

Inverse Parametric Optimization with an Application to Hybrid System Control

Andreas B. Hempel, Paul J. Goulart, and John Lygeros

Abstract

We present a number of results on inverse parametric optimization and its application to hybrid system control. We show that any function that can be written as the difference of two convex functions can also be written as a linear mapping of the solution to a convex parametric optimization problem. We exploit these results in application to the control of systems with piecewise affine dynamics, and show that it is possible to model such systems as optimizing processes. Optimal control problems for such systems can be remodeled as bilevel optimization problems and solved with existing techniques.

I. INTRODUCTION

A parametric optimization problem is one in which the problem data, and consequently the solution, varies with respect to some known parameter, and one wishes to identify a function mapping parameter values to optimal solutions. In general parametric programming problems, one wishes to identify some $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ such that

$$f(p) \in \arg \min_z J(z, p) \tag{1a}$$

$$\text{s.t. } (z, p) \in \Gamma, \tag{1b}$$

where $z \in \mathbb{R}^n$ is the vector of decision variables, $p \in \mathbb{R}^m$ the parameter vector, $J: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ the cost function, and $\Gamma \subseteq \mathbb{R}^n \times \mathbb{R}^m$ the constraint set. The function $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$, which one must identify, maps each admissible parameter to an optimal solution for that parameter, and is called a *solution function*. Methods for constructing a solution function f for (1) have been developed for several common problem classes. A recent overview of algorithms can be found in [1], with earlier results reported in [2].

In control applications, parametric problems in the form (1) are encountered in the optimization problems underlying model predictive control (MPC) of dynamical systems [3], where the parameter p represents the initial state of the system and $p \rightarrow f(p)$ maps the initial state to an associated optimal control sequence. For MPC of linear systems, the problem (1) is typically a parametric linear or quadratic programming (PLP / PQP) problem, for which a solution function f can be computed

The research leading to these results has received funding from the Swiss Secretariat for Education and Research under COST-Action IC0806 and the European Union Seventh Framework Programme under grant agreement number FP7-ICT-257462 (HYCON2).

The authors are with the Automatic Control Laboratory, ETH Zürich, Physikstrasse 3, 8092 Zürich, Switzerland; {hempel, pgoulart, lygeros}@control.ee.ethz.ch.

explicitly in the form of a piecewise affine (PWA) function defined over a polytopic partition of the set of feasible initial states [4], [5]. For MPC of hybrid systems, the underlying optimization problem is typically a parametric mixed-integer linear or quadratic program (MILP / MIQP) [6].

Inverse optimization approaches problem (1) from the opposite direction. Given a solution function f , the task is to find a constraint set $\Gamma \subseteq \mathbb{R}^n \times \mathbb{R}^m$ and a cost function $J: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ such that (1) is satisfied. Inverse optimization problems have appeared in a wide variety of research fields, for example the authors of [7] propose an inverse reinforcement learning algorithm to identify optimization problem data that an observed expert is (unconsciously) using to make (presumably) optimal decisions. In control theory, Kalman posed necessary and sufficient conditions for a given state feedback control law to be optimal for a linear plant in [8]. More recently in [9], the authors showed that every continuous nonlinear control law f can be characterized as the solution to some parametric convex program (PCP) assuming that the parametric value function is also known.

A variety of inverse optimization methods have been proposed, most of which assume at least partial knowledge of the constraint set Γ , reducing the inverse optimization problem to finding a suitable cost function J . The authors of [10] use a finite number of parameter-optimizer-pairs $(f(p_i), p_i)$ and construct a convex optimization problem whose solution provides a suitable cost function J . This approach is applicable to a wide variety of problems with complex cost function parameterizations, but provides only an approximate solution. The authors of [11] consider the special case of inverse linear programming where the left-hand side of the constraints is known. The problem of finding an appropriate cost function vector and right-hand side of the constraints is then formulated as a bilevel optimization problem. If the constraint set is *completely* known, then it has been shown that the inverse linear programming problem itself can be written as a linear program [12].

In this paper, we describe a method for solving the related extended inverse optimization problem. Specifically, for a given function $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ we want to find a PCP (1), a function $\hat{f}: \mathbb{R}^m \rightarrow \mathbb{R}^{\hat{n}}$, and a fixed linear operator $T \in \mathbb{R}^{n \times \hat{n}}$ such that \hat{f} solves the PCP (1) and $f(p) = T\hat{f}(p)$ for any admissible p . We provide a solution to this problem for any f that can be written as the difference of two convex functions, i.e. is a so-called *DC function* [13]. Further, we do so without assuming any a-priori knowledge of the constraint set Γ or the cost function J . In particular, we investigate PWA functions f , which have been shown to be DC [14], [15].

This case is of special interest because our results can be applied to the control of hybrid systems with PWA dynamics by finding a generating problem (1) that has the state transition dynamics as its optimal solution. The resulting *inverse optimization model* of the hybrid system is a special and comparatively compact form of a linear complementarity (LC) model [16]. An optimal control problem using this model is a so-called *mathematical program with equilibrium constraints* (MPEC), which can be solved by appropriate numerical methods [17]. Using an inverse optimization model (instead of a general LC model) leads to optimal control problems that are *bilevel optimization problems* which can be considered as a special case of MPECs [18]. Preliminary versions of some results in this paper were first reported in [15].

Notation: We use superscripts in square brackets to indicate elements of vectors and vector valued functions and matrices, e.g. $f^{[j]}$ is the j^{th} element of f and $F^{[j, \cdot]}$ is the j^{th} row of F . A bold 1 denotes a vector of ones and I an identity matrix of appropriate dimension. The interior and convex

hull of a set $\Omega \subseteq \mathbb{R}^n$ are denoted $\text{int}(\Omega)$ and $\text{conv}(\Omega)$, respectively. The cardinality of a set $\mathcal{I} \subseteq \mathbb{N}$ by $|\mathcal{I}|$ and the epigraph of a function f by $\text{epi}(f)$. All vector-valued inequalities are to be understood component-wise.

II. EXTENDED INVERSE PARAMETRIC NONLINEAR OPTIMIZATION

We consider problems in the form (1) under the additional assumption that Γ is a convex constraint set and $J(\cdot, p)$ is a convex function for any fixed p . Such a problem is referred to as a parametric convex program (PCP). We are interested in its inverse, i.e. given a solution function f we wish to construct convex Γ and J such that f satisfies (1). It is known that such a PCP always exists for any continuous function f equipped with a (possibly artificially chosen) value function [9]. However, such a problem may not be easy to construct or may not be in any conveniently characterized class of convex problems.

We therefore consider a related *extended inverse optimization problem*, in which we construct problem data for a PCP such that the desired function f can be obtained as the image of the solution to the PCP under a linear operator. Using this more general approach we will show that it is possible to represent a large class of functions as solutions to PCPs in a convenient form.

Problem 1 (Extended inverse optimization problem). Given a function $f: \Omega \rightarrow \mathbb{R}^n$ over a convex domain $\Omega \subseteq \mathbb{R}^m$, find some $\hat{n} \in \mathbb{N}$, a cost function $J: \mathbb{R}^{\hat{n}} \times \Omega \rightarrow \mathbb{R}$, a convex constraint set $\Gamma \subseteq \mathbb{R}^{\hat{n}} \times \mathbb{R}^m$, and a fixed matrix $T \in \mathbb{R}^{n \times \hat{n}}$ such that $J(\cdot, p)$ is convex for any fixed p and

$$f(p) = T\hat{f}(p) \quad \forall p \in \Omega, \quad (2)$$

where \hat{f} is a solution to the PCP

$$\hat{f}(p) \in \arg \min_z J(z, p) \quad \text{s.t.} \quad (z, p) \in \Gamma. \quad (3)$$

The *desired solution function* f is the image of the *lifted solution function* \hat{f} to the *generating problem* (3) under the linear operator T . If a solution to Problem 1 can be found with $\hat{n} = n$ and T invertible, then it is also a solution to the original inverse optimization problem subject to a change of variables $y = Tz$.

We provide a solution to Problem 1 for *DC functions*, i.e. functions on a convex domain that can be written as the difference of two convex functions [13]. DC functions and the related field of DC programming allows applications for example in resource allocation and network problems [19] and control [20]. Important examples of DC functions are functions that are twice differentiable on bounded sets, continuous piecewise affine (PWA) functions (see also Section III-A), and compositions of DC and convex functions. Additionally, every continuous function can be approximated to arbitrary degree of accuracy by a DC function [13]. We will call a function $f: \Omega \rightarrow \mathbb{R}^n$ with domain $\Omega \subseteq \mathbb{R}^m$ a *vector-valued DC function* if there exist functions $\gamma: \Omega \rightarrow \mathbb{R}^n$ and $\eta: \Omega \rightarrow \mathbb{R}^n$ such that $f(\cdot) = \gamma(\cdot) - \eta(\cdot)$ and each of the component functions $\gamma^{[i]}(\cdot)$ and $\eta^{[i]}(\cdot)$ are convex.

Our first result shows that the extended inverse optimization problem can be solved exactly for this class of functions:

Theorem 1 (Extended inverse optimization for vector-valued DC functions). *For every vector-valued DC function $f: \Omega \rightarrow \mathbb{R}^n$ there exist a convex set $\Gamma \subseteq \mathbb{R}^{\hat{n}} \times \mathbb{R}^m$, a cost function $J: \mathbb{R}^{\hat{n}} \times \Omega \rightarrow \mathbb{R}$ such that $J(\cdot, p)$ is a convex function for any fixed p , and a function $\hat{f}: \Omega \rightarrow \mathbb{R}^{\hat{n}}$ such that*

$$f(p) = T\hat{f}(p) \quad \forall p \in \Omega,$$

where $T \in \mathbb{R}^{n \times \hat{n}}$ is constant and \hat{f} the unique solution to the PCP

$$\hat{f}(p) \in \arg \min_z J(z, p) \quad \text{s.t.} \quad (z, p) \in \Gamma, \quad (4)$$

with $\hat{n} \leq 2n$.

Proof. Consider the case where $f: \Omega \rightarrow \mathbb{R}$ with $\Omega \subseteq \mathbb{R}^m$ is a scalar-valued DC function, i.e. $f(\cdot) = \gamma(\cdot) - \eta(\cdot)$ for some convex γ, η . Define the function \hat{f} as

$$\hat{f}(p) = \begin{bmatrix} \gamma(p) \\ \eta(p) \end{bmatrix}.$$

Choose $\Gamma := \{(z_1, z_2, p) \in \mathbb{R}^2 \times \mathbb{R}^m \mid p \in \Omega, \gamma(p) \leq z_1, \eta(p) \leq z_2\}$. Since γ and η are convex functions, Γ is a convex set.

Choose $J(z, p) = c_1 z_1 + c_2 z_2$ with any $c_1, c_2 > 0$. The PCP (4) is then separable in z_1 and z_2 and in epigraph form, hence for each admissible p the unique solution is given by $\hat{f}(p)$. The desired solution function f is recovered by applying the difference operator $T = [1 \ -1]$ to \hat{f} .

The extension to a vector-valued DC function $f: \Omega \rightarrow \mathbb{R}^n$ is immediate by applying this procedure to every component function $f^{[i]}: \Omega \rightarrow \mathbb{R}$ and combining the constructed $2n$ generating problems into a separable PCP (4). \square

Note that in case k component functions $f^{[i]}$ are convex or concave the bound on \hat{n} can be tightened to $2n - k$ as it is not necessary to introduce additional decision variables for them. The class of (4) depends on the properties of γ and η , e.g. when both are quadratic functions the generating problem is quadratically constrained.

Instead of the linear cost function employed in the proof of Theorem 1, it is also possible to construct a quadratic J if the decomposition functions $\gamma^{[i]}$ and $\eta^{[i]}$ are lower semi-continuous (all lower-level sets $\{p \in \Omega \mid \gamma^{[i]}(p) \leq c\}$ of $\gamma^{[i]}$ are closed) and proper ($\gamma^{[i]}(p) > -\infty$ for all $p \in \Omega$ and $\gamma^{[i]}(p) < \infty$ for at least one $p \in \Omega$) in addition to being convex. The construction of a quadratic cost function J makes use of the following result:

Lemma 1 (Inverse optimization with quadratic cost function for scalar-valued convex functions). *Suppose that $f: \Omega \rightarrow \mathbb{R}$ is convex, proper, and lower semi-continuous. Then for any $q > 0$ there exist a convex set $\Gamma \subseteq \mathbb{R} \times \mathbb{R}^m$ and a quadratic cost function $J: \Gamma \rightarrow \mathbb{R}$ of the form*

$$J(z, p) := \frac{1}{2}qz^2 + (r^T p)z + sz \quad (5)$$

such that f uniquely satisfies (1).

Proof. Choose the constraint set Γ as the epigraph of f , i.e. $\Gamma := \text{epi}(f) = \{(z, p) \in \mathbb{R} \times \mathbb{R}^m \mid p \in \Omega, f(p) \leq z\}$. Because f is convex, Γ is a convex set.

Because f is convex, proper, and lower semi-continuous it is the pointwise supremum of its affine support functions [21, Thm. 8.13]. Choose $q > 0$, $r \in \mathbb{R}^m$, and $s \in \mathbb{R}$ such that the function $a: \Omega \rightarrow \mathbb{R}$ with expression

$$a(p) = -\frac{1}{q} (r^T p + s) \quad (6)$$

is any one of the affine supports of f .

Note that the unconstrained minimizer for a quadratic cost function (5) with the same parameters q , r , and s is given as

$$z_u^*(p) := \arg \min_z J(z, p) = -\frac{1}{q} (r^T p + s) = a(p)$$

and corresponds exactly to the chosen affine support of f . It remains only to show that with this choice of J the unique solution to (1) is given by f .

Since z_u^* is a support function for f , it follows that $z_u^*(p) \leq f(p)$. For any fixed p , let $(\bar{z}, p) \in \Gamma$ with arbitrary \bar{z} , i.e. $\bar{z} \geq f(p)$. Then there exists some $\lambda \in [0, 1]$ such that $f(p) = \lambda z_u^*(p) + (1 - \lambda)\bar{z}$. From the convexity of $J(z, p)$ for every fixed p , it is straightforward to show that

$$J(f(p), p) = J(\lambda z_u^*(p) + (1 - \lambda)\bar{z}, p) \leq J(\bar{z}, p).$$

Because this holds for any \bar{z} , the constrained optimizer for (1) with J as in (5) is given by $f(p)$. Because $J(\cdot, p)$ is strictly convex for any fixed p , its minimizer is unique. \square

Note that Lemma 1 can also be applied to concave functions f as long as $-f$ is proper and f is upper semi-continuous. The extension to vector-valued DC functions f is immediate by applying Lemma 1 separately to the decompositions $\gamma^{[i]}$ and $\eta^{[i]}$ of the component functions $f^{[i]}$.

The solution constructed in the proof of Theorem 1 is not unique, e.g. one can choose different cost function coefficients for J in (4). An appropriate choice for the parameters will depend on the application, cf. Section IV. However, observe that the solutions to the generating problem we construct in the proof *are* unique and always exactly equal to the desired solution function f .

III. EXTENDED INVERSE QUADRATIC OPTIMIZATION FOR PWA FUNCTIONS

In the remainder of the paper we focus on the specific case of continuous PWA functions. These functions are of particular interest since the solutions to PLP and PQP problems take this form and they are often used to model complex dynamical systems [22].

We will consider PQPs in the form

$$\arg \min_z \frac{1}{2} z^T Q z + p^T R z + s^T z \quad (7a)$$

$$\text{s.t. } Fz + Gp \leq h, \quad p \in \Omega. \quad (7b)$$

Here, $Q = Q^T \succeq 0$ and the set of admissible parameters Ω is a convex polytope. This subsumes PLPs since Q is allowed to be zero.

It is well-known that solutions to strictly convex PQP problems are continuous PWA functions of the parameter [5]. Among all solutions to a given PLP, one can always choose a continuous PWA solution function [23]. In either case, the solution function $f: \Omega \rightarrow \mathbb{R}^n$ is defined over a convex polytope $\Omega \subseteq \mathbb{R}^m$ and can be written as

$$f(p) = f_i(p) := L_i p + l_i \quad \forall p \in \Omega_i, \quad (8)$$

where we refer to $f_i: \Omega_i \rightarrow \mathbb{R}^n$ as the *local function* on the convex polytopical region $\Omega_i \subseteq \mathbb{R}^m$. The regions Ω_i each have non-empty interior and form a partition of the convex polyhedral domain $\Omega \subseteq \mathbb{R}^m$ in the sense that $\text{int}(\Omega_i) \cap \text{int}(\Omega_j) = \emptyset$ for all $i \neq j$ and $\bigcup_{i=1}^{n_r} \Omega_i = \Omega$, where n_r is the total number of regions.

Before considering the inverse problem to (7), we will present some necessary results on PWA functions.

A. PWA functions are DC

It has been established in [14] that scalar-valued continuous PWA functions are DC, i.e. they can always be written as the difference of two continuous and convex PWA functions. A simplified constructive proof of this result was presented in [15]:

Lemma 2 ([14], [15]; continuous PWA functions are DC). *Every continuous PWA function $f: \Omega \rightarrow \mathbb{R}$ defined over a convex polyhedral partition of $\Omega \subseteq \mathbb{R}^m$ can be written as the difference of two continuous convex PWA functions γ and η , i.e. $f(\cdot) = \gamma(\cdot) - \eta(\cdot)$.*

The functions γ and η constructed in the proof of Lemma 2 in [15] are generally defined over a finer partition of the domain Ω than f , i.e. they involve more pieces. Furthermore, the number of regions for γ and η is unknown a priori and only emerges from the construction. In many applications, this may not be desirable.

Here we present a much stronger result than Lemma 2 for an important class of PWA functions $f: \Omega \rightarrow \mathbb{R}$. We start by introducing some notation from [24]. We define the point set P as the collection of n_v vertices p_j of all regions Ω_i , and define the index sets $\mathcal{I}_i \subseteq \{1, \dots, n_v\}$ such that all p_j with $j \in \mathcal{I}_i$ are extreme points of Ω_i . The given partition of Ω is then a polyhedral subdivision of P .

Polyhedral subdivisions of arbitrary point sets P can be generated by assigning a height $\omega^{[j]} \in \mathbb{R}$ to every vertex p_j to obtain the lifted point configuration P_ω with $P_\omega^{[j]} := (p_j, \omega^{[j]}) \in \mathbb{R}^{m+1}$ for all j . A polyhedral subdivision of P is then generated by projecting the lower faces of $\text{conv}(P_\omega)$ onto the first m coordinates. A *lower face* is one that has a ‘non-vertical’ supporting hyperplane \mathcal{H} with P_ω above \mathcal{H} .

A partition generated in this way is called a *regular subdivision*. An important and well-known example of a regular subdivision is the *Delaunay triangulation*, which can be obtained by choosing

the heights $\omega^{[j]} = \|p_j\|_2^2$. Regular subdivisions are of particular interest when considering PWA models, cf. [25], [26].

In order to prove our next result, we introduce the *folding form* $\Psi_{B_i,j}(\omega)$, cf. [24, Def. 5.2.4]. Let $B_i := [p_{r_1} \ \dots \ p_{r_{n+1}}]$ be a collection of affinely independent vertices of the full-dimensional polytope Ω_i . Then

$$\Psi_{B_i,j}(\omega) := \text{sign det} \begin{pmatrix} p_{r_1} & \dots & p_{r_{d+1}} \\ 1 & \dots & 1 \end{pmatrix} \cdot \det \begin{pmatrix} p_{r_1} & \dots & p_{r_{d+1}} & p_j \\ \omega^{[r_1]} & \dots & \omega^{[r_{d+1}]} & \omega^{[j]} \end{pmatrix}$$

indicates the position of $(p_j, \omega^{[j]})$ relative to the hyperplane containing the lifted Ω_i , i.e. $\Psi_{B_i,j}(\omega) > 0$ if $(p_j, \omega^{[j]})$ is above, $\Psi_{B_i,j}(\omega) < 0$ if $(p_j, \omega^{[j]})$ is below, and $\Psi_{B_i,j}(\omega) = 0$ if $(p_j, \omega^{[j]})$ is contained in the hyperplane spanned by the vertices of Ω_i lifted to the heights given by ω . We use this to prove the following result:

Lemma 3 (Convex decomposition of PWA functions over regular subdivisions). *Every continuous PWA function $f: \Omega \rightarrow \mathbb{R}$ defined over a regular subdivision of a bounded polytopic domain $\Omega \subseteq \mathbb{R}^m$ can be written as the difference of two continuous convex PWA functions γ and η , i.e. $f(\cdot) = \gamma(\cdot) - \eta(\cdot)$, where both γ and η are defined over the same regular subdivision.*

Proof. Since f is defined over a regular subdivision, there exists a height vector ω that generates this partition and such that

$$\Psi_{B_i,j}(\omega) \begin{cases} = 0 & \text{if } p_j \in \Omega_i, \\ > 0 & \text{otherwise,} \end{cases}$$

for all $i \in \{1, \dots, n_r\}$ and $j \in \{1, \dots, n_v\}$ [24, Prop. 5.2.6]. Let $\hat{\omega}^{[j]} := \omega^{[j]} + \epsilon \cdot f(p_j)$. Using continuity of the determinant and elementary matrix operations, it is easy to show that one can choose $\epsilon > 0$ small enough such that

$$\Psi_{B_i,j}(\hat{\omega}) = 0 \quad \Leftrightarrow \quad \Psi_{B_i,j}(\omega) = 0 \quad \text{and} \quad \Psi_{B_i,j}(\hat{\omega}) > 0 \quad \Leftrightarrow \quad \Psi_{B_i,j}(\omega) > 0,$$

hence, $\hat{\omega}$ generates the same regular subdivision as ω .

Define $\gamma: \Omega \rightarrow \mathbb{R}$ as the largest convex function whose epigraph contains all point pairs $(p_j, \frac{1}{\epsilon} \hat{\omega}^{[j]})$. Since $\text{epi}(\gamma)$ is polyhedral, γ is PWA. Because $\hat{\omega}$ generates the same partition as ω , γ is defined over the same partition as f .

In a similar manner, define $\eta: \Omega \rightarrow \mathbb{R}$ as the largest convex function whose epigraph contains all the unperturbed point pairs $(p_j, \frac{1}{\epsilon} \omega^{[j]})$, so that η also is a convex PWA function on the regular subdivision generated by ω . It is straightforward to verify that $f(\cdot) = \gamma(\cdot) - \eta(\cdot)$. \square

Verifying whether a given polyhedral subdivision is regular requires the solution of a linear program [24, Lem. 8.2.3]. Because Lemma 3 applies to regular partitions, it covers a number of important cases such as the Delaunay triangulation, Voronoi diagrams, and power diagrams [27].

B. Computing a DC decomposition for PWA functions

Lemma 3 establishes the existence of ‘simple’ convex-concave decompositions for many continuous PWA functions f . Before proceeding to the proof of Theorem 2, we show how to find such a decomposition via solution of a set of linear inequalities.

If $f : \Omega \rightarrow \mathbb{R}$ satisfies the conditions of Lemma 3, it can be written as the difference of $\gamma : \mathbb{R}^m \rightarrow \mathbb{R}$ and $\eta : \mathbb{R}^m \rightarrow \mathbb{R}$ with

$$\begin{aligned}\gamma(p) &:= L_i^\gamma p + l_i^\gamma \quad \forall p \in \Omega_i, \\ \eta(p) &:= L_i^\eta p + l_i^\eta \quad \forall p \in \Omega_i,\end{aligned}$$

where the collection of sets Ω_i constitute the same regular partition of Ω over which f is defined.

The condition that $f(\cdot) = \gamma(\cdot) - \eta(\cdot)$ can be encoded as

$$\forall i: \quad L_i = L_i^\gamma - L_i^\eta \quad \wedge \quad l_i = l_i^\gamma - l_i^\eta. \quad (9)$$

Let the set of tuples (i, j) indicating neighboring regions be

$$\mathcal{I} := \{(i, j) \in \mathbb{N} \times \mathbb{N} \mid \dim(\Omega_i \cap \Omega_j) = n - 1\}.$$

Assuming a halfspace representation $\Omega_i := \{p \in \mathbb{R}^m \mid C_i p \leq d_i\}$ and using Farkas’ Lemma, convexity of γ and η is ensured by the existence of non-negative multipliers $\lambda_{(i,j)}^\gamma$ and $\lambda_{(i,j)}^\eta$ of appropriate dimensions that satisfy

$$\begin{aligned}-d_i^\top \lambda_{(i,j)}^\gamma + l_i^\gamma - l_j^\gamma &\geq 0, \quad C_i^\top \lambda_{(i,j)}^\gamma + (L_i^\gamma - L_j^\gamma)^\top = 0, \\ -d_i^\top \lambda_{(i,j)}^\eta + l_i^\eta - l_j^\eta &\geq 0, \quad C_i^\top \lambda_{(i,j)}^\eta + (L_i^\eta - L_j^\eta)^\top = 0.\end{aligned} \quad (10)$$

Any solution to the system of linear constraints (9) and (10) in the variables L_i^γ , L_i^η , l_i^γ , and l_i^η and the multipliers $\lambda_{(i,j)}^\gamma$ and $\lambda_{(i,j)}^\eta$ provides a convex-concave decomposition of f over the same partition. Regularity of the partition need not be checked beforehand, but by Lemma 3 the system (9) and (10) is always feasible for functions defined over regular partitions.

To obtain appropriately scaled data for γ and η we typically find it convenient to implement (9) and (10) as constraints in a linear program with

$$\min \quad \sum_{i=1}^{n_r} (\|L_i^\gamma\|_1 + |l_i^\gamma|)$$

as its objective. Because of the continuity of f and (9), this has the effect of making all parameters ‘small’.

Although Lemma 3 assumes boundedness of the domain Ω to make use of the properties of the determinant, the set of linear inequalities (9) and (10) can be implemented for PWA functions f over unbounded domains. In case f is not defined over a regular partition, the proof of Lemma 2 provided in [15] provides a constructive alternative method to obtain a convex-concave decomposition. Note that in order to construct such a decomposition of f , regardless of whether Lemma 2 or 3 is used, it is necessary that f is defined over a finite number of partition elements Ω_i , each with non-empty interior.

C. Extended inverse PQP for PWA functions

We will now consider the inverse problem to (7) and present the corresponding result to Theorem 1 for the PQP case:

Theorem 2 (Extended inverse PQP solution). *Every continuous PWA function $f: \Omega \rightarrow \mathbb{R}^n$ can be obtained as the image under a linear operator T of the unique solution $\hat{f}: \Omega \rightarrow \mathbb{R}^{\hat{n}}$ to a PQP in the form (7) with $\hat{n} \leq 2n$. If f is defined over a regular subdivision of Ω with n_r regions, then the number of linear constraints (7b) on z can be bounded by $2nn_r$.*

Note that Theorem 2 is a special case of Theorem 1 where both the solution function f and of the generating optimization problem data Γ and J have a specific structure. The case of a generating linear problem, i.e. $Q = 0$ in (7a), has been considered in [15]. Before proving Theorem 2 for the general PQP case it is prudent to have a closer look at its version for scalar-valued functions f :

Lemma 4 (Extended inverse PQP for continuous scalar-valued PWA functions). *Suppose that $f: \Omega \rightarrow \mathbb{R}$ is continuous and PWA. Then for any $(q_1, q_2) \geq 0$ there exist cost function data (r_i, s_i) and constraint matrices (G_i, g_i) for $i = 1, 2$, and a continuous PWA function $\hat{f}: \Omega \rightarrow \mathbb{R}^2$ such that*

$$f(p) = T\hat{f}(p) \quad \forall p \in \Omega,$$

where $T = [1 \ -1]$ and \hat{f} is the unique solution to the following PQP:

$$\begin{aligned} \hat{f}(p) \in \arg \min_{z_1, z_2} & \quad \frac{1}{2}q_1 z_1^2 + (r_1^\top p) z_1 + s_1 z_1 \\ & \quad + \frac{1}{2}q_2 z_2^2 + (r_2^\top p) z_2 + s_2 z_2 \\ \text{s.t.} & \quad G_1 p + g_1 \leq z_1 \mathbf{1} \\ & \quad G_2 p + g_2 \leq z_2 \mathbf{1} \end{aligned} \quad (11)$$

If, furthermore, f is defined over a regular subdivision of Ω with n_r pieces, the number of constraints in (11) can be bounded by $2n_r$.

Proof. The problem (11) is obviously separable in z_1 and z_2 . By Lemma 2 we can obtain a decomposition $f(p) = \gamma(p) - \eta(p)$ with (γ, η) both continuous, PWA, and convex. Define the function \hat{f} as

$$\hat{f}(p) := \begin{bmatrix} \gamma(p) \\ \eta(p) \end{bmatrix}.$$

Because γ is a convex PWA function it can be written as the pointwise maximum of the finite collection of \hat{n}_r affine local functions

$$\gamma(p) = \max_{i=1 \dots \hat{n}_r} (L_i p + l_i) \quad \forall p \in \Omega.$$

Define $G_1 := [L_1^\top \ \dots \ L_{\hat{n}_r}^\top]^\top$ and $g_1 := [l_1 \ \dots \ l_{\hat{n}_r}]^\top$. Similarly, obtain G_2 and g_2 from the convex PWA function η .

Consider first the problem parameters in (11) indexed by 1, i.e. corresponding to the convex PWA function γ . If $q_1 = 0$, choose r_1 and s_1 such that $r_1^\top p + s_1 > 0$ for all $p \in \Omega$, e.g. $r_1 = 0$ and $s_1 > 0$.

Alternatively, if $q_1 > 0$ and the cost function strictly convex in z_1 , choose an affine support of γ , e.g. one of the local functions γ_i . Let $i \in \{1, \dots, \hat{n}_r\}$ indicate the chosen local function γ_i and define $r_1 := -q_1 L_i^\top$ and $s_1 := -q_1 l_i$.

Note that the unconstrained minimizer for a quadratic cost function $J_1(z, p) = \frac{1}{2}q_1 z^2 + (r_1^\top p)z + s_1$ is given as

$$z_{1,u}^*(p) := \arg \min_z J_1(z, p) = -\frac{1}{q_1} \left(r_1^\top p + s_1 \right) = \gamma_i(p) \quad (12)$$

and corresponds exactly to the chosen affine support of γ .

Because γ is convex, proper, and continuous it is the pointwise supremum of its affine support functions [21, Thm. 8.13]. Since $z_{1,u}^*$ is a support function for γ , it follows that $z_{1,u}^*(p) \leq \gamma(p)$. For any fixed p , let $(\bar{z}_1, p) \in \text{epi}(\gamma)$ with arbitrary \bar{z}_1 , i.e. $\bar{z}_1 \geq \gamma(p)$. Then there exists some $\lambda \in [0, 1]$ such that $\gamma(p) = \lambda z_{1,u}^*(p) + (1 - \lambda)\bar{z}_1$. From the convexity of $J_1(z, p)$ for every fixed p , it is straightforward to show that

$$J_1(\gamma(p), p) = J(\lambda z_{1,u}^*(p) + (1 - \lambda)\bar{z}_1, p) \leq J_1(\bar{z}_1, p). \quad (13)$$

Because this holds for any \bar{z}_1 , the constrained optimizer for $\min_{z \in \text{epi}(\gamma)} J_1(z, p)$ is given by $\gamma(p)$. Because $J_1(\cdot, p)$ is strictly convex for any fixed p , this minimizer is unique.

Analogously, the parameters r_2 and s_2 can be derived from a given q_2 and η . Combining the different parts gives the separable PCP (11) and the desired solution function f follows from applying the difference operator T to \hat{f} .

The number of constraints in (11) corresponds to the sum of the number of regions defining γ and η . In case f is defined over a regular subdivision of Ω , it is possible by Lemma 3 to find γ and η over the same partition, i.e. with the same number of regions, as f . The number of constraints in (11) is then bounded by $2n_r$. \square

While the proof indicates that for general scalar-valued PWA functions Problem 1 can be solved with $\hat{n} = 2$, the comment following the proof of Theorem 1 bears repeating: in case the desired solution function f is convex or concave, a DC decomposition and the introduction of a second decision variable z_2 in (11) is not required.

The proof of Theorem 2 follows from the preceding discussion:

Proof of Theorem 2. Apply Lemma 4 to every component function $f^{[i]}: \Omega \rightarrow \mathbb{R}$ to construct a separable convex PQP. For this separable PQP, \hat{n} will be at most $2n$, which will only occur in the case where all component functions $f^{[i]}$ are neither convex nor concave.

In case f is defined over a regular subdivision, Lemma 3 provides a bound on the number of constraints. \square

Note that, in contrast to [9], our method enables construction of the generating problem without requiring a value function. Furthermore, the quadratic cost functions that are constructed do not require joint convexity in z and p . Our particular choice of cost functions also retain several degrees

of freedom that can be exploited to meet application specific objectives, e.g. scaling the cost function data relative to the constraints.

An alternative lifting construction requiring only one auxiliary variable is proposed in [28], [29]. It uses the ‘height’ of the lifted vertices of the partition and, therefore, requires the computation of a convex hull in \mathbb{R}^{m+n+1} which can quickly become intractable.

For a given continuous PWA function $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ the proof of Theorem 2 produces generating problems with up to $2n$ decision variables. In case it is known that the given f is the solution to an unknown PQP in n variables, the procedure presented in [30] can be used to find a generating problem in n decision variables directly.

IV. INVERSE OPTIMIZATION MODELS FOR HYBRID SYSTEMS

The results of Sections II and III suggest a new modeling approach for a large class of hybrid system models, in particular PWA models [22]. Their discrete-time dynamics are given by

$$x^+ = f(x, u) = A_i x + B_i u + c_i \quad \text{for } (x, u) \in \Omega_i, \quad (14)$$

where $x \in \mathbb{R}^n$ is the current state, $u \in \mathbb{R}^m$ is the control input, $x^+ \in \mathbb{R}^n$ is the successor state and all system matrices are of appropriate dimensions. Note that by defining $p := (x, u)$, the dynamics (14) have the same form as (8). By applying Theorem 2 to the state transition function f of a given continuous PWA model (14), it is possible to obtain problem data for a PQP (possibly with some linear projection operator T) such that the state dynamics are given by the solution to this generating PQP.

A. Obtaining an inverse PQP model of a PWA system

Applying Theorem 2 to a PWA system (14) with

$$p := \begin{bmatrix} x \\ u \end{bmatrix}, \quad L_i := [A_i \quad B_i], \quad l_i := c_i, \quad (15)$$

provides a function $\hat{f}: \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{\hat{n}}$ and problem data such that

$$x^+ = f(p) = T\hat{f}(p), \quad (16a)$$

where \hat{f} is the unique solution to

$$\begin{aligned} \hat{f}(p) &\in \arg \min_z \frac{1}{2} z^\top Q z + p^\top R z + s^\top z \\ &\text{s.t. } Fz + Gp \leq h. \end{aligned} \quad (16b)$$

Here, $z \in \mathbb{R}^{\hat{n}}$, $p \in \mathbb{R}^{n+m}$, $h \in \mathbb{R}^{n_c}$, $T \in \mathbb{R}^{n \times \hat{n}}$, and all other matrices have appropriate dimensions. Note from Theorem 2 that it is always possible to find problem data such that $\hat{n} \leq 2n$.

Let us further assume that $Q = Q^\top \succ 0$, which can be enforced by using quadratic cost functions in the construction from Theorem 2. Accordingly, for every fixed p , $z^* := \hat{f}(p)$ must satisfy the KKT conditions for (16b):

$$Qz^* + R^\top p + s + F^\top \lambda = 0, \quad (17a)$$

$$0 \leq (h - Fz^* - Gp) \perp \lambda \geq 0. \quad (17b)$$

Here, $\lambda \in \mathbb{R}^{n_c}$ are the Lagrange multipliers corresponding to the constraints in (16b).

Together with (16a) and using the parameter substitution from (15) and the fact that Q is invertible, the KKT conditions (17) can be rearranged into the following explicit form:

$$x^+ = -TQ^{-1}R^\top \begin{bmatrix} x \\ u \end{bmatrix} - TQ^{-1}w - TQ^{-1}s \quad (18a)$$

$$v = \left(FQ^{-1}R^\top - G \right) \begin{bmatrix} x \\ u \end{bmatrix} + FQ^{-1}F^\top w + FQ^{-1}s + h \quad (18b)$$

$$0 \leq w \perp v \geq 0 \quad (18c)$$

Note that w are the Lagrange multipliers λ from (17b) and we solved for z^* in (17a).

The explicit form (18) is a so-called *linear complementarity* (LC) model of the system dynamics [16]. It has been proven in [31] that continuous PWA models are equivalent to LC models.

Our method for obtaining an LC model reveals a hidden convexity structure possessed by continuous PWA models, namely that these dynamics can be linked directly to solutions of strictly convex PQP problems. When this method is used to rewrite a PWA model (14) in terms of an optimization problem (16b) and a projection (16a), we call the resulting description an *inverse parametric quadratic programming* (IPQP) model of the dynamic system to emphasize its optimality structure.

As noted by the authors in [31], the standard construction of an LC model from a given PWA model often leads to unnecessarily large LC models and requires an intermediate *mixed-logical dynamical* (MLD) model, cf. [32]. Applying our inverse optimization results to a PWA system model provides a direct method for remodeling PWA systems in terms of an LC model. The cost function parameters in (16b) can be chosen such that the problem data are properly scaled, e.g. to facilitate solving (20).

A reasonable measure of the complexity of the IPQP model is the dimension of the complementarity vector w . This makes IPQP models with few constraints particularly attractive. In case a given PWA system is defined over a regular subdivision, applying Theorem 2 to obtain the IPQP model (18) always leads to a significantly less complex model than the LC model construction from [31]:

Lemma 5 (Complexity of equivalent LC models). *Suppose a given PWA system (14) with n states and m inputs is defined over a regular subdivision with n_r regions. Let $c_i \in \mathbb{N}$ denote the number of inequalities defining region Ω_i . Then the LC model resulting from the procedure in [31] requires*

$$n_r(6n + 1) + \sum_{i=1}^{n_r} c_i + 2 \geq n_r(7n + m + 2) + 2 \quad (19)$$

complementarity variables. Applying Theorem 2 to the same dynamics leads to an IPQP model (18) with $2nn_r$ complementarity variables.

Proof. The construction of an equivalent LC model from a given PWA model outlined in [31] relies on an intermediate mixed logical dynamical (MLD) model [32]. Let n_b denote the number of auxiliary binary variables, n_c the number of auxiliary continuous variables, and n_i the number of inequalities that are part of the MLD dynamics. By (10c) in [31], the LC model will then have $n_w = n_b + 2n_c + n_i$ complementarity variables.

The number of auxiliary binary variables in the MLD model corresponds to the number of regions, i.e. $n_b = n_r$, and for every region one set of states is introduced as auxiliary continuous variables, i.e. $n_c = nn_r$.

From (16a) in [32] it can be seen that all halfspace representations of the regions Ω_i enter the auxiliary inequalities of the MLD description, i.e. $\sum_{i=1}^{n_r} c_i$ inequalities. Equality (16b) from [32] introduces another two inequalities. Finally, the big-M reformulation of the PWA dynamics ((22) in [32]) introduces $4nn_r$ inequalities. Hence, for our discussion $n_i = \sum_{i=1}^{n_r} c_i + 4nn_r + 2$.

The total number of complementarity variables in the LC model constructed from [31] therefore is

$$n_w = \sum_{i=1}^{n_r} c_i + n_r (6n + 1) + 2.$$

Consider the case where all regions are simplices, i.e. compact polytopes described by the smallest possible number of halfspaces. In that case, $c_i = n + m + 1$ and we obtain the lower bound

$$n_w \geq n_r (7n + m + 2) + 2.$$

On the other hand, when applying Theorem 2 to the same PWA system, the number of complementarity constraints \bar{n}_w is given by the number of constraints of the generating problem which is bounded by $2nn_r$, i.e. $\bar{n}_w = 2nn_r$. \square

Theorem 2 enables a systems engineer to obtain automatically a compact and accurate LC model for a given PWA system without the need for special insight or ingenuity in eliminating unnecessary variables.

It is worth noting that the result from [31] is concerned with the existence of an equivalent LC system for any continuous PWA system only and does not try to provide an efficient conversion. While other results exist for the efficient conversion of MLD to PWA models, cf. [33], these will not reduce the left-hand side of (19). The authors are not aware of any attempts other than Lemma 5 at systematically providing compact LC models.

Note that the authors of [32] indicate how the number of binary variables can be reduced to scale logarithmically instead of linearly with the number of regions, but their contribution to the number of complementarity variables in (19) is not large enough for this to have a significant impact on the complexity comparison.

B. Optimal control problems for hybrid system models

In order to outline some advantages of the IPQP model introduced in the preceding section, consider the following finite time-horizon optimal control problem, as commonly encountered in model

predictive control:

$$\begin{aligned}
\min_{\{u_k\}_{k=0}^{N-1}} \quad & x_N^T \bar{P} x_N + \sum_{k=0}^{N-1} (x_k^T \bar{Q} x_k + u_k^T \bar{R} u_k) \\
\text{s.t.} \quad & x_{k+1} = f(x_k, u_k) \quad k = 0, \dots, N-1 \\
& (x_k, u_k) \in \mathcal{S} \quad k = 0, \dots, N-1 \\
& x_N \in \mathcal{T},
\end{aligned} \tag{20}$$

where f corresponds either to the PWA dynamic equation (14) or the IPQP dynamics (18). The set $\mathcal{S} \subseteq \mathbb{R}^n \times \mathbb{R}^m$ describes state/input constraints that must be respected over the horizon $k = 0, \dots, N-1$, with the convex set $\mathcal{T} \subseteq \mathbb{R}^n$ describing a constraint on the terminal state. Finally, $\bar{Q} = \bar{Q}^T \succeq 0$, $\bar{R} = \bar{R}^T \succ 0$ and $\bar{P} = \bar{P}^T \succ 0$ are given cost function matrices.

We assume that the constraint sets \mathcal{S} and \mathcal{T} are convex polytopes, i.e. described by a finite number of affine inequalities. When the PWA dynamic model (14) is used, (20) can be written as a mixed-integer quadratic program (MIQP) [32]. An alternative method in [34] makes use of a different reformulation of the PWA dynamics that casts (20) as a nonconvex, nonlinear optimization problem. Using the equivalent IPQP dynamics (18) leads to a so-called *mathematical program with equilibrium constraints* (MPEC) in complementarity form [35]. Using the implicit IPQP dynamics (16) turns (20) into a *bilevel optimization problem*, which is closely related to an MPEC [18]. Another complementarity approach to (20) for first-order systems has already been explored in [36].

When casting (20) as an MIQP, a ‘big M’ formulation to bound system variables is used [32]. This can lead to badly scaled problem data in (20). The degrees of freedom retained by Theorem 2 in designing the generating cost function for the system dynamics can be used to attenuate or avoid this when writing (20) as an MPEC.

Being able to obtain a compact IPQP model of a given PWA system enables us to tackle (20) with a number of tools that were not previously available. Although the MPEC formulation of (20) is not necessarily easier to solve than the ‘classical’ MIQP formulation, one immediate advantage is that it involves only continuous (and no binary) variables. While in principle MPECs can also be modeled as MIQPs, it is often better to deal only with the continuous variables of the original MPEC [37]. While there is a large body of literature devoted to the solution of MPECs and bilevel optimization problems, we mention interior point methods in particular which have been used successfully to solve MPECs, see [17] and the references therein. Results from [30] indicate that applying these methods to control problems such as (20) can provide significant computational advantages compared to standard MIQP approaches.

V. CONCLUSION AND OUTLOOK

By showing how a large class of functions, in particular continuous PWA functions, can be obtained from convex optimization problems, we can shed a new light on dynamic system modeling. Using our results it is possible to interpret system dynamics as an optimizing process. For any nonlinear system whose state transition function is given by a vector-valued DC function, this modeling idea can be applied.

This approach can provide insight into the modeled process and enables the use of new tools to solve optimal control problems such as (20). Results from [30] indicate that using this new modeling methodology might make it possible to apply advanced optimal control methods such as MPC to fast hybrid systems. Using the inverse optimization interpretation of a given PWA model in an optimal control problem rids us of the need for binary variables and turns the optimal control problem into a continuous optimization problem. A fast implementation of a solution algorithm specifically tailored to the MPEC / bilevel program (20) would enable the application of MPC techniques to control systems for which the standard PWA approach based on mixed-integer programming is prohibitive due to timing constraints.

REFERENCES

- [1] E. N. Pistikopoulos, M. C. Georgiadis, and V. Dua, *Multi-parametric programming*. Weinheim, Germany: Wiley-VCH, 2007.
- [2] B. Bank, J. Guddat, D. Klatte, B. Kummer, and K. Tammer, *Non-linear parametric optimization*. Basel and Boston: Birkhäuser Verlag, 1983.
- [3] J. B. Rawlings and D. Q. Mayne, *Model predictive control: Theory and design*. Madison, Wis: Nob Hill Pub., 2009.
- [4] T. Gal and J. Nedoma, "Multiparametric linear programming," *Management Science*, 18(7):406–422, 1972.
- [5] A. Bemporad, M. Morari, V. Dua, and E. N. Pistikopoulos, "The explicit linear quadratic regulator for constrained systems," *Automatica*, 38(1):3–20, 2002.
- [6] F. Borrelli, "Constrained optimal control for hybrid systems," in *Constrained Optimal Control of Linear and Hybrid Systems*, vol. 290, Berlin, Heidelberg: Springer, 2003:143–171.
- [7] P. Abbeel and A. Y. Ng, "Apprenticeship learning via inverse reinforcement learning," in *Proc. of the 21st Int. Conf. on Machine Learning*, ACM Press, 2004:1.
- [8] R. Kalman, "When is as linear control system optimal?" *Trans. of the ASME. Series D, Journal of Basic Engineering*, 86(1):51–60, 1963.
- [9] M. Baes, M. Diehl, and I. Necoara, "Every continuous nonlinear control system can be obtained by parametric convex programming," *IEEE Trans. on Automatic Control*, 53(8):1963–1967, 2008.
- [10] A. Keshavarz, Y. Wang, and S. Boyd, "Imputing a convex objective function," in *IEEE Int. Symp. on Intelligent Control*, 2011:613–619.
- [11] S. Dempe and S. Lohse, "Inverse linear programming," in *Lecture Notes in Economics and Mathematical Systems*, A. Seeger, Ed., vol. 563, Berlin, Heidelberg: Springer-Verlag, 2006:19–28.
- [12] R. K. Ahuja and J. B. Orlin, "Inverse optimization," *Operations Research*, 49(5):771–783, 2001.
- [13] R. Horst and N. V. Thoai, "DC programming: overview," *Journal of Optimization Theory and Applications*, 103(1):1–43, 1999.
- [14] A. Kripfganz and R. Schulze, "Piecewise affine functions as a difference of two convex functions," *Optimization*, 18(1):23–29, 1987.
- [15] A. Hempel, P. Goulart, and J. Lygeros, "Every continuous piecewise affine function can be obtained by solving a parametric linear program," in *Proc. of the European Control Conf.*, Zurich, CH, 2013:2657–2662.
- [16] W. P. M. H. Heemels and B. Brogliato, "The complementarity class of hybrid dynamical systems," *European Journal of Control*, 9(2-3):322–360, 2003.
- [17] D. Ralph, "Mathematical programs with complementarity constraints in traffic and telecommunications networks," *Philosophical Transactions of the Royal Society A*, 366(1872):1973–1987, 2008.
- [18] B. Colson, P. Marcotte, and G. Savard, "An overview of bilevel optimization," *Ann. of Operations Research*, 153(1):235–256, 2007.
- [19] A. Alvarado, G. Scutari, and J.-S. Pang, "A New Decomposition Method for Multiuser DC-Programming and its Applications," <http://arxiv.org/abs/1308.3521>(), 2013. arXiv: 1308.3521.
- [20] D. M. Raimondo, S. Rivero, C. N. Jones, and M. Morari, "A robust explicit nonlinear MPC controller with input-to-state stability guarantees," in *Proc. of the 18th IFAC World Congress*, 2011:9284–9289.
- [21] R. T. Rockafellar and R. J.-B. Wets, *Variational analysis*. Berlin; New York: Springer, 1998.
- [22] E. Sontag, "Nonlinear regulation: the piecewise linear approach," *IEEE Transactions on Automatic Control*, 26(2):346–358, 1981.

- [23] J. Spjøtvold, P. Tøndel, and T. A. Johansen, "Continuous selection and unique polyhedral representation of solutions to convex parametric quadratic programs," *Journal of Optimization Theory and Applications*, 134(2):177–189, 2007.
- [24] J. A. De Loera, J. Rambau, and F. Santos, *Triangulations - structures for algorithms and applications*, ser. Algorithms and Computation in Mathematics. Berlin; Heidelberg: Springer, 2010, vol. 25.
- [25] G. Ferrari-Trecate, M. Muselli, D. Liberati, and M. Morari, "A clustering technique for the identification of piecewise affine systems," *Automatica*, 39(2):205–217, 2003.
- [26] L. Rodrigues and S. Boyd, "Piecewise-affine state feedback for piecewise-affine slab systems using convex optimization," *Systems & Control Letters*, 54(9):835–853, 2005.
- [27] F. Aurenhammer, "Voronoi diagrams - a survey of a fundamental geometric data structure," *ACM Computing Surveys*, 23(3):345–405, 1991.
- [28] N. A. Nguyen, S. Olaru, O. Rodriguez-Ayerbe, M. Hovd, and I. Necoara, "Inverse parametric convex programming problems via convex liftings," in *Proc. of the 19th IFAC World Congress*, Cape Town, SA, 2014.
- [29] N. A. Nguyen, S. Olaru, P. Rodriguez-Ayerbe, M. Hovd, and I. Necoara, "On the lifting problems and their connections with piecewise affine control law design," in *Proc. of the European Control Conference*, Strasbourg, FR, 2014.
- [30] A. B. Hempel, P. J. Goulart, and J. Lygeros, "Inverse parametric quadratic programming and an application to hybrid control," in *Proc. of the 4th IFAC Conference on NMPC*, Noordwijkerhout, 2012:68–73.
- [31] W. Heemels, B. de Schutter, and A. Bemporad, "Equivalence of hybrid dynamical models," *Automatica*, 37(7):1085–1091, 2001.
- [32] A. Bemporad and M. Morari, "Control of systems integrating logic, dynamics, and constraints," *Automatica*, 35(3):407–427, 1999.
- [33] A. Bemporad, "Efficient Conversion of Mixed Logical Dynamical Systems Into an Equivalent Piecewise Affine Form," *IEEE Transactions on Automatic Control*, 49(5):832–838, 2004.
- [34] B. de Schutter and T. van den Boom, "MPC for continuous piecewise-affine systems," *Systems & Control Letters*, 52(3–4):179–192, 2004.
- [35] Z.-Q. Luo, J.-S. Pang, and D. Ralph, *Mathematical programs with equilibrium constraints*. Cambridge: Cambridge University Press, 1996.
- [36] B. de Schutter, "Optimal control of a class of linear hybrid systems with saturation," *SIAM Journal on Control and Optimization*, 39(3):835–851, 2000.
- [37] B. Baumrucker, J. Renfro, and L. Biegler, "MPEC problem formulations and solution strategies with chemical engineering applications," *Comp. & Chemical Engineering*, 32(12):2903–2913, 2008.