

Computing Switching Surfaces in Optimal Control Based on Triangular Decomposition

Xiaoliang Li^{1,4}, Yanli Huang^{2,4,*}, Zewei Zheng³ and Wanyou Cheng¹

¹ School of Computer Science, Dongguan University of Technology, Dongguan, Guangdong 523808, China.

² School of Computer Science and Software Engineering, Tianjin Polytechnic University, Tianjin 300387, China.

³ School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China.

⁴ Guangxi Key Laboratory of Hybrid Computation and IC Design Analysis, Guangxi University for Nationalities, Nanning 530006, China.

Received 7 May 2014; Accepted (in revised version) 16 October 2014

Available online 5 November 2014

Abstract. Various algorithms for optimal control require the explicit determination of switching surfaces. However, switching strategies may be very complicated, such that the computation of switching surfaces is quite challenging. General methods are proposed here to compute switching surfaces systematically, based on algebraic computational tools such as triangular decomposition. Our methods are highly complex compared to some widely-used numerical options, but they can be made feasible for real-time applications by moving the computational burden off-line. The tutorial-style presentation is intended to introduce potentially powerful symbolic computation methods to system scientists in particular, and an illustrative example of time-optimal control is given to show the effectiveness and generality of our approach.

AMS subject classifications: 13P15, 13P25

Key words: Optimal control, switching surface, triangular decomposition, semi-algebraic system.

1. Introduction

Optimal control has been used in many areas of modern system science such as aeronautics, astronautics, robotics and power electronics [8, 13, 16, 26, 41]. The control algorithms typically require explicit determination of switching surfaces — surfaces where the sign of the control input changes. However, switching strategies may be very complicated in many practical applications, so the computation of switching surfaces becomes quite challenging. Some special approaches to their computation have been developed [1, 18, 24]. Walther *et*

*Corresponding author. Email address: huangyanlibuaa@gmail.com (Yanli Huang)

al. [30] introduced tools from computational algebraic geometry for time-optimal control problems, by transforming the computation of switching surfaces into a combinatorial problem. Their main idea is to first compute Gröbner bases of particular polynomial equations deduced from the original system, and then use Sturm's theorem to determine whether the equations have non-negative solutions. Their approach is somewhat non-systematic and may not be feasible for real-time control due to its high complexity, which motivated the development of more general systematic methods for computing switching surfaces presented here. Another purpose of this article is to introduce powerful methods of symbolic computation (coupled with numerical computation) for system scientists.

Our methods are based on triangular decomposition and related algebraic tools. Like Gröbner bases, triangular decomposition is a main elimination approach for solving systems of multivariate polynomial equations. For example, consider the equations

$$\begin{cases} P_1 = x_1^2 + x_2 + x_3 - 1 = 0, \\ P_2 = x_2^2 + x_3 + x_1 - 1 = 0, \\ P_3 = x_3^2 + x_1 + x_2 - 1 = 0. \end{cases}$$

Under the variable ordering $x_1 < x_2 < x_3$, triangular decomposition of the polynomial set $\mathcal{P} = \{P_1, P_2, P_3\}$ results in

$$\begin{aligned} \mathcal{T}_1 &= [x_1^2 + 2x_1 - 1, x_2 - x_1, x_3 - x_1], & \mathcal{T}_2 &= [x_1 - 1, x_2, x_3], \\ \mathcal{T}_3 &= [x_1, x_2 - 1, x_3], & \mathcal{T}_4 &= [x_1, x_2, x_3 - 1], \end{aligned}$$

such that the union of the zero sets of $\mathcal{T}_1, \dots, \mathcal{T}_4$ is identical to the zero set of \mathcal{P} . The *triangular sets* $\mathcal{T}_1, \dots, \mathcal{T}_4$ are of triangular form. Note that the zero set of a triangular set can readily be obtained by successively computing the zeros of its polynomials. In our algorithms, triangular decomposition is thus used as a preprocessing tool in the analysis of solutions of polynomial equations. Formal notation and properties related to triangular decomposition are provided in subsection 3.1 below.

Since triangular decomposition is such a key aspect of our approach, a brief literature overview may be helpful. In considering differential ideals, Ritt [25] introduced the notion of characteristic set, one of the best known concepts for triangular sets. Several decades later, Wu [37, 38] extended Ritt's work by removing irreducibility requirements of characteristic sets, proposing efficient algorithms for decomposing polynomial sets, and successfully applying them to geometric theorem proof. Wu's method was intensively studied and improved by a number of researchers (e.g. [4, 9–11, 21, 34, 35]), but the zero set of a characteristic set may be empty. To avoid this degeneracy, Kalkbrener [15] and Yang *et al.* [39] independently introduced the notion of a regular set, and proposed methods for decomposing any algebraic variety into finite components represented by regular sets. Wang [31] proposed another method for triangular decomposition, which is considered to be quite efficient. Other relevant work on the triangular decomposition of polynomial sets is discussed in Refs. [3, 17, 23, 33] (and references therein).

Our work presented here is based on the observation that the problem of computing the switching surfaces of the time-optimal control can be translated into identifying whether

there exist (complex and non-negative real) solutions of particular equations, say \mathcal{F} with parameters — cf. Section 2. By applying methods of triangular decomposition, \mathcal{F} can be rendered by a finite number of triangular sets with a simpler form. The projection property of a special class of triangular sets may then be used to directly identify whether \mathcal{F} has complex solutions. However, it is by no means trivial to determine whether or not there exists any non-negative real solution of \mathcal{F} . We propose general algorithms to solve this problem, based on a number of powerful tools from real algebraic geometry such as Sturm’s theorem, real root isolation and cylindrical algebraic decomposition. The border polynomial notation plays an important role, and its formal definition and relevant property are given in subsection 3.2. All the algorithms presented here have been implemented in Maple 17 based on the Epsilon package[†]. (The source code together with experimental results is available from the authors on request.)

In Section 2, as an illustrative example we consider a simple system consisting of a chain of integrators, and address two (complex and real) versions of the problem related to the computation of optimal control. Section 3 summarises our notation, such as for triangular sets and semi-algebraic systems. In Section 4, the complex version of the problem is solved directly, based on the triangular decomposition of a particular polynomial set. Sections 5 and 6 focus on the resolution of the real version of the problem. In Section 5, we propose an approach to determine the switching strategy of the system given in Section 2; and in Section 6, we modify our algorithms for real-time applications by moving the computational burden off-line. Section 7 presents our concluding remarks.

2. Problem Statement

As previously mentioned, our emphasis is on devising systematic general approaches to compute optimal control for a system consisting of a chain of integrators. For convenience of comparison, we consider the same example as in Ref. [30] — viz. the third-order linear system with saturated control input

$$\dot{X}_3(t) = X_2(t), \quad \dot{X}_2(t) = X_1(t), \quad \dot{X}_1(t) = v(t), \quad (2.1)$$

where the dot denotes differentiation with respect to time t and $|v(t)| \leq 1$. It is well known that minimum-time optimal control for this system leads to “bang-bang” control with at most three switchings. However, although we consider such a very simple example, our methods are completely general and can be extended in a straightforward manner to systems of any order.

The objective is to take the minimum time t_f to drive the system (2.1) from any given initial state $(X_1(0), X_2(0), X_3(0)) = (a, b, c)$ to the origin $(X_1(t_f), X_2(t_f), X_3(t_f)) = (0, 0, 0)$. For simplicity, let us first consider the inverse problem — viz. the minimum time to drive the system from $(X_1(0), X_2(0), X_3(0)) = (0, 0, 0)$ to $(X_1(t_f), X_2(t_f), X_3(t_f)) = (a, b, c)$. The equivalence of these two problems is demonstrated at the end of this section.

[†]Available at <http://www-spiral.lip6.fr/~wang/epsilon/>

Suppose that

$$v(t) = \begin{cases} +1, & 0 \leq t < t_1, \\ -1, & t_1 \leq t < t_1 + t_2, \\ +1, & t_1 + t_2 \leq t < t_f = t_1 + t_2 + t_3, \end{cases} \quad (2.2)$$

where $t_1, t_2, t_3 \geq 0$ are the length of successive intervals where $v(t)$ stays constant. Then it is easy to prove that

$$\begin{cases} X_1(t_f) = t_1 - t_2 + t_3, \\ X_2(t_f) = \frac{t_1^2}{2} + t_1 t_2 + \frac{t_3^2}{2} + t_3 t_1 - \frac{t_2^2}{2} - t_2 t_3, \\ X_3(t_f) = \frac{t_1^3}{6} + \frac{t_3^3}{6} + \frac{t_1^2 t_2}{2} + \frac{t_1^2 t_3}{2} + \frac{t_2^2 t_1}{2} + \frac{t_3^2 t_1}{2} + t_1 t_2 t_3 - \frac{t_2^3}{6} - \frac{t_2^2 t_3}{2} - \frac{t_3^2 t_2}{2}. \end{cases}$$

Note that $X_1(t_f) = a$, $X_2(t_f) = b$ and $X_3(t_f) = c$. (The above relations would be more complicated for the direct problem where the initial state is not at the origin.) We consider the following polynomial equations:

$$\begin{cases} F_1 = t_1 - t_2 + t_3 - a = 0, \\ F_2 = \frac{t_1^2}{2} + t_1 t_2 + \frac{t_3^2}{2} + t_3 t_1 - \frac{t_2^2}{2} - t_2 t_3 - b = 0, \\ F_3 = \frac{t_1^3}{6} + \frac{t_3^3}{6} + \frac{t_1^2 t_2}{2} + \frac{t_1^2 t_3}{2} + \frac{t_2^2 t_1}{2} + \frac{t_3^2 t_1}{2} + t_1 t_2 t_3 - \frac{t_2^3}{6} - \frac{t_2^2 t_3}{2} - \frac{t_3^2 t_2}{2} - c = 0, \end{cases} \quad (2.3)$$

involving the variables t_1, t_2, t_3 and parameters a, b, c .

Remark 2.1. Determining the optimal control v of system (2.1) can be transformed to solving system (2.3). If there exists any real solution such that $t_1 \geq 0, t_2 \geq 0, t_3 \geq 0$, then the optimal control v should be set as in (2.2). There are a number of numerical methods for solving polynomial equations — e.g. Newton–Raphson methods [7] and homotopy continuation methods [29]. Numerical methods may suffer from computational instability and the difficulty in identifying signs of the solutions (particularly near zero) — whereas our approach based on (exact) symbolic computation overcomes these shortcomings, since computing the optimal control v then involves determining the solution signs instead of explicitly solving the above equations.

Both the complex and real version of the problem of solving system (2.3) are now considered — viz.

Problem 1. For any given complex numbers a, b and c , do complex solutions of system (2.3) exist?

and

Problem 2. For any given real numbers a, b and c , does system (2.3) have at least one real solution satisfying $t_1 \geq 0, t_2 \geq 0$ and $t_3 \geq 0$?

Remark 2.2. If the answer to Problem 1 is negative, the optimal control v should take the values $-1, +1, -1$, successively. If the answer to Problem 2 is positive, then v should be assumed to be $+1, -1, +1$, and the length of the successive intervals where $v(t)$ stays constant should be the non-negative components t_1, t_2 and t_3 of the solution; otherwise, v should take the values $-1, +1, -1$, successively. Further details are given in Ref. [30].

Suppose that system (2.1) has been driven from the origin to (a, b, c) by setting (2.2). From basic control theory, by reversing the direction of time as well as the sign of control the path is traversed backwards. More precisely, if

$$v(t) = \begin{cases} -1, & 0 \leq t < t_3, \\ +1, & t_3 \leq t < t_3 + t_2, \\ -1, & t_3 + t_2 \leq t < t_3 + t_2 + t_1, \end{cases}$$

then system (2.1) moves back from (a, b, c) to $(0, 0, 0)$. This shows that the direct problem is equivalent to its inverse.

3. Preliminaries

Let us now introduce some notations and properties for the triangular set and semi-algebraic system, for the problems given in the previous section.

3.1. Triangular decomposition methods

In the following, \mathbb{R} and \mathbb{C} denote the real and complex fields, respectively. We use $\mathbb{R}[x_1, \dots, x_n]$ to denote the multivariate polynomial ring over \mathbb{R} with variables ordered as $x_1 < \dots < x_n$. Let $F \in \mathbb{R}[x_1, \dots, x_n]$. We call $lv(F) = \max_{\leq} \{x_i \mid \deg(F, x_i) \neq 0, 1 \leq i \leq n\}$ the *leading variable* of F . The leading coefficient of F , viewed as a univariate polynomial in its leading variable, is called the *initial* of F and denoted by $ini(F)$.

Let \mathcal{P} and \mathcal{Q} be two sets of multivariate polynomials with coefficients in \mathbb{R} . We denote by $\text{Zero}(\mathcal{P})$ the set of all common zeros in \mathbb{C}^n of the polynomials in \mathcal{P} , and by $\text{Zero}(\mathcal{P} / \mathcal{Q})$ the subset of $\text{Zero}(\mathcal{P})$ with elements that do not annihilate any polynomial in \mathcal{Q} .

Definition 3.1 (Triangular Set). An ordered polynomial set $[T_1, \dots, T_r] \subseteq \mathbb{R}[x_1, \dots, x_n] \setminus \mathbb{R}$ is called a *triangular set* if the leading variable of T_i is ordered smaller than that of T_j for any $i < j$.

A triangular set $[T_1, \dots, T_r]$ thus has a simple special structure, and its zeros can easily be obtained by successively solving $T_1 = 0, \dots, T_r = 0$. However, for a generic triangular set \mathcal{T} , it is not guaranteed that the corresponding zero set $\text{Zero}(\mathcal{T} / ini(\mathcal{T}))$ is non-empty. For example, if

$$\mathcal{T} = [x_1^2 - u, x_2^2 + 2x_1x_2 + u, (x_1 + x_2)x_3 + 1]$$

where $u < x_1 < x_2 < x_3$, then $\text{Zero}(\mathcal{T} / ini(\mathcal{T})) = \emptyset$ since any common zero of the first two polynomials annihilates the initial $x + y$ of the last polynomial.

The emptiness of the zero sets of triangular sets would be a problem when counting the real solutions of polynomial equations. To avoid this, triangular sets of other kinds (with better properties) are needed. Such triangular sets include for example *regular sets* [15, 33, 39], *simple sets* [32] and *irreducible triangular sets* [36, 37]. For instance, the definition of regular set is as follows.

Definition 3.2 (Regular Set). A triangular set $[T_1, \dots, T_r]$ is said to be *regular* or called a *regular set* if no regular zero of \mathcal{T}_i annihilates the initial of T_{i+1} for all $i = 1, \dots, r - 1$, where $\mathcal{T}_i = [T_1, \dots, T_i]$ and a regular zero of \mathcal{T}_i is a zero of \mathcal{T}_i such that the variables other than the leading variables of T_1, \dots, T_i are not specific values.

Effective algorithms have been developed by Wu [37, 38], Lazard [17], Kalkbrener [15] and others [2, 12, 14, 19, 23] to decompose any polynomial set \mathcal{P} into finitely many triangular sets $\mathcal{T}_1, \dots, \mathcal{T}_k$ with different properties such that

$$\text{Zero}(\mathcal{P}) = \bigcup_{i=1}^k \text{Zero}(\mathcal{T}_i / \text{ini}(\mathcal{T}_i)), \tag{3.1}$$

where $\text{ini}(\mathcal{T}_i)$ denotes the set of initials of all polynomials in \mathcal{T}_i . Wang [31–33] has also proposed efficient algorithms for computing the triangular decomposition as (3.1). His methods are more general, and can be used to decompose a polynomial system $[\mathcal{P}, \mathcal{Q}]$ (both \mathcal{P} and \mathcal{Q} are polynomial sets) into finite *triangular systems* $[\mathcal{T}_i, \mathcal{S}_i], i = 1, \dots, k$ such that

$$\text{Zero}(\mathcal{P} / \mathcal{Q}) = \bigcup_{i=1}^k \text{Zero}(\mathcal{T}_i / \mathcal{S}_i),$$

where $[\mathcal{T}_i, \mathcal{S}_i]$ could be *fine triangular systems* [31], *regular systems* [33], or *simple systems* [32], respectively corresponding to triangular sets, regular sets, or simple sets. Wang’s algorithms have been implemented in the Maple package Epsilon, which serves as one of our main computational tools here. Let $\mathcal{P} = [x_1x_2^2 + x_3^2, x_1x_3 + x_2]$ and $\mathcal{Q} = \{x_1\}$ with $x_1 < x_2 < x_3$. For example, applying the RegSer function in Epsilon, the polynomial system $[\mathcal{P}, \mathcal{Q}]$ may be decomposed into 2 regular systems:

$$\left[[x_2, x_3], \{x_1, x_1^3 + 1\} \right], \quad \left[[x_1^3 + 1, x_1x_3 + x_2], \{\} \right]. \tag{3.2}$$

We use $\mathcal{P}_{\leq s}$ to denote the subset of \mathcal{P} , where only polynomials with leading variable smaller than and equal to x_s are contained. $\text{Proj}_s Z$ stands for the projection of a zero set Z into the subspace $\mathbb{C}^s := \{(x_1, \dots, x_s) \mid x_1, \dots, x_s \in \mathbb{C}\}$ — e.g. if $Z = \{(0, 1, 2), (5, 6, 7)\}$, then $\text{Proj}_2 Z = \{(0, 1), (5, 6)\}$. The following proposition indicates that the projection of the zero set of a polynomial system can readily be obtained by computing its regular systems.

Proposition 3.1 (Projection Property [36]). *Suppose that $[\mathcal{T}_i, \mathcal{S}_i], i = 1, \dots, k$ are regular systems of the polynomial system $[\mathcal{P}, \mathcal{Q}]$. Then, for any $s = 1, \dots, n$,*

$$\text{Proj}_s \text{Zero}(\mathcal{P} / \mathcal{Q}) = \bigcup_{i=1}^k \text{Zero}((\mathcal{T}_i)_{\leq s} / (\mathcal{S}_i)_{\leq s}).$$

A triangular set in which all polynomials other than the first are linear with respect to their corresponding leading variables is said to be *quasi-linear*. Furthermore, a triangular system $[\mathcal{T}, \mathcal{S}]$ is said to be *quasi-linear* if \mathcal{T} is quasi-linear — e.g. the two regular systems in (3.2) are all quasi-linear by definition. A quasi-linear triangular set \mathcal{T} (or triangular system $[\mathcal{T}, \mathcal{S}]$) has extremely simple structure, with its zeros easily obtained by analysing the first polynomial in \mathcal{T} . The following theorem paves the way for decomposing a polynomial system into a finite number of quasi-linear triangular systems (cf. Ref. [20] for a proof).

Theorem 3.1. *Let $[\mathcal{T}, \mathcal{S}]$ be a regular system with $\mathcal{T} = [T_1(\mathbf{u}, x_1), \dots, T_r(\mathbf{u}, x_1, \dots, x_r)]$, where x_1, \dots, x_r are respectively the leading variables of T_1, \dots, T_r and \mathbf{u} are its parameters. For a random sequence of integers c_2, \dots, c_r , the probability is 1 that all simple systems (under the same ordering $x_1 < \dots < x_r$) of $[\mathcal{T}^*, \mathcal{S}^*]$ are quasi-linear, where \mathcal{T}^* and \mathcal{S}^* are obtained from \mathcal{T} and \mathcal{S} respectively by replacing x_1 with $x_1 + c_2x_2 + \dots + c_r x_r$.*

Triangular systems produced by triangular decomposition are often quasi-linear — cf. (3.2) for example. However, this is not always the case. In practice, to obtain quasi-linear triangular systems of a polynomial system $[\mathcal{P}, \mathcal{Q}]$, one may first decompose $[\mathcal{P}, \mathcal{Q}]$ into regular systems $[\mathcal{T}_i, \mathcal{S}_i]$, $i = 1, \dots, s$. Then for those regular systems, say $[\mathcal{T}_j, \mathcal{S}_j]$, $j = 1, \dots, t$, that are not quasi-linear, from Theorem 3.1 one may randomly choose c_{2j}, \dots, c_{rj} and decompose $[\mathcal{T}_j^*, \mathcal{S}_j^*]$ into simple systems until the resulting triangular systems are all quasi-linear.

3.2. Semi-algebraic system and border polynomial

Definition 3.3 (Semi-algebraic System). A semi-algebraic system is an equation set of the form

$$\begin{cases} F_1(u_1, \dots, u_s, x_1, \dots, x_n) = 0, \\ \vdots \\ F_n(u_1, \dots, u_s, x_1, \dots, x_n) = 0, \\ P_1(u_1, \dots, u_s, x_1, \dots, x_n) \leq 0, \\ \vdots \\ P_r(u_1, \dots, u_s, x_1, \dots, x_n) \leq 0, \end{cases}$$

where F_i and P_j are polynomials over \mathbb{R} with u_1, \dots, u_s as their parameters and x_1, \dots, x_n as their variables, and the symbol \leq represents $>, \geq, <, \leq$ or \neq .

Semi-algebraic systems are often prevalent in practice. Indeed, real solutions of particular semi-algebraic systems can be seen to characterise various problems in science and engineering. For example, Problem 2 in Section 2 can be reduced to determining whether there exists any real solution of $\{F_1 = 0, F_2 = 0, F_3 = 0, t_1 \geq 0, t_2 \geq 0, t_3 \geq 0\}$.

Definition 3.4. Let $A = \sum_{i=0}^m a_i x^i$ and $B = \sum_{j=0}^l b_j x^j$ be two univariate polynomial in x ,

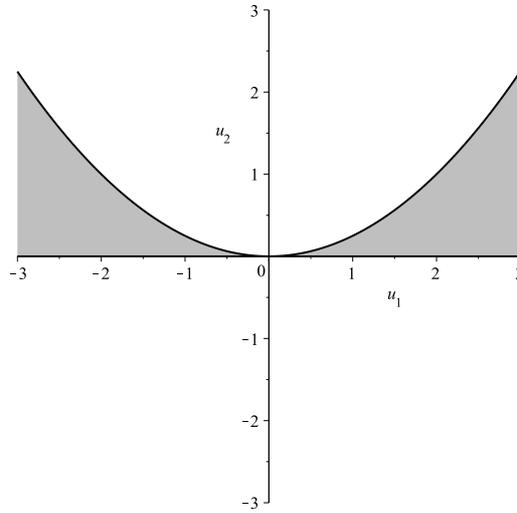


Figure 1: The parameter space for Example 3.1.

Example 3.1. Consider the semi-algebraic system $\{x^2 + u_1x + u_2 = 0, x > 0\}$. We have that $\text{Discr}(x^2 + u_1x + u_2) = u_1^2 - 4u_2$ and $\text{Res}(x^2 + u_1x + u_2, x) = u_2$, so the border polynomial of this system is $u_2(u_1^2 - 4u_2)$. Its zeros divide the parameters space $\{(u_1, u_2) \mid u_1, u_2 \in \mathbb{R}\}$ into 4 separate regions as shown in Fig. 1. From Theorem 3.2, the number of distinct real solutions of the considered semi-algebraic system is invariant, which is obvious for this simple example.

4. Solving Problem 1 using Triangular Decomposition

Let $\mathcal{F} = [F_1, F_2, F_3]$ be the set of polynomials in system (2.3). Decomposing $[\mathcal{F}, \emptyset]$ into regular systems under the variable ordering $a < b < c < t_1 < t_2 < t_3$ by the RegSer function in the Epsilon package, we obtain 6 regular systems $[\mathcal{T}_1, \mathcal{S}_1], \dots, [\mathcal{T}_6, \mathcal{S}_6]$ satisfying

$$\text{Zero}(\mathcal{F}) = \bigcup_{i=1}^6 \text{Zero}(\mathcal{T}_i/\mathcal{S}_i). \tag{4.1}$$

To save space, we give only the first branch $[\mathcal{T}_1, \mathcal{S}_1] = [[T_1, T_2, T_3], \{S_1, S_2\}]$ with

$$\begin{aligned} T_1 &= I_1 t_1^4 + (48 a^3 - 144 ab + 144 c)t_1^3 + (-18 a^4 - 72 b^2 + 72 a^2)t_1^2 \\ &\quad + a^6 - 6 a^4 b - 48 a^3 c + 36 a^2 b^2 + 144 abc - 72 b^3 - 72 c^2, \\ T_2 &= I_2 t_2 + J_2, \quad T_3 = -t_3 + t_2 - t_1 - a, \quad I_1 = -36 a^2 + 72 b, \\ I_2 &= -6 t_1^2 + 3 a^2 - 6 b, \quad J_2 = (-3 a^2 + 6 b)t_1 + 2 a^3 - 6 ab + 6 c, \\ S_1 &= a^2 - 2 b, \quad S_2 = a^6 - 6 a^4 b - 48 a^3 c + 36 a^2 b^2 + 144 abc - 72 b^3 - 72 c^2. \end{aligned} \tag{4.2}$$

For any given value of the parameters a, b, c such that $S_1 \neq 0, S_2 \neq 0$ or simply $S_1 S_2 \neq 0$, the initial I_1 of the polynomial T_1 is non-zero. Furthermore, since $\text{Res}(T_1, I_2) = 1296 S_2^2$,

the initial I_2 of T_2 is also non-zero if $S_1S_2 \neq 0$ and $T_1 = 0$. Thus provided that $S_1S_2 \neq 0$, \mathcal{T}_1 always has complex zeros, which can readily be obtained by solving $T_1 = 0, T_2 = 0$ and $T_3 = 0$ for t_1, t_2 and t_3 , respectively. Hence $\text{Zero}(\mathcal{T}_1/\mathcal{S}_1) \neq \emptyset$, and similarly one can prove that $\text{Zero}(\mathcal{T}_i/\mathcal{S}_i) \neq \emptyset$ for $i = 2, \dots, 6$. We formalise these results in the following proposition:

Proposition 4.1 (Non-emptiness — cf. Ref [36]). *For any regular system $[\mathcal{T}, \mathcal{S}]$, we have $\text{Zero}(\mathcal{T}/\mathcal{S}) \neq \emptyset$.*

To solve Problem 1, we need to project each $\text{Zero}(\mathcal{T}_i/\mathcal{S}_i)$ into the complex parameter space $\mathbb{C}^3 = \{(a, b, c) \mid a, b, c \in \mathbb{C}\}$ and check whether or not the projections cover the entire parameter space. Let us use $[\overline{\mathcal{T}}_i, \overline{\mathcal{S}}_i]$ to denote the regular system corresponding to the projection of $\text{Zero}(\mathcal{T}_i/\mathcal{S}_i)$. From the projection property (Proposition 3.1), these $[\overline{\mathcal{T}}_i, \overline{\mathcal{S}}_i]$ are easily obtained — viz.

$$\begin{aligned}
 [\overline{\mathcal{T}}_1, \overline{\mathcal{S}}_1] &= [[], \{S_1, S_2\}], \\
 [\overline{\mathcal{T}}_2, \overline{\mathcal{S}}_2] &= [[S_1, U_1], \{\}], \\
 [\overline{\mathcal{T}}_3, \overline{\mathcal{S}}_3] &= [[S_1], \{U_2, U_3\}], \\
 [\overline{\mathcal{T}}_4, \overline{\mathcal{S}}_4] &= [[S_2], \{S_1\}], \\
 [\overline{\mathcal{T}}_5, \overline{\mathcal{S}}_5] &= [[S_2], \{S_1\}], \\
 [\overline{\mathcal{T}}_6, \overline{\mathcal{S}}_6] &= [[S_1, U_4], \{\}],
 \end{aligned} \tag{4.3}$$

with $U_1 = a^3 - 6c, \quad U_2 = a^3 - 3ab + 3c,$
 $U_3 = a^6 - 6a^4b - 48a^3c + 144abc - 72c^2, \quad U_4 = a^6 - 12a^3c + 36c^2.$

We now prove that zeros of the above systems cover the entire parameter space \mathbb{C}^3 . It is obvious that

$$\text{Zero}(\overline{\mathcal{T}}_1/\overline{\mathcal{S}}_1) \cup \text{Zero}(\overline{\mathcal{T}}_4/\overline{\mathcal{S}}_4) = \mathbb{C}^3 \setminus \text{Zero}([S_1]),$$

so that

$$\text{Zero}(\overline{\mathcal{T}}_1/\overline{\mathcal{S}}_1) \cup \text{Zero}(\overline{\mathcal{T}}_4/\overline{\mathcal{S}}_4) \cup \text{Zero}(\overline{\mathcal{T}}_3/\overline{\mathcal{S}}_3) = \mathbb{C}^3 \setminus \Lambda,$$

where $\Lambda = \text{Zero}([S_1, U_2]) \cup \text{Zero}([S_1, U_3])$. Furthermore, it can be proved from the PolynomialIdeals package in Maple that

$$\langle S_1, U_1 \rangle \cap \langle S_1, U_4 \rangle \subseteq \sqrt{\langle S_1, U_2 \rangle}, \quad \langle S_1, U_1 \rangle \cap \langle S_1, U_4 \rangle \subseteq \sqrt{\langle S_1, U_3 \rangle},$$

where $\langle \mathcal{P} \rangle$ is the polynomial ideal generated by the polynomial set \mathcal{P} and $\sqrt{\mathcal{I}}$ is the radical of the ideal \mathcal{I} . From basic theory of polynomial algebra,

$$\begin{aligned}
 \text{Zero}([S_1, U_2]) &\subseteq \text{Zero}([S_1, U_1]) \cup \text{Zero}([S_1, U_4]) = \text{Zero}(\overline{\mathcal{T}}_1/\overline{\mathcal{S}}_1) \cup \text{Zero}(\overline{\mathcal{T}}_6/\overline{\mathcal{S}}_6), \\
 \text{Zero}([S_1, U_3]) &\subseteq \text{Zero}([S_1, U_1]) \cup \text{Zero}([S_1, U_4]) = \text{Zero}(\overline{\mathcal{T}}_1/\overline{\mathcal{S}}_1) \cup \text{Zero}(\overline{\mathcal{T}}_6/\overline{\mathcal{S}}_6).
 \end{aligned}$$

Consequently,

$$\bigcup_{i=1}^6 \text{Zero}(\overline{\mathcal{T}}_i/\overline{\mathcal{S}}_i) = \mathbb{C}^3$$

such that the system (2.3) has complex solutions for all parameter assignments.

One may observe that, among all branches in (4.1), only the projection of $\text{Zero}(\mathcal{T}_1/\mathcal{S}_1)$ is of the same dimension as the parameter space \mathbb{C}^3 . We call $[\mathcal{T}_1, \mathcal{S}_1]$ the *main branch* of the regular systems of \mathcal{F} .

5. Algorithm for Optimal Control

We now propose a systematic approach to drive an m -order linear system Φ from any given initial state \bar{u} to the origin in minimum time. In Remark 2.1, it is pointed out that by solving particular polynomial equations one can decide how long to keep a certain constant (+1 or -1) for the optimal control v . However, discrete time implementations are typically used in engineering. Suppose a sample of the state of a system Φ is taken at small time steps, say every 0.001 seconds. From the algorithm Switch presented in subsection 5.1, we know that $\text{Switch}(\Phi, \bar{u})$ should be the first recommended value for the time-optimal control for driving Φ to the origin, so let $v(t) = \text{Switch}(\Phi, \bar{u})$ for $0 \leq t < 0.001$. Under the control of this $v(t)$, the system will move to a new state when $t = 0.001$. By sampling, we get the state and denote it as \bar{u}_1 . Similarly, on setting $v(t) = \text{Switch}(\Phi, \bar{u}_1)$ for $0.001 \leq t < 0.002$, the system will arrive at \bar{u}_2 . In this way, the system Φ would be driven to the target through the path $\bar{u} \rightarrow \bar{u}_1 \rightarrow \bar{u}_2 \rightarrow \dots \rightarrow \mathbf{0}$, and the whole optimal control $v(t)$ could finally be obtained.

5.1. Switching algorithm

For any vector of numbers $\bar{u} = (\bar{u}_1, \dots, \bar{u}_m)$, we use $*|_{\bar{u}}$ to denote the result of $*$ specified at $\mathbf{u} = \bar{u}$, where $*$ could be a polynomial or a polynomial set. In Algorithm 1 we formalise the steps in computing the recommended present value for the time-optimal control to move a system from its current state to the origin. (The algorithm has the system and its current state as the input.) Our method is similar to counting distinct real solutions of a semi-algebraic system [20], which serves as the main computational tools in analysing the multiplicity of competitive equilibria of semi-algebraic economies.

In Algorithm 1, $\text{Simplify}(\{T'_1 = 0, \dots, T'_m = 0, t_1 \geq 0, \dots, t_m \geq 0\})$ translates the semi-algebraic system $\{T'_1 = 0, \dots, T'_m = 0, t_1 \geq 0, \dots, t_m \geq 0\}$ into an equivalent simpler system $\{T'_1(t_1) = 0, A_1(t_1) \geq 0, \dots, A_m(t_1) \geq 0\}$ with a single variable t_1 — cf. Ref. [20] for further details. The operation $\text{Split}(\{T'_1 = 0, A_1 \geq 0, \dots, A_m \geq 0\})$ returns a set of finitely many semi-algebraic systems with only strict inequalities such that the solution set remains the same. Indeed, if the greatest common divisor of T'_1 and A_1 is 1, then $A_1 \geq 0$ can be replaced by $A_1 > 0$; otherwise, the system should be split into two — viz.

$$\left\{ \begin{array}{l} \gcd(T'_1, A_1) = 0, \\ A_2 \geq 0, \\ \vdots \\ A_m \geq 0, \end{array} \right. \quad \left\{ \begin{array}{l} T'_1 / \gcd(T'_1, A_1) = 0, \\ A_1 > 0, \\ A_2 \geq 0, \\ \vdots \\ A_m \geq 0. \end{array} \right.$$

Algorithm 1: $v(0) := \text{Switch}(\Phi, \bar{u})$ **Input:** Φ — an m -order linear system; \bar{u} — the initial state of Φ .**Output:** $v(0)$ — the recommended present value (either +1 or -1) for the time-optimal control of driving Φ from \bar{u} to the origin $\mathbf{0}$. \mathcal{F} := the polynomial set corresponding to Φ with \mathbf{u} as its parameters and t_1, \dots, t_m as its variables (see Section 2 for details);Let $\mathbf{u} < t_1 < \dots < t_m$ and decompose \mathcal{F} into regular systems $[\mathcal{T}_i, \mathcal{S}_i]$, $i = 1, \dots, l$; $[\overline{\mathcal{T}}_i, \overline{\mathcal{S}}_i]$:= the regular system corresponding to $\text{Proj}_m \text{Zero}(\mathcal{T}_i/\mathcal{S}_i)$. Δ := the set of indices i 's such that $\bar{u} \in \text{Zero}(\overline{\mathcal{T}}_i/\overline{\mathcal{S}}_i)$ for $i \in \{1, \dots, l\}$; $\Gamma := \emptyset$;**for** $j \in \Delta$ **do** Suppose that $\mathcal{T}_j|_{\bar{u}} = [T'_1(t_1), \dots, T'_m(t_1, \dots, t_m)]$ is quasi-linear, for
 otherwise we apply the quasi-linearization technique in Section 3.1; $\mathbb{S} := \text{Simplify}(\{T'_1 = 0, \dots, T'_m = 0, t_1 \geq 0, \dots, t_m \geq 0\})$; $\Gamma := \Gamma \cup \text{Split}(\mathbb{S})$;**end****for** $\mathbb{U} \in \Gamma$ **do** Suppose that $\mathbb{U} = \{F(t_1) = 0, P_1(t_1) > 0, \dots, P_s(t_1) > 0\}$; $S := \text{Iso}(F, \prod_{k=1}^s P_k)$; C := the open intervals of the complement of S such that $P_k > 0$ for all
 $k = 1, \dots, s$; **if** $\text{Count}(F, C) > 0$ **then** $v(0) := (-1)^m$; **return**; **end****end** $v(0) := -(-1)^m$;**return**;

Proceeding in the same way for the other ' \geq ' inequalities of the above semi-algebraic systems, we obtain finitely many systems of the form

$$\left\{ \begin{array}{l} F(t_1) = 0, \\ P_1(t_1) > 0, \\ \vdots \\ P_s(t_1) > 0, \end{array} \right.$$

where $\text{gcd}(F, P_k) = 1$ for all $k = 1, \dots, s$. It is obvious that the solutions of all the semi-algebraic systems eventually obtained are the same as those from $\{T'_1(t_1) = 0, A_1(t_1) \geq 0, \dots, A_m(t_1) \geq 0\}$. For any univariate polynomials F and G such that $\text{gcd}(F, G) = 1$, the operation $\text{Iso}(F, G)$ isolates the real zeros of G — e.g. using the modified Uspensky

algorithm [6]. Thus $\text{Iso}(F, G)$ returns a sequence of closed intervals $[f_1, g_1], \dots, [f_m, g_m]$ such that

- f_i, g_i are all rational numbers,
- $f_1 \leq g_1 < f_2 \leq g_2 < \dots < f_m \leq g_m$,
- $[f_i, g_i] \cap [f_j, g_j] = \emptyset$ for $i \neq j$,
- each $[f_i, g_i]$ contains one and only one real zero of G ,
- every $[f_i, g_i]$ covers no real zero of F .[‡]

Finally, for any univariate polynomial F and a set C consisting of finite open intervals, the operation $\text{Count}(F, C)$ returns the number of distinct real solutions of F located in C and Sturm's theorem may be invoked.

Proof of Algorithm 1. As pointed out in Section 2, the values for the optimal control v of the system Φ depend upon whether \mathcal{F} has any real solution such that $t_1 \geq 0, \dots, t_m \geq 0$. It is notable that \mathcal{F} is obtained from the inverse problem (i.e. driving Φ from $\mathbf{0}$ to \mathbf{u}), hence both the order and the sign of the recommended value sequence of v may need to be reversed. Thus if $c = +1, -1, +1$ in the inverse problem for a third-order system, then the recommended values of v in the direct problem are $-1, +1, -1$; on the other hand, if $c = +1, -1, +1, -1$ in the inverse problem for a fourth-order system, we also have $c = +1, -1, +1, -1$ in the direct problem. In Algorithm 1, we set the present value $v(0) = (-1)^m$ if \mathcal{F} has non-negative solutions, and otherwise we set $v(0) = -(-1)^m$.

We now illustrate how Algorithm 1 essentially identifies whether or not \mathcal{F} has non-negative solutions. Obviously,

$$\text{Zero}(\mathcal{F}) = \bigcup_{i=1}^l \text{Zero}(\mathcal{T}_i / \mathcal{S}_i),$$

so that

$$\text{Zero}(\mathcal{F}|_{\bar{u}}) = \bigcup_{j \in \Delta} \text{Zero}(\mathcal{T}_j|_{\bar{u}} / \mathcal{S}_j|_{\bar{u}}),$$

and hence only $\mathcal{T}_j|_{\bar{u}}, j \in \Delta$ need to be considered when counting solutions of \mathcal{F} . Consider the first **for** loop. Suppose that

$$\begin{aligned} & \text{Simplify}(\{T'_1 = 0, \dots, T'_m = 0, t_1 \geq 0, \dots, t_m \geq 0\}) \\ & = \{T'_1(t_1) = 0, A_1(t_1) \geq 0, \dots, A_m(t_1) \geq 0\}. \end{aligned}$$

Then the problem is reduced to determining whether, for each $j \in \Delta$, there exists any real zero of the semi-algebraic system

$$\begin{cases} T'_1 = 0, \\ A_1 \geq 0, \\ \vdots \\ A_m \geq 0. \end{cases}$$

[‡]This can be realized because $\text{gcd}(F, G) = 1$.

From the property of the sub-procedure Split, the problem is equivalent to whether any system in Γ has real solutions. Moreover, we have $\gcd(F, \prod_{k=1}^s P_k) = 1$ for any $\{F(t_1) = 0, P_1(t_1) > 0, \dots, P_s(t_1) > 0\} \in \Gamma$. In the second for loop, suppose that the result of $S := \text{Iso}(F, \prod_{k=1}^s P_k)$ is of the form $[f_1, g_1], \dots, [f_m, g_m]$. Then the complement of S is

$$(-\infty, f_1), (g_1, f_2), \dots, (g_m, +\infty).$$

It is obvious that the signs of P_1, \dots, P_s must be fixed in each of the above open intervals, and can be identified by verifying at a sample point. Consequently, C could readily be obtained and $\text{Count}(F, C)$ counts distinct real solutions of F located in C . If $\text{Count}(F, C) > 0$ for some $\mathbb{U} \in \Gamma$, then \mathbb{U} has real solutions, which proves that \mathcal{F} has at least one real solution such that $t_1 \geq 0, \dots, t_m \geq 0$. □

5.2. Illustrative case

Consider the third-order system (2.1) in Section 2 and the initial state $\bar{u} = (a, b, c) = (1, 1, 1)$. Since $S_1|_{(1,1,1)} = -1 \neq 0$ and $S_2|_{(1,1,1)} = -17 \neq 0$, we have $(1, 1, 1) \in \text{Zero}(\overline{\mathcal{T}}_1/\overline{\mathcal{T}}_1)$. Moreover, it can be verified that $(1, 1, 1) \notin \text{Zero}(\overline{\mathcal{T}}_i/\overline{\mathcal{T}}_i)$ for any $i = 2, \dots, 6$, so from (4.1) we have that $\text{Zero}(\mathcal{F}|_{(1,1,1)}) = \text{Zero}(\mathcal{T}_1|_{(1,1,1)})$. Thus $\Delta = \{1\}$, and it remains to determine whether $\mathcal{T}_1|_{(1,1,1)}$ with non-negative components has at least one real zero. We have $\mathcal{T}_1|_{(1,1,1)} = [T'_1, T'_2, T'_3]$ with

$$\begin{aligned} T'_1 &= 36 t_1^4 + 48 t_1^3 - 18 t_1^2 - 17, \\ T'_2 &= I t_2 + J, \quad I = -6 t_1^2 - 3, \quad J = 3 t_1 + 2, \\ T'_3 &= -t_3 + t_2 - t_1 + 1. \end{aligned}$$

Other than the first, the polynomials in $\mathcal{T}_1|_{(1,1,1)}$ are linear with respect to their leading variables, such that $\mathcal{T}_1|_{(1,1,1)}$ is quasi-linear.

Let us demonstrate how the sub-procedure Simplify($\{T'_1 = 0, T'_2 = 0, T'_3 = 0, t_1 \geq 0, t_2 \geq 0, t_3 \geq 0\}$) works. On solving $T'_3 = 0$ for t_3 , $T'_2 = 0$ for t_2 and then substituting the solutions $t_3 = t_2 - t_1 + 1$, $t_2 = -J/I$ successively into $t_3 \geq 0$ and $t_2 \geq 0$, we obtain $-J/I - t_1 + 1 \geq 0$ and $-J/I \geq 0$, respectively. The problem is thus reduced to determining whether there are real solutions of the following system with only one variable t_1 :

$$\begin{cases} T'_1 = 0, \\ t_1 \geq 0, \\ A/A' \geq 0, \\ B/B' \geq 0, \end{cases}$$

where $A = -J$, $A' = I$, $B = -J - t_1 I + I$ and $B' = I$. Polynomials rather than rational functions are preferred in the computation. Since $I(\bar{t}_1) \neq 0$ for any zero \bar{t}_1 of T'_1 from the definition of a regular system, one could replace the last two inequalities by $A^* = AA' =$

$-IJ \geq 0$ and $B^* = BB' = (-J - t_1 I + I)I \geq 0$, respectively. We arrive at

$$\begin{cases} T'_1 = 0, \\ t_1 \geq 0, \\ A^* \geq 0, \\ B^* \geq 0. \end{cases}$$

Since $\gcd(T'_1, t_1) = 1$, $\gcd(T'_1, A^*) = 1$ and $\gcd(T'_1, B^*) = 1$, we have $\Gamma = \{\mathbb{U}\}$ where

$$\mathbb{U} = \begin{cases} T'_1 = 0, \\ t_1 > 0, \\ A^* > 0, \\ B^* > 0. \end{cases}$$

In order to determine whether system \mathbb{U} has real solutions, let us write $S := \text{Iso}(T'_1, t_1 A^* B^*)$ to obtain the sorted sequence of intervals

$$[-1, -1/2], \quad [0, 0], \quad [3/4, 1].$$

Consequently, the real zeros of T'_1 must lie in

$$(-\infty, -1), \quad (-1/2, 0), \quad (0, 3/4), \quad (1, +\infty).$$

Furthermore, in each of these open intervals the signs of t_1, A^* and B^* are invariant, and can be identified by testing at a sample point in each interval. For example, to determine the sign of A^* on $(-\infty, -1)$, we have that $A^*(-2) = -108 < 0$, so A^* is negative at every point in $(-\infty, -1)$. Proceeding in this way for other intervals, we conclude that the inequality constraints $t_1 > 0$, $A^* > 0$ and $B^* > 0$ are satisfied only on $C = (0, 3/4)$. Finally, by computing the Sturm sequence we obtain that $\text{Count}(T'_1, C) = 1 > 0$.

In conclusion, for $(a, b, c) = (1, 1, 1)$ the system (2.3) has exactly one non-negative solution, so we should set $v(0) = -1$ as the recommended present value for the time-optimal control to drive system (2.1) from $(1, 1, 1)$ to the origin $(0, 0, 0)$.

6. Moving the Computational Burden Off-line

The switching algorithm presented in Section 5 involves the computation of real root isolation and a Sturm sequence, which may be intractable for large systems and near invalid for real-time control. In this section, we modify the switching algorithm to make it available for real-time applications. The key idea is to divide the computation into two phases — the off-line and the on-line. The computational burden is moved off-line as it is only necessary to verify that particular inequalities are satisfied in the on-line stage, and the computation is extremely fast.

6.1. The off-line phase

The example given in Section 2 serves to demonstrate how to compute control strategies in the off-line phase, and we then formalise the steps in an algorithm.

Let us reconsider the main branch $[\mathcal{T}_1, \mathcal{S}_1] = [[T_1, T_2, T_3], \{S_1, S_2\}]$ for regular systems, where S_1, S_2 are polynomials in a, b, c — cf. (4.2). For any $\bar{a}, \bar{b}, \bar{c} \in \mathbb{R}$ such that $S_1|_{(\bar{a}, \bar{b}, \bar{c})} \neq 0$ and $S_2|_{(\bar{a}, \bar{b}, \bar{c})} \neq 0$, from (4.3) we have that $(\bar{a}, \bar{b}, \bar{c}) \notin \text{Zero}(\overline{\mathcal{T}_i}/\overline{\mathcal{S}_i})$ for any $i = 2, \dots, 6$. Consequently, the zeros of $\mathcal{F}|_{(\bar{a}, \bar{b}, \bar{c})}$ should be same as those for $\mathcal{T}_1|_{(\bar{a}, \bar{b}, \bar{c})}$. Since the main branch is already quasi-linear, one can easily solve $T_3 = 0$ for t_3 and $T_2 = 0$ for t_2 . Substituting the solutions $t_3 = t_2 - t_1 - a$, $t_2 = -J_2/I_2$ successively into $t_3 \geq 0$ and $t_2 \geq 0$, we obtain $-J_2/I_2 - t_1 - a \geq 0$ and $-J_2/I_2 \geq 0$, respectively. Thus provided that $S_1 S_2 \neq 0$, the problem is reduced to determining the condition on a, b, c such that there exist real zeros of the semi-algebraic system

$$\begin{cases} T_1 = 0, \\ t_1 \geq 0, \\ C \geq 0, \\ D \geq 0, \end{cases} \quad (6.1)$$

where $C = -I_2 J_2$, $D = (-J_2 - t_1 I_2 + I_2) I_2$. Moreover, let us suppose that

$$\text{Res}(T_1, t_1) \cdot \text{Res}(T_1, C) \cdot \text{Res}(T_1, D) \neq 0.$$

From Proposition 3.2, we know that T_1 has no common zero with t_1 , C and D , so system (6.1) can be transformed further to

$$\begin{cases} T_1 = 0, \\ t_1 > 0, \\ C > 0, \\ D > 0. \end{cases} \quad (6.2)$$

We now construct the border polynomial of (6.2), and its square free part $B = B_1 B_2 B_3 B_4$, where

$$\begin{aligned} B_1 &= S_1 = a^2 - 2b, \\ B_2 &= S_2 = a^6 - 6a^4b - 48a^3c + 36a^2b^2 + 144abc - 72b^3 - 72c^2, \\ B_3 &= a^6 + 6a^4b - 48a^3c + 36a^2b^2 - 144abc + 72b^3 - 72c^2, \\ B_4 &= a^6 - 6a^4b - 144a^3c + 84a^2b^2 + 432abc - 200b^3 - 216c^2, \end{aligned}$$

such that the zeros of B divide the parameter space \mathbb{R}^3 into separated regions. Fig. 2 shows the graphs of $B_1 = 0$, $B_2 = 0$, $B_3 = 0$ and $B_4 = 0$ (shown in blue, green, white and red, respectively.) From Theorem 3.2, the number of real solutions of (6.2) is invariant in a fixed region. Thus one may choose a sample point from each region, which can be done systematically, using for example the cylindrical algebraic decomposition method [5]. The problem then is to determine whether there is any real solution of the result of (6.2)

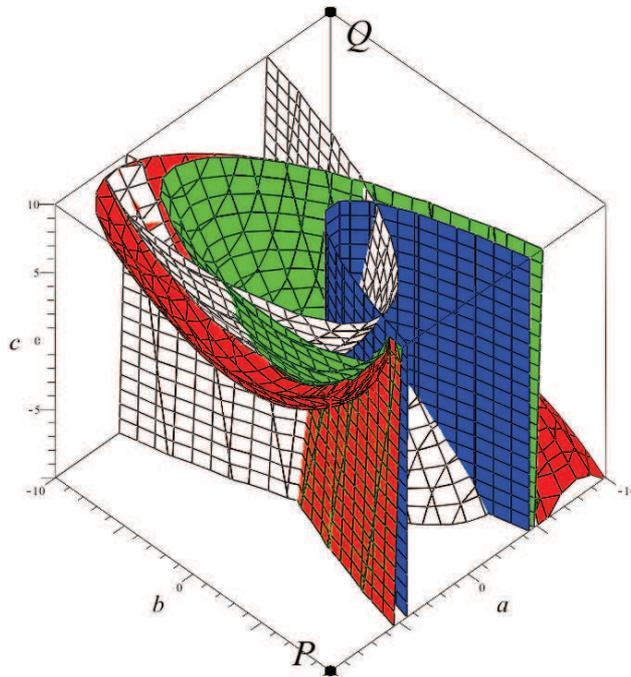


Figure 2: Partitions of the parameter space of (6.2).

specified at these sample points. For example, at $P = (10, 10, -10)$ shown in Fig. 2, the system (6.2) becomes

$$\begin{cases} -2880 t_1^4 + 32160 t_1^3 - 115200 t_1^2 + 1016800 = 0, \\ t_1 > 0, \\ (-120 t_1 + 670)(3 t_1^2 - 120) > 0, \\ (-3 t_1^3 + 30 t_1^2 - 530)(3 t_1^2 - 120) > 0. \end{cases}$$

On applying the method proposed in Section 5, we know that the above system has at least one real solution, so $v(0)$ should be -1 if $(\bar{a}, \bar{b}, \bar{c})$ in the same region as P . Proceeding further in the same way, the recommended present value $v(0)$ for the optimal control therefore can be obtained for other regions.

To succinctly describe the way to set $v(0)$, we give simple and formal representations of these regions, beginning with utilising the sign of $B_i, i = 1, \dots, 4$. For example, as

$$B_1|_P = 80 > 0, \quad B_2|_P = 1016800 > 0, \quad B_3|_P = 2648800 > 0, \quad B_4|_P = 2026400 > 0,$$

we expect that the region where P is located might be described by $B_i > 0, i = 1, \dots, 4$. Unfortunately, this is not the case. Consider the point $Q = (-10, -10, 10)$. We also have $B_i|_Q > 0$ for all $i = 1, \dots, 4$, so the Boolean formula $\bigwedge_{i=1}^4 B_i > 0$ corresponds to at least two different regions. However, we may introduce an additional polynomial $A_1 = a^3 - 3ab + 3c$ that satisfies $A_1|_P = 670 \geq 0$ and $A_1|_Q = -1270 \leq 0$, so the two regions where points P

and Q lie can be described by

$$\bigwedge_{i=1}^4 B_i > 0 \wedge A_1 > 0 \quad \text{and} \quad \bigwedge_{i=1}^4 B_i > 0 \wedge A_1 < 0,$$

respectively. This illustrates that other polynomials could help in the characterisation of different regions. Generally, it would be fairly hard to find these polynomials by hand, but Yang *et al.* [40] have pointed out that they are contained in the so-called *generalised discriminant list* and can be selected by repeated trials. For our example, 4 additional polynomials may be needed — including A_1 above,

$$\begin{aligned} A_2 &= 3a^3 - 7ab + 3c, \\ A_3 &= 117a^8 - 400a^6b + 120a^5c - 496a^4b^2 - 2736a^3bc + 4176a^2b^3 \\ &\quad - 72a^2c^2 + 7296ab^2c - 5808b^4 - 3312bc^2, \end{aligned}$$

and A_4 , a complex polynomial of degree 15 with 27 terms. (We do not give A_4 here, due to limitation in space.)

Given the above preparation, we now describe how to set $v(0)$ for any given $(\bar{a}, \bar{b}, \bar{c})$ such that

$$N = S_1 \cdot S_2 \cdot \text{Res}(T_1, t_1) \cdot \text{Res}(T_1, C) \cdot \text{Res}(T_1, D) \cdot \prod_{i=1}^4 B_i \neq 0.$$

The optimal control $v(0)$ should be set as -1 if and only if system (2.3) has at least one non-negative solution, or if and only if one of the following Boolean formulae holds:

- $B_1 < 0 \wedge B_2 < 0 \wedge B_3 < 0 \wedge B_4 < 0 \wedge A_2 < 0 \wedge A_3 < 0 \wedge A_4 < 0,$
- $0 < B_1 \wedge 0 < B_2 \wedge B_3 < 0 \wedge A_1 < 0 \wedge A_2 < 0 \wedge A_3 < 0 \wedge A_4 < 0,$
- $0 < B_1 \wedge 0 < B_2 \wedge B_3 < 0 \wedge A_1 < 0 \wedge 0 < A_2 \wedge A_4 < 0,$
- $0 < B_1 \wedge 0 < B_2 \wedge 0 < B_3 \wedge 0 < B_4 \wedge A_1 < 0 \wedge 0 < A_2 \wedge A_4 < 0,$
- $0 < B_1 \wedge 0 < B_2 \wedge B_4 < 0 \wedge A_1 < 0 \wedge 0 < A_2 \wedge 0 < A_3 \wedge 0 < A_4,$
- $0 < B_1 \wedge 0 < B_2 \wedge B_3 < 0 \wedge 0 < B_4 \wedge A_1 < 0 \wedge A_3 < 0 \wedge 0 < A_4,$
- $0 < B_1 \wedge B_2 < 0 \wedge B_3 < 0 \wedge B_4 < 0 \wedge 0 < A_1 \wedge A_2 < 0 \wedge A_3 < 0,$
- $0 < B_1 \wedge 0 < B_2 \wedge B_3 < 0 \wedge 0 < B_4 \wedge 0 < A_1 \wedge A_2 < 0 \wedge 0 < A_3 \wedge 0 < A_4,$
- $0 < B_1 \wedge 0 < B_2 \wedge B_3 < 0 \wedge 0 < A_1 \wedge A_2 < 0 \wedge A_3 < 0 \wedge 0 < A_4,$
- $0 < B_1 \wedge B_3 < 0 \wedge B_4 < 0 \wedge 0 < A_1 \wedge A_2 < 0 \wedge 0 < A_3 \wedge 0 < A_4,$
- $0 < B_1 \wedge B_2 < 0 \wedge B_4 < 0 \wedge 0 < A_1 \wedge 0 < A_2 \wedge 0 < A_3 \wedge A_4 < 0,$
- $0 < B_1 \wedge 0 < B_2 \wedge B_3 < 0 \wedge B_4 < 0 \wedge 0 < A_1 \wedge 0 < A_2 \wedge A_3 < 0 \wedge A_4 < 0,$
- $0 < B_1 \wedge 0 < B_2 \wedge B_3 < 0 \wedge 0 < B_4 \wedge 0 < A_1 \wedge 0 < A_2 \wedge 0 < A_3 \wedge 0 < A_4,$
- $0 < B_1 \wedge 0 < B_2 \wedge B_3 < 0 \wedge 0 < B_4 \wedge 0 < A_1 \wedge 0 < A_2 \wedge A_3 < 0 \wedge 0 < A_4,$
- $0 < B_1 \wedge 0 < B_2 \wedge 0 < B_3 \wedge B_4 < 0 \wedge 0 < A_1 \wedge 0 < A_2 \wedge A_3 < 0 \wedge 0 < A_4,$

- $0 < B_1 \wedge 0 < B_2 \wedge 0 < B_3 \wedge 0 < B_4 \wedge 0 < A_1 \wedge 0 < A_2 \wedge A_3 < 0 \wedge A_4 < 0$,
- $0 < B_1 \wedge 0 < B_2 \wedge 0 < B_3 \wedge 0 < A_1 \wedge 0 < A_2 \wedge 0 < A_3 \wedge A_4 < 0$,
- $0 < B_1 \wedge B_3 < 0 \wedge B_4 < 0 \wedge 0 < A_1 \wedge 0 < A_2 \wedge A_3 < 0 \wedge 0 < A_4$,
- $0 < B_1 \wedge 0 < B_3 \wedge B_4 < 0 \wedge 0 < A_1 \wedge 0 < A_2 \wedge A_3 < 0 \wedge A_4 < 0$,
- $0 < B_1 \wedge 0 < B_3 \wedge B_4 < 0 \wedge 0 < A_1 \wedge 0 < A_2 \wedge 0 < A_3 \wedge 0 < A_4$,
- $0 < B_1 \wedge 0 < B_2 \wedge B_3 < 0 \wedge B_4 < 0 \wedge 0 < A_1 \wedge 0 < A_3 \wedge A_4 < 0$,
- $0 < B_1 \wedge 0 < B_2 \wedge B_3 < 0 \wedge 0 < B_4 \wedge 0 < A_1 \wedge A_3 < 0 \wedge A_4 < 0$,
- $0 < B_1 \wedge 0 < B_2 \wedge 0 < B_3 \wedge 0 < B_4 \wedge 0 < A_2 \wedge 0 < A_3 \wedge 0 < A_4$,
- $B_2 < 0 \wedge B_3 < 0 \wedge B_4 < 0 \wedge 0 < A_1 \wedge 0 < A_2 \wedge A_3 < 0 \wedge A_4 < 0$.

The above steps are formalised in Algorithm 2. The termination and correctness of the algorithm are obvious.

Remark 6.1. From our experiments, the runtime can be the bottleneck for Algorithm 2 rather than memory requirements. It should be mentioned that Yang *et al.* [40] proposed a more direct method that may also be used to compute the necessary and sufficient conditions above. Their method avoids the quasi-linearisation process and may be more efficient when the polynomials involved are of high degree.

Both Yang's approach and ours are quite time-consuming for large problems, as the complexities of existing algorithms for cylindrical algebraic decomposition prove to be double-exponential. For the third-order system given in Section 2, Algorithm 2 takes 139 seconds to terminate in Maple 17 running on AMD A8-6500 CPU 3.50 GHz with 20G RAM under Windows 7 OS. Moreover, we found that systems with higher order could not be resolved within 5 hours. However, efficiency is less important as the computation of this stage is performed off-line, and powerful parallel computers may be used to accelerate the process.

Furthermore, Safey El Din [27, 28] and others have proposed a new approach to find sample points of semi-algebraic sets, based on the computation of critical points. The complexity of their method is $\mathcal{O}(d^7 D^{4n})$, where d and D are the parameter number and the degree of the involved system, respectively. Thus it is reasonable to hope for a significant performance boost if the cylindrical algebraic decomposition procedure in our methods is replaced with the critical point computation. This is an interesting issue for further investigation.

Finally, we note that points in the parameter space \mathbb{R}^3 are covered in the sense of Lebesgue measure, except for those points that annihilate N . For these exceptional points, we may add the equation $N = 0$ to (6.1), with a and b might viewed as parameters while c , t_1 , t_2 and t_3 are the variables. Repeating this process, we finally cover all points in the parameter space and so obtain the complete necessary and sufficient conditions on the parameters a , b and c such that system (2.3) has non-negative solutions. (Since somewhat tedious, we have not listed the complete necessary and sufficient conditions.)

Algorithm 2: $\Omega, N := \text{BooleanF}(\Phi)$ **Input:** Φ — an m -order linear system.**Output:** Ω — the necessary and sufficient conditions that the present value $v(0)$ should be $(-1)^m$ for the time-optimal control of driving Φ from \mathbf{u} to the origin $\mathbf{0}$, provided that $N \neq 0$. \mathcal{F} := the polynomial set corresponding to Φ with \mathbf{u} as its parameters and t_1, \dots, t_m as its variables (see Section 2 for details);Let $\mathbf{u} < t_1 < \dots < t_m$ and decompose \mathcal{F} into regular systems $[\mathcal{T}_i, \mathcal{S}_i]$, $i = 1, \dots, l$; Δ := the set of indices i 's such that $[\mathcal{T}_i, \mathcal{S}_i]$ is the main branch; $\Gamma := \emptyset$; $\mathcal{B} := \emptyset$; $N = 1$;**for** $j \in \Delta$ **do** Suppose that $\mathcal{T}_j = [T_1(t_1), \dots, T_m(t_1, \dots, t_m)]$ is quasi-linear, for otherwise we apply the quasi-linearization technique in Section 3.1; $\mathbb{S} := \text{Simplify}(\{T_1 = 0, \dots, T_m = 0, t_1 \geq 0, \dots, t_m \geq 0\})$; $\Gamma := \Gamma \cup \{\mathbb{S}\}$; \mathcal{P} := the set of all the factors of the border polynomial of \mathbb{S} ; $\mathcal{B} := \mathcal{B} \cup \mathcal{P}$; $N := N \cdot \prod_{S \in \mathcal{S}_i} S$;**end** Θ := the set of sample points at all the regions of parameter space divided by zeros of polynomials in \mathcal{B} . Θ^* := the subset of Θ , where at least one $\mathbb{S} \in \Gamma$ has non-negative solutions.Suppose that the signs of all polynomials in \mathcal{B} can exactly characterise the regions corresponding to points in Θ^* , for otherwise we add to \mathcal{B} certain elements in the generalised discriminant list of $\mathbb{S} \in \Gamma$.Suppose that $\mathcal{B} = \{B_1, \dots, B_k\}$; $\Omega := \emptyset$;**for** $\bar{\mathbf{u}} \in \Theta^*$ **do** $\omega := B_1 \lesseqgtr \dots \lesseqgtr B_k \lesseqgtr 0$, where \lesseqgtr could be either $>$ or $<$ depending on $B_i(\bar{\mathbf{u}}) > 0$ or $B_i(\bar{\mathbf{u}}) < 0$; $\Omega := \Omega \cup \{\omega\}$;**end** $N := N \cdot \prod_{B \in \mathcal{B}} B$;**return**;

6.2. The on-line phase

In the on-line stage, we sample the state $\bar{\mathbf{u}}$ of the considered m -order system Φ and verify whether or not the complete necessary and sufficient conditions obtained by the algorithm `BooleanF` are satisfied when $\mathbf{u} = \bar{\mathbf{u}}$ — thus

- if the answer is positive, then the present value $v(0)$ is assumed to be $(-1)^m$;

- otherwise, set $v(0) = -(-1)^m$.

(The system will then be driven to the origin in minimum time.) The computation mainly identifies the signs of certain polynomials, so it can be completed for real-time control.

7. Conclusion

Optimal control is widely used in modern system science. In the context of switching surfaces in optimal control, many problems can be reduced to solving certain semi-algebraic systems. In this article, new methods for time-optimal control were presented and illustrated by a simple example. Complex and real versions for optimal control were considered. Based on triangular decomposition and relevant symbolic computation, our methods are more general and more systematic than those given in Ref. [30] — and since the computation is exact, they are quite different from existing numerical approaches. Their high complexity can be addressed by moving the computational burden off-line, such that the modified version is feasible for real-time control. The future development of faster computer algebra systems and more efficient algorithms for basic operations (including triangular decomposition and cylindrical algebraic decomposition) will also significantly improve the performance of our methods.

Acknowledgments

The authors wish to thank the anonymous referees for their helpful comments. This work has been supported by the National Natural Science Foundation of China (Tianyuan Special Foundation Nos. 11326210 and 11326211), the open fund of Guangxi Key Laboratory of Hybrid Computation and IC Design Analysis (No. HCIC201302) and the Fundamental Research Funds for the Central Universities (No. YWF-14-RSC-103).

References

- [1] M. Athans and P.L. Falb, *Optimal Control: An Introduction to the Theory and Its Applications*, McGraw-Hill, New York (1966).
- [2] F. Chai, X.-S. Gao and C. Yuan, *A characteristic set method for solving Boolean equations and applications in cryptanalysis of stream ciphers*, *J. Systems Science and Complexity* **21**, 191–208 (2008).
- [3] C. Chen and M. Moreno Maza, *Algorithms for computing triangular decompositions of polynomial systems*, *Proc. 36th International Symposium on Symbolic and Algebraic Computation* (A. Leykin, ed.), pp. 83–90, ACM, New York (2011).
- [4] S.-C. Chou and X.-S. Gao, *Ritt-Wu's decomposition algorithm and geometry theorem proving*, *Proc. 10th International Conference on Automated Deduction* (M.E. Stickel, ed.), pp. 207–220, Springer, Berlin (1990).
- [5] G.E. Collins and H. Hong, *Partial cylindrical algebraic decomposition for quantifier elimination*, *J. Symbolic Comp.* **12**, 299–328 (1991).

- [6] G.E Collins and R. Loos, *Real zeros of polynomial*, in *Computer Algebra: Symbolic and Algebraic Computation* (B. Buchberger, G.E. Collins and R. Loos, eds.), pp. 83–94, Springer, New York (1983).
- [7] P. Deuffhard, *Newton Methods for Nonlinear Problems: Affine Invariance and Adaptive Algorithms*, Springer, Berlin (2011).
- [8] J. Fiene and G. Niemeyer, *Toward switching motor control*, *IEEE/ASME Trans. Mechatronics* **11**, 27–34 (2006).
- [9] G. Gallo and B. Mishra, *Efficient algorithms and bounds for Wu-Ritt characteristic sets*, in *Effective Methods in Algebraic Geometry* (T. Mora and C. Traverso, eds.), pp. 119–142. Birkhauser, Berlin (1991).
- [10] G. Gallo and B. Mishra, *Wu-Ritt characteristic sets and their complexity*, in *Discrete and Computational Geometry: Papers from the DIMACS Special Year* (J.E. Goodman, R.D. Pollack and W. Steiger, eds), pp. 111–136, American Mathematical Society, Providence(1991).
- [11] X.-S. Gao and S.-C. Chou. On the dimension of an arbitrary ascending chain. *Chinese Science Bulletin*, 38(5):396–399, 1993.
- [12] X.-S. Gao and Z. Huang, *Characteristic set algorithms for equation solving in finite fields*, *J. Symbolic Comp.* **47**, 655–679 (2012).
- [13] T. Ge and J.S. Chang, *Bang-Bang control class-d amplifiers: power-supply noise*, *IEEE Trans. Circuits and Systems II: Express Briefs* **55**, 723–727 (2008).
- [14] M. Jin, X. Li and D. Wang, *A new algorithmic scheme for computing characteristic sets*, *J. Symbolic Comp.* **51**, 433–449 (2012).
- [15] M. Kalkbrener, *A generalized Euclidean algorithm for computing triangular representations of algebraic varieties*, *J. Symbolic Comp.* **15**, 143–167 (1993).
- [16] B.S. Kumar and A. Ng, *A Bang-Bang control approach to manoeuver spacecraft in a formation with differential drag*, *Proc. AIAA Guidance, Navigation and Control Conference and Exhibit*, pp. 18–21, American Institute of Aeronautics and Astronautics, Reston (2008).
- [17] D. Lazard, *A new method for solving algebraic systems of positive dimension*, *Discrete Applied Math.* **33** 147–160 (1991).
- [18] E.B. Lee and L. Markus, *Foundations of Optimal Control Theory*, Krieger, Malabar (1967).
- [19] X. Li, C. Mou and D. Wang, *Decomposing polynomial sets into simple sets over finite fields: the zero-dimensional case*, *Computers and Mathematics with Applications* **60**, 2983–2997 (2010).
- [20] X. Li and D. Wang, *Computing equilibria of semi-algebraic economies using triangular decomposition and real solution classification*, *ArXiv e-prints arXiv:1308.5029* (2013).
- [21] Y. Li, *Some properties of triangular sets and improvement upon algorithm CharSer*, in *Artificial Intelligence and Symbolic Computation* (J. Calmet, T. Ida, and D. Wang, eds), pp. 82–93, Springer, New York (2006).
- [22] B. Mishra, *Algorithmic Algebra*, Springer, New York (1993).
- [23] M. Moreno Maza, *On triangular decompositions of algebraic varieties*, Technical Report 4/99, NAG, UK. (Presented at the MEGA-2000 Conference, Bath.)
- [24] L.Y. Pao and G.F. Franklin, *Proximate time-optimal control of third-order servomechanisms*, *IEEE Trans. Automatic Control* **38**, 560–580 (1993).
- [25] J.F. Ritt, *Differential Algebra*, American Mathematical Society, Providence (1950).
- [26] M. Rubagotti, M.L. Della Vedova, and A. Ferrara, *Time-optimal sliding-mode control of a mobile robot in a dynamic environment*, *IET Control Theory and Applications* **5**, 1916–1924 (2011).
- [27] M.S. El Din, *Testing sign conditions on a multivariate polynomial and applications*, *Mathematics in Comp. Sc.* **1**, 177–207 (2007).
- [28] M.S. El Din and E. Schost, *Polar varieties and computation of one point in each connected component of a smooth real algebraic set*, *Proc. 2003 International Symposium on Symbolic and*

- Algebraic Computation, pp. 224–231. ACM, New York (2003).
- [29] A.J. Sommese and C.W. Wampler, *The Numerical Solution of Systems of Polynomials Arising in Engineering and Science*, World Scientific, Singapore (2005).
 - [30] U. Walther, T.T. Georgiou and A. Tannenbaum, *On the computation of switching surfaces in optimal control: a Gröbner basis approach*, IEEE Trans. Automatic Control **46**, 534–540 (2001).
 - [31] D. Wang, *An elimination method for polynomial systems*, J. Symbolic Comp. **16**, 83–114 (1993).
 - [32] D. Wang, *Decomposing polynomial systems into simple systems*, J. Symbolic Comp. **25**, 295–314 (1998).
 - [33] D. Wang, *Computing triangular systems and regular systems*, J. Symbolic Comp. **30**, 221–236 (2000).
 - [34] D. Wang, *A generalized algorithm for computing characteristic sets*, in Computer Mathematics: Proc. Fifth Asian Symposium (K. Shirayanagi and K. Yokoyama, eds.), pp. 165–174, Springer, Berlin (2001).
 - [35] D. Wang, *A strategy for speeding-up the computation of characteristic sets.*, Proc. 17th International Symposium on Mathematical Foundations of Computer Science (I.M. Havel and V. Koubek, eds.), pp. 504–510, Springer, Berlin (1992).
 - [36] D. Wang, *Elimination Methods*, Springer, New York (2001).
 - [37] W.-T. Wu, *Basic principles of mechanical theorem proving in elementary geometries*, J. Automated Reasoning **2**, 221–252 (1986).
 - [38] W.-T. Wu, *On zeros of algebraic equations: an application of Ritt principle*, Kexue Tongbao **31**, 1–5 (1986).
 - [39] L. Yang and J.-Z. Zhang, *Searching dependency between algebraic equations: an algorithm applied to automated reasoning*, in *Artificial Intelligence in Mathematics* (J. Johnson, S. McKee and A. Vella, eds.), pp. 147–156, Oxford University Press, Oxford (1994).
 - [40] L. Yang, X. Hou and B. Xia, *A complete algorithm for automated discovering of a class of inequality-type theorems*, Science in China Series F **44**, 33–49 (2001).
 - [41] Y. Zhao and P. Tsiotras, *Time-optimal path following for fixed-wing aircraft*, J. Guidance, Control and Dynamics **36**, 83–95 (2012).