Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Lossless Convexification of Optimal Control Problems with Semi-continuous Inputs

Danylo Malyuta * **Behçet Açıkmeşe** *

* *Dept. of Aeronautics & Astronautics, University of Washington, Seattle, WA 98195 USA (e-mail: {danylo,behcet}@uw.edu)*

Abstract: This paper presents a novel one-shot convex optimization method for finding globally optimal solutions of a class of mixed-integer non-convex optimal control problems. We consider problems with non-convex constraints that restrict the input norms to be either zero or lower- and upper-bounded. The non-convex problem is relaxed to a convex one whose optimal solution is proved to be optimal almost everywhere for the original problem, a procedure known as lossless convexification. The solution relies on second-order cone programming and demonstrates that a meaningful class of optimal control problems with binary variables can be solved reliably and in polynomial time. A rocket landing example with a coupled thrust-gimbal constraint corroborates the effectiveness of the approach.

Keywords: Optimal control, convex optimization, maximum principle, integer programming.

1 Introduction

We present a convex programming solution to a class of optimal control problems with semi-continuous control input norms. Semi-continuity presents a binary non-convexity where a variable $x \in \mathbb{R}$ is constrained $x \in \{0\} \cup [a, b]$ with $0 < a \leq b$ MOSEK ApS (2019). The constraint $az \leq x \leq bz$ with $z \in \{0, 1\}$ models semi-continuity. Practical rocket landing and spacecraft rendezvous path planning problems include such constraints, and can take hours to solve using existing mixed-integer convex programming (MICP) methods Malyuta et al. (2020). In this paper, we propose an algorithm based on lossless convexification that solves these problems to global optimality in seconds.

Lossless convexification is a method for finding the globally optimal solution of non-convex problems using convex optimization. The method relaxes the original problem to a convex one via a slack variable, enabling the use of second-order cone programming (SOCP). The maximum principle is used to prove that the solution of the relaxed problem is globally optimal for the original problem.

Classical lossless convexification deals with non-convexity in the form of an input norm lower-bound. The first result was introduced in Açıkmeşe and Ploen (2007) for minimum-fuel rocket landing and was later expanded to more general non-convex input sets Açıkmeşe and Blackmore (2011). Extensions of the method were introduced in Blackmore et al. (2010); Carson III et al. (2011); Açıkmeşe et al. (2013) to handle minimum-error rocket landing and non-convex pointing constraints. More recently, lossless convexification was shown to handle affine and quadratic state constraints Harris and Açıkmeşe (2013a,b), culminating in Harris and Açıkmeşe (2014).

A recurring assumption of classical lossless convexification is that there is a single input which cannot be turned off. Our interest is in problems that have multiple inputs which may be turned off. When active, the input norm is lower-bounded, making it a semi-continuous variable. This is a richer binary non-convexity than what was handled by classical lossless convexification.

The concept of lossless convexification with binary variables implemented via MICP was explored in Blackmore et al. (2012); Zhang et al. (2017). However, the $\mathcal{NP}$-hard nature of MICP generally makes the approach computationally expensive.

Our main contribution is to extend lossless convexification to a class of mixed-integer non-convex optimal control problems with multiple inputs and semi-continuous input norms. Unlike mixed-integer programming, lossless convexification solves the problem in polynomial time. Compared to Sager (2005); Sager et al. (2007) which develop an iterative scheme, our approach solves the mixed-integer non-convex program in one shot. The approach is amenable to real-time onboard optimization for autonomous systems or for rapid design trade studies.

The paper is organized as follows. Section 2 defines the class of optimal control problems that our method handles. Section 3 proposes our solution method based on lossless convexification. Section 5 proves that our method finds the globally optimal solution based on the necessary conditions of optimality presented in Section 4. Section 6 presents a rocket landing example which corroborates the method's effectiveness for practical path planning applications. Section 7 outlines future work and Section 8 summarizes the result.

Notation: sets are calligraphic, e.g. $\mathcal{S}$. Set $\mathbb{R}^n_-$ denotes the n-dimensional non-positive orthant. The operator $\circ$ denotes the element-wise product. Given a function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$, we use the shorthand $f[t] \equiv f(x(t), y(t))$. In text, functions are referred to by their letter (e.g. f) and conflicts with another variable are to be understood from context. The gradient of f with respect an argument x is denoted $\nabla_x f \in \mathbb{R}^{p \times n}$. Similarly, if f is nonsmooth then its subdifferential with respect to x is $\partial_x f \subseteq \mathbb{R}^{1 \times n}$.

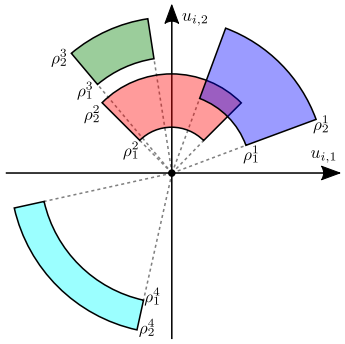


Fig. 1. Input set example for $m = 2$. Multiple overlapping input sets with different bounds can be modelled.

The normal cone at x to $\mathcal{S} \subseteq \mathbb{R}^n$ is denoted $\mathcal{N}_{\mathcal{S}}(x) \subseteq \mathbb{R}^n$. The set $\mathcal{S}^\intercal \triangleq \{x^\intercal : x \in \mathcal{S}\}$. When we refer to an *interval*, we mean some time interval $[t_1, t_2]$ of non-zero duration, i.e. $t_1 < t_2$. We call the Eucledian projection of $y \in \mathbb{R}^n$ onto $\mathcal{S} \subseteq \mathbb{R}^n$ the magnitude of the 2-norm projection of y:

$$\mathcal{P}_{\mathcal{S}}(y) \triangleq \left\| \operatorname{argmin}_{z \in \mathcal{S}} \|y - z\|_2 \right\|_2. \tag{1}$$

2 Problem Definition

This section presents the class of optimal control problems that can be solved via convex optimization by our method. We consider mixed-integer non-convex optimal control problems with linear time-invariant (LTI) dynamics and semi-continuous input norms:

Problem $\mathcal{O}$.

$$\min_{u_i,\gamma_i,t_f} \; \begin{array}{l} m(t_f, x(t_f)) + \int_0^{t_f} \ell(x(t)) + \\ \zeta \sum_{i=1}^M \|u_i(t)\|_2 \mathrm{d}t \end{array} \quad \text{s.t.} \tag{$\mathcal{O}$.a}$$

$$\dot{x}(t) = Ax(t) + B\textstyle\sum_{i=1}^M u_i(t) + w, \; x(0) = x_0, \tag{$\mathcal{O}$.b}$$
$$\gamma_i(t)\rho_1^i \le \|u_i(t)\|_2 \le \gamma_i(t)\rho_2^i \quad i = 1, \dots, M, \tag{$\mathcal{O}$.c}$$
$$\gamma_i(t) \in \{0, 1\} \quad i = 1, \dots, M, \tag{$\mathcal{O}$.d}$$
$$\textstyle\sum_{i=1}^M \gamma_i(t) \le K, \tag{$\mathcal{O}$.e}$$
$$C_i u_i(t) \le 0 \quad i = 1, \dots, M, \tag{$\mathcal{O}$.f}$$
$$x(t) \in \mathcal{X}, \tag{$\mathcal{O}$.g}$$
$$b(x(t_f)) = 0, \tag{$\mathcal{O}$.h}$$

where $x(t) \in \mathbb{R}^n$ is the state, $u_i(t) \in \mathbb{R}^m$ is the i-th input, and $w \in \mathbb{R}^n$ is a known external input. Convex functions $m : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$, $\ell : \mathbb{R}^n \to \mathbb{R}$ and $b : \mathbb{R}^n \to \mathbb{R}^{n_b}$ define the terminal cost, the state running cost and the terminal manifold respectively. The binary coefficient $\zeta \in \{0, 1\}$ toggles the input running cost. The state must lie in the convex set $\mathcal{X} \subseteq \mathbb{R}^n$. The input directions are constrained to polytopic cones called *input pointing sets*:

$$\mathcal{U}_i \triangleq \{u \in \mathbb{R}^m : C_i u \le 0\}, \tag{2}$$

where $C_i \in \mathbb{R}^{p_i \times m}$ is a matrix with $C_{i,j}$ the j-th row. Figure 1 shows the rich input set geometry that is possible with this model.

Assumption 1. Matrices C_i in ($\mathcal{O}$.f) are full row rank.

Assumption 2. The control norm bounds in ($\mathcal{O}$.c) are distinct, i.e. $\rho_1^i < \rho_2^i$.

3 Lossless Convexification

This section presents the two main results, Theorems 1a and 1b, which state that the convex Problem $\mathcal{R}$ finds the global optimum of Problem $\mathcal{O}$ under certain conditions.

The input magnitude in Problem $\mathcal{O}$ is semi-continuous, i.e. $\|u_i(t)\|_2 \in \{0\} \cup [\rho_1^i, \rho_2^i]$. This makes the problem mixed-integer and non-convex, which is readily apparent from Figure 1. Consider the following convex relaxation:

Problem $\mathcal{R}$.

$$\min_{u_i,\gamma_i,\sigma_i,t_f} \; \begin{array}{l} m(t_f, x(t_f)) + \zeta\xi(t_f) + \\ \int_0^{t_f} \ell(x(t))\mathrm{d}t \end{array} \quad \text{s.t.} \tag{$\mathcal{R}$.a}$$

$$\dot{x}(t) = Ax(t) + B\textstyle\sum_{i=1}^M u_i(t) + w, \; x(0) = x_0, \tag{$\mathcal{R}$.b}$$
$$\dot{\xi}(t) = \textstyle\sum_{i=1}^M \sigma_i(t), \tag{$\mathcal{R}$.c}$$
$$\gamma_i(t)\rho_1^i \le \sigma_i(t) \le \gamma_i(t)\rho_2^i \quad i = 1, \dots, M, \tag{$\mathcal{R}$.d}$$
$$\|u_i(t)\|_2 \le \sigma_i(t) \quad i = 1, \dots, M, \tag{$\mathcal{R}$.e}$$
$$0 \le \gamma_i(t) \le 1 \quad i = 1, \dots, M, \tag{$\mathcal{R}$.f}$$
$$\textstyle\sum_{i=1}^M \gamma_i(t) \le K, \tag{$\mathcal{R}$.g}$$
$$C_i u_i(t) \le 0 \quad i = 1, \dots, M, \tag{$\mathcal{R}$.h}$$
$$x(t) \in \mathcal{X}, \tag{$\mathcal{R}$.i}$$
$$b(x(t_f)) = 0. \tag{$\mathcal{R}$.j}$$

Replacing ($\mathcal{O}$.c)-($\mathcal{O}$.d) with ($\mathcal{R}$.d)-($\mathcal{R}$.f) convexifies the input set of Problem $\mathcal{O}$. Figure 2 illustrates an example.

Consider the following conditions which remove degenerate solutions of Problem $\mathcal{R}$ that may be infeasible for Problem $\mathcal{O}$. The conditions use an *adjoint system* whose output $y(t) \in \mathbb{R}^m$ is called the *primer vector*:

$$\dot{\lambda}(t) = -A^\intercal \lambda(t) + v(t), \; v(t) \in \partial\ell(x(t))^\intercal, \tag{3a}$$
$$y(t) = B^\intercal \lambda(t). \tag{3b}$$

It will be seen in Section 5 that we are interested in "how much" $y(t)$ projects onto the i-th input pointing set. This is given by the following input *gain* measure:

$$\Gamma_i(t) \triangleq (\mathcal{P}_{\mathcal{U}_i}(y(t)) - \zeta)\rho_2^i. \tag{4}$$

Condition 1. The adjoint system (3) is strongly observable Trentelman et al. (2001).

Condition 2. The adjoint system (3) and pointing cone geometry ($\mathcal{O}$.f) satisfy either:

(a) $\Gamma_i(t) \neq 0$ a.e. $[0, t_f]$ $\forall i$ s.t. $y(t) \notin \operatorname{int}(\mathcal{N}_{\mathcal{U}_i}(0))$;
(b) on any interval where $\Gamma_i(t) = 0$, $\Gamma_j(t) > 0$ for at least K other inputs.

Condition 3. The adjoint system (3) and pointing cone geometry ($\mathcal{O}$.f) satisfy either:

(a) $\Gamma_i(t) \neq \Gamma_j(t)$ a.e. $[0, t_f]$ $\forall i$ s.t. $y(t) \notin \operatorname{int}(\mathcal{N}_{\mathcal{U}_i}(0))$;
(b) on any interval where $\Gamma_i(t) = \Gamma_j(t)$, there exist K inputs with $\Gamma_k(t) > \Gamma_i(t)$ or $M - K$ inputs where $\Gamma_k(t) < \Gamma_i(t)$.

Condition 4. $\ell[t] + \zeta \sum_{i=1}^M \sigma_i(t) + \nabla_t m[t_f] \neq 0 \; \forall t \in [0, t_f)$.

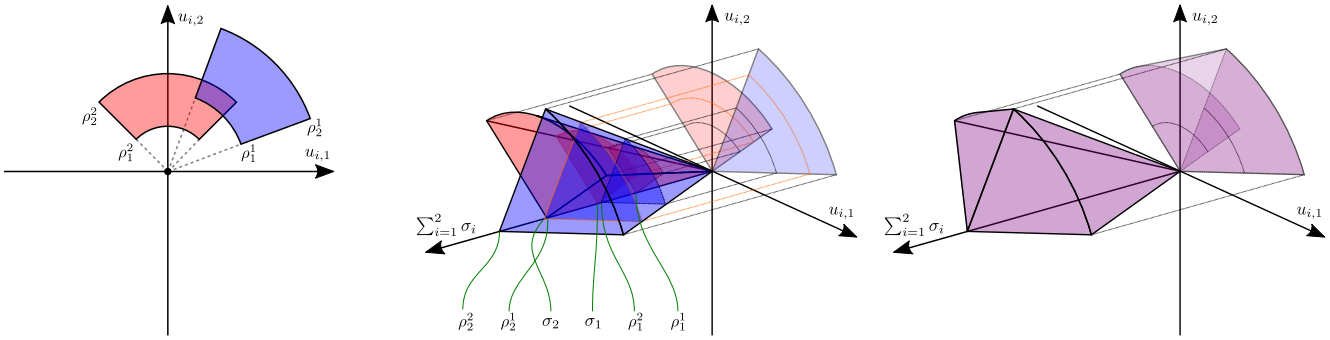
We now state the two main results of this paper, which claim that Problem $\mathcal{R}$ solves Problem $\mathcal{O}$ under certain conditions. The theorems are proved in Section 5.

Theorem 1a. *The solution of Problem $\mathcal{R}$ is globally optimal a.e. $[0, t_f]$ for Problem $\mathcal{O}$ if Conditions 1-4 hold and the state constraint ($\mathcal{O}$.g) is never activated.*

Theorem 1b. *The solution of Problem $\mathcal{R}$ is globally optimal a.e. $[0, t_f]$ for Problem $\mathcal{O}$ if Conditions 1-4 hold and the state constraint ($\mathcal{O}$.g) is activated at discrete times.*

3.1 Discussion on Strong Observability

This section describes Condition 1 and its verification. Strong observability extends the concept of observability



(a) Original non-convex input set defined by ($\mathcal{O}$.c)-($\mathcal{O}$.f).

(b) Non-convexity of individual input sets is removed by relaxing ($\mathcal{O}$.c) to ($\mathcal{R}$.d)-($\mathcal{R}$.e).

(c) Semi-continuity of the input norm is convexified by relaxing ($\mathcal{O}$.d) to ($\mathcal{R}$.f).

Fig. 2. Problem $\mathcal{R}$ convexifies the input set of Problem $\mathcal{O}$, here shown for $M = 2$, $K = 1$ and $m = 2$. The relaxation consists of three steps: a) ($\mathcal{O}$.c)-($\mathcal{O}$.f) originally define a non-convex set of a binary nature; b) by relaxing ($\mathcal{O}$.c) to ($\mathcal{R}$.d)-($\mathcal{R}$.e), individual input sets are convexified; c) by relaxing ($\mathcal{O}$.d) to ($\mathcal{R}$.f), a convex hull is obtained.

to the case of non-zero inputs. A strongly observable system does not have transmission zeroes. To be precise, let us state strong observability in the context of (3).

Definition 1 (Trentelman et al., 2001, Definition 7.8)**.** A point $\lambda_0 \in \mathbb{R}^n$ is *weakly unobservable* if there exists an interval $\mathcal{T} = [\tau_1, \tau_2]$ and an input trajectory $v(t) \in \partial\ell[t]^\intercal$ for $t \in \mathcal{T}$ such that if $\lambda(\tau_1) = \lambda_0$ then the primer vector satisfies $y(t) = 0\ \forall t \in \mathcal{T}$. The set of all weakly unobservable points is denoted $\mathcal{V}$, which is called the weakly unobservable set.

Theorem 2 (Trentelman et al., 2001, Theorem 7.16)**.** *The adjoint system* (3) *is strongly observable if* $\mathcal{V} = \{0\}$.

To verify Condition 1 via simple matrix algebra, it is sufficient to apply the algorithm for computing $\mathcal{V}$ in (Trentelman et al., 2001, Section 7.3) using the following alternative to (3a):

$$\dot{\lambda}(t) = -A^\intercal\lambda(t) + Dv(t), \tag{5}$$

where $v(t) \in \mathbb{R}^n$ and $\operatorname{range} D = \operatorname{span} \bigcup_{t\in[\tau_1,\tau_2]} \partial\ell(x(t))^\intercal$. This conservative approximation assumes that the input can come from a subspace spanned by the subdifferentials. Section 6 uses this approximation to verify Conditions 1-3 for the rocket landing problem.

4 Nonsmooth Maximum Principle

This section states a nonsmooth version of the maximum principle that we shall use for proving Theorems 1a and 1b. Consider the following general optimal control problem:

Problem $\mathcal{G}$.

$$\min_{u,t_f}\ m(t_f, x(t_f)) + \textstyle\int_0^{t_f} \ell(t, u(t), x(t))\mathrm{d}t \ \text{ s.t.} \tag{$\mathcal{G}$.a}$$
$$\dot{x}(t) = f(t, x(t), u(t)), \quad x(0) = x_0, \tag{$\mathcal{G}$.b}$$
$$g(t, u(t)) \le 0, \tag{$\mathcal{G}$.c}$$
$$b(t_f, x(t_f)) = 0. \tag{$\mathcal{G}$.d}$$

where the state trajectory $x(\cdot)$ is absolutely continuous and the control trajectory $u(\cdot)$ is measurable. The dynamics $f : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ are convex and continuously differentiable. The terminal cost $m : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$, the running cost $\ell : \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$, the input constraint $g : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^{n_g}$, and the terminal constraint $b : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^{n_b}$ are convex. Define the terminal manifold as $\mathcal{T} \triangleq \{x \in \mathbb{R}^n : (\mathcal{G}.\text{d}) \text{ holds}\}$ and the Hamiltonian function:

$$H(t, x(t), u(t), \alpha, \psi(t)) \triangleq \alpha\ell[t] + \psi(t)^\intercal f[t], \tag{6}$$

where $\alpha \le 0$ is the *abnormal multiplier* and $\psi(\cdot)$ is the *adjoint variable* trajectory. We now state the nonsmooth maximum principle, due to (Vinter, 2000, Theorem 8.7.1) (see also Clarke (2010); Hartl et al. (1995)), which specifies the necessary conditions of optimality for Problem $\mathcal{G}$.

Theorem 3 (Maximum Principle)**.** *Let* $x(\cdot)$ *and* $u(\cdot)$ *be optimal on the interval* $[0, t_f]$*. There exist a constant* $\alpha \le 0$ *and an absolutely continuous* $\psi(\cdot)$ *such that the following conditions are satisfied:*

(1) Non-triviality:

$$(\alpha, \psi(t)) \neq 0\ \forall t \in [0, t_f]; \tag{7}$$

(2) Pointwise maximum:

$$u(t) = \operatorname*{argmax}_{v\in(\mathcal{G}.\text{c})} H(t, x(t), v, \alpha, \psi(t))\ \ a.e.\ [0, t_f]; \tag{8}$$

(3) The differential equations and inclusions:

$$\dot{x}(t) = \nabla_\psi H[t]^\intercal\ \ a.e.\ [0, t_f], \tag{9a}$$
$$\dot{\psi}(t) \in -\partial_x H[t]^\intercal\ \ a.e.\ [0, t_f], \tag{9b}$$
$$\dot{H}[t] \in \partial_t H[t]\ \ a.e.\ [0, t_f]; \tag{9c}$$

(4) Transversality:

$$\psi(t_f) \in \alpha\partial_x m[t_f]^\intercal + \mathcal{N}_\mathcal{T}(x(t_f)), \tag{10a}$$
$$0 \in H[t_f] + \alpha\partial_t m[t_f] + \mathcal{N}_\mathcal{T}(t_f). \tag{10b}$$

5 Lossless Convexification Proof

This section proves Theorems 1a and 1b. The general outline is as follows. We first prove Theorem 1a by showing that (step 1) the solution of Problem $\mathcal{R}$ is feasible for Problem $\mathcal{O}$, and (step 2) the solution is globally optimal. We then show Theorem 1b via a proof by contradiction in which Theorem 1a is applied on each interval where the state constraint is inactive.

Lemma 1. *The solution of Problem* $\mathcal{R}$ *is feasible a.e.* $[0, t_f]$ *for Problem* $\mathcal{O}$ *if* $x(t) \in \operatorname{int}(\mathcal{X})$ *and Conditions 1-4 hold.*

Proof. The proof uses the maximum principle from Theorem 3. Since there are two states, partition the adjoint variable as $\psi(t) = (\lambda(t) \in \mathbb{R}^n, \eta(t) \in \mathbb{R})$. For Problem $\mathcal{R}$ and $x(t) \in \operatorname{int}(\mathcal{X})$, the adjoint and Hamiltonian dynamics follow from (9b) and (9c):

$$\dot{\lambda}(t) = -A^\intercal\lambda(t) - \alpha v(t),\ v(t) \in \partial\ell[t]^\intercal,\ \text{a.e. } [0, t_f], \tag{11a}$$

$$\dot{\eta}(t) = 0 \text{ a.e. } [0, t_f], \tag{11b}$$
$$\dot{H}[t] = 0 \text{ a.e. } [0, t_f], \tag{11c}$$

Using the subdifferential basic chain rule (Rockafellar and Wets, 1998, Theorem 10.6), the transversality condition (10) yields:

$$\lambda(t_f) = \nabla_x m[t_f]^\intercal \alpha + \nabla_x b[t_f]^\intercal \beta, \tag{12a}$$
$$\eta(t_f) = \alpha\zeta, \tag{12b}$$
$$H[t_f] = -\nabla_t m[t_f]\alpha, \tag{12c}$$

for some $\beta \in \mathbb{R}^{n_b}$. Due to (11b)-(11c), (12b)-(12c) and absolute continuity, we have (Varberg, 1965, Theorem 9):

$$\eta(t) = \alpha\zeta, \ \forall t \in [0, t_f], \tag{13a}$$
$$H[t] = -\nabla_t m[t_f]\alpha, \ \forall t \in [0, t_f]. \tag{13b}$$

We claim that the primer vector $y(t) \neq 0$ a.e. $[0, t_f]$. By contradiction, suppose there exists an interval $[\tau_1, \tau_2] \subseteq [0, t_f]$ for which $y(t) = 0$. Condition 1 implies that $\lambda(\tau_1) = 0$. Due to (13), this implies $\alpha(\ell[\tau_1] + \zeta \sum_{i=1}^M \sigma_i(\tau_1) + \nabla_t m[t_f]) = 0$. Due to Condition 4, it must be that $\alpha = 0$ which implies $(\alpha, \psi(\tau_1)) = 0$. Since this violates non-triviality (7), it must be that $y(t) \neq 0$ a.e. $[0, t_f]$. Having eliminated the pathological case, assume $\alpha < 0$. In particular, since the necessary conditions in Theorem 3 are scale-invariant, we can set $\alpha = -1$ without loss of generality. The pointwise maximum condition (8) implies that the following must hold a.e. $[0, t_f]$:

$$\underset{u_i, \gamma_i, \sigma_i}{\text{argmax}} \ \textstyle\sum_{i=1}^M y(t)^\intercal u_i(t) - \zeta\sigma_i(t) \ \text{ s.t.} \tag{14a}$$
$$\text{constraints } (\mathcal{R}.\text{d})\text{-}(\mathcal{R}.\text{h}) \text{ hold.} \tag{14b}$$

We shall now analyze the optimality conditions of (14). For concise notation, the time argument t shall be omitted. Expressing (14) as a minimization and treating constraints ($\mathcal{R}$.f) and ($\mathcal{R}$.g) implicitly, we can write the Lagrangian of (14) Boyd and Vandenberghe (2004):

$$\mathcal{L}(u_i, \gamma_i, \sigma_i, \lambda^i_{1\ldots4}) = \textstyle\sum_{i=1}^M \zeta\sigma_i - y^\intercal u_i + \lambda^i_1(\|u_i\|_2 - \sigma_i) + \lambda^i_2(\gamma_i\rho^i_1 - \sigma_i) + \lambda^i_3(\sigma_i - \gamma_i\rho^i_2) + {\lambda^i_4}^\intercal C_i u_i, \tag{15}$$

where $\lambda^i_j \geq 0$ are Lagrange multipliers satisfying the following complementarity conditions:

$$\lambda^i_1(\|u_i\|_2 - \sigma_i) = 0, \tag{16a}$$
$$\lambda^i_2(\gamma_i\rho^i_1 - \sigma_i) = 0, \tag{16b}$$
$$\lambda^i_3(\sigma_i - \gamma_i\rho^i_2) = 0, \tag{16c}$$
$$\lambda^i_4 \circ C_i u_i = 0. \tag{16d}$$

Next, the Lagrange dual function is given by:

$$\begin{aligned} g(\lambda^i_{1\ldots4}) &= \inf_{u_i, \gamma_i, \sigma_i} \mathcal{L}(u_i, \gamma_i, \sigma_i, \lambda^i_{1\ldots4}) \\ &= \textstyle\sum_{i=1}^M \inf_{\sigma_i} \left[(\zeta + \lambda^i_3 - \lambda^i_2 - \lambda^i_1)\sigma_i\right] - \\ &\quad \textstyle\sum_{i=1}^M \sup_{u_i} \left[(y - C_i^\intercal \lambda^i_4)^\intercal u_i - \lambda^i_1\|u_i\|_2\right] + \\ &\quad \textstyle\inf_{(\mathcal{R}.\text{f}),(\mathcal{R}.\text{g})} \sum_{i=1}^M (\lambda^i_2\rho^i_1 - \lambda^i_3\rho^i_2)\gamma_i. \end{aligned} \tag{17}$$

The dual function bounds the primal optimal cost from above. A non-trivial upper-bound requires:

$$\|y - C_i^\intercal \lambda^i_4\|_2 \leq \lambda^i_1, \tag{18a}$$
$$\zeta + \lambda^i_3 - \lambda^i_2 - \lambda^i_1 = 0, \tag{18b}$$

where the first inequality is akin to the $\|\cdot\|_2$ conjugate function (Boyd and Vandenberghe, 2004, Example 3.26). However, note that if (18a) is strict then $\|u_i\|_2 = 0$ is optimal, which is trivially feasible for Problem $\mathcal{O}$. Substituting (18b) into (18a) gives the following condition for non-trivial solutions:

$$\|y - C_i^\intercal \lambda^i_4\|_2 = \zeta + \lambda^i_3 - \lambda^i_2. \tag{19}$$

Further simplification is possible by recognizing that a non-trivial solution implies $\gamma_i > 0$. By Assumption 2, (16b) and (16c), this means $\lambda^i_2 > 0$ and $\lambda^i_3 > 0$ cannot occur simultaneously. Furthemore, (17) reveals that $\gamma_i > 0$ is not sub-optimal if and only if $\lambda^i_2\rho^i_1 - \lambda^i_3\rho^i_2 \leq 0$. By this reasoning, $\lambda^i_2 = 0$ and $\lambda^i_3 \geq 0$ are necessary for optimality. Thus, (19) simplifies to:

$$\|y - C_i^\intercal \lambda^i_4\|_2 = \zeta + \lambda^i_3. \tag{20}$$

Next, note that at optimality the left-hand side of (20) equals the Eucledian projection onto $\mathcal{U}_i$, i.e. $\|y - C_i^\intercal \lambda^i_4\|_2 = \mathcal{P}_{\mathcal{U}_i}(y)$. This can be shown by contradiction using Assumption 1, (16d) and that $u_i = \|u_i\|_2(y - C_i^\intercal \lambda^i_4)/\|y - C_i^\intercal \lambda^i_4\|_2$ in (17). Note that the degenerate case of $u_i \neq 0$ and $\|y - C_i^\intercal \lambda^i_4\|_2 = 0$ is eliminated in the discussion below which leverages Condition 2. Thus (20) simplifies to the following relationship, which we call the *characteristic equation* of non-trivial solutions to Problem $\mathcal{R}$:

$$\mathcal{P}_{\mathcal{U}_i}(y) = \zeta + \lambda^i_3. \tag{21}$$

Substituting (21) into (17) yields:

$$g(\lambda^i_{1\ldots4}) = -\sup_{(\mathcal{R}.\text{f}),(\mathcal{R}.\text{g})} \textstyle\sum_{i=1}^{K'} (\mathcal{P}_{\mathcal{U}_i}(y) - \zeta)\rho^i_2\gamma_i, \tag{22}$$

where we assume that the characteristic equation (21) does not hold for $i = K' + 1, \ldots, M$ such that $\gamma_{i>K'} = 0$. To facilitate discussion, define the i-th input *gain* as in (4). Note that $\Gamma_i \geq 0$ due to (21). Thus (22) becomes:

$$g(\lambda^i_{1\ldots4}) = -\sup_{(\mathcal{R}.\text{f}),(\mathcal{R}.\text{g})} \textstyle\sum_{i=1}^{K'} \Gamma_i\gamma_i. \tag{23}$$

Without loss of generality, assume a descending ordering $\Gamma_i \geq \Gamma_j$ for $i > j$. Let $K'' \triangleq \min\{K, K'\}$. By inspection of (23), the condition:

$$\Gamma_{K''} > 0 \ \wedge \ \Gamma_{K''} > \Gamma_{K''+1}, \tag{24}$$

is sufficient to ensure that it is optimal to set

$$\gamma_i = \begin{cases} 1 & \text{if } i \leq K'', \\ 0 & \text{otherwise.} \end{cases} \tag{25}$$

The lemma holds if (24) holds a.e. $[0, t_f]$. This is assured by Conditions 2 and 3. Condition 2 case (a) assures $\Gamma_{K''} > 0$ a.e. $[0, t_f]$. If on some interval $\Gamma_k = 0$, Condition 2 case (b) assures that $k > K''$. If $K'' < K$ then due to $\Gamma_{K''} > 0$ and the definition of K', it must be that $\Gamma_{K''+1} = 0 \Rightarrow \Gamma_{K''} > \Gamma_{K''+1}$. Else if $K'' = K$, Condition 3 case (a) assures $\Gamma_K > \Gamma_{K+1}$ a.e. $[0, t_f]$. If on some interval $\Gamma_k = \Gamma_{k+1}$, Condition 3 case (b) assures that $k \neq K$.

Thus, (24) holds a.e. $[0, t_f]$ and the lemma is proved. From (25), the structure of the optimal solution is **bang-bang with at most K inputs active** a.e. $[0, t_f]$. □

Lemma 1 guarantees that Problem $\mathcal{R}$ produces a feasible solution of Problem $\mathcal{O}$. We will now show that this solution is globally optimal, thus proving Theorem 1a.

Proof of Theorem 1a. The solution of Problem $\mathcal{R}$ is feasible a.e. $[0, t_f]$ for Problem $\mathcal{O}$ due to Lemma 1. Furthermore, if $\zeta = 0$ then the cost functions of Problems $\mathcal{O}$ and $\mathcal{R}$ are the same. This is also true when $\zeta = 1$ because

Lemma 1 guarantees that $\|u_i(t)\|_2 = \sigma_i(t)$. The optimal costs thus satisfy $J^*_{\mathcal{O}} \leq J^*_{\mathcal{R}}$. However, any solution of Problem $\mathcal{O}$ is feasible for Problem $\mathcal{R}$ by setting $\sigma_i(t) = \|u_i(t)\|_2$, thus $J^*_{\mathcal{R}} \leq J^*_{\mathcal{O}}$. Therefore $J^*_{\mathcal{R}} = J^*_{\mathcal{O}}$ so the Problem $\mathcal{R}$ solution is globally optimal for Problem $\mathcal{O}$ a.e. $[0, t_f]$. □

Theorem 1a implies that Problem $\mathcal{O}$ is solved in polynomial time by an SOCP solver applied to Problem $\mathcal{R}$. This can be done efficiently with several numerically reliable SOCP solvers Dueri et al. (2014). Therefore the class of $\mathcal{NP}$-hard problems defined by Problem $\mathcal{O}$ is in fact of $\mathcal{P}$ complexity if $x(t) \in \text{int}(\mathcal{X})$ and Conditions 1-4 hold.

5.1 The Case of Active State Constraints

So far it has been assumed that the state constraint ($\mathcal{O}$.g) is inactive. This section guarantees lossless convexification in a limited setting when ($\mathcal{O}$.g) is activated at a discrete set of times. To begin, define the *interior time* and *contact time* sets as follows:

$$\mathcal{T}_i \triangleq \{t \in (0, t_f) : x(t) \in \text{int}(\mathcal{X})\}, \tag{26a}$$

$$\mathcal{T}_c \triangleq [0, t_f] \setminus \mathcal{T}_i. \tag{26b}$$

A point τ of $\mathcal{T}_c$ is called an *isolated point* if there exists a neighborhood of τ not containing other points of $\mathcal{T}_c$ Stein and Shakarchi (2005). A set of isolated points is called a *discrete set* and any discrete subset of a Eucledian space has measure zero Açıkmeşe and Blackmore (2011). We can now prove Theorem 1b.

Proof of Theorem 1b. The proof is similar to (Açıkmeşe and Blackmore, 2011, Corollary 3). To begin, let $\Sigma_O = \{t^*_f, x^*, \xi^*, u^*_i, \gamma^*_i, \sigma^*_i\}$ be the *original solution* returned by Problem $\mathcal{R}$, which achieves the optimal cost value $J^*_{\mathcal{R}}$. Since $\mathcal{T}_c$ is a discrete set, for any consecutive contact times $\tau_1 < \tau_2$ there exists a large enough real $a > 0$ such that $\tau_1 + 1/a < \tau_2 - 1/a$. Let $\tau_e = \tau_1 + 1/a$ and $\tau_f = \tau_2 - 1/a$. Now consider solving Problem $\mathcal{R}$ over $[\tau_e, \tau_e + \Delta\tau]$ with $t_f = \Delta\tau$, $x_0 = x(\tau_e)$, $b[t_f] = x(\Delta\tau) - x(\tau_f)$. Call the solution to this problem the *subproblem solution* $\Sigma_S = \{\tilde{\Delta}\tau, \tilde{x}, \tilde{\xi}, \tilde{u}_i, \tilde{\gamma}_i, \tilde{\sigma}_i\}$, and let J^*_S be the achieved optimal cost. We claim that the corresponding portion of Σ_O must also achieve J^*_S. If it does not, the *modified solution* $\Sigma_M = \{\hat{t}_f, \hat{x}, \hat{\xi}, \hat{u}_i, \hat{\gamma}_i, \hat{\sigma}_i\}$ such that $\hat{t}_f = t^*_f + \tilde{\Delta}\tau - (\tau_f - \tau_e)$ and $\{\hat{x}, \hat{\xi}, \hat{u}_i, \hat{\gamma}_i, \hat{\sigma}_i\} =$

$$\begin{cases} \{x^*(t), \xi^*(t), u^*_i(t), \gamma^*_i(t), \sigma^*_i(t)\} & \text{for } t \in [0, \hat{t}_f] \setminus \\ & [\tau_e, \tau_e + \tilde{\Delta}\tau], \\ \{\tilde{x}(t), \tilde{\xi}(t), \tilde{u}_i(t), \tilde{\gamma}_i(t), \tilde{\sigma}_i(t)\} & \text{for } t \in [\tau_e, \tau_e + \tilde{\Delta}\tau], \end{cases}$$

is also feasible for Problem $\mathcal{R}$ and achieves a lower cost than $J^*_{\mathcal{R}}$, which contradicts that the $[\tau_e, \tau_f]$ segment of Σ_O is optimal. Thus, Σ_S must be optimal for the original problem. By Theorem 1a, Σ_S must be globally optimal for Problem $\mathcal{O}$. Since a is arbitrarily large, Σ_S must be optimal for Problem $\mathcal{O}$ over $t \in (t_1, t_2)$. Let $\mathcal{T}_c = \{\tau_i,\ i = 1, 2, \dots\}$, $\tau_i < \tau_{i+1}\ \forall i$. Hence $\text{int}(\mathcal{T}_i) = \bigcup_i (\tau_i, \tau_{i+1})$ and Σ_O is globally optimal for Problem $\mathcal{O}$ a.e. $\mathcal{T}_i$. Since $\mathcal{T}_c$ is a discrete set, $\text{cl}(\mathcal{T}_i) = [t_0, t_f]$ and so the Problem $\mathcal{R}$ solution is globally optimal for Problem $\mathcal{O}$ a.e. $[0, t_f]$. □

6 Numerical Example

This section shows how rocket landing trajectories can be generated much faster via Problem $\mathcal{R}$ than MICP. Python source code for this example is available online [1]. Consider the in-plane rocket dynamics:

$$\dot{x}(t) = A(\omega)x(t) + B\sum_{i=1}^{M} u_i(t) + w, \tag{27}$$

where the vehicle is treated as a point mass with $x(t) = (r(t), v(t)) \in \mathbb{R}^4$ the position and velocity state and $\omega \in \mathbb{R}$ the planet rotation rate, which is assumed to be constant and perpendicular to the trajectory plane [2]. The input $u_i(t) \in \mathbb{R}^2$ represents an acceleration imparted on the rocket by a gimballed thruster. The LTI matrices are:

$$A(\omega) = \begin{bmatrix} 0 & I \\ \omega^2 I & 2\omega S \end{bmatrix},\ B = \begin{bmatrix} 0 \\ I \end{bmatrix}, w = \begin{bmatrix} 0 \\ \omega^2 l + g \end{bmatrix}, \tag{28}$$

where $S = [0\ 1; -1\ 0] \in \mathbb{R}^{2\times 2}$, $I \in \mathbb{R}^{2\times 2}$ is identity, $l \in \mathbb{R}^2$ is the landing pad position with respect to the planet's center of rotation, and $g \in \mathbb{R}^2$ is the gravity vector. Note that the dynamics assume constant mass and gravity for concision, but both can be made variable within the lossless convexification framework Açıkmeşe and Ploen (2007); Blackmore et al. (2012).

The rocket is equipped with a single gimballed thruster which operates in two modes: 1) low-thrust high-gimbal, and 2) high-thrust low-gimbal. A maximum gimbal angle range of $\theta_i \in (0, \pi)$ is enforced via ($\mathcal{O}$.f) by setting:

$$C_i = \begin{bmatrix} -\cos(\theta_i/2) & -\sin(\theta_i/2) \\ \cos(\theta_i/2) & -\sin(\theta_i/2) \end{bmatrix}. \tag{29}$$

We also impose a glide slope constraint as in Blackmore et al. (2010) which prevents the rocket from approaching the ground too closely prior to touchdown:

$$\mathcal{X} = \{x = (r, v) \in \mathbb{R}^4 : \hat{e}_y^{\mathsf{T}} r \geq \|r\|_2 \sin(\gamma_{gs})\}, \tag{30}$$

where $\hat{e}_y = (0, 1) \in \mathbb{R}^2$ is the unit vector along the altitude axis. We choose the following parameters, corresponding to a Martian divert maneuver similar to Açıkmeşe and Ploen (2007):

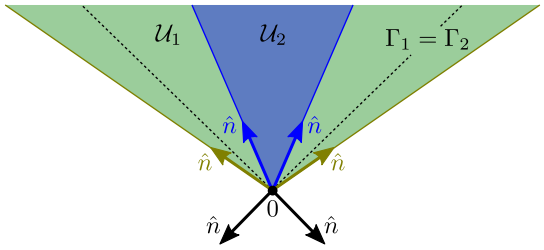
$$M = 2,\ K = 1,\ \omega = 2\pi/88775\ \text{rad}\,\text{s}^{-1},\ \rho^1_1 = 4\ \text{m}\,\text{s}^{-2},$$
$$\rho^2_1 = 8\ \text{m}\,\text{s}^{-2},\ \rho^1_2 = 8\ \text{m}\,\text{s}^{-2},\ \rho^2_2 = 12\ \text{m}\,\text{s}^{-2},\ \theta_1 = 120\ ^\circ,$$
$$\theta_2 = 10\ ^\circ,\ \gamma_{gs} = 10\ ^\circ,\ l = (0, 3396.2)\ \text{km},\ \zeta \in \{0, 1\},$$
$$g = (0, -3.71)\ \text{m}\,\text{s}^{-2},\ m[t_f] = (1 - \zeta)t_f \xi_{\max}/t_{f,\max},$$
$$\ell(x(t)) = 10^{-3}\xi_{\max}(|r_1(t)|\tan(\gamma_{gs}) + |r_2(t)|)/h_0,$$
$$r(0) = (1500, h_0)\ \text{m},\ v(0) = (50, -70)\ \text{m}\,\text{s}^{-1},$$
$$r(t_f) = (0, 0)\ \text{m},\ v(t_f) = (0, 0)\ \text{m}\,\text{s}^{-1},$$

where $t_{f,\max} = 100$ s is the time of flight upper-bound and $\xi_{\max} = t_{f,\max}\rho^2_2$ is the maximum input integral cost. The optimal cost is verified to be unimodal in t_f such that golden search can be applied to find the optimal t_f Blackmore et al. (2010); Kochenderfer and Wheeler (2019). The initial altitude above ground level (AGL) h_0 and $\zeta \in \{0, 1\}$ are independent variables that we shall vary. When $\zeta = 0$, we solve for a minimum-time trajectory, while for $\zeta = 1$ we solve for a minimum-fuel trajectory.

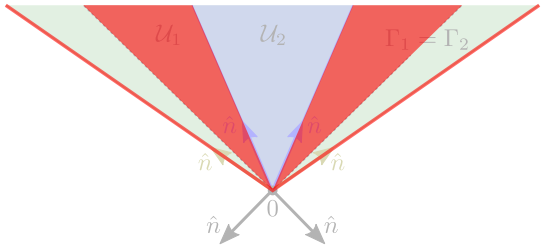
The problem satisfies Conditions 1-4 under a few mild assumptions. Because the glide slope (30) maintains the rocket above zero altitude, $\ell[t] > 0\ \forall t \in [0, t_f)$ such that Condition 4 holds irrespective of m. To check Condition 1, recognize that for our choice of ℓ:

[1] https://github.com/dmalyuta/lcvx

[2] This is done for simplicity in order to keep the motion planar. A general 3-dimensional angular velocity vector can also be considered.



(a) Illustration of the six vectors $\hat{n}$ that $\dot{y}(t)$ must not be normal to for Conditions 2 and 3 to hold.



(b) If the normality check fails, the optimal input could point in the directions highlighted in red.

Fig. 3. Illustrated verification of Condition 2 and 3 when $\zeta = 0$. If $\dot{y}(t)$ can evolve normal to any vector in (a), the input can point in the directions shown in (b) while violating ($\mathcal{O}$.d).

$$\partial\ell[t]^\intercal = D\partial_r\ell[t]^\intercal,\ D \triangleq \begin{bmatrix} I \\ 0 \end{bmatrix}. \tag{31}$$

Following the discussion in Section 3.1, we confirm that the LTI system $\{-A^\intercal, D, B^\intercal, 0\}$ is strongly observable, hence Condition 1 holds. To check Conditions 2 and 3, we need to make the following assumption because replacing $\partial_r\ell[t]^\intercal$ with $\mathbb{R}^2$ is too conservative.

Assumption 3. The downrange and altitude are non-zero almost everywhere, i.e. $r_1(t) \neq 0$ and $r_2(t) \neq 0$ a.e. $[0, t_f]$.

Leveraging Assumption 3 yields a piecewise constant input to the adjoint system:

$$\partial_r\ell[t]^\intercal = \frac{10^{-3}\xi_{\max}}{h_0}\left\{\begin{bmatrix}\tan(\gamma_{gs}) \\ 1\end{bmatrix}, \begin{bmatrix}-\tan(\gamma_{gs}) \\ 1\end{bmatrix}\right\}. \tag{32}$$

Leveraging (32), consider the following LTI system where a constant input is modelled as a static state, yielding an augmented state $\lambda'(t) \in \mathbb{R}^6$:

$$\dot{\lambda}'(t) = \begin{bmatrix}-A^\intercal & D \\ 0 & 0\end{bmatrix}\lambda'(t) = A'\lambda'(t), \tag{33a}$$

$$y(t) = \begin{bmatrix}B^\intercal & 0\end{bmatrix}\lambda'(t) = C'\lambda'(t). \tag{33b}$$

When $\zeta = 0$, checking Conditions 2 and 3 reduces to ensuring that $\dot{y}(t) = C'A'\lambda'(t)$ cannot evolve perpendicular to certain constant vectors $\hat{n} \in \mathbb{R}^2$. The values of $\hat{n}$ that need to be checked are illustrated in Figure 3a. To verify Conditions 2 and 3, we check the observability properties of the pair $\{A', \hat{n}^\intercal C'A'\}$. Let $V_{\hat{n}}$ be a matrix whose columns span the unobservable subspace. It turns out for the rocket landing problem that $A'V_{\hat{n}} = 0\ \forall\hat{n}$. Conditions 2 and 3 can thus be violated only by a constant primer vector. If this occurs, the input is constrained to point in the directions shown in Figure 3b. Notice that this constrains the downrange acceleration to always have the same sign. The following assumption requires the rocket to experience both acceleration *and* deceleration. The assumption is satisfied if, for example, the rocket is initially travelling away from the landing site and has to reverse its velocity.

Table 1. Optimal cost and solver runtime when solving Problem $\mathcal{R}$ versus MICP. Dashes show when MICP took too long to converge (> 10 min per iteration).

h_0 [m]	ζ	$J^*_{\mathcal{R}}$	J^*_{MICP}	$t_{\mathcal{R}}$ [s]	t_{MICP} [s]
650	0	636.2	–	2.9	–
650	1	374.5	–	2.4	–
800	0	577.7	577.8	2.4	232.3
800	1	350.8	350.9	2.3	269.9
1000	0	548.9	–	3.9	–
1000	1	333.7	333.7	2.3	566.8
1500	0	493.4	–	2.5	–
1500	1	316.1	316.1	2.2	177.3
3000	0	558.0	558.0	2.5	73.1
3000	1	323.0	323.1	1.8	505.9

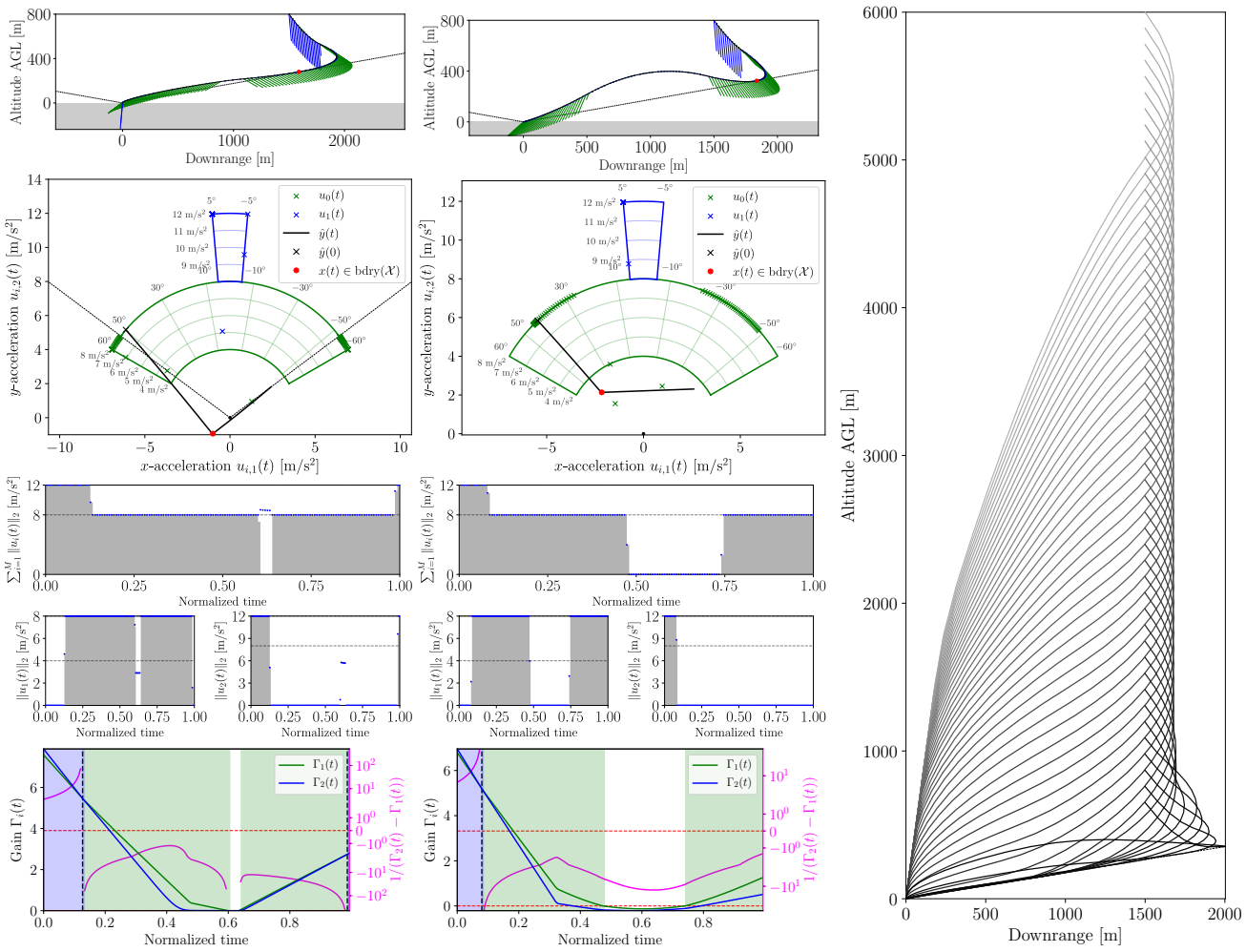
Assumption 4. The downrange acceleration $\sum_{i=1}^{M} u_{i,1}(t)$ changes sign at least once over $[0, t_f]$.

The assumption is sufficient for Theorem 1a but not Theorem 1b, because a discontinuity in $\dot{y}(t)$ may occur at $t \in \mathcal{T}_c$ (26b) Hartl et al. (1995). If state constraints are activated, a "sufficiently rich" gimbal history may be assumed or Conditions 2 and 3 may be verified *a posteriori*, i.e. the solution is lossless if they hold.

When $\zeta = 1$, Condition 2 requires $\|y(t)\|_2 \neq 1$ a.e. $[0, t_f]$. Modal shape analysis for the pair $\{A', C'\}$ reveals that, given a constant input in (32), $\|y(t)\|_2 = 1$ for an interval is only possible if $y(t)$ is constant. This is eliminated by Assumption 4 with the same caveat about state constraint activation. Checking Condition 3 is not possible *a priori* when $\zeta = 1$. The condition is verified *a posteriori*.

The dynamics (27) are discretized via zeroth-order hold on a uniform temporal grid of 150 nodes. Python 2.7.15 and ECOS 2.0.7.post1 Domahidi et al. (2013) are used on a Ubuntu 18.04.1 64-bit platform with a 2.5 GHz Intel Core i5-7200U CPU and 8 GB of RAM. The solution and runtime are compared to a MICP formulation where ($\mathcal{O}$.d) is implemented directly as a binary constraint using Gurobi 8.1 Gurobi Optimization (2018).

Figure 4 shows the resulting state, input and input gain trajectories. Let us first discuss Figures 4a and 4b. The top row shows the overall trajectory, from which we note that Assumptions 3 and 4 are satisfied. The second and third rows show that the input norm is feasible almost everywhere for Problem $\mathcal{O}$. In particular, the thrust magnitude is bang-bang as predicted in Lemma 1. The intermediate thrusts occuring at the rising and falling edges in the third row are discretization artifacts. Recall that the lossless convexification guarantee is only "almost everywhere" in nature. These artifacts have been observed since the early days of lossless convexification theory Açıkmeşe and Ploen (2007). Note the kink that occurs in the $y(t)$ trajectory in the second row, which coincides with the glide slope state constraint activation as highlighted by the red dot in the first row. Looking at the third row, $\sigma_i(t) \neq \|u_i(t)\|_2$ as expected when $\zeta = 0$ and both inputs are off, since there is no cost incentive to minimize $\sigma_i(t)$. Note that optimality nevertheless requires $u_i(t) = 0$, as predicted by Lemma 1. Finally, the fourth row shows the



(a) Landing from $h_0 = 800$ m AGL, $\zeta = 0$. Time of flight $t_f = 46.93$ s.

(b) Landing from $h_0 = 800$ m AGL, $\zeta = 1$. Time of flight $t_f = 53.97$ s.

(c) Trajectory sweep over $h_0 \in [650, 6000]$ m AGL, $\zeta = 0$ and $N = 30$.

Fig. 4. Landing trajectories computed by Problem $\mathcal{R}$. Green shows the high-gimbal low-thrust mode and blue shows the low-gimbal high-thrust mode. In (a) and (b), the top row shows the position trajectory with overlaid thrusts ($-u_i(t)$). Dotted lines show glide slope (30). The second row shows the input with the (normalized) primer vector (3b). Dotted lines show the equal-gain manifold $\Gamma_1(t) = \Gamma_2(t)$. The third row shows the input magnitude history. The bottom row shows each input's gain (4) and their difference. The background colour shows when the corresponding input is active. In (c), landing trajectories are shown for a sweep over the initial altitude AGL.

$\Gamma_i(t)$ trajectories. As predicted by (25), when $\Gamma_i(t) > \Gamma_j(t)$, optimality forces input $\gamma_i(t) = 1$ and $\gamma_j(t) = 0$.

Table 1 compares the achieved optimal cost and solver runtimes of lossless convexification versus a direct MICP implementation of ($\mathcal{O}$.d). One can see that the optimal cost values are quasi-identical, with some slightly lower values for lossless convexification due to the "intermediate thrusts" discussed above. More importantly, solving Problem $\mathcal{R}$ is up to two orders of magnitude faster than using MICP. This is expected, since SOCP has polynomial time complexity in the problem size while MICP has exponential time complexity. Furthermore, MICP was not able to find a trajectory in several cases (the computation was aborted when runtime exceeded 10 min for a single golden search iteration). The third column of Figure 4 shows a sequence of 50 landing trajectories for a sweep over $h_0 \in [650, 6000]$ m AGL. Computing this sequence of 50 trajectories with $N = 150$ takes 130 s, which is less than the average MICP solution time for a single trajectory.

7 Future Work

Future work consists of expanding the class of problems that can be handled. This includes considering different input norm types in ($\mathcal{O}$.a) and ($\mathcal{O}$.c), time-varying dynamics in ($\mathcal{O}$.b), a lower-bound $L \leq \sum_{i=1}^{M} \gamma_i(t)$ in ($\mathcal{O}$.e), a constraint on the input rate of change $\dot{u}_i(t)$, persistently active state constraints in ($\mathcal{O}$.g), and removing the discretization artifacts observed in Section 6. A minor caveat of the Lemma 1 proof is that conditions which are proven to hold "almost everywhere" are assumed not to fail on nowhere dense sets of positive measure (e.g. the fat Cantor set) Morgan II (1990). We do not expect this pathology to occur for any practical problem, and in the future we seek to rigorously eliminate this artifact.

8 Conclusion

This paper presented a lossless convexification method for solving a class of optimal control problems with semi-continuous input norms. By relaxing the problem to a convex one and proving that the relaxed solution is globally optimal for the original problem, solutions can be found via one-shot convex optimization in polynomial time. The resulting algorithm is amenable to real-time onboard implementation and can also be used to accelerate design trade studies.

References

Açıkmeşe, B. and Blackmore, L. (2011). Lossless convexification of a class of optimal control problems with non-convex control constraints. *Automatica*, 47(2), 341–347. doi: 10.1016/j.automatica.2010.10.037.

Açıkmeşe, B., Carson III, J.M., and Blackmore, L. (2013). Lossless convexification of nonconvex control bound and pointing constraints of the soft landing optimal control problem. *IEEE Transactions on Control Systems Technology*, 21(6), 2104–2113. doi: 10.1109/tcst.2012.2237346.

Açıkmeşe, B. and Ploen, S.R. (2007). Convex programming approach to powered descent guidance for Mars landing. *Journal of Guidance, Control, and Dynamics*, 30(5), 1353–1366. doi: 10.2514/1.27553.

Blackmore, L., Açıkmeşe, B., and Carson III, J.M. (2012). Lossless convexification of control constraints for a class of nonlinear optimal control problems. *Systems & Control Letters*, 61(8), 863–870. doi: 10.1016/j.sysconle.2012.04.010.

Blackmore, L., Acikmese, B., and Scharf, D.P. (2010). Minimum-landing-error powered-descent guidance for Mars landing using convex optimization. *Journal of Guidance, Control, and Dynamics*, 33(4), 1161–1171. doi: 10.2514/1.47202.

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press.

Carson III, J.M., Açıkmeşe, B., and Blackmore, L. (2011). Lossless convexification of powered-descent guidance with non-convex thrust bound and pointing constraints. In *Proceedings of the 2011 American Control Conference*. IEEE. doi: 10.1109/acc.2011.5990959.

Clarke, F. (2010). The Pontryagin maximum principle and a unified theory of dynamic optimization. *Proceedings of the Steklov Institute of Mathematics*, 268(1), 58–69. doi: 10.1134/s0081543810010062.

Domahidi, A., Chu, E., and Boyd, S. (2013). ECOS: An SOCP solver for embedded systems. In *European Control Conference (ECC)*, 3071–3076. IEEE.

Dueri, D., Zhang, J., and Açıkmeşe, B. (2014). Automated custom code generation for embedded, real-time second order cone programming. *IFAC Proceedings Volumes*, 47(3), 1605–1612. doi: 10.3182/20140824-6-za-1003.02736.

Gurobi Optimization, L. (2018). Gurobi optimizer reference manual.

Harris, M.W. and Açıkmeşe, B. (2013a). Lossless convexification for a class of optimal control problems with linear state constraints. In *52nd IEEE Conference on Decision and Control*. IEEE. doi: 10.1109/cdc.2013.6761017.

Harris, M.W. and Açıkmeşe, B. (2013b). Lossless convexification for a class of optimal control problems with quadratic state constraints. In *2013 American Control Conference*. IEEE. doi: 10.1109/acc.2013.6580359.

Harris, M.W. and Açıkmeşe, B. (2014). Lossless convexification of non-convex optimal control problems for state constrained linear systems. *Automatica*, 50(9), 2304–2311. doi: 10.1016/j.automatica.2014.06.008.

Hartl, R.F., Sethi, S.P., and Vickson, R.G. (1995). A survey of the maximum principles for optimal control problems with state constraints. *SIAM Review*, 37(2), 181–218. doi: 10.1137/1037043.

Kochenderfer, M.J. and Wheeler, T.A. (2019). *Algorithms for Optimization*. The MIT Press, Cambridge, Massachusetts.

Malyuta, D., Reynolds, T., Szmuk, M., Acikmese, B., and Mesbahi, M. (2020). Fast trajectory optimization via successive convexification for spacecraft rendezvous with integer constraints. In *AIAA Scitech 2020 Forum*. American Institute of Aeronautics and Astronautics. doi: 10.2514/6.2020-0616.

Morgan II, J.C. (1990). *Point Set Theory*. CRC Press.

MOSEK ApS (2019). *MOSEK Modeling Cookbook*, 3.1 edition.

Rockafellar, R.T. and Wets, R.J.B. (1998). *Variational Analysis*. Springer Berlin Heidelberg. doi: 10.1007/978-3-642-02431-3.

Sager, S. (2005). *Numerical methods formixedinteger optimal control problems*. Ph.D. thesis, Heidelberg University, Germany.

Sager, S., Bock, H.G., and Reinelt, G. (2007). Direct methods with maximal lower bound for mixed-integer optimal control problems. *Mathematical Programming*, 118(1), 109–149. doi: 10.1007/s10107-007-0185-6.

Stein, E.M. and Shakarchi, R. (2005). *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*. Princeton University Press.

Trentelman, H.L., Stoorvogel, A.A., and Hautus, M. (2001). *Control Theory for Linear Systems*. Springer.

Varberg, D.E. (1965). On absolutely continuous functions. *The American Mathematical Monthly*, 72(8), 831. doi: 10.2307/2315025.

Vinter, R. (2000). *Optimal Control*. Birkhauser.

Zhang, Z., Wang, J., and Li, J. (2017). Lossless convexification of nonconvex MINLP on the UAV path-planning problem. *Optimal Control Applications and Methods*, 39(2), 845–859. doi: 10.1002/oca.2380.