi}^j(i) = \boldsymbol{\phi}^j(i^+) = \begin{cases} \boldsymbol{0}, & \text{if } i = 0, \tag{23a} \\[2mm] \dfrac{\partial \boldsymbol{h}^i\big(\boldsymbol{y}(i^-|\boldsymbol{\theta},\boldsymbol{\zeta}),\boldsymbol{\zeta}\big)}{\partial \boldsymbol{x}}\boldsymbol{\phi}^j(i^-) \\[3mm] + \dfrac{\partial \boldsymbol{h}^i\big(\boldsymbol{y}(i^-|\boldsymbol{\theta},\boldsymbol{\zeta}),\boldsymbol{\zeta}\big)}{\partial \zeta_j}, & \text{if } i = 1, \ldots, m, \quad \text{(23b)} \end{cases}$$

where $(\boldsymbol{\theta},\boldsymbol{\zeta}) \in \Theta \times \mathcal{W}$. Let $\boldsymbol{\phi}^j(\cdot|\boldsymbol{\theta},\boldsymbol{\zeta})$ denote the solution of (22)-(23).

We have the following important result.

Theorem 4. *For each pair $(\boldsymbol{\theta},\boldsymbol{\zeta}) \in \Theta \times \mathcal{W}$,*

$$\frac{\partial \boldsymbol{y}(s|\boldsymbol{\theta},\boldsymbol{\zeta})}{\partial \theta_k} = \boldsymbol{\psi}^k(s|\boldsymbol{\theta},\boldsymbol{\zeta}), \qquad s \in [0,m], \qquad k = 1, \ldots, m, \tag{24}$$

and

$$\frac{\partial \boldsymbol{y}(s|\boldsymbol{\theta},\boldsymbol{\zeta})}{\partial \zeta_j} = \boldsymbol{\phi}^j(s|\boldsymbol{\theta},\boldsymbol{\zeta}), \qquad s \in [0,m], \qquad j = 1, \ldots, r. \tag{25}$$

Proof. Let $(\boldsymbol{\theta},\boldsymbol{\zeta}) \in \Theta \times \mathcal{W}$, $k \in \{1, \ldots, m\}$, and $j \in \{1, \ldots, r\}$ be arbitrary but fixed. For simplicity, we write $\boldsymbol{y}$ instead of $\boldsymbol{y}(\cdot|\boldsymbol{\theta},\boldsymbol{\zeta})$.

It is clear from (8)-(9) that $\boldsymbol{y}(s)$ does not depend on θ_k for $s \in [0, k-1)$. Thus,

$$\frac{\partial \boldsymbol{y}(s)}{\partial \theta_k} = \boldsymbol{0}, \qquad s \in [0, k-1),$$

and

$$\frac{d}{ds}\left\{\frac{\partial \boldsymbol{y}(s)}{\partial \theta_k}\right\} = \boldsymbol{0}, \qquad s \in [0, k-1). \tag{26}$$

Furthermore, by (8),

$$\boldsymbol{y}(s) = \boldsymbol{y}(i-1) + \int_{i-1}^{s} \theta_i \boldsymbol{f}^i\big(\boldsymbol{y}(\eta), \boldsymbol{\zeta}\big)d\eta, \quad s \in (i-1, i), \quad i = 1, \ldots, m. \quad (27)$$

When $i \geq k$, differentiating this equation with respect to θ_k yields

$$\frac{\partial \boldsymbol{y}(s)}{\partial \theta_k} = \frac{\partial \boldsymbol{y}(i-1)}{\partial \theta_k} + \int_{i-1}^{s} \theta_i \frac{\partial \boldsymbol{f}^i\big(\boldsymbol{y}(\eta), \boldsymbol{\zeta}\big)}{\partial \boldsymbol{x}} \frac{\partial \boldsymbol{y}(\eta)}{\partial \theta_k} d\eta$$
$$+ \int_{i-1}^{s} \delta_{k,i} \boldsymbol{f}^i\big(\boldsymbol{y}(\eta), \boldsymbol{\zeta}\big) d\eta, \quad s \in (i-1, i), \quad i = k, \ldots, m.$$

Thus,

$$\frac{d}{ds}\left\{ \frac{\partial \boldsymbol{y}(s)}{\partial \theta_k} \right\} = \theta_i \frac{\partial \boldsymbol{f}^i\big(\boldsymbol{y}(s), \boldsymbol{\zeta}\big)}{\partial \boldsymbol{x}} \frac{\partial \boldsymbol{y}(s)}{\partial \theta_k}$$
$$+ \delta_{k,i} \boldsymbol{f}^i\big(\boldsymbol{y}(s), \boldsymbol{\zeta}\big), \quad s \in (i-1, i), \quad i = k, \ldots, m. \quad (28)$$

Combining equations (26) and (28) gives

$$\frac{d}{ds}\left\{ \frac{\partial \boldsymbol{y}(s)}{\partial \theta_k} \right\} = \hat{\delta}_{k,i} \theta_i \frac{\partial \boldsymbol{f}^i\big(\boldsymbol{y}(s), \boldsymbol{\zeta}\big)}{\partial \boldsymbol{x}} \frac{\partial \boldsymbol{y}(s)}{\partial \theta_k}$$
$$+ \delta_{k,i} \boldsymbol{f}^i\big(\boldsymbol{y}(s), \boldsymbol{\zeta}\big), \quad s \in (i-1, i), \quad i = 1, \ldots, m. \quad (29)$$

Now, differentiating (9) with respect to θ_k gives

$$\frac{\partial \boldsymbol{y}(i)}{\partial \theta_k} = \frac{\partial \boldsymbol{y}(i^+)}{\partial \theta_k} = \begin{cases} \boldsymbol{0}, & \text{if } i = 0, \\ \dfrac{\partial \boldsymbol{h}^i\big(\boldsymbol{y}(i^-), \boldsymbol{\zeta}\big)}{\partial \boldsymbol{x}} \dfrac{\partial \boldsymbol{y}(i^-)}{\partial \theta_k}, & \text{if } i = 1, \ldots, m. \end{cases} \quad (30)$$

Equations (29) and (30) show that $\partial \boldsymbol{y}/\partial \theta_k$ is the unique solution of (20)-(21). Thus, (24) must hold.

We can prove equation (25) in a similar way. First, differentiating (27) with respect to ζ_j gives

$$\frac{\partial \boldsymbol{y}(s)}{\partial \zeta_j} = \frac{\partial \boldsymbol{y}(i-1)}{\partial \zeta_j} + \int_{i-1}^{s} \theta_i \frac{\partial \boldsymbol{f}^i\big(\boldsymbol{y}(\eta), \boldsymbol{\zeta}\big)}{\partial \boldsymbol{x}} \frac{\partial \boldsymbol{y}(\eta)}{\partial \zeta_j} d\eta$$
$$+ \int_{i-1}^{s} \theta_i \frac{\partial \boldsymbol{f}^i\big(\boldsymbol{y}(\eta), \boldsymbol{\zeta}\big)}{\partial \zeta_j} d\eta, \quad s \in (i-1, i), \quad i = 1, \ldots, m.$$

Hence,

$$\frac{d}{ds}\left\{ \frac{\partial \boldsymbol{y}(s)}{\partial \zeta_j} \right\} = \theta_i \frac{\partial \boldsymbol{f}^i\big(\boldsymbol{y}(s), \boldsymbol{\zeta}\big)}{\partial \boldsymbol{x}} \frac{\partial \boldsymbol{y}(s)}{\partial \zeta_j} + \theta_i \frac{\partial \boldsymbol{f}^i\big(\boldsymbol{y}(s), \boldsymbol{\zeta}\big)}{\partial \zeta_j},$$
$$s \in (i-1, i), \quad i = 1, \ldots, m. \quad (31)$$

Furthermore, differentiating (9) with respect to ζ_j gives

$$\frac{\partial \boldsymbol{y}(i)}{\partial \zeta_j} = \begin{cases} \boldsymbol{0}, & \text{if } i = 0, \\ \dfrac{\partial \boldsymbol{h}^i\big(\boldsymbol{y}(i^-), \boldsymbol{\zeta}\big)}{\partial \boldsymbol{x}} \dfrac{\partial \boldsymbol{y}(i^-)}{\partial \zeta_j} + \dfrac{\partial \boldsymbol{h}^i\big(\boldsymbol{y}(i^-), \boldsymbol{\zeta}\big)}{\partial \zeta_j}, & \text{if } i = 1, \dots, m. \end{cases} \tag{32}$$

Equations (31) and (32) show that $\partial \boldsymbol{y}/\partial \zeta_j$ is the solution of (22)-(23), which proves (25). □

We now derive formulae for the partial derivatives of $\tilde{J}$. By Theorem 4 and the chain rule,

$$\begin{aligned} \frac{\partial \tilde{J}(\boldsymbol{\theta}, \boldsymbol{\zeta})}{\partial \theta_k} &= \sum_{i=1}^{m} \Bigg\{ \frac{\partial \Psi_i\big(\boldsymbol{\zeta}, \boldsymbol{y}(i^-), \boldsymbol{y}(i^+)\big)}{\partial \boldsymbol{x}^-} \frac{\partial \boldsymbol{y}(i^-)}{\partial \theta_k} \\ &\qquad\qquad + \frac{\partial \Psi_i\big(\boldsymbol{\zeta}, \boldsymbol{y}(i^-), \boldsymbol{y}(i^+)\big)}{\partial \boldsymbol{x}^+} \frac{\partial \boldsymbol{y}(i^+)}{\partial \theta_k} \Bigg\} \\ &= \sum_{i=1}^{m} \Bigg\{ \frac{\partial \Psi_i\big(\boldsymbol{\zeta}, \boldsymbol{y}(i^-), \boldsymbol{y}(i^+)\big)}{\partial \boldsymbol{x}^-} \boldsymbol{\psi}^k(i^-) \\ &\qquad\qquad + \frac{\partial \Psi_i\big(\boldsymbol{\zeta}, \boldsymbol{y}(i^-), \boldsymbol{y}(i^+)\big)}{\partial \boldsymbol{x}^+} \boldsymbol{\psi}^k(i^+) \Bigg\}, \end{aligned} \tag{33}$$

where $\boldsymbol{y} \triangleq \boldsymbol{y}(\cdot|\boldsymbol{\theta}, \boldsymbol{\zeta})$, $\boldsymbol{\psi}^k \triangleq \boldsymbol{\psi}^k(\cdot|\boldsymbol{\theta}, \boldsymbol{\zeta})$, and $\partial/\partial \boldsymbol{x}^-$ and $\partial/\partial \boldsymbol{x}^+$ denote differentiation with respect to the state before and after the ith switch, respectively.

Similarly,

$$\begin{aligned} \frac{\partial \tilde{J}(\boldsymbol{\theta}, \boldsymbol{\zeta})}{\partial \zeta_j} &= \sum_{i=1}^{m} \Bigg\{ \frac{\partial \Psi_i\big(\boldsymbol{\zeta}, \boldsymbol{y}(i^-), \boldsymbol{y}(i^+)\big)}{\partial \zeta_j} + \frac{\partial \Psi_i\big(\boldsymbol{\zeta}, \boldsymbol{y}(i^-), \boldsymbol{y}(i^+)\big)}{\partial \boldsymbol{x}^-} \frac{\partial \boldsymbol{y}(i^-)}{\partial \zeta_j} \\ &\qquad\qquad + \frac{\partial \Psi_i\big(\boldsymbol{\zeta}, \boldsymbol{y}(i^-), \boldsymbol{y}(i^+)\big)}{\partial \boldsymbol{x}^+} \frac{\partial \boldsymbol{y}(i^+)}{\partial \zeta_j} \Bigg\} \\ &= \sum_{i=1}^{m} \Bigg\{ \frac{\partial \Psi_i\big(\boldsymbol{\zeta}, \boldsymbol{y}(i^-), \boldsymbol{y}(i^+)\big)}{\partial \zeta_j} + \frac{\partial \Psi_i\big(\boldsymbol{\zeta}, \boldsymbol{y}(i^-), \boldsymbol{y}(i^+)\big)}{\partial \boldsymbol{x}^-} \boldsymbol{\phi}^j(i^-) \\ &\qquad\qquad + \frac{\partial \Psi_i\big(\boldsymbol{\zeta}, \boldsymbol{y}(i^-), \boldsymbol{y}(i^+)\big)}{\partial \boldsymbol{x}^+} \boldsymbol{\phi}^j(i^+) \Bigg\}. \end{aligned} \tag{34}$$

The following algorithm for computing $\tilde{J}$ and its partial derivatives is based on equations (33) and (34).

Algorithm 1. Input a pair $(\boldsymbol{\theta}, \boldsymbol{\zeta}) \in \Theta \times \mathcal{W}$.

(i) Solve the impulsive systems (8)-(9), (20)-(21), and (22)-(23) to obtain $\boldsymbol{y}(\cdot|\boldsymbol{\theta}, \boldsymbol{\zeta})$, $\boldsymbol{\psi}^k(\cdot|\boldsymbol{\theta}, \boldsymbol{\zeta})$, and $\boldsymbol{\phi}^j(\cdot|\boldsymbol{\theta}, \boldsymbol{\zeta})$.

(ii) Use $\boldsymbol{y}(\cdot|\boldsymbol{\theta}, \boldsymbol{\zeta})$ to compute $\tilde{J}$.

(iii) Use $\boldsymbol{y}(\cdot|\boldsymbol{\theta}, \boldsymbol{\zeta})$, $\boldsymbol{\psi}^k(\cdot|\boldsymbol{\theta}, \boldsymbol{\zeta})$, and $\boldsymbol{\phi}^j(\cdot|\boldsymbol{\theta}, \boldsymbol{\zeta})$ to compute $\partial \tilde{J}(\boldsymbol{\theta}, \boldsymbol{\zeta})/\partial \theta_k$ and $\partial \tilde{J}(\boldsymbol{\theta}, \boldsymbol{\zeta})/\partial \zeta_j$ according to equations (33) and (34).

5   Examples

In this section, we consider two example problems. To solve these examples, we wrote a Fortran program that implements the approach described in Sections 3 and 4. This program transforms Problem P into Problem $\tilde{P}$ and then solves Problem $\tilde{P}$ by combining Algorithm 1 with the optimization subroutine NLPQLP [18] and the differential equation solver LSODA [5].

Example 5.1: A Nonlinear Impulsive System

Consider the following impulsive system:

$$
\dot{x}_1 = \begin{cases} 0.01x_1^2 + 2.02x_1x_2 - 0.99x_2^2 - 2x_1 + 4x_2 + 1, & \text{if } 0 < t < 1.8, \\ 1.01x_1^2 + 0.02x_1x_2 + 0.01x_2^2 - 2x_1 + 4x_2 + 1, & \text{if } 1.8 < t < 2, \end{cases} \tag{35a}
$$

$$
\dot{x}_2 = \begin{cases} \begin{aligned} 0.01x_1x_2 + 1.01x_2^2 + 1.01x_1x_3 - 0.99x_2x_3 \\ -3x_1 - x_2 + 2x_3 + 1, \end{aligned} & \text{if } 0 < t < 1.8, \\ \begin{aligned} 1.01x_1x_2 + 0.01x_2^2 + 0.01x_1x_3 + 0.01x_2x_3 \\ -3x_1 - x_2 + 2x_3 + 1, \end{aligned} & \text{if } 1.8 < t < 2, \end{cases} \tag{35b}
$$

$$
\dot{x}_3 = \begin{cases} 0.01x_2^2 + 2.02x_2x_3 - 0.99x_3^2 - 6x_2 + 1, & \text{if } 0 < t < 1.8, \\ 1.01x_2^2 + 0.02x_2x_3 + 0.01x_3^2 - 6x_2 + 1, & \text{if } 1.8 < t < 2, \end{cases} \tag{35c}
$$

and

$$
x_1(0) = 0.1, \tag{36a}
$$
$$
x_2(0) = 0, \tag{36b}
$$
$$
x_3(0) = 25. \tag{36c}
$$

Suppose that there are $m - 1$ switching times $\tau_1, \ldots, \tau_{m-1}$ satisfying

$$
0 = \tau_0 < \tau_1 < \cdots < \tau_{m-1} = 1.8.
$$

We impose the following state jump conditions:

$$
x_1(\tau_i) = \begin{cases} \dfrac{4x_1(\tau_i^-) + x_1(\tau_i^-)x_3(\tau_i^-) - x_2(\tau_i^-)^2}{4x_1(\tau_i^-) - 4x_2(\tau_i^-) + x_3(\tau_i^-) + 4}, & i = 1, \ldots, m-2, \\ x_1(\tau_i^-), & i = m-1, \end{cases} \tag{37a}
$$

$$
x_2(\tau_i) = \begin{cases} \dfrac{4x_2(\tau_i^-) + 2x_1(\tau_i^-)x_3(\tau_i^-) - 2x_2(\tau_i^-)^2}{4x_1(\tau_i^-) - 4x_2(\tau_i^-) + x_3(\tau_i^-) + 4}, & i = 1, \ldots, m-2, \\ x_2(\tau_i^-), & i = m-1, \end{cases} \tag{37b}
$$

$$
x_3(\tau_i) = \begin{cases} \dfrac{4x_3(\tau_i^-) + x_1(\tau_i^-)x_3(\tau_i^-) - x_2(\tau_i^-)^2}{4x_1(\tau_i^-) - 4x_2(\tau_i^-) + x_3(\tau_i^-) + 4}, & i = 1, \ldots, m-2, \\ x_3(\tau_i^-), & i = m-1. \end{cases} \tag{37c}
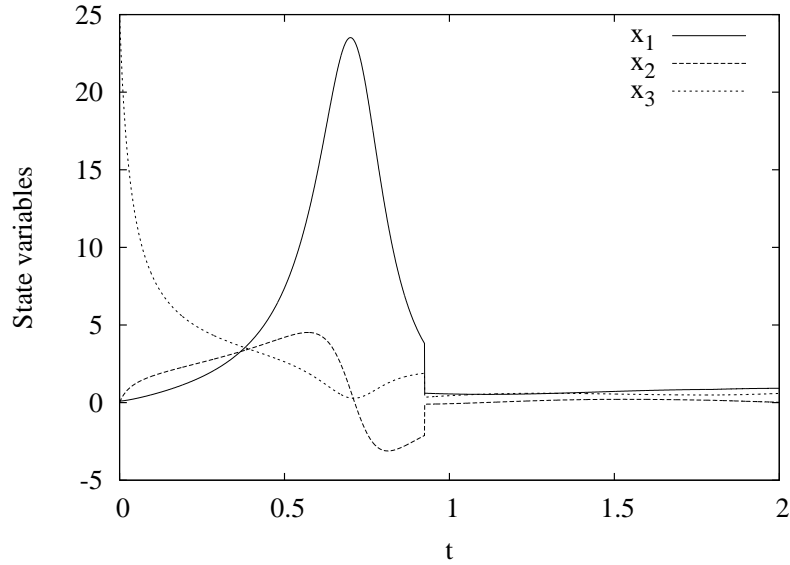$$

Figure 1: Optimal state trajectories for Example 5.1 with $m = 3$.

Furthermore, we assume

$$\tau_i - \tau_{i-1} \geq 0.1, \qquad i = 1, \ldots, m - 1. \tag{38}$$

The problem is to choose the switching times $\tau_1, \ldots, \tau_{m-2}$ to minimize the cost function

$$J = x_1(2)^2 + 2x_2(2)^2 + x_3(2)^2$$

subject to the dynamics (35)-(36), the state jump conditions (37), and the constraints (38). This problem (for $m = 4$) was solved in [12] using the optimal control software MISER 3.2 [9]. MISER also uses NLPQLP; however, MISER computes the cost function's gradient by solving a costate system backwards in time. This is quite different to our new method, which only involves forward integration.

We solved this problem for $m = 3$ and $m = 4$. When $m = 3$, the optimal solution is

$$\tau_1^* = 0.9252, \qquad J^* = 1.2040.$$

When $m = 4$, the optimal solution is

$$\tau_1^* = 1.0972, \qquad \tau_2^* = 1.7000, \qquad J^* = 0.6844.$$

The optimal state trajectories for $m = 3$ and $m = 4$ are shown in Figures 1 and 2, respectively.
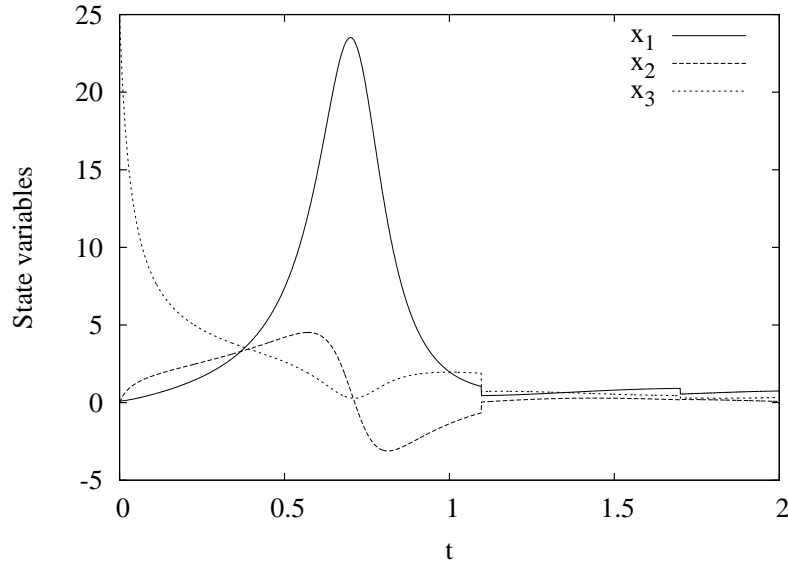
Figure 2: Optimal state trajectories for Example 5.1 with $m = 4$.

Note that our results are superior to those given in [12] (for $m = 4$). Indeed, the optimal solution reported in [12] is

$$\tau_1^* = 0.8141, \qquad \tau_2^* = 0.9634, \qquad J^* = 1.1041.$$

Example 5.2: Optimal Shrimp Harvesting

In [22], the following model for shrimp population growth is discussed:

$$\dot{x}_1(t) = -0.03x_1(t), \tag{39a}$$
$$\dot{x}_2(t) = 3.5 - 0.00001x_1(t)x_2(t), \tag{39b}$$

and

$$x_1(0) = 40000, \tag{40a}$$
$$x_2(0) = 1, \tag{40b}$$

where t is the time in weeks, $x_1(t)$ is the number of shrimp at time t, and $x_2(t)$ is the average weight of shrimp (in grams) at time t.

Shrimp are harvested at times $t = \tau_i$, $i = 1, \ldots, m$. The number of shrimp drops instantaneously at the harvest times, but the average weight doesn't change. Hence, we have the following state jump conditions:

$$x_1(\tau_i^+) = x_1(\tau_i^-) - \zeta_i x_1(\tau_i^-), \tag{41a}$$
$$x_2(\tau_i^+) = x_2(\tau_i^-), \tag{41b}$$
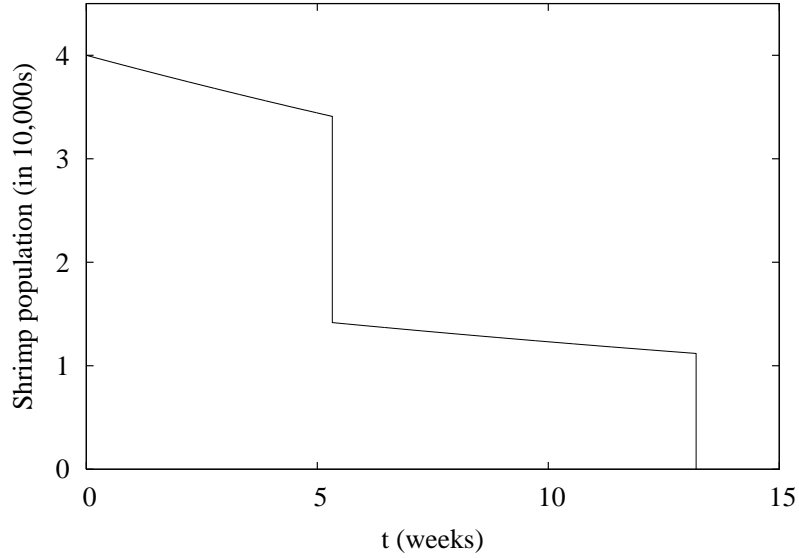
Figure 3: Optimal shrimp population for $m = 2$.

where ζ_i denotes the fraction of total shrimp harvested at time $t = \tau_i$.

At least 1% of the total shrimp stock must be harvested at each harvesting time. Hence,

$$0.01 \leq \zeta_i \leq 1, \qquad i = 1, \ldots, m. \tag{42}$$

The revenue obtained by harvesting a fraction ζ_i of the total shrimp stock at time $t = \tau_i$ is

$$px_2(\tau_i^-)\zeta_i x_1(\tau_i^-) - h,$$

where $p \triangleq \$0.008$ is the price per gram of shrimp and $h \triangleq \$50$ is the fixed cost of harvesting.

We assume that the final harvesting time is $\tau_m = T = 13.2$. At this time, all remaining shrimp are harvested. Hence, $\zeta_m = 1$. We also assume that

$$\tau_i - \tau_{i-1} \geq 0.01, \qquad i = 1, \ldots, m. \tag{43}$$

The problem is to find $\tau_1, \ldots, \tau_{m-1}$ and $\zeta_1, \ldots, \zeta_{m-1}$ to maximize the revenue function

$$J = \sum_{i=1}^{m} \left(px_2(\tau_i^-)\zeta_i x_1(\tau_i^-) - h \right)$$

subject to the dynamics (39)-(40), the state jump conditions (41), and the constraints (42) and (43).

The optimal solution for this problem with $m = 2$ is

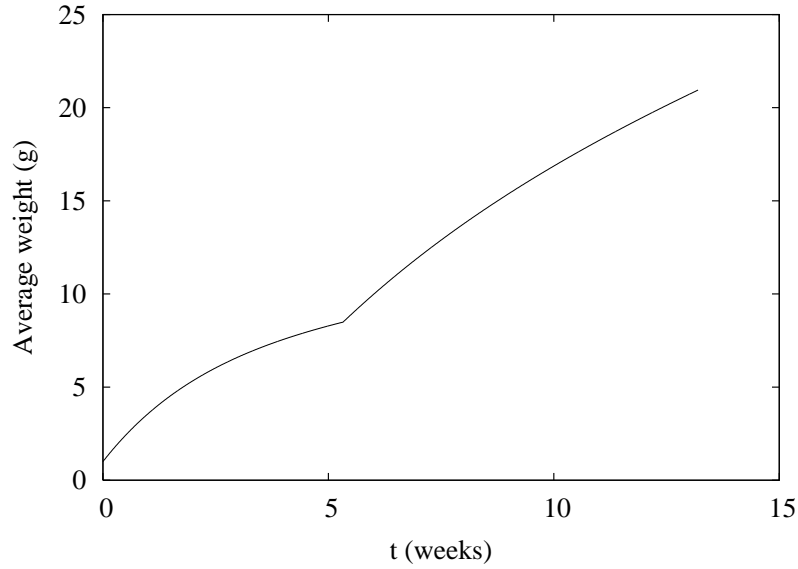$$\tau_1^* = 5.330, \qquad \zeta_1^* = 0.584, \qquad J^* = 3128.$$

Figure 4: Optimal average weight for $m = 2$.

This means that we can obtain a maximum revenue of \$3128 by harvesting 58.4% of the shrimp stock at time $t = 5.33$, and harvesting the remaining shrimp stock at the final time $t = 13.2$. The state variables corresponding to this optimal solution are plotted in Figures 3 and 4.

When $m = 3$, the optimal solution is

$$\tau_1^* = 4.270, \quad \tau_2^* = 7.810, \quad \zeta_1^* = 0.388, \quad \zeta_2^* = 0.454, \quad J^* = 3189.$$

The optimal state variables corresponding to this solution are shown in Figures 5 and 6. Note that the optimal revenue for $m = 3$ is greater than the optimal revenue for $m = 2$. Hence, harvesting three times is more profitable than harvesting two times.
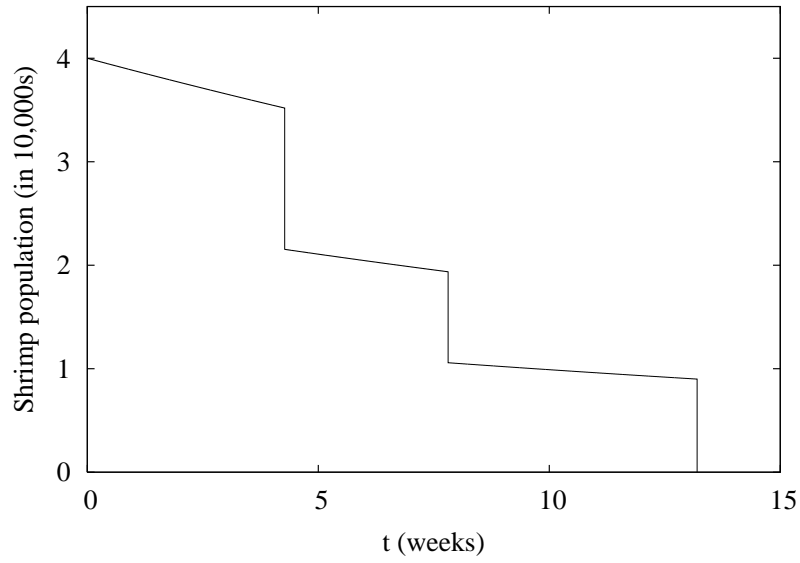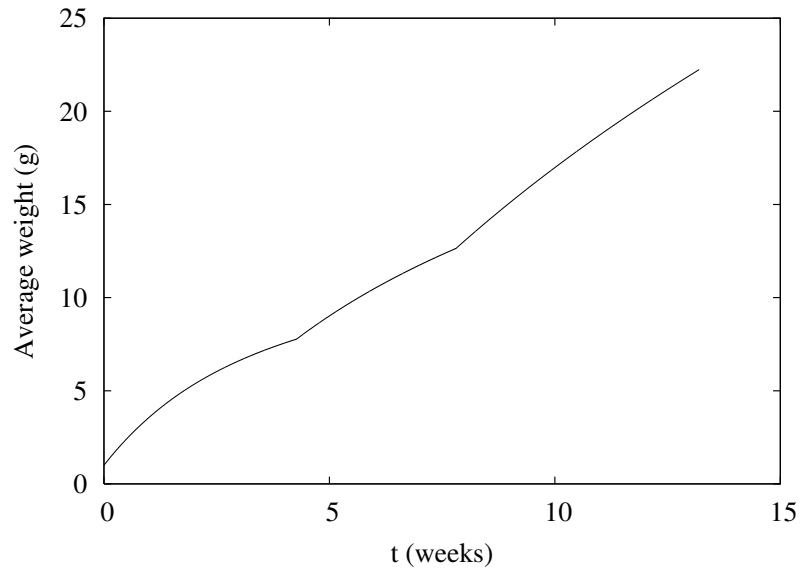
The optimal solution for $m = 4$ is

$$\tau_1^* = 3.854, \qquad \tau_2^* = 6.120, \qquad \tau_3^* = 9.110,$$

$$\zeta_1^* = 0.289, \qquad \zeta_2^* = 0.323, \qquad \zeta_3^* = 0.374,$$

with optimal revenue

$$J^* = 3172.$$

Hence, harvesting four times results in a decrease in maximum revenue.

Figure 5: Optimal shrimp population for $m = 3$.

Figure 6: Optimal average weight for $m = 3$.

6   References

[1] N. U. Ahmed, Elements of Finite-Dimensional Systems and Control Theory, Essex: Longman Scientific and Technical, 1988.

[2] M. Boccadoro, Y. Wardi, M. Egerstedt, and E. Verriest, Optimal control of switching surfaces in hybrid dynamical systems, *Discrete Event Dynamic Systems: Theory and Applications*, **15**, (2005) 433-448.

[3] C. Gao, E. Feng, Z. Wang, and Z. Xiu, Nonlinear dynamical systems of biodissimilation of glycerol to 1,3-Propanediol and their optimal controls, *Journal of Industrial and Management Optimization*, **1**, (2005) 377-388.

[4] C. Gao, K. Li, E. Feng, and Z. Xiu, Nonlinear impulsive system of fed-batch culture in fermentative production and its properties, *Chaos, Solitons and Fractals*, **28**, (2006) 271-277.

[5] A. C. Hindmarsh, Large ordinary differential equation systems and software, *IEEE Control Systems Magazine*, **2**, (1982) 24-30.

[6] C. Y. F. Ho, B. W. K. Ling, Y. Q. Liu, P. K. S. Tam, and K. L. Teo, Optimal PWM control of switched-capacitor DC-DC power converters via model transformation and enhancing control techniques, *IEEE Transactions on Circuits and Systems—I: Regular Papers*, **55**, (2008) 1382-1391.

[7] P. Howlett, Optimal strategies for the control of a train, *Automatica*, **32**, (1996) 519-532.

[8] P. Howlett, The optimal control of a train, *Annals of Operations Research*, **98**, (2000) 65-87.

[9] L. S. Jennings, M. E. Fisher, K. L. Teo, and C. J. Goh, MISER 3: Solving optimal control problems—an update, *Advances in Engineering Software and Workstations*, **13**, (1991) 190-196.

[10] H. W. J. Lee, K. L. Teo, V. Rehbock, and L. S. Jennings, Control parameterization enhancing technique for time optimal control problems, *Dynamic Systems and Applications*, **6**, (1997) 243-262.

[11] R. Li, K. L. Teo, K. H. Wong, and G. R. Duan, Control parameterization enhancing transform for optimal control of switched systems, *Mathematical and Computer Modelling*, **43**, (2006) 1393-1403.

[12] Y. Liu, K. L. Teo, L. S. Jennings, and S. Wang, On a class of optimal control problems with state jumps, *Journal of Optimization Theory and Applications*, **98**, (1998) 65-82.

[13] R. Loxton, K. L. Teo, and V. Rehbock, Computational method for a class of switched system optimal control problems, *IEEE Transactions on Automatic Control*, **54**, (2009) 2455-2460.

[14] R. Loxton, K. L. Teo, V. Rehbock, and W. K. Ling, Optimal switching instants for a switched-capacitor DC/DC power converter, *Automatica*, **45**, (2009) 973-980.

[15] D. G. Luenberger, Linear and Nonlinear Programming, second edition, New York: Springer, 2006.

[16] J. Nocedal and S. R. Wright, Numerical Optimization, New York: Springer, 1999.

[17] T. Ruby and V. Rehbock, Numerical solutions of optimal switching control problems, in Optimization and Control with Applications, L. Qi, K. L. Teo, and X. Yang, Eds. New York: Springer, 2005.

[18] K. Schittkowski, NLPQLP: A Fortran Implementation of a Sequential Quadratic Programming Algorithm with Distributed and Non-Monotone Line Search, version 2.24, user's guide, Bayreuth: University of Bayreuth, 2007.

[19] S. F. Woon, Global algorithms for nonlinear discrete optimization and discrete-valued optimal control problems, Ph.D. dissertation, Curtin University of Technology, Perth, December 2009.

[20] C. Z. Wu and K. L. Teo, Global impulsive optimal control computation, *Journal of Industrial and Management Optimization*, **2**, (2006) 435-450.

[21] X. Xu and P. J. Antsaklis, Optimal control of switched systems based on parameterization of the switching instants, *IEEE Transactions on Automatic Control*, **49**, (2004) 2-16.

[22] R. Yu and P. Leung, Optimal partial harvesting schedule for aquaculture operations, *Marine Resource Economics*, **21**, (2006) 301-315.