Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 7, No. 1, pp. 63–81, January 2011

A NEW COMPUTATIONAL METHOD FOR A CLASS OF FREE TERMINAL TIME OPTIMAL CONTROL PROBLEMS

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

Abstract: We develop a numerical method for solving an optimal control problem whose terminal time is not fixed, but is instead determined by a state-dependent stopping criterion. The main idea of this method is to approximate the control by a piecewise constant function whose values and switching times are decision variables to be determined optimally. The optimal control problem then becomes an optimization problem with a finite number of decision variables. We develop a novel method for computing the gradient of the cost function in this approximate problem. On this basis, the approximate problem can be solved using any gradient-based optimization technique. We use this approach to solve an aeronautical control problem involving a gliding projectile. We also prove several important convergence results that justify our approximation scheme.

Key words: *nonlinear optimal control, control parameterization, gradient-based optimization, nonlinear programming*

Mathematics Subject Classification: *49J15, 90C30*

1 Introduction

In this paper, we consider an optimal control problem for a system of ordinary differential equations whose terminal time is determined by a stopping criterion. This stopping criterion is defined by a smooth surface in the state space; when the state trajectory hits this surface, the system stops. Hence, changing the input to the system not only influences its state trajectory, but also changes the time horizon over which it evolves. The problem is to choose the control input function in an optimal manner.

This type of optimal control problem was introduced in [8], where a control policy was sought for a gliding projectile launched from an aircraft. The control function in this context is the glider's angle of attack; the state represents its horizontal and vertical coordinates. The problem is to vary the glider's angle of attack during flight so that the glider covers as much ground as possible before crashing. In [8], a numerical method is discussed for solving this problem. Another numerical method is discussed in [9]. Both of these methods are based on control parameterization, whereby the control is approximated by a piecewise constant function with pre-fixed switching times. Under this approximation scheme, the original optimal control problem is reduced to an approximate optimization problem with a finite number of decision variables. The approximate problem can then be solved using a standard optimization method, such as a quasi-Newton method [5, 6].

The numerical methods discussed in [8, 9] have several disadvantages. First, the switching times for the approximate controls are pre-specified and cannot be varied adaptively during the optimization process. Thus, a very fine discretization of the time horizon may be necessary to ensure accurate results. Second, the methods for computing the gradient of the cost function involve integrating two systems of differential equations—the state system and the so-called *costate system*—successively in different directions. Since the state and costate systems are integrated in opposite directions, it is impossible to ensure that their knot sets coincide (unless a crude integration technique with fixed step lengths is used). This is a major problem, because the costate system actually depends on the solution of the state system. Therefore, the state needs to be interpolated when the costate system is being solved, which compromises accuracy.

Furthermore, the numerical method proposed in [9] involves not one, but two successive approximations of the original optimal control problem. In fact, in this method, for every primary approximate problem a sequence of secondary approximate problems needs to be solved. The method in [9] is therefore very complex and intensive.

In this paper, we propose a new computational method for solving the optimal control problem formulated in [8, 9]. This method, like those in [8, 9], is based on control parameterization. The major difference is that we use a more flexible piecewise constant approximation of the control. More specifically, we allow *both* the control heights and the control switching times to be decision variables. Hence, the control switching times do not need to be specified beforehand and are instead determined optimally. Furthermore, inspired by work in [3, 4, 11], we develop a new scheme for computing the cost function's gradient. Our new scheme involves integrating an auxiliary dynamical system forward in time, simultaneously with the state system. Accordingly, state interpolation is not required. This makes our new method much easier to implement than those in [8, 9].

2 Problem Statement

Consider the following nonlinear control system:

$$\dot{\mathbf{x}}(t) = \mathbf{f}\big(\mathbf{x}(t), \mathbf{u}(t)\big), \qquad t \geq 0, \tag{2.1}$$

and

$$\mathbf{x}(0) = \mathbf{x}^0, \tag{2.2}$$

where $\mathbf{x}(t) \in \mathbb{R}^n$ is the system state at time t; $\mathbf{u}(t) \in \mathbb{R}^r$ is the control input at time t; $\mathbf{f} : \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^n$ is a given function; and $\mathbf{x}^0 \in \mathbb{R}^n$ is a given initial state.

Define

$$\Upsilon := \big\{\, \mathbf{u} = [u_1, \ldots, u_r]^T \in \mathbb{R}^r : \; a_\varsigma \leq u_\varsigma \leq b_\varsigma, \; \varsigma = 1, \ldots, r \,\big\},$$

where a_ς and b_ς are given real numbers such that $a_\varsigma < b_\varsigma$. Any measurable function $\mathbf{u} : [0, \infty) \to \mathbb{R}^r$ such that $\mathbf{u}(t) \in \Upsilon$ for almost all $t \in [0, \infty)$ is called an *admissible control.* Let $\mathcal{U}$ denote the class of all such admissible controls.

We assume that the following two conditions are satisfied.

Assumption 2.1. The function $\mathbf{f}$ is continuously differentiable.

Assumption 2.2. There exists a real number $K > 0$ such that

$$\big\|\mathbf{f}(\mathbf{x}, \mathbf{u})\big\| \leq K(1 + \|\mathbf{x}\|), \qquad (\mathbf{x}, \mathbf{u}) \in \mathbb{R}^n \times \Upsilon,$$

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

It follows from Theorem 3.1.6 of [1] that the system (2.1)-(2.2) has a unique solution corresponding to each admissible control $\mathbf{u} \in \mathcal{U}$. We denote this solution by $\mathbf{x}(\cdot|\mathbf{u})$. The function $\mathbf{x}(\cdot|\mathbf{u})$ is absolutely continuous, satisfies the dynamics (2.1) almost everywhere on $[0, \infty)$, and satisfies the initial condition (2.2).

Define a functional $T : \mathcal{U} \to [0, \infty)$ as follows:

$$T(\mathbf{u}) := \inf \big\{ t \in (0, \infty) : \ \Phi(\mathbf{x}(t|\mathbf{u})) = 0 \big\},$$

where $\Phi : \mathbb{R}^n \to \mathbb{R}$ is a given continuously differentiable function such that $\Phi(\mathbf{x}^0) > 0$. Clearly,

$$\Phi\big(\mathbf{x}(T(\mathbf{u})|\mathbf{u})\big) = 0. \tag{2.3}$$

Furthermore, $T(\mathbf{u})$ is the first positive time at which the state trajectory hits the surface

$$\big\{ \mathbf{x} \in \mathbb{R}^n : \ \Phi(\mathbf{x}) = 0 \big\}.$$

We assume that the control system (2.1)-(2.2) stops when $t = T(\mathbf{u})$. Hence, $T(\mathbf{u})$ is called the *stopping time* or *terminal time* corresponding to the admissible control $\mathbf{u} \in \mathcal{U}$.

We assume that the following condition is satisfied.

Assumption 2.3. There exists a real number $T_{\max} > 0$ such that

$$T_{\max} = \sup \big\{ T(\mathbf{u}) : \ \mathbf{u} \in \mathcal{U} \big\}.$$

We now define the following optimal control problem.

Problem P. Choose an admissible control $\mathbf{u} \in \mathcal{U}$ such that the cost functional

$$J(\mathbf{u}) := \Psi\big(\mathbf{x}(T(\mathbf{u})|\mathbf{u})\big), \tag{2.4}$$

where $\Psi : \mathbb{R}^n \to \mathbb{R}$ is a given continuously differentiable function, is minimized over $\mathcal{U}$.

Remark 2.4. We can easily incorporate an integral term of the form

$$\int_0^{T(\mathbf{u})} \mathcal{L}\big(\mathbf{x}(t|\mathbf{u}), \mathbf{u}(t)\big) dt, \tag{2.5}$$

where $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}$, into the cost functional (2.4). This is done by augmenting the dynamic system (2.1)-(2.2) with the following auxiliary dynamics:

$$\dot{v}(t) = \mathcal{L}\big(\mathbf{x}(t), \mathbf{u}(t)\big), \qquad t \geq 0,$$

and

$$v(0) = 0.$$

Clearly, the value of v at the stopping time is equal to the integral cost (2.5).

3 Control Parameterization

In general, Problem P is too complicated to solve analytically. We will instead approximate it by a finite-dimensional optimization problem.

Let $p \geq 2$ be a fixed integer and define corresponding sets Ξ and Θ as follows:

$$\Xi := \prod_{i=1}^{p} \Upsilon$$

and

$$\Theta := \big\{\boldsymbol{\theta} = [\theta_1, \ldots, \theta_{p-1}]^T \in \mathbb{R}^{p-1} : \ \theta_i \geq 0, \ i = 1, \ldots, p-1\big\}.$$

Notice that Ξ is the set of all tuples $(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^p)$ such that $\boldsymbol{\sigma}^i \in \Upsilon$, $i = 1, \ldots, p$.

Now, for each pair $(\boldsymbol{\sigma}, \boldsymbol{\theta}) \in \Xi \times \Theta$, define a corresponding control function $\mathbf{u}^p(\cdot|\boldsymbol{\sigma}, \boldsymbol{\theta}) : [0, \infty) \to \mathbb{R}^r$ as follows:

$$\mathbf{u}^p(t|\boldsymbol{\sigma}, \boldsymbol{\theta}) := \sum_{i=1}^{p} \boldsymbol{\sigma}^i \chi_{\mathcal{I}_i(\boldsymbol{\theta})}(t), \qquad t \in [0, \infty), \tag{3.1}$$

where $\mathcal{I}_i(\boldsymbol{\theta}) := [t_{i-1}(\boldsymbol{\theta}), t_i(\boldsymbol{\theta}))$,

$$t_i(\boldsymbol{\theta}) := \begin{cases} 0, & \text{if } i = 0, \\ \displaystyle\sum_{j=1}^{i} \theta_j, & \text{if } i \in \{1, \ldots, p-1\}, \\ \infty, & \text{if } i = p, \end{cases}$$

and $\chi_{\mathcal{I}} : \mathbb{R} \to \mathbb{R}$ is the indicator function defined by

$$\chi_{\mathcal{I}}(t) := \begin{cases} 1, & \text{if } t \in \mathcal{I}, \\ 0, & \text{otherwise.} \end{cases}$$

Clearly, for each $\boldsymbol{\theta} \in \Theta$,

$$t_{i-1}(\boldsymbol{\theta}) \leq t_i(\boldsymbol{\theta}), \qquad i = 1, \ldots, p.$$

Since the function $\mathbf{u}^p(\cdot|\boldsymbol{\sigma}, \boldsymbol{\theta})$ changes its value at $t_i(\boldsymbol{\theta})$, $i = 1, \ldots, p-1$, these times are called *switching times*.

We immediately see that for each $(\boldsymbol{\sigma}, \boldsymbol{\theta}) \in \Xi \times \Theta$, the piecewise constant function $\mathbf{u}^p(\cdot|\boldsymbol{\sigma}, \boldsymbol{\theta})$ is an admissible control for Problem P. Accordingly, we may define

$$\mathbf{x}^p(\cdot|\boldsymbol{\sigma}, \boldsymbol{\theta}) := \mathbf{x}(\cdot|\mathbf{u}^p(\cdot|\boldsymbol{\sigma}, \boldsymbol{\theta})),$$

$$T^p(\boldsymbol{\sigma}, \boldsymbol{\theta}) := T(\mathbf{u}^p(\cdot|\boldsymbol{\sigma}, \boldsymbol{\theta})),$$

and

$$J^p(\boldsymbol{\sigma}, \boldsymbol{\theta}) := J(\mathbf{u}^p(\cdot|\boldsymbol{\sigma}, \boldsymbol{\theta})) = \Psi\big(\mathbf{x}^p(T^p(\boldsymbol{\sigma}, \boldsymbol{\theta})|\boldsymbol{\sigma}, \boldsymbol{\theta})\big).$$

Thus, when the controls are restricted to those described by equation (3.1), Problem P becomes the following optimization problem.

Problem P(p). Choose a pair $(\boldsymbol{\sigma}, \boldsymbol{\theta}) \in \Xi \times \Theta$ to minimize the objective function J^p over $\Xi \times \Theta$.

Remark 3.1. If $(\boldsymbol{\sigma}^*, \boldsymbol{\theta}^*)$ is an optimal solution of Problem P(p), then $\mathbf{u}^p(\cdot|\boldsymbol{\sigma}^*, \boldsymbol{\theta}^*)$ is a suboptimal control for Problem P.

4 Solving Problem P(p)

Problem P(p) is a nonlinear optimization problem whose decision variables are the components of $\boldsymbol{\sigma}$ and $\boldsymbol{\theta}$. To solve this problem using a gradient-based optimization technique, we need a method for computing the gradient of the cost function J^p. The purpose of this section is to develop such a method.

Consider Problem P(p) for a fixed integer $p \geq 2$. Define

$$\Gamma := \left\{ (\boldsymbol{\sigma}, \boldsymbol{\theta}) \in \Xi \times \Theta : \ t_{p-1}(\boldsymbol{\theta}) < T^p(\boldsymbol{\sigma}, \boldsymbol{\theta}) \right\}.$$

Thus, $(\boldsymbol{\sigma}, \boldsymbol{\theta}) \in \Gamma$ if and only if the corresponding control $\mathbf{u}^p(\cdot|\boldsymbol{\sigma}, \boldsymbol{\theta})$ changes value $p-1$ times *before* the end of the time horizon.

We now prove the following important result.

Theorem 4.1. *Let $(\boldsymbol{\sigma}, \boldsymbol{\theta}) \in \Xi \times \Theta$ and suppose that $T^p(\boldsymbol{\sigma}, \boldsymbol{\theta}) > 0$. Then there exists a corresponding $(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\theta}}) \in \Gamma$ such that*

$$J^p(\boldsymbol{\sigma}, \boldsymbol{\theta}) = J^p(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\theta}}).$$

Proof. Let $\boldsymbol{\sigma} \in \Xi$ and $\boldsymbol{\theta} \in \Theta$ be arbitrary but fixed. Since the result follows immediately when $(\boldsymbol{\sigma}, \boldsymbol{\theta}) \in \Gamma$, we assume that $(\boldsymbol{\sigma}, \boldsymbol{\theta}) \notin \Gamma$. Thus, since $T^p(\boldsymbol{\sigma}, \boldsymbol{\theta})$ is strictly positive, there exists an integer $\kappa \in \{1, \ldots, p-1\}$ such that

$$t_{\kappa-1}(\boldsymbol{\theta}) < T^p(\boldsymbol{\sigma}, \boldsymbol{\theta}) \leq t_\kappa(\boldsymbol{\theta}).$$

We consider two cases: (i) $\kappa = 1$; and (ii) $\kappa \geq 2$.

We start with Case (i). In this case,

$$0 = t_0(\boldsymbol{\theta}) < T^p(\boldsymbol{\sigma}, \boldsymbol{\theta}) \leq t_1(\boldsymbol{\theta}).$$

Define vectors $\hat{\boldsymbol{\sigma}}^i \in \mathbb{R}^r$, $i = 1, \ldots, p$, and $\hat{\boldsymbol{\theta}} \in \mathbb{R}^{p-1}$ as follows:

$$\hat{\boldsymbol{\sigma}}^i := \boldsymbol{\sigma}^1, \qquad i = 1, \ldots, p,$$

and

$$\hat{\theta}_i := \frac{T^p(\boldsymbol{\sigma}, \boldsymbol{\theta})}{p}, \qquad i = 1, \ldots, p-1.$$

It is clear that $\hat{\boldsymbol{\sigma}} := (\hat{\boldsymbol{\sigma}}^1, \ldots, \hat{\boldsymbol{\sigma}}^p) \in \Xi$ and $\hat{\boldsymbol{\theta}} \in \Theta$. Moreover,

$$t_{p-1}(\hat{\boldsymbol{\theta}}) = \sum_{j=1}^{p-1} \hat{\theta}_j = \sum_{j=1}^{p-1} \frac{T^p(\boldsymbol{\sigma}, \boldsymbol{\theta})}{p} < T^p(\boldsymbol{\sigma}, \boldsymbol{\theta}). \tag{4.1}$$

Now, we immediately see that

$$\mathbf{u}^p(t|\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\theta}}) = \mathbf{u}^p(t|\boldsymbol{\sigma}, \boldsymbol{\theta}), \qquad t \in [0, T^p(\boldsymbol{\sigma}, \boldsymbol{\theta})).$$

Hence,

$$\mathbf{x}^p(t|\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\theta}}) = \mathbf{x}^p(t|\boldsymbol{\sigma}, \boldsymbol{\theta}), \qquad t \in [0, T^p(\boldsymbol{\sigma}, \boldsymbol{\theta})],$$

and

$$T^p(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\theta}}) = T^p(\boldsymbol{\sigma}, \boldsymbol{\theta}).$$

Substituting this into (4.1) shows that $(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\theta}}) \in \Gamma$. Furthermore, this implies that

$$J^p(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\theta}}) = J^p(\boldsymbol{\sigma}, \boldsymbol{\theta}).$$

Thus, $(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\theta}}) \in \Gamma$ is the required pair.

We now consider Case (ii), when $\kappa \geq 2$. Define vectors $\hat{\boldsymbol{\sigma}}^i \in \mathbb{R}^r$, $i = 1, \ldots, p$, and $\hat{\boldsymbol{\theta}} \in \mathbb{R}^{p-1}$ by

$$\hat{\boldsymbol{\sigma}}^i := \begin{cases} \boldsymbol{\sigma}^1, & \text{if } i = 1, \ldots, p-\kappa+1, \\ \boldsymbol{\sigma}^{i-p+\kappa}, & \text{if } i = p-\kappa+2, \ldots, p, \end{cases}$$

and

$$\hat{\theta}_i := \begin{cases} \theta_1/(p-\kappa+1), & \text{if } i = 1, \ldots, p-\kappa+1, \\ \theta_{i-p+\kappa}, & \text{if } i = p-\kappa+2, \ldots, p-1. \end{cases}$$

It is clear that $\hat{\boldsymbol{\sigma}} := (\hat{\boldsymbol{\sigma}}^1, \ldots, \hat{\boldsymbol{\sigma}}^p) \in \Xi$ and $\hat{\boldsymbol{\theta}} \in \Theta$. Furthermore, although the working is rather tedious, it is not too difficult to show that

$$\mathbf{u}^p(t|\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\theta}}) = \mathbf{u}^p(t|\boldsymbol{\sigma}, \boldsymbol{\theta}), \qquad t \in [0, T^p(\boldsymbol{\sigma}, \boldsymbol{\theta})).$$

Hence,

$$\mathbf{x}^p(t|\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\theta}}) = \mathbf{x}^p(t|\boldsymbol{\sigma}, \boldsymbol{\theta}), \qquad t \in [0, T^p(\boldsymbol{\sigma}, \boldsymbol{\theta})], \tag{4.2}$$

and

$$T^p(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\theta}}) = T^p(\boldsymbol{\sigma}, \boldsymbol{\theta}). \tag{4.3}$$

Equations (4.2) and (4.3) imply that

$$J^p(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\theta}}) = J^p(\boldsymbol{\sigma}, \boldsymbol{\theta}).$$

It remains to show that $(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\theta}}) \in \Gamma$. Indeed, we have

$$\begin{aligned} t_{p-1}(\hat{\boldsymbol{\theta}}) = \sum_{j=1}^{p-1} \hat{\theta}_j &= \sum_{j=1}^{p-\kappa+1} \frac{\theta_1}{p-\kappa+1} + \sum_{j=p-\kappa+2}^{p-1} \theta_{j-p+\kappa} \\ &= \sum_{j=1}^{\kappa-1} \theta_j = t_{\kappa-1}(\boldsymbol{\theta}) < T^p(\boldsymbol{\sigma}, \boldsymbol{\theta}) = T^p(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\theta}}). \end{aligned}$$

Thus, $(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\theta}}) \in \Gamma$, as required. □

Remark 4.2. The proof of Theorem 4.1 is constructive; it shows how to construct $(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\theta}})$ from $(\boldsymbol{\sigma}, \boldsymbol{\theta})$.

Now, for each $k = 1, \ldots, p$ and $\varsigma = 1, \ldots, r$, consider the following auxiliary dynamic system:

$$\dot{\boldsymbol{\phi}}^{k,\varsigma}(t) = \rho_{k,i} \frac{\partial \mathbf{f}\big(\mathbf{x}^p(t|\boldsymbol{\sigma}, \boldsymbol{\theta}), \boldsymbol{\sigma}^i\big)}{\partial \mathbf{x}} \boldsymbol{\phi}^{k,\varsigma}(t) + \delta_{k,i} \frac{\partial \mathbf{f}\big(\mathbf{x}^p(t|\boldsymbol{\sigma}, \boldsymbol{\theta}), \boldsymbol{\sigma}^i\big)}{\partial u_\varsigma}, \qquad t \in \mathcal{I}_i(\boldsymbol{\theta}), \quad i = 1, \ldots, p, \tag{4.4}$$

and

$$\boldsymbol{\phi}^{k,\varsigma}(0) = \mathbf{0}, \tag{4.5}$$

where $(\boldsymbol{\sigma}, \boldsymbol{\theta}) \in \Xi \times \Theta$ and

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

and

$$\rho_{k,i} := \begin{cases} 1, & \text{if } k \leq i, \\ 0, & \text{otherwise.} \end{cases}$$

Let $\boldsymbol{\phi}^{k,\varsigma}(\cdot|\boldsymbol{\sigma},\boldsymbol{\theta})$ denote the solution of (4.4)-(4.5) corresponding to $(\boldsymbol{\sigma},\boldsymbol{\theta}) \in \Xi \times \Theta$.

We now show that the partial derivatives of J^p with respect to the decision variables σ_ς^k, $k = 1,\dots,p$, $\varsigma = 1,\dots,r$, can be expressed in terms of $\boldsymbol{\phi}^{k,\varsigma}(\cdot|\boldsymbol{\sigma},\boldsymbol{\theta})$.

Theorem 4.3. *Let $(\boldsymbol{\sigma},\boldsymbol{\theta}) \in \Gamma$. Furthermore, let $T^p := T^p(\boldsymbol{\sigma},\boldsymbol{\theta})$, $\mathbf{x}^p := \mathbf{x}^p(\cdot|\boldsymbol{\sigma},\boldsymbol{\theta})$, and $\boldsymbol{\phi}^{k,\varsigma} := \boldsymbol{\phi}^{k,\varsigma}(\cdot|\boldsymbol{\sigma},\boldsymbol{\theta})$. Then for each $k = 1,\dots,p$ and $\varsigma = 1,\dots,r$,*

$$\frac{\partial J^p(\boldsymbol{\sigma},\boldsymbol{\theta})}{\partial \sigma_\varsigma^k} = \frac{\partial \Psi(\mathbf{x}^p(T^p))}{\partial \mathbf{x}} \boldsymbol{\phi}^{k,\varsigma}(T^p) + \alpha_{k,\varsigma} \frac{\partial \Psi(\mathbf{x}^p(T^p))}{\partial \mathbf{x}} \mathbf{f}\big(\mathbf{x}^p(T^p),\boldsymbol{\sigma}^p\big),$$

where

$$\alpha_{k,\varsigma} := -\frac{\partial \Phi(\mathbf{x}^p(T^p))}{\partial \mathbf{x}} \boldsymbol{\phi}^{k,\varsigma}(T^p) \left[\frac{\partial \Phi(\mathbf{x}^p(T^p))}{\partial \mathbf{x}} \mathbf{f}\big(\mathbf{x}^p(T^p),\boldsymbol{\sigma}^p\big) \right]^{-1}.$$

Proof. Let $(\boldsymbol{\sigma},\boldsymbol{\theta}) \in \Gamma$, $k \in \{1,\dots,p\}$, and $\varsigma \in \{1,\dots,r\}$ be arbitrary but fixed. For simplicity, we write t_i instead of $t_i(\boldsymbol{\theta})$ and $\mathcal{I}_i$ instead of $\mathcal{I}_i(\boldsymbol{\theta})$. Since $(\boldsymbol{\sigma},\boldsymbol{\theta})$ is fixed, these simplifications will not cause confusion.

Recall that

$$J^p(\boldsymbol{\sigma},\boldsymbol{\theta}) = \Psi\big(\mathbf{x}^p(T^p|\boldsymbol{\sigma},\boldsymbol{\theta})\big).$$

Differentiating this equation with respect to σ_ς^k gives

$$\frac{\partial J^p(\boldsymbol{\sigma},\boldsymbol{\theta})}{\partial \sigma_\varsigma^k} = \frac{\partial \Psi(\mathbf{x}^p(T^p))}{\partial \mathbf{x}} \frac{\partial \mathbf{x}^p(T^p)}{\partial \sigma_\varsigma^k} + \frac{\partial \Psi(\mathbf{x}^p(T^p))}{\partial \mathbf{x}} \mathbf{f}\big(\mathbf{x}(T^p),\boldsymbol{\sigma}^p\big) \frac{\partial T^p}{\partial \sigma_\varsigma^k}. \tag{4.6}$$

Now, for each $i = 1,\dots,p$, it follows from (2.1)-(2.2) that

$$\mathbf{x}^p(t) = \mathbf{x}^p(t_{i-1}) + \int_{t_{i-1}}^t \mathbf{f}\big(\mathbf{x}^p(s),\boldsymbol{\sigma}^i\big) ds, \qquad t \in \mathcal{I}_i. \tag{4.7}$$

If $i > k$, then differentiating (4.7) with respect to σ_ς^k yields

$$\frac{\partial \mathbf{x}^p(t)}{\partial \sigma_\varsigma^k} = \frac{\partial \mathbf{x}^p(t_{i-1})}{\partial \sigma_\varsigma^k} + \int_{t_{i-1}}^t \frac{\partial \mathbf{f}\big(\mathbf{x}^p(s),\boldsymbol{\sigma}^i\big)}{\partial \mathbf{x}} \frac{\partial \mathbf{x}^p(s)}{\partial \sigma_\varsigma^k} ds, \qquad t \in \mathcal{I}_i. \tag{4.8}$$

On the other hand, if $i = k$, then differentiating (4.7) with respect to σ_ς^k gives

$$\begin{aligned} \frac{\partial \mathbf{x}^p(t)}{\partial \sigma_\varsigma^k} = \frac{\partial \mathbf{x}(t_{i-1})}{\partial \sigma_\varsigma^k} &+ \int_{t_{i-1}}^t \frac{\partial \mathbf{f}\big(\mathbf{x}^p(s),\boldsymbol{\sigma}^i\big)}{\partial \mathbf{x}} \frac{\partial \mathbf{x}^p(s)}{\partial \sigma_\varsigma^k} ds \\ &+ \int_{t_{i-1}}^t \frac{\partial \mathbf{f}\big(\mathbf{x}^p(s),\boldsymbol{\sigma}^i\big)}{\partial u_\varsigma} ds, \quad t \in \mathcal{I}_i. \end{aligned} \tag{4.9}$$

Since $\boldsymbol{\sigma}^k$ is the value of $\mathbf{u}^p(\cdot|\boldsymbol{\sigma},\boldsymbol{\theta})$ on the subinterval $\mathcal{I}_k$, it does not affect the state before $\mathcal{I}_k$. Hence, if $i < k$, then

$$\frac{\partial \mathbf{x}^p(t)}{\partial \sigma_\varsigma^k} = \mathbf{0}, \qquad t \in \mathcal{I}_i. \tag{4.10}$$

We can combine (4.8)-(4.10) into one equation as follows:

$$\begin{aligned} \frac{\partial \mathbf{x}^p(t)}{\partial \sigma_\varsigma^k} = \rho_{k,i} \frac{\partial \mathbf{x}(t_{i-1})}{\partial \sigma_\varsigma^k} &+ \int_{t_{i-1}}^t \rho_{k,i} \frac{\partial \mathbf{f}\big(\mathbf{x}^p(s),\boldsymbol{\sigma}^i\big)}{\partial \mathbf{x}} \frac{\partial \mathbf{x}^p(s)}{\partial \sigma_\varsigma^k} ds \\ &+ \int_{t_{i-1}}^t \delta_{k,i} \frac{\partial \mathbf{f}\big(\mathbf{x}^p(s),\boldsymbol{\sigma}^i\big)}{\partial u_\varsigma} ds, \quad t \in \mathcal{I}_i, \quad i = 1,\dots,p. \end{aligned}$$

Differentiating this equation with respect to time gives

$$\frac{d}{dt}\left\{\frac{\partial \mathbf{x}^p(t)}{\partial \sigma_\varsigma^k}\right\} = \rho_{k,i}\frac{\partial \mathbf{f}\big(\mathbf{x}^p(t),\boldsymbol{\sigma}^i\big)}{\partial \mathbf{x}}\frac{\partial \mathbf{x}^p(t)}{\partial \sigma_\varsigma^k} + \delta_{k,i}\frac{\partial \mathbf{f}\big(\mathbf{x}^p(t),\boldsymbol{\sigma}^i\big)}{\partial u_\varsigma}, \qquad t \in \mathcal{I}_i, \quad i = 1,\ldots,p. \tag{4.11}$$

Furthermore,

$$\frac{\partial \mathbf{x}^p(0)}{\partial \sigma_\varsigma^k} = \frac{\partial}{\partial \sigma_\varsigma^k}\left\{\mathbf{x}^0\right\} = \mathbf{0}. \tag{4.12}$$

Equations (4.11)-(4.12) show that $\partial \mathbf{x}^p(\cdot)/\partial \sigma_\varsigma^k$ is the unique solution of (4.4)-(4.5). Hence,

$$\frac{\partial \mathbf{x}^p(t)}{\partial \sigma_\varsigma^k} = \boldsymbol{\phi}^{k,\varsigma}(t), \qquad t \in [0,\infty). \tag{4.13}$$

Now, recall from (2.3) that

$$\Phi(\mathbf{x}^p(T^p)) = 0.$$

Differentiating this equation with respect to σ_ς^k yields

$$\frac{\partial \Phi(\mathbf{x}^p(T^p))}{\partial \mathbf{x}}\frac{\partial \mathbf{x}^p(T^p)}{\partial \sigma_\varsigma^k} + \frac{\partial \Phi(\mathbf{x}^p(T^p))}{\partial \mathbf{x}}\mathbf{f}\big(\mathbf{x}^p(T^p),\boldsymbol{\sigma}^p\big)\frac{\partial T^p}{\partial \sigma_\varsigma^k} = 0.$$

Therefore, by using (4.13),

$$\frac{\partial T^p}{\partial \sigma_\varsigma^k} = -\frac{\partial \Phi(\mathbf{x}^p(T^p))}{\partial \mathbf{x}}\boldsymbol{\phi}^{k,\varsigma}(T^p)\left[\frac{\partial \Phi(\mathbf{x}^p(T^p))}{\partial \mathbf{x}}\mathbf{f}\big(\mathbf{x}^p(T^p),\boldsymbol{\sigma}^p\big)\right]^{-1} = \alpha_{k,\varsigma}. \tag{4.14}$$

Substituting equations (4.13) and (4.14) into equation (4.6) completes the proof. □

We will now derive formulae for computing the partial derivatives of J^p with respect to θ_k, $k = 1,\ldots,p-1$.

For each $k = 1,\ldots,p-1$, consider the following auxiliary dynamic system:

$$\dot{\boldsymbol{\psi}}^k(t) = (\rho_{k,i} - \delta_{k,i})\frac{\partial \mathbf{f}\big(\mathbf{x}^p(t|\boldsymbol{\sigma},\boldsymbol{\theta}),\boldsymbol{\sigma}^i\big)}{\partial \mathbf{x}}\boldsymbol{\psi}^k(t), \qquad t \in \mathcal{I}_i, \qquad i = 1,\ldots,p, \tag{4.15}$$

and, for each $i = 1,\ldots,p-1$,

$$\lim_{t\to t_i(\boldsymbol{\theta})+}\boldsymbol{\psi}^k(t) = \lim_{t\to t_i(\boldsymbol{\theta})-}\boldsymbol{\psi}^k(t) + \rho_{k,i}\Big\{\mathbf{f}\big(\mathbf{x}^p(t_i(\boldsymbol{\theta})|\boldsymbol{\sigma},\boldsymbol{\theta}),\boldsymbol{\sigma}^i\big) - \mathbf{f}\big(\mathbf{x}^p(t_i(\boldsymbol{\theta})|\boldsymbol{\sigma},\boldsymbol{\theta}),\boldsymbol{\sigma}^{i+1}\big)\Big\}, \tag{4.16}$$

and

$$\boldsymbol{\psi}^k(0) = \mathbf{0}, \tag{4.17}$$

where $(\boldsymbol{\sigma},\boldsymbol{\theta}) \in \Xi \times \Theta$. Let $\boldsymbol{\psi}^k(\cdot|\boldsymbol{\sigma},\boldsymbol{\theta})$ denote the solution of (4.15)-(4.17) corresponding to the pair $(\boldsymbol{\sigma},\boldsymbol{\theta}) \in \Xi \times \Theta$.

The next theorem shows that the partial derivatives of J^p with respect to θ_k, $k = 1,\ldots,p-1$, can be expressed in terms of $\boldsymbol{\psi}^k(\cdot|\boldsymbol{\sigma},\boldsymbol{\theta})$.

Theorem 4.4. *Let $(\boldsymbol{\sigma},\boldsymbol{\theta}) \in \Gamma$. Furthermore, let $T^p := T^p(\boldsymbol{\sigma},\boldsymbol{\theta})$, $\mathbf{x}^p := \mathbf{x}^p(\cdot|\boldsymbol{\sigma},\boldsymbol{\theta})$, and $\boldsymbol{\psi}^k := \boldsymbol{\psi}^k(\cdot|\boldsymbol{\sigma},\boldsymbol{\theta})$. Then for each $k = 1,\dots,p-1$,*

$$\frac{\partial J^p(\boldsymbol{\sigma},\boldsymbol{\theta})}{\partial \theta_k} = \frac{\partial \Psi(\mathbf{x}^p(T^p))}{\partial \mathbf{x}}\boldsymbol{\psi}^k(T^p) + \beta_k \frac{\partial \Psi(\mathbf{x}^p(T^p))}{\partial \mathbf{x}}\mathbf{f}\big(\mathbf{x}^p(T^p),\boldsymbol{\sigma}^p\big),$$

where

$$\beta_k := -\frac{\partial \Phi(\mathbf{x}^p(T^p))}{\partial \mathbf{x}}\boldsymbol{\psi}^k(T^p)\left[\frac{\partial \Phi(\mathbf{x}^p(T^p))}{\partial \mathbf{x}}\mathbf{f}\big(\mathbf{x}^p(T^p),\boldsymbol{\sigma}^p\big)\right]^{-1}.$$

Proof. Let $(\boldsymbol{\sigma},\boldsymbol{\theta}) \in \Gamma$ and $k \in \{1,\dots,p-1\}$ be arbitrary but fixed. As in the proof of Theorem 4.3, we will simplify the notation by writing t_i instead of $t_i(\boldsymbol{\theta})$ and $\mathcal{I}_i$ instead of $\mathcal{I}_i(\boldsymbol{\theta})$.

We first differentiate J^p with respect to θ_k:

$$\frac{\partial J^p(\boldsymbol{\sigma},\boldsymbol{\theta})}{\partial \theta_k} = \frac{\partial \Psi(\mathbf{x}^p(T^p))}{\partial \mathbf{x}}\frac{\partial \mathbf{x}^p(T^p)}{\partial \theta_k} + \frac{\partial \Psi(\mathbf{x}^p(T^p))}{\partial \mathbf{x}}\mathbf{f}\big(\mathbf{x}^p(T^p),\boldsymbol{\sigma}^p\big)\frac{\partial T^p}{\partial \theta_k}. \tag{4.18}$$

Next, for each $\epsilon > 0$, define

$$\boldsymbol{\theta}^\epsilon := \boldsymbol{\theta} + \epsilon \mathbf{e}^k,$$

where $\mathbf{e}^k$ is the kth standard unit vector in $\mathbb{R}^{p-1}$. Then

$$t_i(\boldsymbol{\theta}^\epsilon) = \begin{cases} t_i(\boldsymbol{\theta}), & \text{if } i = 1,\dots,k-1, \\ t_i(\boldsymbol{\theta}) + \epsilon, & \text{if } i = k,\dots,p. \end{cases} \tag{4.19}$$

Let $t \in \mathcal{I}_i(\boldsymbol{\theta})$ for some $i \leq k$. Then $t < t_k(\boldsymbol{\theta})$, and thus we can find a real number $\epsilon' > 0$ such that whenever $|\epsilon| < \epsilon'$,

$$t < t_k(\boldsymbol{\theta}) + \epsilon = t_k(\boldsymbol{\theta}^\epsilon). \tag{4.20}$$

From (4.19), we see that the first $k-1$ switching times of $\mathbf{u}(\cdot|\boldsymbol{\sigma},\boldsymbol{\theta}^\epsilon)$ and $\mathbf{u}(\cdot|\boldsymbol{\sigma},\boldsymbol{\theta})$ coincide, and from (4.20), we see that when $|\epsilon| < \epsilon'$ these two controls switch for the kth time *after* time t. Thus,

$$\mathbf{u}(s|\boldsymbol{\sigma},\boldsymbol{\theta}^\epsilon) = \mathbf{u}(s|\boldsymbol{\sigma},\boldsymbol{\theta}), \qquad s \in [0,t], \qquad |\epsilon| < \epsilon'.$$

Consequently, when ϵ is of sufficiently small magnitude,

$$\mathbf{x}^p(t|\boldsymbol{\sigma},\boldsymbol{\theta}^\epsilon) = \mathbf{x}^p(t|\boldsymbol{\sigma},\boldsymbol{\theta}),$$

where we recall that $t \in \mathcal{I}_i(\boldsymbol{\theta})$, $i \leq k$, was chosen arbitrarily. Therefore,

$$\frac{\partial \mathbf{x}^p(t)}{\partial \theta_k} = \lim_{\epsilon \to 0} \frac{\mathbf{x}^p(t|\boldsymbol{\theta}^\epsilon,\boldsymbol{\sigma}) - \mathbf{x}^p(t|\boldsymbol{\sigma},\boldsymbol{\theta})}{\epsilon} = \mathbf{0}, \qquad t \in \mathcal{I}_i, \qquad i = 1,\dots,k. \tag{4.21}$$

Differentiating this equation with respect to time yields

$$\frac{d}{dt}\left\{\frac{\partial \mathbf{x}^p(t)}{\partial \theta_k}\right\} = \mathbf{0}, \qquad t \in \mathcal{I}_i, \qquad i = 1,\dots,k. \tag{4.22}$$

We now consider the case when $i > k$.

It follows from (2.1)-(2.2) that

$$\begin{aligned} \mathbf{x}^p(t) = \mathbf{x}^0 &+ \sum_{j=1}^{k-1} \int_{t_{j-1}}^{t_j} \mathbf{f}\big(\mathbf{x}^p(s),\boldsymbol{\sigma}^j\big)ds + \int_{t_{k-1}}^{t_k} \mathbf{f}\big(\mathbf{x}^p(s),\boldsymbol{\sigma}^k\big)ds \\ &+ \sum_{j=k+1}^{i-1} \int_{t_{j-1}}^{t_j} \mathbf{f}\big(\mathbf{x}^p(s),\boldsymbol{\sigma}^j\big)ds + \int_{t_{i-1}}^{t} \mathbf{f}\big(\mathbf{x}^p(s),\boldsymbol{\sigma}^i\big)ds, \quad t \in \mathcal{I}_i. \end{aligned}$$

Differentiating this equation with respect to θ_k, and then using (4.21), gives

$$\begin{aligned}\frac{\partial \mathbf{x}^p(t)}{\partial \theta_k} &= \mathbf{f}\big(\mathbf{x}^p(t_k), \boldsymbol{\sigma}^k\big) + \sum_{j=k+1}^{i-1} \int_{t_{j-1}}^{t_j} \frac{\partial \mathbf{f}\big(\mathbf{x}^p(s), \boldsymbol{\sigma}^j\big)}{\partial \mathbf{x}} \frac{\partial \mathbf{x}^p(s)}{\partial \theta_k} ds \\ &+ \sum_{j=k+1}^{i-1} \Big\{ \mathbf{f}\big(\mathbf{x}^p(t_j), \boldsymbol{\sigma}^j\big) - \mathbf{f}\big(\mathbf{x}^p(t_{j-1}), \boldsymbol{\sigma}^j\big) \Big\} \\ &+ \int_{t_{i-1}}^{t} \frac{\partial \mathbf{f}\big(\mathbf{x}^p(s), \boldsymbol{\sigma}^i\big)}{\partial \mathbf{x}} \frac{\partial \mathbf{x}^p(s)}{\partial \theta_k} ds - \mathbf{f}\big(\mathbf{x}^p(t_{i-1}), \boldsymbol{\sigma}^i\big), \quad t \in \mathcal{I}_i. \end{aligned} \tag{4.23}$$

Differentiating this equation with respect to time yields

$$\frac{d}{dt}\left\{ \frac{\partial \mathbf{x}^p(t)}{\partial \theta_k} \right\} = \frac{\partial \mathbf{f}\big(\mathbf{x}^p(t), \boldsymbol{\sigma}^i\big)}{\partial \mathbf{x}} \frac{\partial \mathbf{x}^p(t)}{\partial \theta_k}, \qquad t \in \mathcal{I}_i, \qquad i = k+1, \ldots, p. \tag{4.24}$$

We can combine equations (4.22) and (4.24) into one equation as follows:

$$\frac{d}{dt}\left\{ \frac{\partial \mathbf{x}^p(t)}{\partial \theta_k} \right\} = (\rho_{k,i} - \delta_{k,i}) \frac{\partial \mathbf{f}\big(\mathbf{x}^p(t), \boldsymbol{\sigma}^i\big)}{\partial \mathbf{x}} \frac{\partial \mathbf{x}^p(t)}{\partial \theta_k}, \quad t \in \mathcal{I}_i, \quad i = 1, \ldots, p. \tag{4.25}$$

Now, it is clear from (4.21) that

$$\lim_{t \to t_i+} \frac{\partial \mathbf{x}^p(t)}{\partial \theta_k} = \lim_{t \to t_i-} \frac{\partial \mathbf{x}^p(t)}{\partial \theta_k}, \qquad i = 1, \ldots, k-1. \tag{4.26}$$

Furthermore, we see from (4.23) that

$$\lim_{t \to t_i+} \frac{\partial \mathbf{x}^p(t)}{\partial \theta_k} = \lim_{t \to t_i-} \frac{\partial \mathbf{x}^p(t)}{\partial \theta_k} + \mathbf{f}\big(\mathbf{x}^p(t_i), \boldsymbol{\sigma}^i\big) - \mathbf{f}\big(\mathbf{x}^p(t_i), \boldsymbol{\sigma}^{i+1}\big), \qquad i = k, \ldots, p-1. \tag{4.27}$$

Combining (4.26) and (4.27) yields

$$\lim_{t \to t_i+} \frac{\partial \mathbf{x}^p(t)}{\partial \theta_k} = \lim_{t \to t_i-} \frac{\partial \mathbf{x}^p(t)}{\partial \theta_k} + \rho_{k,i} \Big\{ \mathbf{f}\big(\mathbf{x}^p(t_i), \boldsymbol{\sigma}^i\big) - \mathbf{f}\big(\mathbf{x}^p(t_i), \boldsymbol{\sigma}^{i+1}\big) \Big\}, \qquad i = 1, \ldots, p-1. \tag{4.28}$$

Furthermore,

$$\frac{\partial \mathbf{x}^p(0)}{\partial \theta_k} = \frac{\partial}{\partial \theta_k} \Big\{ \mathbf{x}^0 \Big\} = \mathbf{0}. \tag{4.29}$$

Equations (4.25), (4.28), and (4.29) show that $\partial \mathbf{x}^p(\cdot)/\partial \theta_k$ is the solution of (4.15)-(4.17). Hence,

$$\frac{\partial \mathbf{x}^p(t)}{\partial \theta_k} = \boldsymbol{\psi}^k(t), \qquad t \in [0, \infty). \tag{4.30}$$

Now, recall from equation (2.3) that

$$\Phi(\mathbf{x}^p(T^p)) = 0.$$

Differentiating this equation with respect to θ_k yields

$$\frac{\partial \Phi(\mathbf{x}^p(T^p))}{\partial \mathbf{x}} \boldsymbol{\psi}^k(T^p) + \frac{\partial \Phi(\mathbf{x}^p(T^p))}{\partial \mathbf{x}} \mathbf{f}\big(\mathbf{x}^p(T^p), \boldsymbol{\sigma}^p\big) \frac{\partial T^p}{\partial \theta_k} = 0.$$

Thus, by (4.30),

$$\frac{\partial T^p}{\partial \theta_k} = -\frac{\partial \Phi(\mathbf{x}^p(T^p))}{\partial \mathbf{x}} \boldsymbol{\psi}^k(T^p) \left[\frac{\partial \Phi(\mathbf{x}^p(T^p))}{\partial \mathbf{x}} \mathbf{f}\big(\mathbf{x}^p(T^p), \boldsymbol{\sigma}^p\big) \right]^{-1} = \beta_k. \tag{4.31}$$

Substituting equations (4.30) and (4.31) into equation (4.18) completes the proof. □

Theorems 4.3 and 4.4 give formulae for computing the partial derivatives of J^p, but they are only applicable when the pair $(\boldsymbol{\sigma}, \boldsymbol{\theta})$ belongs to the set Γ. Nevertheless, when $(\boldsymbol{\sigma}, \boldsymbol{\theta}) \notin \Gamma$, we can use the procedure shown in the proof of Theorem 4.1 to generate a new pair in Γ that also has an objective value of $J^p(\boldsymbol{\sigma}, \boldsymbol{\theta})$. The gradient of J^p at this new pair can then be calculated using the formulae in Theorems 4.3 and 4.4, and this gradient can subsequently be used to compute a descent direction. This is the main idea of the following algorithm for solving Problem P(p), which combines the formulae in Theorems 4.3 and 4.4, the procedure in the proof of Theorem 4.1, and a gradient-based optimization method.

Algorithm 4.5. Input an initial pair $(\boldsymbol{\sigma}, \boldsymbol{\theta}) \in \Xi \times \Theta$.

(i) If $(\boldsymbol{\sigma}, \boldsymbol{\theta}) \notin \Gamma$, then go to Step (ii).
Otherwise, go to Step (iv).

(ii) Construct $(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\theta}}) \in \Gamma$ from $(\boldsymbol{\sigma}, \boldsymbol{\theta})$ using the procedure in the proof of Theorem 4.1.

(iii) Set $(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\theta}}) \to (\boldsymbol{\sigma}, \boldsymbol{\theta})$.

(iv) Compute the partial derivatives $\partial J^p(\boldsymbol{\sigma}, \boldsymbol{\theta})/\partial \sigma_\varsigma^k$ and $\partial J^p(\boldsymbol{\sigma}, \boldsymbol{\theta})/\partial \theta_k$ using the formulae in Theorems 4.3 and 4.4.

(v) If the optimality conditions are satisfied, then stop; $(\boldsymbol{\sigma}, \boldsymbol{\theta})$ is an optimal solution of Problem P(p).
Otherwise, use the gradient information obtained in Step (iv) to determine an appropriate search direction in $\Xi \times \Theta$.

(vi) Obtain a new pair $(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\theta}}) \in \Xi \times \Theta$ by performing a line search along the direction calculated in Step (v).

(vii) Set $(\bar{\boldsymbol{\sigma}}, \bar{\boldsymbol{\theta}}) \to (\boldsymbol{\sigma}, \boldsymbol{\theta})$ and go to Step (i).

5 Convergence Results

Problem P(p) can be solved using Algorithm 4.5, after which a suboptimal control for Problem P can be constructed according to Remark 3.1. By repeating these steps for increasing values of p, we can generate a sequence of suboptimal controls. In this section, we will present two convergence results linking these suboptimal controls with an optimal control of Problem P.

Since p is no longer fixed, we now denote Ξ by Ξ^p, Θ by Θ^p, $t_i(\boldsymbol{\theta})$ by $t_i^p(\boldsymbol{\theta})$, and $\mathcal{I}_i(\boldsymbol{\theta})$ by $\mathcal{I}_i^p(\boldsymbol{\theta})$.

We first recall the following result (Lemma 6.4.1 of [10]).

Lemma 5.1. *Let $\mathbf{u} \in \mathcal{U}$ be an admissible control for Problem P. Then there exists a sequence of admissible controls $\{\mathbf{u}^p\}_{p=1}^{\infty} \subset \mathcal{U}$ that converges to $\mathbf{u}$ almost everywhere on $[0, T_{\max}]$. Furthermore, for each $p \geq 1$, $\mathbf{u}^p$ can be expressed as*

$$\mathbf{u}^p(t) = \sum_{i=1}^{p} \boldsymbol{\sigma}^{p,i} \chi_{[t_{i-1}^p, t_i^p)}(t), \qquad t \in [0, \infty),$$

where $\boldsymbol{\sigma}^{p,i} \in \Upsilon$, $i = 1, \ldots, p$, $t_0^p = 0$, $t_p^p = \infty$, *and* $t_{i-1}^p < t_i^p$, $i = 1, \ldots, p$.

We also assume that the following condition is satisfied.

Assumption 5.2. For each admissible control $\mathbf{u} \in \mathcal{U}$, there exists a corresponding real number $\bar{\omega} > 0$ such that

$$\Phi\big(\mathbf{x}(T(\mathbf{u}) + \omega|\mathbf{u})\big) < 0, \qquad \omega \in (0, \bar{\omega}).$$

Lemma 5.3. *Let* $\{\mathbf{u}^p\}_{p=1}^\infty \subset \mathcal{U}$ *be a sequence of admissible controls converging to* $\mathbf{u} \in \mathcal{U}$ *almost everywhere on* $[0, T_{\max}]$*. Then the following results hold:*

1. $\mathbf{x}(\cdot|\mathbf{u}^p) \to \mathbf{x}(\cdot|\mathbf{u})$ *uniformly on* $[0, T_{\max}]$ *as* $p \to \infty$*;*
2. $T(\mathbf{u}^p) \to T(\mathbf{u})$ *as* $p \to \infty$*; and*
3. $J(\mathbf{u}^p) \to J(\mathbf{u})$ *as* $p \to \infty$*.*

Proof. 1. This result is a simple extension of Lemma 6.4.3 in [10].

2. Suppose, to the contrary, that $\{T(\mathbf{u}^p)\}_{p=1}^\infty$ does not converge to $T(\mathbf{u})$. Then there exists a real number $\gamma > 0$ and a subsequence $\{T(\mathbf{u}^{p_j})\}_{j=1}^\infty$ such that

$$\big|T(\mathbf{u}^{p_j}) - T(\mathbf{u})\big| \geq \gamma, \qquad j \geq 1. \tag{5.1}$$

By Assumption 2.3,

$$0 \leq T(\mathbf{u}^{p_j}) \leq T_{\max}, \qquad j \geq 1.$$

Thus, by virtue of the Bolzano-Weierstrass Theorem, and by passing to a subsequence if necessary, we may assume that $\{T(\mathbf{u}^{p_j})\}_{j=1}^\infty$ converges to a real number $\bar{T} \in [0, T_{\max}]$. It is clear from (5.1) that

$$\big|\bar{T} - T(\mathbf{u})\big| \geq \gamma > 0. \tag{5.2}$$

Hence, $\bar{T} \neq T(\mathbf{u})$.

Now, let $\epsilon > 0$ be arbitrary but fixed. By part 1, there exists an integer $j' \geq 1$ such that

$$\big\|\mathbf{x}(T(\mathbf{u}^{p_j})|\mathbf{u}^{p_j}) - \mathbf{x}(T(\mathbf{u}^{p_j})|\mathbf{u})\big\| < \frac{\epsilon}{2}, \qquad j \geq j'.$$

Furthermore, since $T(\mathbf{u}^{p_j}) \to \bar{T}$ as $j \to \infty$, and $\mathbf{x}(\cdot|\mathbf{u})$ is continuous, there exists another integer $j'' \geq 1$ such that

$$\big\|\mathbf{x}(T(\mathbf{u}^{p_j})|\mathbf{u}) - \mathbf{x}(\bar{T}|\mathbf{u})\big\| < \frac{\epsilon}{2}, \qquad j \geq j''.$$

Thus, for each integer $j \geq \max\{j', j''\}$,

$$\begin{aligned}\big\|\mathbf{x}(T(\mathbf{u}^{p_j})|\mathbf{u}^{p_j}) - \mathbf{x}(\bar{T}|\mathbf{u})\big\| &\leq \big\|\mathbf{x}(T(\mathbf{u}^{p_j})|\mathbf{u}^{p_j}) - \mathbf{x}(T(\mathbf{u}^{p_j})|\mathbf{u})\big\| \\ &\quad + \big\|\mathbf{x}(T(\mathbf{u}^{p_j})|\mathbf{u}) - \mathbf{x}(\bar{T}|\mathbf{u})\big\| \\ &< \epsilon/2 + \epsilon/2 = \epsilon.\end{aligned}$$

Since $\epsilon > 0$ was chosen arbitrarily, this inequality shows that

$$\lim_{j\to\infty} \mathbf{x}(T(\mathbf{u}^{p_j})|\mathbf{u}^{p_j}) = \mathbf{x}(\bar{T}|\mathbf{u}).$$

Thus, since Φ is continuous,

$$\lim_{j\to\infty} \Phi\big(\mathbf{x}(T(\mathbf{u}^{p_j})|\mathbf{u}^{p_j})\big) = \Phi\big(\mathbf{x}(\bar{T}|\mathbf{u})\big). \tag{5.3}$$

However, recall from equation (2.3) that

$$\Phi\big(\mathbf{x}(T(\mathbf{u}^{p_j})|\mathbf{u}^{p_j})\big) = 0, \qquad j \geq 1. \tag{5.4}$$

Combining (5.3) and (5.4) gives

$$\Phi\big(\mathbf{x}(\bar{T}|\mathbf{u})\big) = 0.$$

Therefore, since $T(\mathbf{u}) \neq \bar{T}$,

$$T(\mathbf{u}) < \bar{T}$$

and so by (5.2),

$$T(\mathbf{u}) + \gamma \leq \bar{T}. \tag{5.5}$$

Now, define

$$\upsilon := \min\big\{\bar{\omega}/2, \gamma/4\big\},$$

where $\bar{\omega} > 0$ is the real number in Assumption 5.2 corresponding to the admissible control $\mathbf{u}$. Then $\upsilon < \gamma/2$, and so (5.5) implies that there exists an integer $M \geq 1$ such that

$$T(\mathbf{u}) + \upsilon < T(\mathbf{u}) + \frac{\gamma}{2} \leq T(\mathbf{u}^{p_j}), \qquad j \geq M. \tag{5.6}$$

Furthermore, $\upsilon < \bar{\omega}$ and thus by Assumption 5.2,

$$\Phi\big(\mathbf{x}(T(\mathbf{u}) + \upsilon|\mathbf{u})\big) < 0.$$

By part 1,

$$\lim_{j\to\infty} \Phi\big(\mathbf{x}(T(\mathbf{u}) + \upsilon|\mathbf{u}^{p_j})\big) = \Phi\big(\mathbf{x}(T(\mathbf{u}) + \upsilon|\mathbf{u})\big) < 0.$$

Thus, there exists an integer $N \geq 1$ such that

$$\Phi\big(\mathbf{x}(T(\mathbf{u}) + \upsilon|\mathbf{u}^{p_j})\big) < 0, \qquad j \geq N.$$

Since $\Phi(\mathbf{x}^0) > 0$ (see Section 2), this implies that

$$T(\mathbf{u}^{p_j}) \leq T(\mathbf{u}) + \upsilon, \qquad j \geq N. \tag{5.7}$$

When $j \geq \max\{M, N\}$, inequalities (5.6) and (5.7) imply

$$T(\mathbf{u}) + \upsilon < T(\mathbf{u}) + \upsilon,$$

a contradiction. This completes the proof.

3. By parts 1 and 2,

$$\lim_{p\to\infty} \mathbf{x}(T(\mathbf{u}^p)|\mathbf{u}^p) = \mathbf{x}(T(\mathbf{u})|\mathbf{u}). \tag{5.8}$$

Thus, since Ψ is continuous,

$$\lim_{p\to\infty} \Psi\big(\mathbf{x}(T(\mathbf{u}^p)|\mathbf{u}^p)\big) = \Psi\big(\mathbf{x}(T(\mathbf{u})|\mathbf{u})\big),$$

which completes the proof.

□

We now present two important convergence results.

Theorem 5.4. *Suppose that Problem P has an optimal control* $\mathbf{u}^*$*. For each integer* $p \geq 2$*, let* $\mathbf{u}^{p,*}$ *denote the suboptimal control constructed from the solution of Problem P*(p) *according to Remark 3.1. Then*

$$\lim_{p\to\infty} J(\mathbf{u}^{p,*}) = J(\mathbf{u}^*).$$

Proof. Let $\epsilon > 0$ be arbitrary. By Lemma 5.1, there exists a sequence of admissible controls $\{\mathbf{u}^{*,p}\}_{p=1}^{\infty}$ converging to $\mathbf{u}^*$ almost everywhere on $[0, T_{\max}]$. Thus, by Lemma 5.3, part 3, there exists an integer $p_1 \geq 1$ such that

$$J(\mathbf{u}^{*,p}) \leq J(\mathbf{u}^*) + \epsilon, \qquad p \geq p_1. \tag{5.9}$$

Now, recall from Lemma 5.1 that for each $p \geq 1$, $\mathbf{u}^{*,p}$ can be written as

$$\mathbf{u}^{*,p}(t) = \sum_{i=1}^{p} \boldsymbol{\sigma}^{*,p,i} \chi_{[t_{i-1}^{*,p}, t_i^{*,p})}(t),$$

where $\boldsymbol{\sigma}^{*,p,i} \in \Upsilon$, $i = 1, \ldots, p$, $t_0^{*,p} = 0$, $t_p^{*,p} = \infty$, and $t_{i-1}^{*,p} < t_i^{*,p}$, $i = 1, \ldots, p$. For each integer $p \geq 2$, define

$$\boldsymbol{\sigma}^{*,p} := (\boldsymbol{\sigma}^{*,p,1}, \ldots, \boldsymbol{\sigma}^{*,p,p})$$

and

$$\boldsymbol{\theta}^{*,p} := [\theta_1^{*,p}, \ldots, \theta_{p-1}^{*,p}]^T,$$

where

$$\theta_i^{*,p} := t_i^{*,p} - t_{i-1}^{*,p}, \qquad i = 1, \ldots, p-1.$$

Clearly, $(\boldsymbol{\sigma}^{*,p}, \boldsymbol{\theta}^{*,p}) \in \Xi^p \times \Theta^p$ for each $p \geq 2$. Moreover,

$$t_i^p(\boldsymbol{\theta}^{*,p}) = t_i^{*,p}, \qquad i = 1, \ldots, p.$$

This implies that for each $p \geq 2$, we can express $\mathbf{u}^{*,p}$ as

$$\mathbf{u}^{*,p}(t) = \sum_{i=1}^{p} \boldsymbol{\sigma}^{*,p,i} \chi_{\mathcal{I}_i^p(\boldsymbol{\theta}^{*,p})}(t) = \mathbf{u}^p(t|\boldsymbol{\sigma}^{*,p}, \boldsymbol{\theta}^{*,p}).$$

Thus,

$$J^p(\boldsymbol{\sigma}^{*,p}, \boldsymbol{\theta}^{*,p}) = J(\mathbf{u}^{*,p}), \qquad p \geq 2.$$

Therefore,

$$J(\mathbf{u}^{p,*}) \leq J^p(\boldsymbol{\sigma}^{*,p}, \boldsymbol{\theta}^{*,p}) = J(\mathbf{u}^{*,p}), \qquad p \geq 2. \tag{5.10}$$

In view of (5.9) and (5.10), we see that when $p \geq \max\{p_1, 2\}$,

$$J(\mathbf{u}^*) \leq J(\mathbf{u}^{p,*}) \leq J(\mathbf{u}^*) + \epsilon.$$

Since $\epsilon > 0$ was chosen arbitrarily, this shows that $J(\mathbf{u}^{p,*}) \to J(\mathbf{u}^*)$ as $p \to \infty$, completing the proof. □

Theorem 5.5. *Let* $\mathbf{u}^*$ *and* $\mathbf{u}^{p,*}$ *be as defined in Theorem 5.4. If* $\{\mathbf{u}^{p,*}\}_{p=2}^{\infty}$ *converges to a function* $\bar{\mathbf{u}}$ *almost everywhere on* $[0, T_{\max}]$*, then* $\bar{\mathbf{u}}$ *is an optimal control for Problem P.*

Proof. We immediately see that $\bar{\mathbf{u}}$ is admissible. Furthermore, by part 3 of Lemma 5.3,

$$\lim_{p\to\infty} J(\mathbf{u}^{p,*}) = J(\bar{\mathbf{u}}). \tag{5.11}$$

On the other hand, Theorem 5.4 implies that

$$\lim_{p\to\infty} J(\mathbf{u}^{p,*}) = J(\mathbf{u}^*). \tag{5.12}$$

Combining (5.11) and (5.12) gives

$$J(\bar{\mathbf{u}}) = J(\mathbf{u}^*),$$

which shows that $\bar{\mathbf{u}} \in \mathcal{U}$ is an optimal control for Problem P. □

$\boxed{6}$ An Applied Aeronautical Control Problem

As an example, we consider the glider control problem introduced in [8]. This problem involves a gliding projectile that is described by the following dynamic system:

$$\dot{x}_1(t) = x_3(t)\cos(x_4(t)), \tag{6.1a}$$

$$\dot{x}_2(t) = x_3(t)\sin(x_4(t)), \tag{6.1b}$$

$$\dot{x}_3(t) = -\big(K_{D0} + K_{D2}\alpha^2(t)\big)x_3^2(t) - g\sin(x_4(t)), \tag{6.1c}$$

$$\dot{x}_4(t) = K_L\alpha(t)x_3(t) - \frac{g}{x_3(t)}\cos(x_4(t)), \tag{6.1d}$$

and

$$x_1(0) = 0.0, \tag{6.2a}$$

$$x_2(0) = 0.0, \tag{6.2b}$$

$$x_3(0) = 370.0, \tag{6.2c}$$

$$x_4(0) = 1.5, \tag{6.2d}$$

where $x_1(t)$ is the glider's horizontal distance from the launch point at time t; $x_2(t)$ is the glider's height at time t; $x_3(t)$ is the glider's speed at time t; $x_4(t)$ is the angle between the glider's velocity vector and the horizon at time t; $\alpha(t)$ is the glider's angle of attack at time t; and $K_{D0} = 3.289 \times 10^{-5}$ m^{-1}, $K_{D2} = 1.133 \times 10^{-3}$ m^{-1}, $K_L = 3.289 \times 10^{-3}$ m^{-1}, and $g = 9.80$ m s^{-2} are constants.

Let $T > 0$ be the time at which the glider crashes. Then T is the first positive time at which the following equation is satisfied:

$$x_2(T) = 0. \tag{6.3}$$

Thus,

$$x_2(t) > 0, \qquad t \in (0, T). \tag{6.4}$$

Assumption 2.3 is clearly satisfied here, because the glider will eventually crash, regardless of its angle of attack.

The optimal control problem is as follows: choose the angle of attack α in such a way that the glider's range, $x_1(T)$, where T is the first positive time satisfying (6.3)-(6.4), is maximized.

We wrote a Fortran program, which implements the method discussed in Section 4, for solving Problem P(p) for the optimal glider control problem described above. This program uses the differential equation solver LSODAR (see [2]) to solve the state and auxiliary systems, and NLPQLP (see [7]) to perform the optimization. We started the program with the following initial control (which is suggested in [8]):

$$\alpha(t) = \begin{cases} 0, & \text{if } t \leq 40, \\ 0.17, & \text{if } t > 40. \end{cases} \tag{6.5}$$

We first used the program to solve Problem P(4) (that is, Problem P(p) with $p = 4$ subintervals). This was done on a MacBook Pro with a 2.4GHz Intel Core 2 Duo processor and 4GB of RAM. The program terminated after 116.79 seconds of computation time. Under the optimal control that was obtained, the glider reaches a maximum range of 45,782 meters and glides for 378.68 seconds.

To put these results in perspective, we wrote another Fortran program that implements the method in [8] (with $p = 25$ subintervals). This program also uses LSODAR to solve the state system (and LSODA to solve the costate system), and NLPQLP to perform the optimization. It was run on the same computer and with the same initial control (equation (6.5)) as the first program. The program terminated after 818.01 seconds of computation time, with a maximum range of 45,622 meters and an optimal flight time of 360.66 seconds. These results are slightly worse than those obtained using the first program (which implements our new method), and took much longer to generate. As expected, allowing the switching times to be decision variables is a major advantage of our new method: we only needed four subintervals to generate better results than what was obtained using the method in [8] with 25 subintervals.

NLPQLP also converged much better when our new method was used to construct the gradients. In fact, NLPQLP's line search at each iteration normally used fewer than five function evaluations. In contrast, NPQLP usually needed more than ten function evaluations when the method in [8] was used to generate the gradients. Again, this is what we expected—because our new method allows the control switching times to be decision variables, NLPQLP has more flexibility to update the control, and therefore converges rapidly.

Two important things about the optimal glider control problem are worth mentioning. First, the problem is highly nonlinear, and thus both our new method and the method in [8] are very sensitive to the initial guess chosen for the control. Second, the gradient of the cost function in the approximate Problem P(p) is very large and needs to be normalized. NLPQLP's line search procedure does not perform correctly on this problem unless the gradient is normalized.

In an attempt to improve the glider's range, we used our solution to Problem P(4) as the initial guess for Problem P(8). Then, solving Problem P(8) using our program gave a maximum range of 46,093 meters and an optimal flight time of 374.41 seconds. These results are superior to the maximum range of 45,490 meters reported in [8].

Our results are displayed in Figures 1-2. In these figures, "old method, $p = 25$" refers to the method in [8] with $p = 25$ subintervals, "new method, $p = 4$" refers to the method discussed in this paper with $p = 4$ subintervals, and "new method, $p = 8$" also refers to the method in this paper but with $p = 8$ subintervals.

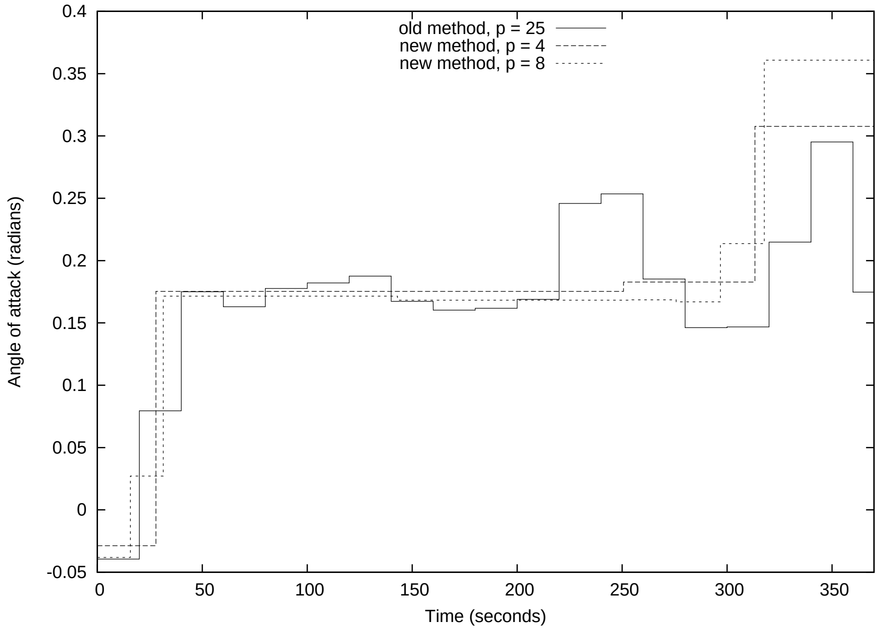

Figure 1: The optimal angle of attack.

7 Conclusion

In this paper, we have developed a new computational method for solving an optimal control problem whose terminal time is governed by a state-dependent stopping criterion. Thus, the terminal time in this optimal control problem is "free"—it is not fixed and is instead determined implicitly as the state system is being solved. Extending the optimal control method developed in this paper to more complex systems, which may have state and control constraints and more general stopping conditions, is an interesting topic for future research.

References

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

[2] A. Hindmarsh, Large ordinary differential equation systems and software. *IEEE Control Syst. Mag.* 2 (1982) 24–30.

[3] C.Y. Kaya and J.L. Noakes, Computational method for time-optimal switching control. *J. Optim. Theory Appl.* 117 (2003) 69–92.

[4] R.C. Loxton, K.L. Teo and V. Rehbock, Optimal control problems with multiple characteristic time points in the objective and constraints, *Automatica J. IFAC* 44 (2008) 2923–2929.

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

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

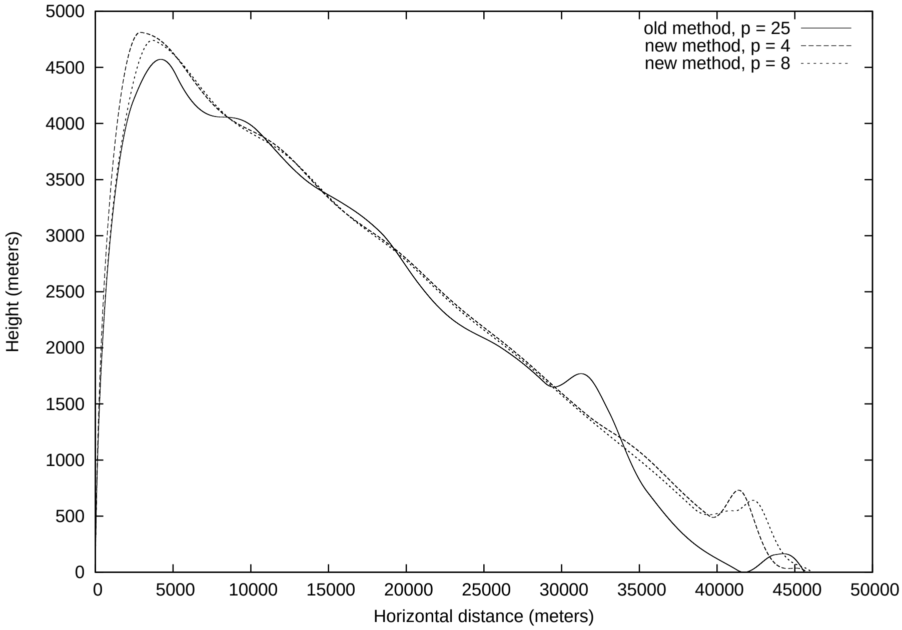

Figure 2: The glider's optimal trajectory.

[7] K. Schittkowski, *NLPQLP: A Fortran Implementation of a Sequential Quadratic Programming Algorithm with Distributed and Non-Monotone Line Search*, User's Guide Version 2.0, University of Bayreuth, 2004.

[8] K.L. Teo, G. Jepps, E.J. Moore and S. Hayes, A computational method for free time optimal control problems, with application to maximizing the range of an aircraft-like projectile, *J. Aust. Math. Soc.* Series B 28 (1987) 393–413.

[9] K.L. Teo, C.J. Goh and C.C. Lim, A computational method for a class of dynamical optimization problems in which the terminal time is conditionally free, *IMA J. Math. Control Inform.* 6 (1989) 81–95.

[10] K.L. Teo, C.J. Goh and K.H. Wong, *A Unified Computational Approach to Optimal Control Problems*, Longman Scientific and Technical, Essex, 1991.

[11] T.L. Vincent, W.J. Grantham, *Optimality in Parametric Systems*, Wiley, New York, 1981.

Manuscript received 24 September 2008
revised 14 November 2009
accepted for publication 4 January 2010

Qun Lin
Department of Mathematics and Statistics, Curtin University of Technology
GPO Box U1987 Perth, Western Australia 6845
E-mail address: `Q.Lin@curtin.edu.au`

Ryan Loxton
Department of Mathematics and Statistics, Curtin University of Technology
GPO Box U1987 Perth, Western Australia 6845
E-mail address: R.Loxton@curtin.edu.au

Kok Lay Teo
Department of Mathematics and Statistics, Curtin University of Technology
GPO Box U1987 Perth, Western Australia 6845
E-mail address: K.L.Teo@curtin.edu.au

Yong Hong Wu
Department of Mathematics and Statistics, Curtin University of Technology
GPO Box U1987 Perth, Western Australia 6845
E-mail address: Y.Wu@curtin.edu.au