

Mixed-Integer Formulations for Optimal Control of Piecewise-Affine Systems

Tobia Marcucci Massachusetts Institute of Technology Cambridge, Massachusetts tobiam@mit.edu Russ Tedrake Massachusetts Institute of Technology Cambridge, Massachusetts russt@mit.edu

ABSTRACT

In this paper we study how to formulate the optimal control problem for a piecewise-affine dynamical system as a mixed-integer program. Problems of this form arise typically in hybrid Model Predictive Control (MPC), where at every time step an open-loop optimal control sequence is computed via numerical optimization and applied to the system in a moving horizon fashion. Not surprisingly, the efficiency in the formulation of the underlying mathematical program has a crucial influence on computation times, and hence on the applicability of hybrid MPC to high-dimensional systems.

We leverage on modern concepts and results from the fields of mixed-integer and disjunctive programming to conduct a comprehensive analysis of this formulation problem. Among the outcomes enabled by this novel perspective is the derivation of multiple highly-efficient formulations of the control problem, each of which represents a different tradeoff between the two most important features of a mixed-integer program: the size and the strength. First in theory, then through a numerical example, we show how all the proposed methods outperform the traditional approach employed in MPC, enabling the solution of larger-scale problems.

CCS CONCEPTS

• Computing methodologies → Computational control theory; • Theory of computation → Mixed discrete-continuous optimization;

KEYWORDS

Optimal control, model predictive control, piecewise-affine systems, hybrid systems, mixed-integer programming, disjunctive programming, mixed-integer formulations, perspective formulations

ACM Reference Format:

Tobia Marcucci and Russ Tedrake. 2019. Mixed-Integer Formulations for Optimal Control of Piecewise-Affine Systems. In 22nd ACM International Conference on Hybrid Systems: Computation and Control (HSCC '19), April 16–18, 2019, Montreal, QC, Canada. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3302504.3311801

HSCC '19, April 16–18, 2019, Montreal, QC, Canada

© 2019 Association for Computing Machinery. ACM ISBN 978-1-4503-6282-5/19/04...\$15.00

https://doi.org/10.1145/3302504.3311801

1 INTRODUCTION

PieceWise-Affine (PWA) systems are a very flexible framework for modeling a great variety of phenomena. Originally they were introduced as a generalization of linear systems and as an approximation of more general classes of nonlinear systems [35]; later they became very popular as a more intuitive equivalent of several other families of hybrid systems [21], such as Mixed Logical Dynamical (MLD) systems. Their structure makes them very suitable for different types of analysis (e.g., stability [23], observability and controllability [6]) as well as optimal control [8, 32]. PWA systems have seen many applications [14], to mention a few: automotive control [12], power electronics [18], temporal logic control [41], and they also have recently spread in robotics [20, 25, 34]. Nowadays, together with MLD, they are the most popular modeling framework for hybrid systems in Model Predictive Control (MPC) [13, 14].

Here we are interested in numerical optimal control of PWA systems. In particular, we study the process of recasting, in a form suitable for an optimization software, the logical implications in the PWA dynamics (i.e., different affine dynamics apply in different portions of the state and input space). The outcome of this process is a Mixed-Integer Convex Program (MICP), which can be effectively solved to global optimality via Branch-and-Bound (B&B). Even if efficient disjunctive models of various families of hybrid systems have been proposed in the literature [29, 37], to these days, when formulating an optimal control problem, PWA systems are generally cast in MLD form through the method described in [8, 13]; which typically leads to the synthesis of very inefficient MICPs.

We start this work examining the structure of the optimal control problem: interpreting the PWA dynamics as a disjunctive polytopic set that links the state evolution and the control actions across time, we show how this problem can be naturally interpreted as a disjunctive program. This point of view allows us to leverage on both modern and well-established tools from the fields of mixed-integer and disjunctive programming [3, 15, 40] to efficiently rephrase our problem as a MICP. We then propose multiple problem formulations, each of which is a different tradeoff between the two critical features of a MICP: its size (number of optimization variables and constraints) and its strength (the closeness of the solution of the MICP to the one of its convex relaxation). We show how, in principle, it is possible to span the whole spectrum between a very compact but weak formulation of the problem and a formulation whose convex relaxation is exact, but whose size is generally unmanageable. Finally, we compare both on a theoretical and a practical level the formulations we propose with the traditional approach employed in MPC, showing that the higher efficiency of our formulations can make the difference between solving a problem to global optimality in few minutes and not finding a feasible solution in hours.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

HSCC '19, April 16-18, 2019, Montreal, QC, Canada

Tobia Marcucci and Russ Tedrake

1.1 Related Works

An efficient convex relaxation for switched-affine optimal control based on the use of perspective functions has been recently proposed in [26]; one of the problem formulations presented here (in Section 5.2) follows closely this approach, extending it to the class of systems we consider and specializing this method to the most common objective functions employed in optimal control. Convex relaxations of a similar family of problems have been analyzed in [2, 30], these works however are focused on semidefinite relaxations for linear systems with discrete control inputs.

A strong formulation of the problem is just the first step to decrease solution times: several algorithms have been presented in the literature for the solution of MICPs arising in hybrid MPC [1, 7, 16]. Recently, a heuristic technique to quickly find approximate solutions for this class of problems has been presented in [38] and an algorithm with local convergence guarantees based on operator splitting has been developed in [17]. Also other approaches, not based on mixed-integer optimization, have been proposed for the optimal control of PWA systems, for example explicit MPC [11] or nonlinear-programming techniques [22]. The applicability of the first, however, is limited to very low-dimensional systems, whereas the second assumes a continuity property that not all the PWA systems of interest present and does not provide convergence and global-optimality guarantees.

1.2 Notation

We denote with \mathbb{R} the real numbers and with, e.g., $\mathbb{R}_{>0}$ nonnegative reals. The same notation is used for integers \mathbb{Z} . Writing $x \in \mathbb{R}^n$ we tacitly assume $n \in \mathbb{Z}_{\geq 0}$. For two vectors $x \in \mathbb{R}^{n_x}$ and $y \in \mathbb{R}^{n_y}$, $(x, y) \in \mathbb{R}^{n_x + n_y}$ represents their concatenation. $\mathbb{S}_{>0}^n$ ($\mathbb{S}_{>0}^n$) represents the set of positive semidefinite (definite) symmetric matrices of dimension *n*; for $A \in \mathbb{S}_{\geq 0}^{n}$ we represent its Cholesky decomposition as $A = (A^{1/2})^{\top} A^{1/2}$. For a set $S \subseteq \mathbb{R}^{n}$, we denote its closure as cl(S), its convex hull as conv(S), and its recession cone as $S_{\infty} := \{ d \in \mathbb{R}^n \mid x + \alpha d \in S, \forall x \in S, \forall \alpha \in \mathbb{R}_{\geq 0} \}$. Given $S \subseteq \mathbb{R}^{n_x + n_y}$, proj_x $(S) := \{x \in \mathbb{R}^{n_x} \mid \exists y \in \mathbb{R}^{n_y} : (x, y) \in S\}$ denotes the orthogonal projection of S onto the subspace \mathbb{R}^{n_x} . $gr(f) := \{(x, y) \mid y = f(x)\}$ represents the graph of the function f. With diag(x_1, \ldots, x_n) we refer to the $n \times n$ diagonal matrix whose elements are x_1, \ldots, x_n and, whenever clear from the context, we will denote with 0 (1) a matrix of appropriate dimensions filled with zeros (ones). All physical units may be assumed to be expressed in the MKS system.

2 THE OPTIMAL CONTROL PROBLEM FOR PIECEWISE-AFFINE SYSTEMS

Denoting with $x \in \mathbb{R}^{n_x}$ the state of the system and with $u \in \mathbb{R}^{n_u}$ the control input, we consider the collection of polytopes (i.e., bounded polyhedra) $\mathcal{D}^i := \{(x, u) \mid F^i x + G^i u \leq h^i\}$ for *i* in the index set \mathcal{I} . The dynamics of a discrete-time PWA system can be expressed as

$$x_{+} = A^{i}x + B^{i}u + c^{i} \quad \text{for any} \quad i \in \mathcal{I} \mid (x, u) \in \mathcal{D}^{i}, \qquad (1)$$

where $x_+ \in \mathbb{R}^{n_x}$ is the state at the next time step and *i* the current mode of the system. For brevity we will often represent the stateupdate law (1) in the compact form $x_+ = \psi(x, u) : \mathcal{D} \mapsto \mathbb{R}^{n_x}$, where $\mathcal{D} := \bigcup_{i \in I} \mathcal{D}^i$, and similarly we will denote each affine piece of this map as $\psi^i(x, u) := A^i x + B^i u + c^i : \mathcal{D}^i \mapsto \mathbb{R}^{n_x}$. In the following we will assume the origin (x, u) = 0 to be an equilibrium point and (1) to be *well posed*, i.e., $\psi^i(x, u) = \psi^j(x, u)$, $\forall (x, u) \in \mathcal{D}^i \cap \mathcal{D}^j, \forall (i, j) \in I^2$.

In order to fully expose the geometry of these systems, in this paper we employ a compact description of (1) in terms of its graph

$$\operatorname{gr}(\psi) := \{(x, u, x_+) \mid (x, u) \in \mathcal{D}, x_+ = \psi(x, u)\} = \bigcup_{i \in I} \operatorname{gr}(\psi^i).$$
(2)

Since each set $gr(\psi^i)$ is a polytope, whose halfspace representation $\{(x, u, x_+) \mid P^i(x, u, x_+) \leq q^i\}$ is given by

$$P^{i} := \begin{bmatrix} A^{i} & B^{i} & -I \\ -A^{i} & -B^{i} & I \\ F^{i} & G^{i} & 0 \end{bmatrix}, \quad q^{i} := \begin{bmatrix} -c^{i} \\ c^{i} \\ h^{i} \end{bmatrix},$$
(3)

we can interpret the PWA dynamics (1) as a *disjunctive polytopic set* that constraints the state evolution across time steps. (This geometric interpretation is shared with the recent work [17].)

The main application of optimal control for PWA systems comes from the field of hybrid MPC, a numerical technique that enables the design of optimal feedback controllers for systems in the form (1). Given the current state \bar{x} , in order to regulate the system to the origin, an optimal control problem with the following structure (or a slight variation of it) is solved

$$\min_{\substack{(u_t)_{t=0}^{T-1}, (x_t)_{t=0}^T}} \varphi_T(x_T) + \sum_{t=0}^{I-1} \varphi(x_t, u_t)$$
(4a)

su

bject to
$$x_0 = \bar{x}$$
, (4b)

$$(x_t, u_t, x_{t+1}) \in \operatorname{gr}(\psi), \quad t = 0, \dots, T-1, \quad (4c)$$

 $x_T \in \mathcal{X}_T, \quad (4d)$

where $T \in \mathbb{Z}_{>0}$ is the time horizon of the controller, $\varphi : \mathbb{R}^{n_x+n_u} \mapsto \mathbb{R}$ is a convex stage cost, $\varphi_T : \mathbb{R}^{n_x} \mapsto \mathbb{R}$ is a convex terminal cost, and $X_T := \{x \mid F_T x \leq h_T\}$ is a polyhedral terminal set. (Note that state and input constraints can be enforced through the definition of the domains \mathcal{D}^i .) The outcome of (4) is an open-loop optimal control sequence $(u_t^*)_{t=0}^{T-1}$ with the related state trajectory $(x_t^*)_{t=0}^T$. In MPC only the first action u_0^* is applied to the system; then, after a time step, the state is measured again and (4) is solved in a moving-horizon fashion with the new value of \bar{x} . Asymptotic stability of the closed-loop system can be ensured with a proper choice of the terminal cost $\varphi_T(x)$ and the terminal set X_T [24].

The source of difficulty in problem (4) is the disjunctive constraint (4c), which requires us to take discrete decisions (in which mode to be at each time step). In this paper we analyze the different paths we can follow to efficiently cast this decision problem in the form of a MICP.

3 MIXED-INTEGER FORMULATIONS OF DISJUNCTIVE SETS

In this section we introduce some basic concepts and techniques from the field of disjunctive programming, the interested reader can refer to [3, 40] for more details.

We start considering a *disjunctive convex set* $S := \bigcup_{i \in I} S^i$, where each set $S^i \subset \mathbb{R}^{n_z}$ is convex and closed, but non necessarily bounded, and I is an index set. We can imagine S to force some of the variables of our (for the rest convex) optimization problem to lie in at least one of the regions S^i ; our goal is to derive a Mixed-Integer (MI) formulation of such a set. Let us then define more precisely what a MI formulation of a set is.

Definition 3.1 (Mixed-integer formulation of a set). Consider a set $Q \subset \mathbb{R}^{n_z}$. We say that the set

$$\mathcal{L}_{Q} := \{ (z \in \mathbb{R}^{n_{z}}, \lambda \in \mathbb{R}^{n_{\lambda}}, \mu \in \mathbb{Z}^{n_{\mu}}) \mid l(z, \lambda, \mu) \le v \},$$
(5)

is a *mixed-integer formulation* of Q if $\operatorname{proj}_z(\mathcal{L}_Q) = Q$. In case the function $l : \mathbb{R}^{n_z + n_\lambda + n_\mu} \mapsto \mathbb{R}^{n_l}$ is Linear (Convex) we will call the related formulation MIL (MIC).

Given a MIC formulation of the disjunctive set S, our problem can be stated as a MICP. State-of-the-art solvers for MICP are based on the B&B algorithm [28]. In the application of this method, there are two fundamental features of the formulation we employ that affect overall solution times. The first is the size, i.e., the number of auxiliary variables n_λ , n_μ and constraints n_l : the higher n_λ and n_l the slower will be the solution of each convex subproblem, the higher n_μ the more branches the B&B tree will have and, consequently, the more convex subproblems we will need to solve. The second critical feature of a MI formulation is its strength.

Definition 3.2 (Relaxation of a MI formulation). Given a MI formulation \mathcal{L}_Q of a set Q, we define its *relaxation* as the set¹

$$\tilde{\mathcal{L}}_{Q} := \{ (z \in \mathbb{R}^{n_{z}}, \lambda \in \mathbb{R}^{n_{\lambda}}, \mu \in \mathbb{R}^{n_{\mu}}) \mid l(z, \lambda, \mu) \le \upsilon \}.$$
(6)

Definition 3.3 (Relative strength of two MI formulations). Given a set Q and two MI formulations of it, \mathcal{L}_Q and \mathcal{L}'_Q , we say that the first formulation is *stronger* than the second (and the second is *weaker* than the first) if $\text{proj}_z(\tilde{\mathcal{L}}_Q) \subset \text{proj}_z(\tilde{\mathcal{L}}'_Q)$.

Stronger formulations facilitate the detection of infeasible branches and result in tighter bounds in the B&B algorithm: this eases the pruning process and, consequently, decreases overall solution times.

For a MIC formulation of a set Q, the set $\operatorname{proj}_z(\tilde{\mathcal{L}}_Q)$ is convex and by definition contains Q, this implies that a formulation such that $\operatorname{proj}_z(\tilde{\mathcal{L}}_Q) = \operatorname{conv}(Q)$ is the strongest MIC formulation we can obtain. MIC formulations with this property (or, more broadly, for which $\operatorname{proj}_z(\tilde{\mathcal{L}}_Q) = \operatorname{cl}\operatorname{conv}(Q)$) are called *sharp*. Note that this makes it meaningful to speak about the strength of a MIC formulation in an absolute sense, as the tightness with which the projection of its relaxation covers the set $\operatorname{conv}(Q)$.

In the following we present the most popular MI formulations for disjunctive sets and we apply them to the set $gr(\psi)$ to express the PWA dynamics (1) in MIL form.

3.1 Big-M Formulation

The big-M is the most common MI formulation for disjunctive sets, here we describe a recent improvement of it [39, 40].

Consider the disjunctive set S union of multiple polyhedra $S^i := \{z \mid P^i z \leq q^i\} \neq \emptyset$ with common recession cone S^i_{∞} . We define $M^{ij} := \max_{z \in S^j} P^i z - q^i, \forall (i, j) \in I^2$. Denoting with n_{P^i} the number of rows of P^i , we have a total of $|I| \sum_{i \in I} n_{P^i}$ big-M parameters and, since the polyhedra S^i are not empty and have a common recession cone, each maximization is a feasible Linear Program (LP) and its optimal value is finite [40, Proposition 6.1].

PROPOSITION 3.4. The set of z and $\mu^i \in \{0, 1\}$, with $i \in I$, for which the conditions $\sum_{i \in I} \mu^i = 1$ and $P^i z \leq q^i + \sum_{j \in I \setminus \{i\}} \mu^j M^{ij}$, $\forall i \in I$, are verified, is a MIL formulation of S.

PROOF. Consider $\mu^j = 1$ and, consequently, $\mu^i = 0$, $\forall i \in I \setminus \{j\}$. We have $P^j z \leq q^j$ and $P^i z \leq q^i + M^{ij}$, the second of which is redundant by definition of M^{ij} .

The main advantage of the this method is the absence of auxiliary continuous variables; on the other hand, even this improved big-M formulation which uses $|\mathcal{I}|^2$ big-M vectors, as opposed to the $|\mathcal{I}|$ of the traditional one, is not sharp [40, Example 12].

3.1.1 Big-M Formulation of the PWA Dynamics. Applying the big-M method to derive a MIL formulation of the disjunctive set $gr(\psi)$ from (2) we obtain, for each couple $(i, j) \in I^2$, a big-M vector

$$\begin{bmatrix} M_1^{ij} \\ M_2^{ij} \\ M_3^{ij} \end{bmatrix} \coloneqq \max_{(x,u)\in\mathcal{D}^j} \begin{bmatrix} (A^i - A^j)x + (B^i - B^j)u + c^i - c^j \\ (A^j - A^i)x + (B^j - B^i)u + c^j - c^i \\ F^ix + G^iu - h^i \end{bmatrix}$$
(7)

and the constraints

Æ

$$A^{i}x + B^{i}u + c^{i} - x_{+} \leq \sum_{j \in I \setminus \{i\}} \mu^{j} M_{1}^{ij}, \quad \forall i \in I,$$
 (8a)

$$x_{+} - A^{i}x - B^{i}u - c^{i} \le \sum_{j \in I \setminus \{i\}} \mu^{j} M_{2}^{ij}, \qquad \forall i \in I,$$
(8b)

$$F^{i}x + G^{i}u \le h^{i} + \sum_{j \in I \setminus \{i\}} \mu^{j} M_{3}^{ij}, \qquad \forall i \in I, \qquad (8c)$$

$$\sum_{i \in I} \mu^i = 1.$$
(8d)

Note that, the substitution of the set $\{(x, u, x_+) \mid \exists \{\mu^i \in \{0, 1\}\}_{i \in I} :$ (8)} to gr(ψ) in (4c) is already sufficient to express (4) as a MICP.

3.2 Convex-Hull Formulation

We now describe the convex-hull method that, as the name suggests, is a technique to derive sharp formulations of disjunctive sets. Here we assume S to be the union of sets defined by quadratic inequalities $S^i := \{z \mid z^\top H^i_j z + (p^i_j)^\top z \le q^i_j, \forall j \in \mathcal{J}_i\} \neq \emptyset, H^i_j \in \mathbb{S}^{n_z}_{\ge 0}$. (This is a more general class of sets than the one considered in Section 3.1, the need for this increase in generality will be clear in the Section 5.2 where we will describe perspective formulations.) The following theorem characterizes the closure of the convex hull of S.

THEOREM 3.5. $z \in \text{cl conv}(S)$ if and only if there exist $z^i \in \mathbb{R}^{n_z}$ and $\mu^i \in \mathbb{R}_{\geq 0}$, with $i \in I$, such that

$$(z^{i})^{\top} H_{j}^{i} z^{i} + \mu^{i} (p_{j}^{i})^{\top} z^{i} \leq (\mu^{i})^{2} q_{j}^{i}, \qquad \forall j \in \mathcal{J}_{i}, \forall i \in I, \qquad (9a)$$
$$(p_{i}^{i})^{\top} z^{i} \leq \mu^{i} q_{i}^{i}, \qquad \forall j \in \mathcal{J}_{i}, \forall i \in I, \qquad (9b)$$

$$(p_j^{\iota})^{\iota} z^{\iota} \le \mu^{\iota} q_j^{\iota}, \qquad \forall j \in \mathcal{J}_i, \forall i \in I, \qquad (9b)$$

$$(1, z) = \sum (\mu^i, z^i). \qquad (9c)$$

$$1, z) = \sum_{i \in I} (\mu^{i}, z^{i}).$$
(9c)

PROOF. This result is a particular case of [9, Proposition 3.3.5], whose proof can readily adapted to this case expressing each quadratic set S^i as the intersection of Second-Order Cones (SOCs).²

¹In this paper we are only interested in *binary* MI formulations, i.e. $\mu \in \{0, 1\}^{n\mu}$. In the relaxation of these we will assume μ to take values in the interval $[0, 1]^{n\mu}$.

² An even more general statement of Theorem 3.5 can be found in [15], where the closed convex hull of the union of generic convex sets is characterized using perspective functions.

Theorem 3.5 contains as a special case the well-known result from Balas [5, Theorem 2.1] for which the closed convex hull of a disjunctive polyhedral set can be expressed as the projection of an higher-dimensional polyhedron. In fact, if $H_j^i = 0$ we have: $S^i := \{z \mid P^i z \leq q^i\}$ (where we identified p_j^i and q_j^i with the *j*th rows of P^i and q^i), equation (9a) is redundant, and condition (9b) can be expressed as $P^i x^i \leq \mu^i q^i$, $\forall i \in I$.

Noticing that the recession cone of the quadratic set S^i can be expressed as $S^i_{\infty} = \{z \mid H^i_j z = 0, (p^i_j)^\top z \leq 0, \forall j \in \mathcal{J}_i\}$ [10, Proposition 1.5.7], we have the following corollary.

COROLLARY 3.6. Assume that the sets S^i have a common recession cone S^i_{∞} , then the set defined by (9) is a sharp MIC formulation of S.

PROOF. Since (9a) is the affine transformation of the rotated SOC { $(y \in \mathbb{R}^{n_y}, t \in \mathbb{R}_{\geq 0}, s \in \mathbb{R}_{\geq 0}) \mid y^\top y \leq ts$ } under the map $(y, t, s) = ((H_j^i)^{1/2} z^i, \mu^i, \mu^i q_j^i - (p_j^i)^\top z^i)$, the set defined by (9) is convex. Consider $\mu^j = 1$ and, consequently, $\mu^i = 0, \forall i \in I \setminus \{j\}$. From (9a) we get $z^j \in S^j$ whereas from (9a) and (9b) it follows $z^i \in S_{\infty}^i$ and, since by assumption $S_{\infty}^i = S_{\infty}^j$, (9c) implies that $z \in S^j$. The resulting MI formulation is sharp by definition. \Box

3.2.1 Convex-Hull Formulation of the PWA Dynamics. We now apply the convex-hull formulation to the disjunctive set $gr(\psi)$, to do this we consider the special case of Theorem 3.5 for which $H_j^i = 0$. After removing the $|\mathcal{I}|$ auxiliary copies of x_+ , we get the concise MIL formulation

$$F^{i}x^{i} + G^{i}u^{i} \le \mu^{i}h^{i}, \qquad \forall i \in I,$$
(10a)

$$(1, x, u, x_{+}) = \sum_{i \in I} (\mu^{i}, x^{i}, u^{i}, A^{i}x^{i} + B^{i}u^{i} + \mu^{i}c^{i}).$$
(10b)

Also in this case, substituting the set $\{(x, u, x_+) \mid \exists \{\mu^i \in \{0, 1\}, x^i \in \mathbb{R}^{n_x}, u^i \in \mathbb{R}^{n_u}\}_{i \in I} : (10)\}$ to $\operatorname{gr}(\psi)$ in (4c), the overall problem (4) would already become a MICP.

4 FORMULATION STRENGTHENING

In the previous section we have analyzed MI formulations for disjunctive sets, here we extend the analysis to MI formulations of disjunctive optimization problems. We start considering the more generic and less structured version of (4)

min
$$f(z)$$
 subject to $z \in \bigcap_{t \in \mathcal{T}} \bigcup_{i \in I_t} S_t^i$, (11)

where \mathcal{T} and I_t are index sets, and we define $S_t := \bigcup_{i \in I_t} S_t^i$ and $S := \bigcap_{t \in \mathcal{T}} S_t$. We assume $f : \mathbb{R}^{n_z} \mapsto \mathbb{R}$ to be a convex function and the sets $S_t^i \neq \emptyset$ to be compact and convex. In the language of disjunctive programming the sets S and S_t are said to be, respectively, in *regular form* (intersection of unions of sets) and in *disjunctive normal form* (single union of sets) [4, 33].

In analogy with the definitions from Section 3, any MI Program (MIP) equivalent³ to problem (11) is called a MI formulation of (11), its relaxation is the convex program obtained by dropping the integrality constraints, and the strength of a formulation quantifies the gap between the optimal value of the MIP and the optimal value

of its relaxation. With a little abuse of terminology, we will also call a MI formulation of (11) sharp in case the latter gap is zero.

Clearly, a MI formulation of (11) can be derived just by finding a MI formulation of each set S_t , however this is not the only option. We now describe two techniques to rearrange problem (11) in such a way that the application of the methods presented in Section 3 leads to stronger problem formulations. As we will see, the combination of these two will allow us, in principle, to get arbitrarily strong formulations, all the way up to the synthesis of a sharp MICP.

4.1 Basic Step

We start observing that for two disjunctive convex sets S_t and S_s in normal form, with $(t, s) \in T^2$, the intersection

$$S_t \cap S_s = \bigcup_{(i,j) \in I_t \times I_s} S_t^i \cap S_s^j$$
(12)

is still a disjunctive convex set in normal form and, moreover,

$$\operatorname{conv}(\mathcal{S}_t \cap \mathcal{S}_s) \subseteq \operatorname{conv}(\mathcal{S}_t) \cap \operatorname{conv}(\mathcal{S}_s).$$
(13)

This suggests that, to increase the strength of a MI formulation of (11), instead of reformulating each set S_t independently, we could first intersect some of these and then reformulate the resulting set. The operation in (12) is called *basic step* and its repeated application leads to stronger and stronger MI problem formulations (a so-called *hierarchy* of relaxations) until when, after $|\mathcal{T}| - 1$ steps, the entire constraint set of (11) is cast in disjunctive normal form [4, 33]. Clearly the effectiveness of this operation depends on the relative volume of the two sets in (13), but also on the number of empty sets generated in (12): in case most of these intersections result in non-empty sets, in fact, the iteration of this step would lead to an exponential growth of the number of disjunctions and, consequently, to MI formulations of unmanageable size.

4.2 **Perspective Formulation**

As seen in Section 4.1, through the application of $|\mathcal{T}| - 1$ basic steps, problem (11) is cast in disjunctive normal form

$$f^* := \min f(z)$$
 subject to $z \in \mathcal{S} := \bigcup_{i \in I} \mathcal{S}^i$. (14)

In this subsection we show how to derive a sharp MI formulation for this family of problems.

Defining $\bar{S}^i := \{(z, s) \mid z \in S^i, s \ge f(z)\}$, we restate (14) in the form $f^* = \min_{(z,s)\in\bar{S}} s$, where $\bar{S} := \bigcup_{i\in I} \bar{S}^i$. In case of a sharp formulation of \bar{S} , from the linearity of the objective function, it now follows that⁴ (see Figure 1)

$$f^* = \min_{\substack{(z,s) \in \operatorname{conv}(\tilde{S})}} s = \min_{\substack{(z,s,\lambda,\mu) \in \tilde{\mathcal{L}}_{\tilde{S}}}} s.$$
(15)

Our original problem (14) is hence replaced with an equivalent convex program whose optimal value is f^* . In the general convex case this process entails the use of perspective functions to describe the set conv(\bar{S}), for this reason, a MI formulation of this type is generally called a *perspective formulation* of problem (14) [15, 19].

In Section 6 we will use the following proposition to compare different MI formulations of problem (4) (see Figure 1).

³Two optimization problems are said equivalent if either they are both infeasible or they have the same optimal cost.

⁴ Assuming each set S^i to be compact, conv (\tilde{S}) is closed, hence we can omit to explicitly take the closure of it in (15).



Figure 1: Illustration of the perspective formulation and the equivalence between a disjunctive and a continuous convex program. The derivation of a (sharp) MI formulation of S and the minimization of f over its convex relaxation can result in a weak lower bound on f^* (white cross). The use of the perspective formulation, on the other hand, leads to a convex program with optimal value equal to f^* (white star).

PROPOSITION 4.1. Let $s^*(z) := \min_{s \mid (z,s) \in \operatorname{conv}(\bar{S})} s$. For all $z \in \operatorname{conv}(S)$ we have $s^*(z) \ge f(z)$.

PROOF. Given $z \in \text{conv}(S)$, $s^*(z)$ is the smallest s for which $\exists \{\zeta^i, \sigma^i, \mu^i \in \mathbb{R}_{\geq 0}\}_{i \in I}$ such that $(z, s) = \sum_{i \in I} \mu^i(\zeta^i, \sigma^i)$, with $(\zeta^i, \sigma^i) \in \bar{S}^i, \sum_{i \in I} \mu^i = 1$. Hence, for $s = s^*(z)$, it holds $\sigma^i = f(\zeta^i)$ and, by convexity of f, it follows $s^*(z) = \sum_{i \in I} \mu^i f(\zeta^i) \ge f(z)$. \Box

5 STRENGTH/SIZE COMPROMISE IN THE OPTIMAL CONTROL PROBLEM

In the previous section we have shown how, in principle, the combination of the basic step and a perspective formulation allows us to span the whole spectrum between a small (but weak) and a sharp (but generally very large) MI formulation of any problem in the form (11), and in particular our control problem (4). Here we come back to problem (4), and we discuss under which circumstances the application of these techniques is actually beneficial.

5.1 Basic Step

In our control problem we have three groups of constraints: the initial conditions (4b), the PWA model (4c), and the terminal constraint (4d). All of these are disjunctive constraints (with the first and the last degenerate cases, unions of a single set) and, in this sense, problem (4) is a particular case of (11). As pointed out in Section 4.1, the effectiveness of the basic step depends the possibility to generate smaller sets when intersecting two disjunctive constraints and, since this can happen only if two constraints enforce limits on the same set of variables, the temporal structure of our control problem allows us to focus only on the following three cases.

5.1.1 Initial Conditions. Intersecting the initial conditions (4b) with the dynamics of the first time step, we get the disjunction

$$(x_0, u_0, x_1) \in \bigcup_{i \in \mathcal{I}} \left(\operatorname{gr}(\psi^i) \cap (\{\bar{x}\} \times \mathbb{R}^{n_u + n_x}) \right).$$
(16)

Even if this seems a harmless process, it has a significant pitfall: the use of the current state of the system \bar{x} implies that this operation has to be performed online. Imagine, for example, to apply the

big-M formulation to the set in (16): the computation of the big-Ms would require the online solution of $2n_x |\mathcal{I}|^2 + |\mathcal{I}| \sum_{i \in \mathcal{I}} n_{F^i}$ LPs (with n_{F^i} number of rows of F^i) that, depending on the application, could take a non-negligible amount of time. In case of the convex-hull method the operation in (16) has less severe, but more subtle, consequences: the parameter \bar{x} would be multiplied by the binary indicators μ^i and would hence affect the constraint matrix of the problem. This can be unappealing since it forces to assemble part of the problem and perform the most expensive linear algebra online, and also makes the warm start across time steps more involved.

5.1.2 Piecewise-Affine Model. The application of the basic step (12) to a couple of consecutive constraints in the form (4c) has a very handy interpretation. Consider the PWA system

$$y_{+} = A^{ij}y + B^{ij}v + c^{ij} \quad \text{for any} \quad (i,j) \in \mathcal{I}^{2} \mid (y,v) \in \mathcal{D}^{ij}, \quad (17)$$

where, $\forall (i,j) \in I^2$, we define $\mathcal{D}^{ij} := \{(y,v) \mid F^{ij}y + G^{ij}v \le h^{ij}\},$

$$A^{ij} := \begin{bmatrix} 0 & A^{i} \\ 0 & A^{j}A^{i} \end{bmatrix}, B^{ij} := \begin{bmatrix} B^{i} & 0 \\ A^{j}B^{i} & B^{j} \end{bmatrix}, c^{ij} := \begin{bmatrix} c^{i} \\ c^{j} + A^{j}c^{i} \end{bmatrix},$$
(18a)
$$r^{ij} = \begin{bmatrix} 0 & F^{i} \\ 0 & F^{i} \end{bmatrix}, c^{ij} = \begin{bmatrix} G^{i} & 0 \\ 0 & I^{ij} \end{bmatrix},$$
(18b)

$$F^{ij} := \begin{bmatrix} 0 & F^{j}A^{i} \end{bmatrix}, G^{ij} := \begin{bmatrix} F^{j}B^{i} & G^{j} \end{bmatrix}, h^{ij} := \begin{bmatrix} h^{j} - F^{j}c^{i} \end{bmatrix}.$$
 (18b)
If, for a given time step *t*, we interpret *y* as (x_{t-1}, x_t) , *v* as (u_t, u_{t+1}) , and y_{+} as (x_{t+1}, x_{t+2}) , it is easy to verify that enforcing (17) is equiv-

and y_+ as (x_{t+1}, x_{t+2}) , it is easy to verify that enforcing (17) is equivalent to require $(x_t, u_t, x_{t+1}, u_{t+1}, x_{t+2})$ to lie in the intersection of two consecutive constraints (4c). After a straightforward modification of the initial conditions, the objective function, and the terminal set, we can therefore formulate a problem equivalent to (4) where the original PWA system is replaced by (17) and the horizon of the problem is cut in half.⁵ A PWA system analogous to (17) can be easily defined in case we decide to adopt this procedure to *condense* more than two stages.

The basic step can be beneficial in case many of the sets \mathcal{D}^{ij} result to be empty (a check that can be performed through a LP), otherwise the increase in the size of the formulation is not worth the higher strength and the lower combinatorial complexity of the MICP we synthesize. A deeper analysis of this tradeoff is given in Section 7.1 in a case of a simple optimal control problem.

We conclude emphasizing that, even in case we decide to apply this technique, since the structure of (17) is identical to the one of (1), both the big-M and the convex-hull formulations (as well as the perspective formulation we present in Section 5.2) can be applied without any adjustment.

5.1.3 Terminal Constraint. Intersecting the terminal set with the dynamics of the terminal stage we get the disjunctive set

$$(x_{T-1}, u_{T-1}, x_T) \in \bigcup_{i \in \mathcal{I}} \left(\operatorname{gr}(\psi^i) \cap (\mathbb{R}^{n_x + n_u} \times \mathcal{X}_T) \right).$$
(19)

With the big-M method this operation requires the computation of specialized big-Ms for the terminal stage but, on the other hand, this computation can be done offline. The convex-hull formulation deals with this more transparently: we get again (10) with the additional constraints $F_T x_T^i \leq \mu_T^i h_T$, $\forall i \in I$. In both cases, the only drawback is that we generate |I| copies of the terminal constraint which, in some circumstances, can be defined by many inequalities [13].

⁵Note that the assumption in Section 2 on the boundedness of the sets \mathcal{D}^i can be safely relaxed requiring the sets $\operatorname{gr}(\psi^i)$ to share a common recession cone.

5.2 Perspective Formulation

In Section 4 we have seen how, in principle, any program in regular form (e.g. (4)) can be cast in normal form, and how to derive a sharp MI formulation for the latter family of problems. This recasting process, however, has generally the price of drastically enlarging the size of our optimization problem. The multistage structure of (4), on the other hand, allows us to take advantage of the principles discussed in Section 4.2 without the need of expressing the problem in normal form [26]. In this subsection we focus on the three most popular objectives in MPC and we show how in these cases we are able to synthesize very tractable instances of our problem.

5.2.1 *Linear Objective.* We consider $\varphi(x, u) := ||Qx||_{\infty} + ||Ru||_{\infty}$ for some weight matrices Q, R of appropriate dimensions. We introduce the variables $s_x \in \mathbb{R}_{\geq 0}$, $s_u \in \mathbb{R}_{\geq 0}$ and we substitute the stage cost $\varphi(x, u)$ with $s_x + s_u$. The disjunction (4c) is now replaced by

$$(x, u, x_{+}, s_{x}, s_{u}) \in \bigcup_{i \in I} \{ (x, u, x_{+}, s_{x}, s_{u}) \mid (x, u, x_{+}) \in \operatorname{gr}(\psi^{i}), 1s_{x} \ge \pm Qx, 1s_{u} \ge \pm Ru \}.$$
(20)

Applying the convex hull formulation to this new disjunctive set, we get again (10) plus the additional constraints

$$(s_x, s_u) = \sum_{i \in \mathcal{I}} (s_x^i, s_u^i), \tag{21a}$$

$$1s_x^i \ge \pm Qx^i, \quad 1s_u^i \ge \pm Ru^i, \qquad \forall i \in \mathcal{I},$$
 (21b)

where $s_x^i \in \mathbb{R}_{\geq 0}$ and $s_u^i \in \mathbb{R}_{\geq 0}$, $\forall i \in I$. Note that the disjunctive set in (20) is actually unbounded, but since the directions of recession do not depend on *i*, we can still apply the convex-hull method (cf. Corollary 3.6). It is worth emphasizing that in this case the problem formulation is MIL, as it would have been applying the big-M or the convex-hull formulation.

In case of 1-norm cost function $\varphi(x, u) := \|Qx\|_1 + \|Ru\|_1$, the process is almost identical. Now the slack variables $s_x \in \mathbb{R}_{\geq 0}^{n_x}$ and $s_u \in \mathbb{R}_{\geq 0}^{n_u}$ are vectors, hence the stage cost is substituted by $1^{\top}s_x + 1^{\top}s_u$ and in (20) we have $s_x \geq \pm Qx$, $s_u \geq \pm Ru$. The convex-hull method now leads to (10), (21a), and $s_x^i \geq \pm Qx^i$, $s_u^i \geq \pm Ru^i$, $\forall i \in I$. Also in this case the resulting optimization problem is MIL.

5.2.2 *Quadratic Objective.* In case of quadratic cost $\varphi(x, u) := x^{\top}Qx + u^{\top}Ru$, with $Q \in \mathbb{S}_{\geq 0}^{n_x}$ and $R \in \mathbb{S}_{\geq 0}^{n_u}$, we introduce $s \in \mathbb{R}_{\geq 0}$ which takes the place of the stage cost $\varphi(x, u)$, and the disjunctive constraint (4c) becomes

$$(x, u, x_+, s) \in \bigcup_{i \in I} \{(x, u, x_+, s) \mid (x, u, x_+) \in \operatorname{gr}(\psi^i), s \ge x^\top Q x + u^\top R u\}.$$
(22)

Applying the convex hull formulation we get again (10) and

$$s = \sum_{i \in \mathcal{I}} s^{i}, \quad \mu^{i} s^{i} \ge (x^{i})^{\top} Q x^{i} + (u^{i})^{\top} R u^{i}, \quad \forall i \in \mathcal{I},$$
(23)

where $s^i \in \mathbb{R}_{\geq 0}, \forall i \in \mathcal{I}$.

Also in this case the unboundedness of (22) does not represent an issue since all the sets have common recession cone. The overall optimization problem is in this case a MISOC program, which is still a MICP, but is harder to solve than the MI quadratic program we would have got using the big-M or the convex-hull method.

6 COMPARISON OF THE PROBLEM FORMULATIONS

We finally compare the proposed formulations: as a benchmark we consider the traditional approach in MCP, i.e., the derivation of a equivalent MLD representation of (1) as described in [8, Section 3.1].

Because of the auxiliary continuous variables required by the MLD formulation, we cannot provide a rigorous proof that the big-M method from Section 3.1.1 is strictly stronger, even if in practice this turns out to be true (see Section 7.2). Nevertheless, our big-M approach has three clear advantages: the use of a specialized big-M for each couple $(i, j) \in I^2$ makes this formulation generally much stronger than the MLD, this formulation does not require auxiliary continuous variables (as opposed to the $|I|n_x$ per stage of the MLD), and, finally, this formulation has a smaller size $(n_x$ equalities and $2|I|n_x$ inequalities per stage less than the MLD).

The convex-hull method in Section 3.2.1 is sharp by definition, hence it is the strongest possible MIC formulation of the PWA dynamics (1), in particular, stronger than the big-M and the MLD; furthermore, it does not require the computation of any big-M. However, this has the price of introducing $|\mathcal{I}|(n_x + n_u)$ auxiliary continuous variables per stage. The number of constraints required by the two approaches is similar.

For what concerns the perspective formulation from Section 5.2, we have that any feasible solution of the relaxation of problem (4) reformulated with the convex-hull method will still be feasible in this case (and vice versa), but the contribution to the total cost of each stage will now be greater than the one of the previous method (cf. Proposition 4.1). Hence the relaxation of the optimal control problem, reformulated with this technique, will lead to a higher lower bound on the solution of (4) than all the other approaches. This at the cost of increasing the size of the problem and, in case of quadratic objective, introducing the SOC constraint in (23).

7 EXAMPLES

We conclude testing the problem formulations we presented on two examples. We consider two mechanical systems that have to be controlled exploiting the contact forces arising from the impacts with the environment. Optimal control through contacts is nowadays one of the central themes in robotics, especially in locomotion and manipulation: state-of-the-art approaches to this problem typically rely on nonlinear-optimization methods [27, 31], calling for the development of more reliable techniques.

7.1 Cart with Elastic Walls

In this subsection we analyze a toy example to discuss the effects of the basic step presented in Section 5.1. We consider the mechanical system depicted in Figure 2: the task is to regulate the cart to the origin by exploiting the impacts that occur when colliding with the walls. We denote with x_1 the horizontal position of the cart and with x_2 its velocity, the control input is the force *u*. A simple discrete-time PWA model of the system is the following: ⁶

$$x_{+} = \begin{cases} (x_{1} + hx_{2}, x_{2} + hu/m) & \text{if } |x_{1} + hx_{2}| \le d - r, \\ (x_{1}, -\varepsilon x_{2}) & \text{if } |x_{1} + hx_{2}| > d - r, \end{cases}$$
(24)

⁶Domains described by strict inequalities are approximated with closed sets introducing a small numeric tolerance as in [13, Section 16.2.1].



Figure 2: Regulation to the origin of a cart with elastic walls.

where *h* is the discretization step, *m* is the mass of the cart, $\varepsilon > 0$ is the coefficient of restitution, *d* and *r* are geometric parameters (see Figure 2). The system is hence a double integrator whose velocity is changed of sign and scaled by ε whenever a collision is detected. The total number of modes is $|\mathcal{I}| = 3$: i = 1 no contact, i = 2 contact with the left wall, and i = 3 contact with the right wall.

If we assume that $x_1 \le d - r$ and $x_1 + hx_2 > d - r$ (i.e., the cart is on the left of the right wall and is going to collide with it), then we must have $x_2 > 0$ and, at the next time step, $x_{1+} = x_1 \le d - r$, $x_{2+} = -\varepsilon x_2 < 0$, which implies $x_{1+} + hx_{2+} < d - r$. This shows that any mode sequence which requires the system to be in mode 2 or 3 for two consecutive time steps is infeasible, independently from the control input *u*. Similarly, we see that, because of the inertia of the system, any reasonable choice of the problem parameters would also make infeasible any mode sequence which presents a 2 and a 3 within a small time window (we assume this window to be wider than 5 time steps). These observations make (24) a system very suitable for the application of the basic step from Section 5.1.2.

In Table 1 we analyze the effects of the basic step on system (24). For an increasing number of basic steps (i.e., for an increasing number of stages condensed in a single PWA system (17)) we report the size of the MI formulation we get applying the convex-hull method from Section 3.2.1; in order to make these values comparable, we normalize them to a generic time horizon *T*. For example, a single application of the basic step (condensation of two stages) would lead to a new PWA system (17) with five modes $(\{(i, j) \mid \mathcal{D}^{ij} \neq \emptyset\} = \{(1, 1), (1, 2), (1, 3), (2, 1), (3, 1)\})$ and the optimal control problem for this system would have, e.g., 2.5*T* binary variables as opposed to the 3*T* of the original problem.

Table 1 clearly highlights how approaching this problem one stage per time would add a fictitious combinatorial complexity to the system dynamics. Even in case of an optimal control problem with a short horizon, e.g. T = 10, modeling each time step independently we would get an optimization problem with worst-case complexity proportional to $|\mathcal{I}|^T = 3^{10} = 59,049$. On the other hand, grouping for example the stages five by five, we would only have $25^{T/5} = 1.90^{10} = 625$ potential binary assignments. This is on top of the fact that each application of the basic step leads to a stronger MI formulation in case of use of the convex-hull method (cf. (13)). On the other hand, Table 1 also reveals that every application of the basic step entails a substantial increase in the size of the MICP.

In conclusion, an effective application of the basic step relies on our ability to find the right compromise between the strength, the worst-case complexity, and the size of the MICP we formulate. This tradeoff is in general non-trivial and, because of the many factors which influence computation times, very case dependent.

Table 1: Size and complexity of a generic optimal control problem with horizon *T* for the system in Figure 2 after the application of the basic step. Results are obtained applying the convex-hull method from Section 3.2.1 and assuming the presence of box constraints on the state and the input.

Condensed stages	1	2	3	4	5
Modes of condensed system	3	5	9	15	25
$1/T \cdot \text{Binary variables}$	3	2.5	3	3.75	5
$1/T \cdot \text{Continuous variables}$	12	18	30	48	78
$1/T \cdot \text{Equality constraints}$	6	5.5	5.33	5.25	5.2
1/T · Inequality constraints	22	38	68.67	115	192
$\sqrt[T]{\text{Binary assignments}}$	3	2.24	2.08	1.97	1.90



Figure 3: Benchmark problem: rotate a two-dimensional ball by 180° . The ball is subject to the gravity force and can collide with both the paddle (bottom) and the ceiling (top), the control input is the translational acceleration of the paddle.

7.2 Ball and Paddle

7.2.1 Task Description. In this subsection we consider the problem, illustrated in Figure 3, of flipping a ball through the use of a paddle. The control input $u \in \mathbb{R}^2$ is the translational acceleration of the paddle (which cannot rotate); the state $x \in \mathbb{R}^{10}$ of the system is composed by: x_1, x_2 position of the ball, x_3 angle of the ball, x_4, x_5 position of the paddle, and their time derivatives. The task is to rotate the ball from the initial state $\bar{x}_3 = \pi$, $\bar{x}_k = 0 \forall k \in \{1, \ldots, 10\} \setminus \{3\}$ to the origin. The motion of the ball is governed by the gravity and the ceiling. The dynamics is discretized using the semi-implicit Euler method, resulting in

$$x_{k+} = x_k + hx_{(k+5)+},$$
 $k = 1, \dots, 5,$ (25a)

$$x_{6+} = x_6 + h(f_{pt} - f_{ct})/m,$$
(25b)

$$x_{7+} = x_7 + h(f_{rr} - f_{rr} - mq)/m$$
(25c)

$$y_{7+} = x_7 + h(f_{pn} - f_{cn} - mg)/m,$$
 (25c)

$$x_{8+} = x_8 + hr(f_{pt} + f_{ct})/f,$$
 (25d)

$$x_{9+} = x_9 + nu_1, \tag{25e}$$

$$x_{10+} = x_{10} + hu_2, \tag{25f}$$

where r = 0.1, m = 1, $j = \frac{2}{5}mr^2$ are the radius, the mass, and the moment of inertia of the ball, g = 10 is the gravity acceleration, h = 0.05 is the discretization step, d = 0.4 and l = 0.6 are geometric parameters (see Figure 3). f represents a contact force, with the subscripts p and c denoting the paddle and the ceiling and t and n denoting the tangential and the normal components.

HSCC '19, April 16-18, 2019, Montreal, QC, Canada



Figure 4: Solution of the open-loop optimal control problem (4) with initial condition $\bar{x}_3 = \pi$ and ∞ -norm objective.

Contact phenomena are modeled with the inelastic time-stepping scheme from [36], where the complementarity conditions are solved ahead of time to explicitly express the contact forces as PWA functions of the state and the inputs. Friction coefficients are set to $\mu = 0.2$. Overall we get a PWA system with 7 modes: no contact, ball in contact with the paddle (sticking, sliding left or right), and ball in contact with the ceiling (sticking, sliding left or right). Together with the guards defined by the contact constraints, we enforce the limits $\pm x \leq (0.3, 0.2, 1.2\pi, 0.3, 0.15, 2, 2, 10, 2, 2)$ and $\pm u \leq (30, 30)$. Moreover, since the paddle and the ceiling have limited width, in case at the next time step the ball is going to hit, e.g., the paddle, we require $|x_{1+} - x_{4+}| \le l/2$. Independently from the norm used in the cost function, we set Q = diag(1, 1, 0.01, 1, 1, 1, 1, 0.01, 1, 1)and R = diag(0.01, 0.001). The terminal set X_T forces the system to be in the origin after T = 20 time steps. We apply the basic step to the terminal set (19) but we do not condense the PWA dynamics as in (17), since in this case, even for a time window of two steps, all the mode transitions are feasible.

Despite the simplicity of the system, the control problem described above is quite hard to solve. Analyzing the dynamics (25) it can be noticed that the presence of the ceiling is essential to accomplish the rotation task. Indeed, assuming $f_{\rm ct}$ to be zero for the whole motion, the horizontal and the angular dynamics of the ball would not be independently controllable, even assuming to have direct control on $f_{\rm pt}$ (see (25b) and (25d)); hence we could not both rotate the ball and place it back at the center of the paddle. This severe lack of controllability makes the detection of an initial feasible trajectory a complex operation, preventing many branches from being cut in the early stages of the B&B algorithm. In this sense, the strength of the MI formulation we employ plays here a fundamental role in guiding the growth of the B&B tree, and preventing the algorithm from getting stuck in the exploration of countless dead-end branches.

7.2.2 Numeric Results. Figure 4 depicts the optimal trajectory obtained solving a single open-loop control problem in case of, e.g., ∞ -norm objective (the trajectories obtained with 1-norm or quadratic cost are almost indistinguishable). The optimal mode sequence $(i_t^*)_{t=0}^{T-1}$ is represented in Figure 5.

To compare the strength of the MI formulations we propose, in Figure 6 we report the objective values obtained with the different formulations for a decreasing degree of relaxation: for t = 0, ..., T-1 we solve the convex subproblem obtained by fixing the mode of system (25) to its optimal value (obtained solving the problem above) for the first *t* time steps, then the resulting cost is normalized with the optimal objective of the MICP. (Progressively fix the values of

Tobia Marcucci and Russ Tedrake



Figure 5: Optimal mode sequence for the open-loop problem (4) with initial condition $\bar{x}_3 = \pi$ and ∞ -norm objective.

the binaries in chronological order is a common branching heuristic in B&B solvers for hybrid MPC [7, 16].) As a benchmark we also present the results obtained with the MLD formulation from [8] and, to make the comparison fair, we set all the big-M parameters in it to their tightest value. The hierarchies from Section 6 are confirmed: our big-M method leads to a stronger formulation than the MLD, the convex hull is stronger than the previous ones, and with the perspective formulation we obtain the highest lower bounds.

In Table 2 we report the optimal objectives obtained with the different formulations for multiple initial conditions. We vary the horizontal position \bar{x}_1 and angle \bar{x}_3 of the ball in the interval $[0, l/2] \times [-\pi, \pi]$ and we fix all the other states to zero, we then relax all the binary constraints in the problem, and we solve the resulting convex program (which is hence the same problem we would have at the root of the B&B tree). All the proposed methods result in better lower bounds than the MLD formulation, and also in this case the numeric results confirm the hierarchies from Section 6.

In Table 3 we consider a single open-loop instance of problem (4) and we report the performances of the MI formulations in case of different cost functions.⁷ In this case, the best compromise between size and strength is the convex-hull formulation, with which we are able to solve globally all the three problems within the time limit of one hour. The perspective formulation, in case of linear objective, reduces the number of subproblems solved in the B&B algorithm but, because of its larger size, overall leads to an increase in solution times. On the other hand, in case of quadratic objective, the SOC constraints make the problem much harder and we are barely able to find a feasible solution within the time limit. Despite of its reduced size, in none of the cases the big-M formulation converges to the global optimum before the time limit. Finally, the MLD formulation presents the worst performances and in none of the cases is able to find a feasible solution within the time limit.

In conclusion, we point out that computation times in Table 3 are still far from practical control rates, making the synthesis of a pure MPC controller for a system of this size still unrealistic. Nonetheless, we underline how several sampling-based control techniques (e.g., approximate explicit hybrid MPC [25, 34]) could greatly benefit from the higher efficiency of the proposed formulations.

7.2.3 Source Code. The source code for the results presented in this subsection is available at https://github.com/TobiaMarcucci/pympc/tree/hscc19.

⁷Problems are solved with the commercial solver Gurobi 8.0 with default options on a 3.50 GHz Intel Core i7-5930K with 32 GB of RAM.



Figure 6: Ratio between the objective of the convex relaxations and the objective of the MICP in case of different cost functions and MI formulations. The binary variables in the MICP are fixed to their optimal value for the first $t \in \{0, ..., T-1\}$ time steps and relaxed for the remaining steps in order to get a convex subproblem. Formulation acronyms: perspective formulation (PF), convex hull (CH), big-M (BM), MLD equivalent (MLD).

Table 2: First row: samples of the optimal-value function of the MICP in case of different objective functions for multiple initial conditions (horizontal position of the ball $\bar{x}_1 \in [0, l/2]$ on the x axis, angle of the ball $\bar{x}_3 \in [-\pi, \pi]$ on the y axis). Remaining rows: lower bound obtained at the root-node relaxation of the B&B tree (note the different scales of the level plots).



8 CONCLUSIONS

In this paper we have presented a comprehensive analysis of the problem of formulating a MICP for the optimal control of PWA systems. We have shown how the structure of these dynamical systems allows us to take advantage of several disjunctive-programming modeling techniques to derive highly-efficient formulations of the control problem under analysis. Furthermore, this observation also allowed us to borrow some well-known metrics in mixed-integer programming to rigorously define a hierarchy of the strength of the methods we analyzed. We presented multiple formulations of the problem and we tested them on a challenging numerical example, showing that all the proposed methods outperform the standard approach employed in MPC, enabling the solution of large-scale problems that could not have been tackled with previous techniques. Table 3: Performances of the MI formulations in the solution of a single instance of problem (4). "Solution gap" denotes the difference between the highest lower bound and the cost of the best feasible solution found by the B&B algorithm (0 if the global optimum has been found, ∞ if no feasible solution has been found). "t.l." stands for time limit (set to 1 h).

		∞-n. obj.	1-n. obj.	Quad. obj.
PF	Solution gap (%)	0	0	98.7
	Solution time (s)	439	464	t.l.
	B&B nodes (k)	56	145	14
СН	Solution gap (%)	0	0	0
	Solution time (s)	318	212	3537
	B&B nodes (k)	95	157	1,999
BM	Solution gap (%)	92.4	91.1	93.4
	Solution time (s)	t.l.	t.l.	t.l.
	B&B nodes (k)	4,039	3,721	8,477
MLD	Solution gap (%)	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	∞	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
	Solution time (s)	t.l.	t.l.	t.l.
	B&B nodes (k)	2,605	2,928	1,718

ACKNOWLEDGMENTS

The authors would like to acknowledge the support from the Grass Instruments Company and the Department of the Navy, Office of Naval Research, Award No. N00014-18-1-2210. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Office of Naval Research.

REFERENCES

- Daniel Axehill and Anders Hansson. 2008. A dual gradient projection quadratic programming algorithm tailored for model predictive control. In Decision and Control, 2008. CDC 2008. 47th IEEE Conference on. IEEE, 3057–3064.
- [2] Daniel Axehill, Lieven Vandenberghe, and Anders Hansson. 2010. Convex relaxations for mixed integer predictive control. Automatica 46, 9 (2010), 1540–1545.
- [3] Egon Balas. 1979. Disjunctive programming. In Annals of Discrete Mathematics. Vol. 5. Elsevier, 3–51.
- [4] Egon Balas. 1985. Disjunctive programming and a hierarchy of relaxations for discrete optimization problems. SIAM Journal on Algebraic Discrete Methods 6, 3 (1985), 466–486.
- [5] Egon Balas. 1998. Disjunctive programming: Properties of the convex hull of feasible points. Discrete Applied Mathematics 89, 1-3 (1998), 3–44.
- [6] Alberto Bemporad, Giancarlo Ferrari-Trecate, and Manfred Morari. 2000. Observability and controllability of piecewise affine and hybrid systems. *IEEE transactions on automatic control* 45, 10 (2000), 1864–1876.
- [7] Alberto Bemporad, Domenico Mignone, and Manfred Morari. 1999. An efficient branch and bound algorithm for state estimation and control of hybrid systems. In *Proceedings of the European Control Conference*. Citeseer, 557–562.
- [8] Alberto Bemporad and Manfred Morari. 1999. Control of systems integrating logic, dynamics, and constraints. *Automatica* 35, 3 (1999), 407–427.
- [9] Ahron Ben-Tal and Arkadi Nemirovski. 2001. Lectures on modern convex optimization: analysis, algorithms, and engineering applications. Siam.
- [10] Dimitri P Bertsekas, Angelia Nedi, Asuman E Ozdaglar, et al. 2003. Convex analysis and optimization. (2003).
- [11] Francesco Borrelli, Mato Baotić, Alberto Bemporad, and Manfred Morari. 2005. Dynamic programming for constrained optimal control of discrete-time linear hybrid systems. *Automatica* 41, 10 (2005), 1709–1721.
- [12] Francesco Borrelli, Alberto Bemporad, Michael Fodor, and Davor Hrovat. 2006. An MPC/hybrid system approach to traction control. *IEEE Transactions on Control Systems Technology* 14, 3 (2006), 541–552.
- [13] Francesco Borrelli, Alberto Bemporad, and Manfred Morari. 2017. Predictive control for linear and hybrid systems. Cambridge University Press.
- [14] Eduardo F Camacho, Daniel R Ramírez, Daniel Limón, D Muñoz De La Peña, and Teodoro Alamo. 2010. Model predictive control techniques for hybrid systems.

Annual reviews in control 34, 1 (2010), 21-31.

- [15] Sebastián Ceria and João Soares. 1999. Convex programming for disjunctive convex optimization. *Mathematical Programming* 86, 3 (1999), 595–614.
- [16] Damian Frick, Alexander Domahidi, and Manfred Morari. 2015. Embedded optimization for mixed logical dynamical systems. *Computers & Chemical Engineering* 72 (2015), 21–33.
- [17] Damian Frick, Angelos Georghiou, Juan L Jerez, Alexander Domahidi, and Manfred Morari. 2016. Low-complexity method for hybrid MPC with local guarantees. arXiv preprint arXiv:1609.02819 (2016).
- [18] Tobias Geyer, Georgios Papafotiou, and Manfred Morari. 2008. Hybrid model predictive control of the step-down DC-DC converter. IEEE Transactions on Control Systems Technology 16, 6 (2008), 1112–1124.
- [19] Oktay Günlük and Jeff Linderoth. 2012. Perspective reformulation and applications. In Mixed Integer Nonlinear Programming. Springer, 61–89.
- [20] Weiqiao Han and Russ Tedrake. 2017. Feedback design for multi-contact push recovery via LMI approximation of the piecewise-affine quadratic regulator. In Humanoid Robotics (Humanoids), 2017 IEEE-RAS 17th International Conference on. IEEE, 842–849.
- [21] Wilhemus PMH Heemels, Bart De Schutter, and Alberto Bemporad. 2001. Equivalence of hybrid dynamical models. Automatica 37, 7 (2001), 1085–1091.
- [22] Andreas B Hempel, Paul J Goulart, and John Lygeros. 2015. Inverse parametric optimization with an application to hybrid system control. *IEEE Transactions on automatic control* 60, 4 (2015), 1064–1069.
- [23] Mikael Johansson and Anders Rantzer. 1997. Computation of piecewise quadratic Lyapunov functions for hybrid systems. In *Control Conference (ECC), 1997 European.* IEEE, 2005–2010.
- [24] Mircea Lazar, WPMH Heemels, Siep Weiland, and Alberto Bemporad. 2006. Stabilizing model predictive control of hybrid systems. *IEEE Trans. Automat. Control* 51, 11 (2006), 1813–1818.
- [25] Tobia Marcucci, Robin Deits, Marco Gabiccini, Antonio Bicchi, and Russ Tedrake. 2017. Approximate hybrid model predictive control for multi-contact push recovery in complex environments. In *Humanoid Robotics (Humanoids), 2017* IEEE-RAS 17th International Conference on. IEEE, 31–38.
- [26] Nicholas Moehle and Stephen Boyd. 2015. A perspective-based convex relaxation for switched-affine optimal control. Systems & Control Letters 86 (2015), 34-40.
- [27] Igor Mordatch, Emanuel Todorov, and Zoran Popović. 2012. Discovery of complex behaviors through contact-invariant optimization. ACM Transactions on Graphics (TOG) 31, 4 (2012), 43.
- [28] George L. Nemhauser and Laurence A. Wolsey. 1988. Integer and Combinatorial Optimization. Wiley-Interscience, New York, NY, USA.
- [29] Jan Oldenburg and Wolfgang Marquardt. 2008. Disjunctive modeling for optimal control of hybrid systems. *Computers & chemical engineering* 32, 10 (2008), 2346-2364.
- [30] Pablo A Parrilo and Sanjay Lall. 2003. Semidefinite programming relaxations and algebraic optimization in control. *European Journal of Control* 9, 2-3 (2003), 307–321.
- [31] Michael Posa, Cecilia Cantu, and Russ Tedrake. 2014. A direct method for trajectory optimization of rigid bodies through contact. *The International Journal* of Robotics Research 33, 1 (2014), 69–81.
- [32] Anders Rantzer and Mikael Johansson. 1997. Piecewise linear quadratic optimal control. In American Control Conference, 1997. Proceedings of the 1997, Vol. 3. IEEE, 1749–1753.
- [33] Juan P Ruiz and Ignacio E Grossmann. 2012. A hierarchy of relaxations for nonlinear convex generalized disjunctive programming. *European Journal of Operational Research* 218, 1 (2012), 38–47.
- [34] Sadra Sadraddini and Russ Tedrake. 2018. Sampling-based Polytopic Trees for Approximate Optimal Control of Piecewise Affine Systems. arXiv preprint arXiv:1809.09716 (2018).
- [35] Eduardo Sontag. 1981. Nonlinear regulation: The piecewise linear approach. IEEE Transactions on automatic control 26, 2 (1981), 346–358.
- [36] David Stewart and Jeffrey C Trinkle. 2000. An implicit time-stepping scheme for rigid body dynamics with coulomb friction. In *Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference on*, Vol. 1. IEEE, 162–169.
- [37] Olaf Stursberg and Sebastian Panek. 2002. Control of switched hybrid systems based on disjunctive formulations. In International Workshop on Hybrid Systems: Computation and Control. Springer, 421–435.
- [38] Reza Takapoui, Nicholas Moehle, Stephen Boyd, and Alberto Bemporad. 2017. A simple effective heuristic for embedded mixed-integer quadratic programming. *Internat. J. Control* (2017), 1–11.
- [39] Francisco Trespalacios and Ignacio E Grossmann. 2015. Improved Big-M reformulation for generalized disjunctive programs. *Computers & Chemical Engineering* 76 (2015), 98–103.
- [40] Juan Pablo Vielma. 2015. Mixed integer linear programming formulation techniques. SIAM Rev. 57, 1 (2015), 3–57.
- [41] Boyan Yordanov, Jana Tumova, Ivana Cerna, Jiří Barnat, and Calin Belta. 2012. Temporal logic control of discrete-time piecewise affine systems. *IEEE Trans. Automat. Control* 57, 6 (2012), 1491–1504.