Facial Input Decompositions for Robust Peak and Reachable Set Estimation under Polyhedral Uncertainty

Jared Miller$^1$, Mario Sznaier$^1$

May 6, 2022

Abstract

This work bounds extreme values of state functions and approximates reachable sets for a class of input-affine continuous-time systems that are affected by polyhedral-bounded uncertainty. Instances of these systems may arise in data-driven peak estimation, in which the state function must be bounded for all systems that are consistent with a set of state-derivative data records corrupted under L-infinity bounded noise. Existing occupation measure-based methods form a convergent sequence of outer approximations to the true peak value or reachable set volume, given an initial set, by solving a hierarchy of semidefinite programs in increasing size. These techniques scale combinatorially in the number of state variables and uncertain parameters. We present tractable algorithms for peak and reachable set estimation that scale linearly in the number of faces of the uncertainty-bounding polytope rather than combinatorially in the number of uncertain parameters by leveraging convex duality and a theorem of alternatives (facial decomposition). The sequence of decomposed semidefinite programs will converge to the true optimal value under mild assumptions (convergence and smoothness of dynamics).

1 Introduction

Robust peak estimation aims to bound all possible values of a function $p(x(t))$ of the state $x$ of an uncertain dynamical system along the trajectories that start from an initial set $X_0 \subseteq X \subset \mathbb{R}^n$. This problem with admissible uncertainty processes $w(t)$ remaining in a set $W$ in times $[0,T]$ may be expressed as,

$$P^* = \max_{t \in [0,T], x_0 \in X_0, w(t)} p(x(t \mid x_0, w(t)))$$ (1)

$$\dot{x}(t) = f(t, x(t), w(t)), \ w(t) \in W \ \forall t \in [0,T].$$

This paper considers a continuous-time input-affine dynamical system of the form:

$$\dot{x}(t) = f(t, x(t), w(t)) = f_0(t, x) + \sum_{\ell=1}^{L} w_\ell(t) f_\ell(t, x).$$ (2)

The uncertainty set $W$ is restricted to a compact non-empty polytope described by,

$$W = \{ w \mid Aw \leq b \} \quad A \in \mathbb{R}^{m \times L}, \ b \in \mathbb{R}^m.$$ (3)

The uncertainty set dimension $L$ and number of affine constraints $m$ are each assumed to be finite.

The peak estimation problem in (1) is a particular instance of an optimal control problem with zero running cost and a free terminal time. Infinite-dimensional Linear Programs (LP) in occupation measures were developed for optimal control in [1]. These infinite-dimensional programs were truncated into a converging sequence of finite-dimensional Linear Matrix Inequalities in increasing size through the moment-Sum of Squares (SOS) hierarchy in [2].

$^1$J. Miller and M. Sznaier are with the Robust Systems Lab, ECE Department, Northeastern University, Boston, MA 02115. (e-mails: miller.jare@northeastern.edu, msznaier@coe.neu.edu).

J. Miller and M. Sznaier were partially supported by National Science Foundation (NSF) grants CNS–1646121, CMMI–1638234, ECCS–1808381 and CNS–2038493, and Air Force Office of Scientific Research (AFOSR) grant FA9550-19-1-0005. This material is based upon research supported by the Chateaubriand Fellowship of the Office for Science & Technology of the Embassy of France in the United States.
Infinite-dimensional LPs in measures for peak estimation were formulated in [3] and were solved through a gridded discretization of the infinite-dimensional LP and through Markov Chain Martingale techniques. The work in [4] applied the converging moment-SOS hierarchy to the dual peak estimation problem in terms of a continuous auxiliary function $v(t,x)$. Peak estimation was extended to systems with dynamical uncertainty in [5], which includes the class of systems considered in Eq. (2). Infinite-dimensional LPs have also been applied to perform reachable set approximation, such as from outside in [6] and from inside in [7].

This paper is a sequel to [5], which formulates a peak estimation program for general uncertainties $w(t) \in \mathcal{W}$ possibly including switching structure (vertex decomposition of $\mathcal{W}$). The approach in [5] yields a convex optimization problem for polyhedral $\mathcal{W}$, but its computational complexity scales in a combinatorial manner as $L$ increases. To address this issue, this paper uses a facial decomposition of the polyhedral $\mathcal{W}$ in (3) with respect to input-affine dynamics (2) to eliminate the variables $w$ and generate tractable Semidefinite Programs (SDPs) for peak estimation. The facial decomposition of $\mathcal{W}$ arises from a theorem of alternatives in robust optimization [8] and convex duality [9]. Specific cases of facial decompositions when $\mathcal{W}$ is a unit box were discussed in [10, 11]; this paper allows for general convex polytopes $\mathcal{W}$ to be considered.

Letting $\deg(f)$ be the degree of the polynomial dynamics function $f(t,x,w)$, the following table lists the size of the largest Positive Semidefinite (Gram) Matrix for a peak estimation problem under polyhedral uncertainty with parameters $d = 4$, $L = 10$, $n = 2$, $\deg(f) = 3$ (Lie constraint to be observed in Section 6.2) in the degree-$d$ SOS tightenings of the robust peak estimation programs.

| Table 1: Size of largest Lie constraint Gram Matrix (Peak Estimation) |
|----------------------------------------------------------|
| Previous work in [5] | $(1+n+L+d+[\deg(f)/2]-1) = 8568$ |
| This paper | $(1+n+d+\max\{\deg(f)/2\}-1) = 56$ |

Imposing that a symmetric matrix of size 8568 is positive semidefinite is intractable in numerical solvers such as Mosek and Sedum. The polytope $\mathcal{W}$ has 7534 vertices and 33 faces. A vertex decomposition would require 7534 PSD constraints of size 56, while the equivalent facial decomposition imposes 33 + 1 PSD constraints of size 56.

Polytopic uncertainty sets in (3) may arise from data-driven settings, in which a noisy set of state-derivative observations are collected. The uncertainty terms $w(t)$ may represent either external inputs or unknown parameters of the uncertain dynamical system. Data $D = \{(t_k,x_k,y_k)\}_{k=1}^{N}$ are acquired where $y(t_k) = \dot{x}(t_k) + \eta_k$ with respect to dynamics (2) for $L_{\infty}$ bounded error terms $\eta_k$. These errors $\eta_k$ are intended to model the approximation error when $\dot{x}(t)$ is computed numerically using finite differences.

The contributions of this paper are,

- Application of a theorem of alternatives to simplify a Lie constraint for continuous-time systems
- Extending facial input decompositions to general polytopes $\mathcal{W}$
- Reduction in computational complexity of SDPs
- Presentation and demonstration of convex data-driven peak and reachable set estimation programs

This paper is organized as follows: Section 2 reviews preliminaries including notation, peak estimation, and the data-driven uncertainty formalism. Section 3 splits up the Lie derivative constraint through a facial decomposition for arbitrary polytopic sets $\mathcal{W}$ using a Theorem of Alternatives. The decomposed peak estimation program for polytopic time-varying uncertainty is formulated in Section 4. The data-driven framework and its polytopic description of $\mathcal{W}$ is covered in Section 5. Examples of polytopic-decomposed peak estimation problems in the context of data-driven system analysis are presented in Section 6. The reachable set problem under polyhedral uncertainty is decomposed and solved in Section 7. Section 8 presents some conclusions and briefly discusses future work. Appendix A contains a proof that multiplier functions associated with a certificate of Lie constraint nonnegativity may be chosen to be continuous, and Appendix B shows that the multiplier functions may be approximated by polynomials.
2 Preliminaries

2.1 Acronyms

| Acronym | Description |
|---------|-------------|
| LP      | Linear Program |
| PSD     | Positive Semidefinite |
| SDP     | Semidefinite Program |
| SIR     | Susceptible, Infected, Removed |
| SOS     | Sum of Squares |

2.2 Notation

The set of real numbers is \( \mathbb{R} \), and the \( n \)-dimensional Euclidean space is \( \mathbb{R}^n \). Two vectors \( x, y \in \mathbb{R}^n \) have the relation \( x \geq y \) if each element satisfies \( x_i \geq y_i \) for all \( i = 1, \ldots, n \). The inner product between two vectors in \( \mathbb{R}^n \) is \( x \cdot y = \sum_i x_i y_i \). A matrix \( Q \in \mathbb{R}^{n \times n} \) is PSD (\( Q \succeq 0 \)) if the associated quadratic form satisfies \( \forall x \in \mathbb{R}^n : x^T Q x \geq 0 \). The set of polynomials with real coefficients in indeterminate values \( x \) of degree at most \( d \) is \( \mathbb{R}[x]_{\leq d} \). The set of polynomials in degree at most \( d \) is \( \mathbb{R}[x]_{\leq d} \). The set of continuous functions with continuous first derivatives is \( C^1(X) \subset C(X) \).

2.3 Sum-of-Squares Hierarchy

A polynomial \( p(x) \in \mathbb{R}[x] \) is SOS if it may be (non-uniquely) decomposed into the sum \( p(x) = \sum_i q_i(x)^2 \) for a finite number of polynomial terms \( q_i(x) \in \mathbb{R}[x] \forall i = 1, \ldots, N \). The cone of SOS polynomials is written as \( \Sigma[x] \), and this cone is a subcone of the set of continuous functions \( C(X) \). Every compact basic semialgebraic set \( K \) satisfies the Archimedean condition if there exists a finite \( R \) and SOS polynomials \( \{ \sigma_i(x) \}_{i=0}^N \) such that,

\[
R^2 - \|x\|^2 = \sigma_0(x) + \sum_{i=1}^N \sigma_i(x) g_i(x). \quad (4)
\]

Every compact basic semi-algebraic set \( K \) may be made Archimedean by appending the redundant ball constraint \( R^2 - \|x\|^2 \geq 0 \) to the description of \( K \).

The constrained optimization problem of \( P^* = \min_{x \in \mathbb{K}} c(x) \) may be expressed through a nonnegativity constraint as,

\[
P^* = \max_{\gamma \in \mathbb{K}} \gamma, \quad c(x) - \gamma \geq 0 \quad \forall x \in \mathbb{K} \quad (5)
\]

M. Putinar introduced an algebraic certificate (Putinar Positivestellensatz) that is necessary and sufficient condition for a function \( p(x) \) to be positive over the Archimedean basic semi-algebraic set \( \mathbb{K} \),

\[
p(x) = \sigma_0(x) + \sum_{i=1}^N \sigma_i(x) g_i(x)
\]

\[
\sigma(x) \in \Sigma[x], \quad \sigma_i(x) \in \Sigma[x] \quad (6)
\]

The degree-\( d \) SOS tightening of problem \( (5) \) uses a bounded-degree Putinar Psatsz in place of the nonnegativity constraint \( (5) \),

\[
d^* = \sup \gamma \quad (7a)
\]

\[
c(x) - \gamma = \sigma_0(x) + \sum_{i=1}^N \sigma_i(x) g_i(x) \quad (7b)
\]

\[
\sigma_i(x) \in \Sigma[x] \quad \forall i = 0, \ldots, N_c \quad (7c)
\]
Problem \( \text{(7)} \) is a Semidefinite Program (SDP) in terms of elements of the finite-dimensional Gram matrices \( Q_0 \ldots Q_N \) describing the Putinar multipliers \( \sigma_0 \ldots \sigma_N \). The objective values between programs \( \text{(5)} \) and \( \text{(7)} \) are related by \( d_0^* \leq P^* \), and the chain of lower bounds obeys \( d_0^* \leq d_1^* \leq d_2^* \leq \ldots \) as the degree of the SOS tightening increases \( \text{[14]} \). When \( K \) is Archimedean, the objective value \( d_0^* \) will converge to \( P^* \) as the relaxation degree \( d \) approaches \( \infty \). The SOS (moment-SOS) hierarchy is the process of forming SDPs in increasing degree by replacing polynomial nonnegativity constraints by SOS constraints. The size of a Gram matrix in \( n \) variables and at degree \( d \) is \( N = (n + d) \approx n^d \). The per-iteration complexity of an Interior Point Method for solving an SDP up to \( \epsilon \)-accuracy with an \( N \)-dimensional SDP matrix and \( M \) affine constraints is \( O(N^3M + N^2M^2 + M^3) \) \( \text{[15]} \). This runtime is polynomial in \( n \) for fixed \( d \), and is combinatorial as \( d \) increases in the SOS hierarchy.

### 2.4 Peak Estimation and Uncertainty

In this paper, we make the following assumptions:

**A1** The time horizon \( T \) is finite.

**A2** The state \( x \in X \subset \mathbb{R}^n \) and initial condition \( x_0 \in X_0 \subseteq X \) lie in compact sets.

**A3** The dictionary functions \( f_0(t,x) \) and \( f_\ell(t,x) \) for all \( \ell = 1, \ldots, L \) are each assumed to be Lipschitz.

**A4** The uncertainty set \( W \) is a compact polytope with a non-empty interior.

**A5** At least one optimal trajectory with \( P^* = p(x(t^* \mid x_0^*, w^*(t))) \) satisfies \( t^* \in [0,T], \ x_0^* \in X_0 \), and \( x(t' \mid x_0^*) \in X, \ w^*(t') \in W \) for all \( t' \in [0, t^*] \).

**Remark 1.** The combination of A1 and A2 imply that system \( \text{(2)} \) does not have finite escape time. Further, if the function \( p(x) \) is continuous, this assumption implies that \( P^* \) is bounded above.

Figure \( \text{[1]} \) illustrates an example of peak estimation under uncertainty. The system dynamics are a modification of the Flow system from \( \text{[16]} \) with a new time-varying uncertainty term \( w(t) \in [-0.5, 0.5] \),

\[
\dot{x}(t) = [x_2(t); -x_1(t) - x_2(t) + (1 + w(t))x_3(t)/3].
\]  

(8)

Sample trajectories in Figure \( \text{[1]} \) (cyan curves) start within the initial set \( X_0 = \{ x \mid (x_1 - 1.5)^2 + x_2^2 \leq 0.4^2 \} \) (black circle) and follow Flow dynamics for a time horizon of \( T = 5 \). It is desired to lower-bound the minimum value of the vertical coordinate \( x_2 \).

\[ {\text{Figure 1: Plot of Uncertain Flow system (8) trajectories}} \]
The auxiliary function $v(t, x)$ must decrease along trajectories generated by all possible admissible disturbance processes $w(t)$. The objectives $P^* = d^*$ between programs (11) and (9) are equal when $[0, T] \times X \times W$ is compact, $p(x)$ is bounded below, and $f(t, x, w)$ is Lipschitz.

The infinite-dimensional LP in (9) may be approximated through the moment-SOS hierarchy as discussed in Section 2.3. When the set $x$ satisfies the Archimedean property, the sequence of SOS tightenings $d^* \geq d^*_t \geq \ldots$ will converge as $\lim_{t \to \infty} d^*_t = P^*$. The order-4 moment-SOS peak estimate to program (9) yields a bound $x_2(t) \geq -0.7862$ along trajectories starting from $X_0$ for time $t \in [0, 5]$, as shown in Figure 1 by the red line.

3 Decomposed Lie Constraint

The decomposable expression of interest for polytopic uncertainty of the form (3), is the Lie derivative constraint in (9c). An auxiliary function $v(t, x) \in C^1([0, T] \times X)$ must be non-increasing along trajectories of the disturbance-affine dynamical system $f$. The Lie derivative may be expanded into,

$$\mathcal{L}_f v(t, x) = \partial_v v(t, x) + f(t, x, w) \cdot \nabla_x v(t, x)$$

$$= \mathcal{L}_f v(t, x) + \sum_{\ell=1}^L w_{\ell} f_{\ell}(t, x) \cdot \nabla_x v(t, x),$$

and the Lie derivative constraint in (9c) is,

$$\mathcal{L}_f v(t, x) \leq 0 \quad \forall (t, x, w) \in [0, T] \times X \times W. \quad (12)$$

Convex duality may be used to decompose Equation (12) from a nonnegativity constraint over $(t, x, w)$ into a constraint that no longer depends on $w$ (reducing the maximum size of the Gram matrices in the SOS programs). This theory is based on the robust optimization work of [8], and work applying duality to control constraints includes [17] [18].

If the Lie constraint (12) holds, then there does not exist a point $(t, x, w)$ such that $\mathcal{L}_f(t, x, w)v(t, x)$ takes on positive values. The following program is therefore infeasible, where $w$ lies in the polytope $W = \{w \mid Aw \leq b\}$ from [3]:

$$\text{find } (t, x, w) \in [0, T] \times X \times \mathbb{R}^L$$

$$\mathcal{L}_f v(t, x) > 0$$

$$= Aw \geq -b \quad (13c)$$

3.1 One Strict Inequality

Let $R \subset X$ be a region formed by $R = (F(x) > 0) \cap \{f_j(x) \geq 0\}_{j=1}^m$ for one strict inequality and $m$ non-strict inequalities. Problem (13) is a particular instance of this problem with $F = \mathcal{L}_f v(t, x)$ and $f_j = [-Aw + b]$. Multipliers $Z > 0$ and $\{\zeta_j \geq 0\}_{j=1}^m$ can be defined to form the weighted sum,

$$S(x; Z, \zeta) = ZF(x) + \sum_{j=1}^m \zeta_j f_j(x). \quad (14)$$

Theorem 3.1. The set $R$ is empty if there exists a choice of $Z > 0$, $\{\zeta_j \geq 0\}_{j=1}^m$ such that,

$$g(Z, \zeta) = \sup_x S(x; Z, \zeta) \leq 0. \quad (15)$$

Proof. Every point $x \in R$ will satisfy $F(x) > 0$ due to the definition of $R$. The weighted sum $ZF(x) > 0$ is likewise positive in $R$ because $Z > 0$ (strict inequality). The term $\sum_{j=1}^m \zeta_j f_j(x) \geq 0$ will be nonnegative for all points in $R$ given that $\zeta_j, f_j(x) \geq 0, \forall \ell \in 1, \ldots, L$. The resultant sum $S(x; Z, \zeta)$ from Eq. (14) will be positive for all points in $R$.

The region $R$ is empty if there exists a choice of $Z > 0$ and $\zeta_j \geq 0$ such that $S(x) \leq 0$. The addition of a positive value and a sum of nonnegative values $S(x)$ cannot possibly be nonpositive, so such a $Z > 0$ and $\zeta_j \geq 0$ serves as a certificate of emptiness. A proof of the theorem of alternatives for all non-strict inequalities or all strict inequalities is available in Section 5.8 of [3].
Remark 2. The expression for $S(x)$ may be divided through by the nonzero positive quantity $Z$. A new version of $S(x)/Z \to S(x)$ and $\zeta_j/Z \to \zeta_j \geq 0$ may be defined such that the rescaling satisfies,

$$S(x) = F(x) + \sum_{j=1}^{m} \zeta_j f_j(x) \quad (16)$$

### 3.2 Lie Constraint

Program (13c) may be decomposed through convex alternatives as shown in the previous subsection. Define nonnegative (possibly discontinuous) dual variables $\zeta_j(t,x)$ for each linear constraint $j = 1, \ldots, m$ in (13c). A Lagrangian $\mathcal{L}(w; \zeta, t, x)$ for problem (13) with a given auxiliary function $v(t,x)$ is:

$$\mathcal{L} = \mathcal{L}_f v(t, x) + \zeta(t,x)^T (b - Aw) \quad (17a)$$

$$= \mathcal{L}_f v(t, x) + b^T \zeta(t, x) \quad (17b)$$

$$+ \sum_{\ell=1}^{L} w_\ell \left( f_\ell \cdot \nabla_x v(t, x) - \sum_{j=1}^{m} A_{j\ell} \zeta_j(t, x) \right)$$

The dual function $g(\zeta, t, x; v) = \sup_{w \in \mathbb{R}^L} \mathcal{L}(w; \zeta, t, x)$ of this Lagrangian takes on values,

$$\begin{cases} 
\mathcal{L}_f v + b^T \zeta & \forall \ell : f_\ell \cdot \nabla_x v(t, x) = \sum_{j=1}^{m} A_{j\ell} \zeta_j \\
\infty & \text{else} \end{cases} \quad (18)$$

Problem (13) is infeasible if the dual function $g(\zeta, t, x)$ is nonpositive for every $(t, x)$ by the proof in Section 5.8 of [9]. The dual problem of finding a $\zeta(t, x)$ that infinites $\mathcal{L}(w; \zeta, t, x)$ is,

$$\begin{align*}
\text{find} & \quad \zeta_j \geq 0 & \forall j \\
\mathcal{L}_f v + b^T \zeta(t, x) & \leq 0 & \forall j \\
\sum_{j=1}^{m} A_{j\ell} \zeta_j(t, x) & = f_\ell(t, x) \cdot \nabla_x v(t, x) & \forall \ell \quad (19a) \\
\end{align*}$$

The nonnegative multiplier functions $\zeta_j(t, x)$ are constant in the input $w_\ell$ for each $\ell = 1, \ldots, L$. The pair of programs (13) and (19) are strong alternatives because the function $\mathcal{L}_f v(t, x)$ is affine (convex) in $w$, the set $W$ is nonempty (feasible point $w \in W$), and the constraints in $Aw \leq b$ are convex in $w$. Feasibility of (12) and (19) are therefore equivalent for a given $v(t, x)$, and (19) no longer involves $w$. The multipliers $\zeta_j(t, x)$ may be chosen to be continuous when $[0, T] \times X \times W$ is compact ($\zeta_j(t, x) \in C_+([0, T] \times X)$) as proven in Appendix A.

## 4 Polyhedral Uncertainty Peak Problem

This section will present peak estimation programs where the Lie constraint (9c) is decomposed into (19).

### 4.1 Polytopic Decomposition (Function Program)

The polytopic-decomposed peak program only changes the Lie derivative constraint (9c), forming the program,

$$d^* = \min_{\gamma \in \mathbb{R}} \gamma \quad (20a)$$

$$\gamma \geq v(0, x) \quad \forall x \in X_0 \quad (20b)$$

$$\mathcal{L}_f v(t, x) + b^T \zeta(t, x) \leq 0 \quad \forall (t, x) \in [0, T] \times X \quad (20c)$$

$$(A^T)_{j\ell} \zeta(t, x) = f_\ell \cdot \nabla_x v(t, x) \quad \forall \ell = 1, \ldots, L \quad (20d)$$

$$v(t, x) \geq p(x) \quad \forall (t, x) \in [0, T] \times X \quad (20e)$$

$$v(t, x) \in C^1([0, T] \times X) \quad (20f)$$

$$\zeta_j(t, x) \in C_+([0, T] \times X) \quad \forall j = 1, \ldots, m \quad (20g)$$

The slacks $\zeta_j(t, x)$ for $j = 1, \ldots, m$ are concatenated into the vector of nonnegative functions $\zeta(t, x)$. The expression $(A^T)_{j\ell} \zeta(t, x)$ may be read as $\sum_{j=1}^{m} A_{j\ell} \zeta_j(t, x)$ for each $j = 1, \ldots, L$. 

Theorem 4.1. The objectives $d^*$ between programs (9) and (20) are equal when A1-5 hold.

Proof. Tightness is assured by application of the Theorem of Alternatives in (19). The Lie constraint (9c) is replaced by constraints (20c)-(20d). The functions $\zeta$ may be restricted to be continuous as per Appendix A.

Theorem 4.2. The objectives from (20) and (1) are equal when assumptions A1-A5 are satisfied.

Proof. Problems (1) and (9) have equal objectives under assumptions A1-A5 by Theorem 2.1 of [1]. It is therefore proven that $P^* = d^*$ from problems (20) and (1) by application of Theorem 4.1.

4.2 Polytopic Decomposition (SOS Program)

Problem (20) may be approximated through SOS programming. Assume that the functions $\{f_\ell\}_{\ell=0}^L$ and $p(x)$ are each polynomial, the time horizon $T$ is finite, and that $X, X_0$ are Archimedean basic semialgebraic sets with the following descriptions,

$$X_0 = \{x \mid g_0^t(x) \forall k = 1, \ldots, N_c^0\}$$
$$X = \{x \mid g_k(x) \forall k = 1, \ldots, N_c\}.$$ (21) (22)

The shorthand notations $\forall j$ and $\forall \ell$ may be expanded into $\forall j = 1, \ldots, m$ and $\forall \ell = 1, \ldots, L$ in the following program. In addition, the expression $c(x) \in \Sigma[1] \leq 2d$ will denote that the degree $\leq 2d$ Putinar certificate of nonnegativity (6) over the Archimedean domain $X$ in (22). Define the dynamics degree $d'$ as $d' = \max\{\text{deg}(f_\ell)/2\} + d - 1$. The degree-$d'$ SOS tightening of program (20) forms the SDP

$$d'_d = \inf_{\gamma \in \mathbb{R}}$$
$$\gamma - \nu(0, x) \in \Sigma[X_0] \leq 2d$$
$$-\mathcal{L}_f v(t, x) - b^T \zeta(t, x) \in \Sigma[[0, T] \times X] \leq 2d'$$
$$v(t, x) - p(x) \in \Sigma[[0, T] \times X] \leq 2d$$
$$\forall \ell : (A^T)_\ell \zeta(t, x) = f_\ell(t, x) \cdot \nabla_x v(t, x)$$
$$v(t, x) \in \mathbb{R}[t, x] \leq 2d$$
$$\forall j : \zeta_j(t) \in \Sigma[[0, T] \times X] \leq 2d'$$ (23a) (23b) (23c) (23d) (23e) (23f) (23g)

Constraints (23b)-(23e) are linear equality constraints involving coefficient vectors of $v, \zeta$ and elements of the Gram matrices of each multiplier $\sigma$ in the Putinar expressions (6). The degree-$d'$ SOS tightening of programs (9) and (20) each involve a polynomial auxiliary function $v(t, x)$ of degree $2d$.

Theorem 4.3. The objective of (23) will approach (20) as the degree increases with $\lim_{d \to \infty} d'_d = d^*$ when $X_0$ is nonempty, assuming that $[0, T] \times X \times W$ is Archimedean.

Proof. The proof of this theorem using Stone-Weierstrass approximations is contained in Appendix B.

5 Data-Driven Setting

An unknown ground truth ODE system $\dot{x} = F(t, x)$ is observed by a noisy measurement process. The corrupted measurement model with bounded noise term $\eta$ satisfying $\|\eta\|_\infty \leq \epsilon$ is $\dot{x}_\text{observed} = y = F(t, x) + \eta$.

Figure 2 plots 40 noisy observations of the Flow system in the disk with center $(1.5, 0)$ and radius 0.4. The blue arrows of each plot are the true noiseless system dynamics, and the orange arrows are noise-corrupted estimates of the derivatives $\dot{x}$. These records only possess $\epsilon = 0.5$ noise in the $x_2$ coordinate and have perfect knowledge of $\dot{x}_1 = x_2$.

Let $\dot{x} = f(t, x; \omega) = f_0(t, x) + \sum_{\ell=1}^L w_\ell f_\ell(t, x)$ be a continuous ODE depending affinely on a set of parameters $\omega \in \mathbb{R}^L$. The function $f_0$ encodes prior knowledge about system dynamics, and the basis
functions \( f_\ell \) are a dictionary to describe unknown dynamics. Let \( \mathcal{D} \) be a set of tuples \( \mathcal{D}_k = (t_k, x_k, y_k) \) for \( k = 1, \ldots, N \), noisy observations, which may arise from multiple time sequences and traces of the same system. The set of system parameters \( w \) for each system \( k \) is described by

\[
W = \{ w \in \mathbb{R}^L | \forall k : \|y_k - f(t_k, x_k; w)\|_\infty \leq \epsilon \}.
\]

The \( L_\infty \) term in \( W \)'s constraint associated with a single data record \( \mathcal{D}_k = (t_k, x_k, y_k) \) is,

\[
\|y_k - f(t_k, x_k; w)\|_\infty = \|y_k - f_0(t_k, x_k) - \sum_{\ell=1}^L w_\ell f_\ell(t_k, x_k)\|_\infty
\]

The \( L_\infty \) norm constraint in (25) can be broken up into \( n \) absolute value constraints in the states \( x_i \),

\[
|y_{ik} - f_{i0}(t_k, x_k) - \sum_{\ell=1}^L w_\ell f_{i\ell}(t_k, x_k)| \leq \epsilon.
\]

The absolute value constraint can be decomposed into its positive and negative sides:

\[
\begin{align*}
y_{ik} - f_{i0}(t_k, x_k) - \sum_{\ell=1}^L w_\ell f_{i\ell}(t_k, x_k) & \leq \epsilon \\
y_{ik} - f_{i0}(t_k, x_k) - \sum_{\ell=1}^L w_\ell f_{i\ell}(t_k, x_k) & \geq -\epsilon.
\end{align*}
\]

With further reformulation by taking the \( w_\ell \) terms to the left side, the constraints are,

\[
\begin{align*}
-\sum_{\ell=1}^L w_\ell f_{i\ell}(t_k, x_k) & \leq \epsilon - y_{ik} + f_{i0}(t_k, x_k) \\
\sum_{\ell=1}^L w_\ell f_{i\ell}(t_k, x_k) & \leq \epsilon + y_{ik} - f_{i0}(t_k, x_k).
\end{align*}
\]

New terms that are constant in \( w_\ell \) can be defined,

\[
\begin{align*}
z_{ik\ell} &= [-f_{i\ell}(t_k, x_k); f_{i\ell}(t_k, x_k)] \\
d_{ik} &= [\epsilon - y_{ik} + f_{i0}(t_k, x_k) ; \epsilon + y_{ik} - f_{i0}(t_k, x_k)].
\end{align*}
\]

The polytope \( W \) for use in the peak estimation program (1) is the set, (refined from Equation (28)),

\[
W = \left\{ w \in \mathbb{R}^L \mid \sum_{\ell=1}^L z_{ik\ell} w_\ell \leq d_{ik} \right\}.
\]

**Remark 3.** Assumption A4 is satisfied when each corrupting noise term \( \eta_k \) is \( L_\infty \)-norm bounded and sufficiently many samples \( (t_k, x_k, y_k) \) are taken.

Data driven peak estimation involves solving the SOS program (23) with respect to the set \( W \) in (30). The set \( W \) is described by \( m = 2nN_s \) affine constraints, and in practice, many of these representing constraints are redundant. Elimination of redundant constraints (such as through the LP method of [19]) reduces the dimensionality of the multiplier term \( \zeta(t, x) \).
6 Data-Driven Peak Examples

Code for peak estimation under polytopic uncertainty (SOS program (23)) are available in https://github.com/Jarmill/data_driven_occ. All routines were written in MATLAB 2021a, and dependencies include YALMIP [20] and MOSEK [21] (or a YALMIP-compatible SDP solver). Redundant constraints in W were identified through the LP method of [19]. Uniform sampling in the convex polytope W to generate the process w(t) was performed by the hit-and-run method [22] as implemented by [23].

6.1 SIR System

A demonstration of the polytopic uncertainty arising from the data driven setting may take place on an epidemic example. A basic compartmental epidemic model involves three states: $S$ (susceptible), $I$ (infected), and $R$ (removed). The population is assumed to be normalized such that $S + I + R = 1$. Temporal dynamics of the Susceptible, Infected, Removed (SIR) system with parameters $(\beta, \gamma)$ are,

$$
S' = -\beta SI \\
I' = \beta SI - \gamma I
$$

The $R$ trajectory with the dynamics $R' = \gamma I$ may be recovered by the relation $R = 1 - S - I$.

Figure 3a plots 100 true and $\epsilon = 0.1$-corrupted observations of the SIR system with a ground truth of $\beta = 0.4$, $\gamma = 0.1$. Each data-record enforces 4 constraints (positive and negative sides for $S$ and $I$), and there are 400 affine constraints in total. The 5-sided polytope $\Theta_D$ is plotted in Figure 3b along with its Chebyshev center in the asterisk at $(\beta_{cheb}, \gamma_{cheb}) = (0.0977, 0.4003)$ (the Chebyshev center of a polytope is the center of the inscribed sphere with maximum radius).

Only 5 out of the 400 constraints describing $\Theta_D$ in this SIR example are non-redundant. The active constraints in Figure 4 are dotted black lines, all other inactive constraints are the gray dotted lines. The polytope $\Theta_D$ observed in Figure 3b is bordered by solid black lines.
Peak estimation to bound the maximum value of the infected population $I$ is performed on the system with observations in Fig. 3a. The $L = 2$ uncertain parameters are $(\beta, \gamma)$, and the consistency set aligning with the observed data is the 5-sided polytope in Fig. 3. The peak estimate over a time horizon of $T = 40$ days is $I_{max} \leq 0.511$ at an order-$3$ SOS tightening in (23).

6.2 Flow System

This subsection will consider the case where $\dot{x}_2$ in the Flow system (8) is modeled by a cubic polynomial in all possible monomials of $(x_1, x_2)$,

$$\dot{x} = [x_2, \text{cubic}(x_1, x_2)].$$

The polynomial model cubic$(x_1, x_2)$ has $L = \binom{2+3}{3} = 10$ free parameters, and the $x_1$ dynamics of $\dot{x}_1 = x_2$ are perfectly known. The parameters of Table 1 originate from this Flow peak estimation problem. The $N = 40$ observed data points sampled from the circle $\{x \mid (x_1 - 1.5)^2 + x_2^2 \leq 0.4^2\}$ are shown in Figure 2, yielding $2N = 80$ affine constraints. This $L = 10$-dimensional consistency polytope $\Theta_D = W$ has has 33 faces and 7534 vertices (80 – 33 = 47 redundant affine constraints) with an $\epsilon$ of 0.5 in the coordinate $x_2$. A time horizon of $T = 5$ and a valid region of $X = \{x \mid \|x\|_2^2 \leq 8\}$ is considered in these Flow examples.

Trajectories of the flow system start at the point $X_0 = (1.5, 0)$ in Figure 6a. The first 4 SOS bounds of Program (23) to maximize $p(x) = -x_2$ starting at the point $X_0$ are $d_{1:4}^* = [2.828, 2.448, 1.018, 0.8407]$. 

![Figure 4: Active and inactive constraints describing $\Theta_D$ in Figure 3b](image1)

![Figure 5: Maximum value of $I(t)$ over a time horizon of $T = 40$, with unknown $(\beta, \gamma)$](image2)
Trajectories in Figure (6b) begin in the disk with radius 0.4 and center (1.5, 0). The first four SOS bounds starting from this circular $X_0$ are $d_{1:4}^* = [2.828, 2.557, 1.245, 0.894]$.

![Order 4 bound = 0.841](image1)

(a) $X_0 = (1.5, 0)$

![Order 4 bound = 0.894](image2)

(b) $X_0 = \{x \mid (x_1 - 1.5)^2 + x_2 \leq 0.4^2\}$

Figure 6: Minimizing $x_2$ on Flow system at order-4 SOS tightening

6.3 Twist System

This subsection involves peak estimation applied to the Twist system in [24],

$$\dot{x}_i(t) = \sum_j B_{ij}^1 x_j(t) - B_{ij}^3 (4x_j^3(t) - 3x_j(t))/2,$$

(33)

$$B^1 = \begin{bmatrix} -1 & 1 & 1 \\ -1 & 0 & -1 \\ 0 & 1 & -2 \end{bmatrix} \quad B^3 = \begin{bmatrix} -1 & 0 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$  

(34)

The twist system is parameterized by $3 \times 3$ matrices ($B^1, B^3$) in (34). The matrix $B^1$ introduces linear terms and the matrix $B^3$ multiplies against cubic polynomials. Figure 7 plots 100 observations of the Twist system with a noise bound of $\epsilon = 0.5$. The 100 observations induce 600 affine constraints in parameters.

![100 Noisy Observations with $\epsilon=0.5$](image3)

Figure 7: 100 observations of Twist system
The peak estimation task on this twist system is to find the maximum value of \( p(x) = x_3 \) in a time horizon of \( T = 8 \), starting from the initial point \( X_0 = [-1, 0, 0] \). The set considered is \( X = \{ x \mid -1 \leq x_1, x_2 \leq 1, 0 \leq x_3 \leq 1 \} \).

Figure 8 plots trajectories where either \( B^1 \) or \( B^3 \) are unknown. Each case has \( L = 9 \) uncertain parameters. The polytope \( W \) for unknown \( B^1 \) in Figure 8a has 30 faces, and the peak bounds are \( d^{1:3} = [1.000, 1.000, 0.8189] \). The uncertainty polytope \( W \) for unknown \( B^3 \) in Figure 8b has 30 faces, with peak bounds of \( d^{1:3} = [1.000, 0.9050, 0.8174] \). Note how trajectories spiral closer to the origin in Fig. 8b as compared to trajectories in Fig. 8a. The maximal Gram matrix size of each Lie constraint at degree-3 is 2380 pre-decomposition and 70 post-decomposition.

Figure 8: Twist (33) system where either \( B^1 \) or \( B^3 \) are unknown.

Both parameters \((B^1, B^3)\) are unknown in Figure 9. This uncertainty polytope in \( L = 18 \) parameters has 70 nonredundant faces. The SOS bounds of maximizing \( x_3 \) in this system are \( d^{1:2} = [1.000, 0.9703] \). The maximal Gram matrix size for the degree-2 Lie constraint is 2300 pre-decomposition and is 35 post-decomposition. The MATLAB/YALMIP process became unresponsive on the experimental platform when attempting to compile the degree-3 SOS approximation program on the unknown \((B^1, B^3)\) twist system.

Figure 9: Twist (33) with unknown \((B^1, B^3)\)
7 Reachability Set Estimation

The data-driven techniques can be applied to reachability set estimation. The reachability set \( X_T \) is the set of all \( x \) that can be reached at time index \( t = T \) for trajectories starting in the set \( X_0 \),

\[
X_T = \{ x(T \mid x_0) \mid x_0 \in X_0, \ x'(t) = f(t, x) \}. \tag{35}
\]

7.1 Deterministic Problem

The methods in [6] propose the following volume maximization problem to find the reachable set \( X_T \):

\[
P^* = \sup_{X_T \subset X} \text{vol}(X_T) \tag{36a}
\]

\[
\forall \tilde{x} \in X_T, \exists x_0 \in X_0, \ w(t) \in W \mid \tilde{x} = x(T \mid x_0, w(t)) \tag{36b}
\]

\[
x'(t) = f(t, x) \quad \forall t \in [0, T] \tag{36c}
\]

An infinite-dimensional linear program in continuous functions \( v(t, x) \) and \( w(x) \) may be developed to outer-approximate the reachable set \( X_T \) [6],

\[
d^* = \inf \int_X \phi(x)dx \tag{37a}
\]

\[
v(0, x) \leq 0 \quad \forall x \in X_0 \tag{37b}
\]

\[
\phi(x) + v(T, x) \geq 1 \quad \forall x \in X \tag{37c}
\]

\[
\mathcal{L} f v(t, x) \leq 0 \quad \forall (t, x, w) \in [0, T] \times X \times W \tag{37d}
\]

\[
v(t, x) \in C^1([0, T] \times X) \tag{37e}
\]

\[
\phi(x) \in C^1([0, T] \times X) \tag{37f}
\]

At a degree-\( d \) LMI relaxation, the set \( \{ x \in X \mid \phi(x) \geq 1 \} \) is an outer approximation to the reachable set with volume bounds yielding the bounds \( d^*_d \geq d^{d+1}_d \geq P^* = \text{vol}(X_T) \). This sublevel set will converge in volume to the region of attraction’s (excluding sets of measure 0) as \( d \to \infty \). The level set approximations will be valid except for possibly a set with Lebesgue measure zero. Inner approximations to the region of attraction can be performed through the methods in [7].

7.2 Robust Function Program

The Lie derivative constraint in (37d) may be decomposed through application of the convex Theorem of Alternatives. The robust peak-estimation problem is,

\[
d^* = \min \int_X \phi(x)dx \tag{38a}
\]

\[
v(0, x) \leq 0 \quad \forall x \in X_0 \tag{38b}
\]

\[
\mathcal{L} f_0 v(t, x) + b^T \zeta(t, x) \leq 0 \quad \forall (t, x) \in [0, T] \times X \tag{38c}
\]

\[
(A^T) \ell \zeta(t, x) = (f_x \cdot \nabla_x) v(t, x) \quad \forall \ell = 1, \ldots, L \tag{38d}
\]

\[
\phi(x) + v(T, x) + \geq 1 \quad \forall x \in X \tag{38e}
\]

\[
v(t, x) \in C^1([0, T] \times X) \tag{38f}
\]

\[
\phi(x) \in C^1(X) \tag{38g}
\]

\[
\zeta_j(t, x) \in C^1([0, T] \times X) \quad \forall j = 1, \ldots, m \tag{38h}
\]

Constraints (38c) and (38d) replace the Lie derivative constraint (37d). The set \( \{ x \mid \phi(x) \geq 1 \} \) remains an outer approximation for the robust reachable set. Arguments in Appendices [A] and [B] may be modified to prove that continuous functions \( \zeta \) (and polynomial approximations thereof) may be chosen to solve Program (38).
7.3 Reachable Set Example

Figure 10 illustrates reachable set estimation on the Twist system from (33) for a time horizon of $T = 8$ by SOS relaxations to program (38). The 100 observations from this system are pictured in Figure 7, yielding a $L = 9$-dimensional polytope with 34 non-redundant faces. As the order of relaxation to program (38) increases from 3 to 4, the red region (level set of $\omega(x)$) tightens to the spiraling attractor region of the $T = 8$ reachable set.

![Order 3 volume = 1.18](image1)

(a) Order 3 relaxation

![Order 4 volume = 0.756](image2)

(b) Order 4 relaxation

Figure 10: Reachable set estimation of twist (33) system where $B^1$ is known and $B^3$ are unknown.

8 Conclusion

This paper formulates and applies a polytopic facial decomposition in order to simplify the computational complexity of peak and reachable set estimation approximation programs. The Lie derivative constraint $L_f(t, x, w) \leq 0$ may be decomposed in a peak estimation setting by applying a theorem of alternatives with respect to a facial description of the polytope $W$. The peak estimation problem may be approximated by a tractable sequence of semidefinite programs arising from SOS relaxations. The data-driven setting with $L_\infty$ bounded noise is a specific instance in which facial decompositions may be applied. Other applications of the polytopic face-decomposition for continuous-time input-affine dynamical systems include optimal control [10], distance-to-unsafe-set estimation [24], and maximum controlled invariant set estimation [25].

Future work includes exploiting network and sparsity structures, finding an alternatives-based decomposed formulation that will work for discrete-time data-driven trajectory analysis, and developing a streaming algorithm with warm starts that will allow for iterative refinement of peak estimates after acquisition of new data.

Acknowledgements

The authors would like to thank Didier Henrion and Milan Korda for their advice and support.

References

[1] R. Lewis and R. Vinter, “Relaxation of Optimal Control Problems to Equivalent Convex Programs,” *Journal of Mathematical Analysis and Applications*, vol. 74, no. 2, pp. 475–493, 1980.

[2] D. Henrion, J. B. Lasserre, and C. Savorgnan, “Nonlinear optimal control synthesis via occupation measures,” in *2008 47th IEEE Conference on Decision and Control*. IEEE, 2008, pp. 4749–4754.
[3] M. J. Cho and R. H. Stockbridge, “Linear Programming Formulation for Optimal Stopping Problems,” *SIAM Journal on Control and Optimization*, vol. 40, no. 6, pp. 1965–1982, 2002.

[4] G. Fantuzzi and D. Goluskin, “Bounding Extreme Events in Nonlinear Dynamics Using Convex Optimization,” *SIAM Journal on Applied Dynamical Systems*, vol. 19, no. 3, pp. 1823–1864, 2020.

[5] J. Miller, D. Henrion, M. Sznaier, and M. Korda, “Peak Estimation for Uncertain and Switched Systems,” *arXiv preprint arXiv:2103.13017*, 2021.

[6] D. Henrion and M. Korda, “Convex computation of the region of attraction of polynomial control systems,” *IEEE Transactions on Automatic Control*, vol. 59, no. 2, p. 297–312, Feb 2014. [Online]. Available: [http://dx.doi.org/10.1109/TAC.2013.2283095](http://dx.doi.org/10.1109/TAC.2013.2283095)

[7] M. Korda, D. Henrion, and C. N. Jones, “Inner approximations of the region of attraction for polynomial dynamical systems,” *IFAC Proceedings Volumes*, vol. 46, no. 23, pp. 534–539, 2013.

[8] A. Ben-Tal, D. Den Hertog, and J.-P. Vial, “Deriving robust counterparts of nonlinear uncertain inequalities,” *Mathematical programming*, vol. 149, no. 1-2, pp. 265–299, 2015.

[9] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex Optimization*. Cambridge university press, 2004.

[10] A. Majumdar, R. Vasudevan, M. M. Tobenkin, and R. Tedrake, “Convex Optimization of Nonlinear Feedback Controllers via Occupation Measures,” *The International Journal of Robotics Research*, vol. 33, no. 9, pp. 1209–1230, 2014.

[11] M. Korda, D. Henrion, and C. N. Jones, “Controller design and value function approximation for nonlinear dynamical systems,” 2015.

[12] G. Blekherman, “There are significantly more nonegative polynomials than sums of squares,” *Israel Journal of Mathematics*, vol. 153, no. 1, pp. 355–380, 2006.

[13] M. Putinar, “Positive Polynomials on Compact Semi-algebraic Sets,” *Indiana University Mathematics Journal*, vol. 42, no. 3, pp. 969–984, 1993.

[14] J. B. Lasserre, *Moments, Positive Polynomials And Their Applications*, ser. Imperial College Press Optimization Series. World Scientific Publishing Company, 2009.

[15] F. Alizadeh, “Interior Point Methods in Semidefinite Programming with Applications to Combinatorial Optimization,” *SIAM J. Optim.*, vol. 5, no. 1, pp. 13–51, 1995.

[16] S. Prajna and A. Jadbabaie, “Safety Verification of Hybrid Systems UsingBarrier Certificates,” in *International Workshop on Hybrid Systems: Computation and Control*. Springer, 2004, pp. 477–492.

[17] Y. Cheng, M. Sznaier, and C. Lagoa, “Robust Superstabilizing Controller Design from Open-Loop Experimental Input/Output Data,” *IFAC-PapersOnLine*, vol. 48, no. 28, pp. 1337–1342, 2015.

[18] T. Dai and M. Sznaier, “A Moments Based Approach to Designing MIMO Data Driven Controllers for Switched Systems,” in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 5652–5657.

[19] R. Caron, J. McDonald, and C. Ponic, “A degenerate extreme point strategy for the classification of linear constraints as redundant or necessary,” *Journal of Optimization Theory and Applications*, vol. 62, no. 2, pp. 225–237, 1989.

[20] J. Löfberg, “YALMIP : A Toolbox for Modeling and Optimization in MATLAB,” in *In Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[21] M. ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 9.2.*, 2020. [Online]. Available: [https://docs.mosek.com/9.2/toolbox/index.html](https://docs.mosek.com/9.2/toolbox/index.html)
A Continuity of Multipliers

Let \((v, \tilde{\zeta})\) be a certificate proving that \(L_f v \leq 0\) by Equation (19), where \(v\) is \(C^1\)-continuous and the non-negative multipliers \(\tilde{\zeta}_j \forall j \in 1, \ldots, m\) may be discontinuous over the space \([0, T] \times \mathbb{X}\). This appendix will prove that there exists a set of multipliers \(\zeta\) with \(\zeta_j \in C_+(\mathbb{R}^+ \times \mathbb{X}^+) \forall j \in 1, \ldots, m\) such that \((v, \zeta)\) is a nonpositivity certificate for the same function \(v(t, x)\) under assumptions A1-A5.

Let \((t, x)\) be a generic point in \([0, T] \times \mathbb{X}\). Further define the function \(q(t, x) : [0, T] \times \mathbb{X} \to \mathbb{R}^L\) as,

\[
q_\ell(t, x) = f_\ell(t, x) \cdot \nabla_x v(t, x) \quad \forall \ell \in 1, \ldots, L.
\]  

The set of plausible values of \(\zeta(t, x)\) certifying that \(L_f v \leq 0\) for fixed \((t, x, v)\) may be described as the set \(S\) (formulated from (19)),

\[
S = \{ z \in \mathbb{R}^m \mid z \geq 0, b^T z \leq -L_f_0 v(t, x), A^T z = q(t, x) \}. 
\]  

The set \(S\) is the solution space of a set of linear equality and inequality constraints in \(z\),

\[
S = \left\{ z \in \mathbb{R}^m \mid \begin{bmatrix} -I \\ b^T \end{bmatrix} z \leq \begin{bmatrix} 0 \\ -L_f_0 v(t, x) \end{bmatrix}, \quad A^T z = q(t, x) \right\}. 
\]  

Since \(v(t, x)\) and \(q(t, x)\) are continuous, from Theorem 2.2 in [26] it follows that the set valued mapping \(\Xi : (t, x) \to S\) is lower semicontinuous. Since \(\Xi\) has closed and convex images, it follows from Michael's Theorem (Theorem 9.1.2 in [27]) that \(\Xi\) has a continuous selection \(\zeta(x, t)\). Since from assumption A4 the set \(S\) is compact, the following minimal selection for \(\zeta\) is continuous:

\[
m(\Xi(t, x)) \doteq \{ \zeta(x, t) \in \Xi(t, x) : \|\zeta(x, t)\| = \min_{y \in \Xi(t, x)} \|y\| \}
\]  

(Prop. 9.3.2 in [27]).

B Polynomial Approximation

This appendix uses arguments from [4] to prove that Problem (20) may be approximated with \(\epsilon\)-accuracy by polynomial auxiliary and multiplier functions. It is assumed that the set \(\Omega = [0, T] \times \mathbb{X} \times \mathbb{W}\) is compact in order to invoke the Stone-Weierstrass theorem.
B.1 Polynomial Auxiliary

Let $\epsilon > 0$ be an optimality bound, and let $v(t, x) \in C^1([0, T] \times X)$ be an auxiliary function that satisfies constraints (9c) and (9d) with,

$$\sup_{x \in X_0} v(0, x) \leq p^* + \epsilon. \quad (42)$$

Define the $C^0$-norm over $\Omega$ as,

$$\|\phi\|_{C^0(\Omega)} = \sup_{(t, x, w) \in \Omega} |\phi(t, x, w)| \quad \forall \phi \in C^0(\Omega). \quad (43)$$

The $i$th coordinate of dynamics $\dot{x} = F(t, x, w) = f_0(t, x) + \sum_{\omega=1}^{L} \omega_t f_x(t, x)$ from (2) is indexed by $F_i(t, x, w)$.

A tolerance $\delta > 0$ may be chosen as (Equation 4.10 of [4]),

$$\delta < \frac{\epsilon}{\max \{2, 2T, 2T\|F_1\|_{C^0(\Omega)}, \ldots, \|F_n\|_{C^0(\Omega)}\}}. \quad (44)$$

A Stone-Weierstrass approximation may be performed to find a polynomial $w \in \mathbb{R}[t, x]$ such that $\sup_{t,x}|w(t, x) - v(t, x)| < \delta$ uniformly. The perturbed auxiliary function,

$$V(t, x) = w(t, x) + \epsilon(1 - t/(2T)), \quad (45)$$

satisfies the following strict inequalities from [9] (equation 4.12 in [4]),

$$p^* + (5/2)\epsilon > V(0, x) \quad \forall x \in X_0 \quad (46a)$$

$$L_{F(t,x,w)}V(t,x) < 0 \quad \forall (t, x, w) \in \Omega \quad (46b)$$

$$V(t, x) < p(x) \quad \forall (t, x) \in [0, T] \times X \quad (46c)$$

There exists some finite $d$ such that the polynomial $V(t, x)$ with an optimal solution of at most $p^* + (5/2)\epsilon$ has degree $d$ [4].

B.2 Polynomial Multipliers

The polynomial function $L_F V(t, x) \in \mathbb{R}[t, x, w]$ is negative in constraint (46b). It therefore possesses a certificate of nonpositivity by Equation (19) with multiplier functions $(\zeta_j(t, x))_{j=1}^{m}$. These multipliers may be chosen to be continuous, as proved in the above Appendix A.

For reference, define $Q$ as the polynomial vector with $Q(t, x) = f(t, x) \cdot \nabla_x V(t, x)$. The matrix $A$ has full column rank given that $W$ is compact. Assume that the left kernel $N(A^T)$ has dimension $K$, and the space $N(A^T)$ may be spanned by $K$ unit-norm linearly independent vectors $\{\nu_k\}_{k=1}^K$ stored in the matrix $\nu$. The set of solutions $\zeta^*$ to the linear equality constraint $(A^T)_{i} = q(t, x)$ may be parameterized by continuous functions $\phi_k \in C^0([0, T] \times X)$ with,

$$\zeta^*(t, x) = A(A^T A)^{-1} Q(t, x) + \nu \phi(t, x). \quad (47)$$

The polynomial function $A(A^T A)^{-1} Q(t, x)$ is the projection of $Q$ onto the row space of $A$. Let $\{h_k(t, x)\}_{k=1}^K$ be polynomials such that $\sup_{t,x}|h_k(t, x) - \phi_k(t, x)| \leq \tau_1$ for some $\tau_1 > 0$. A parameter $\tau_2 > 0$ may be chosen to form the polynomial multiplier approximant,

$$\zeta^p(t, x) = A(A^T A)^{-1} Q(t, x) + \nu h(t, x) + \tau_2 1. \quad (48)$$

The continuous multiplier vector and its approximant are related by,

$$\begin{align*}
\zeta^p(t, x) &= \zeta^*(t, x) + \nu (h(t, x) - \phi(t, x)) + \tau_2 1 \\
&\leq \zeta^*(t, x) - \|\nu\|_\infty \tau_1 + \tau_2 1. \quad (49a) \\
&\leq \zeta^*(t, x) - \|\nu\|_\infty \tau_1 + \tau_2 1. \quad (49b)
\end{align*}$$
Each coordinate in $\zeta^p$ will be in $[0, T] \times X$ positive if $\|\nu\|_\infty \tau_1 < \tau_2$ given that $\zeta^*_t(t, x) \geq 0$ in $[0, T] \times X$.

The left hand side of constraint (19b)'s polynomial approximant certifying nonpositivity of the Lie term with $V(t, x)$ may be posed as,

\[
\mathcal{L}_{f_0} V(t, x) + b^T \zeta^p(t, x) = \mathcal{L}_{f_0} V(t, x) + b^T (\zeta^*_t(t, x) + \nu(h(t, x) - \phi(t, x)) + 1\tau_2) \leq \mathcal{L}_{f_0} V(t, x) + b^T \zeta^*_t(t, x) + \|\text{diag}(b)^T \nu\|_\infty \tau_1 + b^T \mathbf{1} \tau_2
\]

Letting $\Phi = \sup_{(t, x) \in [0, T] \times X} \mathcal{L}_{f_0} V(t, x) + b^T \zeta^*_t(t, x)$, the constraint on parameters $(\tau_1, \tau_2)$ from the Lie term is that,

\[
\Phi + \|\text{diag}(b)^T \nu\|_\infty \tau_1 + b^T \mathbf{1} \tau_2 \leq 0.
\]

Choosing a pair $(\tau_1, \tau_2) > 0$ that satisfies (51) and $\tau_2 > \|\nu\|_\infty \tau_1$ is therefore sufficient for admissible polynomial multipliers $\zeta^p$ to exist. The functions $(V(t, x), \zeta^p(t, x))$ are all polynomial with a peak estimate of $p^* + (5/2)\epsilon$, and this optimality gap will shrink as $\epsilon$ approaches arbitrarily close to zero. The polynomials $(V(t, x), \zeta^p(t, x))$ are bounded degree for each $\epsilon > 0$, so there exists some degree $d$ in the SOS relaxation to (23) such that $(V(t, x), \zeta^p(t, x))$ will be feasible solutions.