Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 55-25 (2022) 55–60

Facial Input Decompositions for Robust Peak Estimation under Polyhedral Uncertainty

Jared Miller * **Mario Sznaier** *

* *ECE Department, Northeastern University, Boston, MA 02115. (Emails: miller.jare@northeastern.edu, msznaier@coe.neu.edu).*

Abstract: This work bounds extreme values of state functions 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, 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 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 peak value under mild assumptions (convergence and smoothness of dynamics).

Keywords: Robust Optimization in Control, Uncertain Systems, Sum-of-Squares, Computational Methods in Control, Safety Certificates

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,

$$P^* = \max_{t \in [0,T],\, x_0 \in X_0,\, w(t)} p(x(t \mid x_0, w(t))) \tag{1}$$
$$\dot{x}(t) = f(t, x(t), w(t)),\ w(t) \in W \quad \forall t \in [0, T].$$

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

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

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

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

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 (LPs) in occupation measures were developed for optimal control in (Lewis and Vinter, 1980). 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 (Henrion et al., 2008).

Infinite-dimensional LPs in measures for peak estimation were formulated in (Cho and Stockbridge, 2002) and were solved through a gridded discretization of the infinite-dimensional LP and through Markov Chain Martingale techniques. The work in (Fantuzzi and Goluskin, 2020) 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 (Miller et al., 2021), which includes the class of systems considered in Eq. (2). Infinite-dimensional LPs have also been applied to perform reachable set approximation (Henrion and Korda, 2013).

This paper is a sequel to (Miller et al., 2021), which formulates a peak estimation program for general uncertainties $w(t) \in W$ possibly including switching structure (vertex decomposition of W). The approach in (Miller et al., 2021) yields a convex optimization problem for polyhedral 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 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 W arises from a theorem of alternatives in robust optimization (Ben-Tal et al., 2015) and convex duality (Boyd

[1] J. Miller and M. Sznaier were partially supported by NSF grants CNS–1646121, CMMI–1638234, ECCS–1808381 and CNS–2038493, and 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.

10.1016/j.ifacol.2022.09.323

et al., 2004). Specific cases of facial decompositions when W is a unit box were discussed in (Majumdar et al., 2014; Korda et al., 2015); this paper allows for general compact and convex polytopes 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 (PSD) (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) in the degree-d SOS tightenings of the robust peak estimation programs.

Table 1. Size of largest Lie constraint Gram Matrix

Miller et al. (2021)	$\binom{1+n+L+d+\lceil \deg(f)/2\rceil-1}{1+n+L} = 8568$
This paper	$\binom{1+n+d+\max_\ell \lceil \deg(f_\ell)/2\rceil-1}{1+n} = 56$

Imposing that a symmetric matrix of size 8568 is PSD is intractable in numerical solvers such as Mosek and Sedumi. The polytope W has 7534 vertices and 33 faces. A vertex decomposition would require 7534 PSD constraints of size 56, while the equivalent facial decomposition imposes $33+1=34$ 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 $\mathcal{D} = \{(t_k, x_k, y_k)\}_{k=1}^{N_s}$ are acquired where $y(t_k) \doteq \dot{x}(t_k) + \eta_k$ with respect to dynamics (2) for L_∞ bounded error terms η_k. These errors η_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 W
- Reduction in computational complexity of SDPs
- Presentation and demonstration of convex data-driven peak 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 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 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. Section 7 presents some conclusions and briefly discusses future work. An extended version of this paper (including applications to reachable set estimation, further examples, and proofs of function continuity) is available at `https://arxiv.org/abs/2112.14838`.

2. PRELIMINARIES

2.1 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 $x, y \in \mathbb{R}^n$ is $x \cdot y = x^T y = \sum_i x_i y_i$. 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 is $\mathbb{R}[x]$. The degree of a polynomial $p \in \mathbb{R}[x]$ is $\deg(p)$. If p is a vector of polynomials such that $p_i \in \mathbb{R}[x]\ \forall i = 1, \ldots, n$, then $\deg(p) = \max_i \deg(p_i)$. The set of polynomials in degree at most d is $\mathbb{R}[x]_{\leq d}$.

The set of continuous functions over a space X is $C(X)$. Its subcone of nonnegative continuous functions over X is $C_+(X) \subset C(X)$. The set of continuous functions with continuous first derivatives is $C^1(X) \subset C(X)$.

2.2 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 nonnegative polynomials. For every $p \in \Sigma[x]$, there exists a $Q \succeq 0$ and a polynomial vector $m(x)$ such that $p(x) = m(x)^T Q m(x)$. The ratio of the volume between SOS and nonnegative polynomials for fixed degree d approaches zero as $n \to \infty$ (Blekherman, 2006).

A basic semiagebraic set is a set formed by the locus of a finite number of bounded degree inequality constraints, such as $\mathbb{K} = \{x \mid g_i(x) \geq 0\ \forall i = 1, \ldots, N_c\}$. The set $\mathbb{K}$ satisfies the Archimedean condition if there exists a finite R and SOS polynomials $\{\sigma_i(x)\}_{i=0}^{N_c}$ such that,

$$R^2 - \|x\|_2^2 = \sigma_0(x) + \textstyle\sum_{i=1}^{N_c} \sigma_i(x) g_i(x). \tag{4}$$

Every compact basic semialgebraic set $\mathbb{K}$ may be made Archimedean by appending the redundant ball constraint $R^2 - \|x\|_2^2 \geq 0$ to the description of $\mathbb{K}$.

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

$$P^* = \max_{\gamma\in\mathbb{R}} \ \gamma, \quad c(x) - \gamma \geq 0 \qquad \forall x \in \mathbb{K} \tag{5a}$$

M. Putinar introduced an algebraic certificate (Putinar Positivestellensatz) that is a necessary and sufficient condition for a function $p(x)$ to be positive over the Archimedean basic semialgebraic set $\mathbb{K}$ (Putinar, 1993),

$$\begin{aligned} &p(x) = \sigma_0(x) + \textstyle\sum_{i=1}^{N_c} \sigma_i(x) g_i(x) \\ &\sigma(x) \in \Sigma[x] \qquad \sigma_i(x) \in \Sigma[x] : i = 0, \ldots, N_c. \end{aligned} \tag{6}$$

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

$$d_d^* = \sup \ \gamma \tag{7a}$$

$$c(x) - \gamma = \sigma_0(x) + \textstyle\sum_{i=1}^{N_c} \sigma_i(x) g_i(x) \tag{7b}$$

$$\sigma_i(x) \in \Sigma[x]_{2d} \quad \forall i = 0, \ldots, N_c \tag{7c}$$

Problem (7) is a Semidefinite Program (SDP) in terms of elements of the finite-dimensional Gram matrices $Q_0 \dots Q_{N_c}$ describing the Putinar multipliers $\sigma_0 \dots \sigma_{N_c}$. The objective values between programs (5a) and (7) are related by $d_d^* \leq P^*$, and the chain of lower bounds obeys $d_d^* \leq d_{d+1}^* \leq d_{d+2}^* \dots$ as the degree of the SOS tightening increases (Lasserre, 2009). When $\mathbb{K}$ is Archimedean, the objective value d_d^* will converge to P^* as the relaxation degree d approaches ∞. 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 $N = \binom{n+d}{d} \approx n^d$. The per-iteration complexity of an Interior Point Method for solving an SDP up to ϵ-accuracy with an N-dimensional SDP matrix and M affine constraints is $O(N^3M + N^2M^2 + M^3)$ (Alizadeh, 1995). This runtime is polynomial in n for fixed d, and is combinatorial as d increases in the SOS hierarchy.

2.3 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 $\ell = 1, \dots, L$ 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 (2) does not have finite escape time. Further, if the function $p(x)$ is continuous, this assumption implies that P^* is bounded above

Figure 1 illustrates an example of peak estimation under uncertainty. The system dynamics are a modification of the Flow system from (Prajna and Jadbabaie, 2004) with a new time-varying uncertainty term $w(t) \in [-0.5, 0.5]$,

$$\dot{x}(t) = [x_2(t);\ -x_1(t) - x_2(t) + (1 + w(t))x_1^3(t)/3]. \quad (8)$$

Sample trajectories in Figure 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.

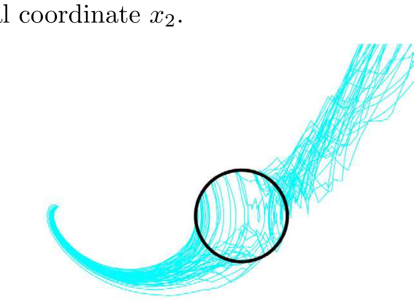


Fig. 1. Plot of Uncertain Flow system (8) trajectories

The Lie derivative of a test function $v(t, x) \in C^1([0, T] \times X)$ along dynamics $\dot{f}(t, x, w)$ is,

$$\mathcal{L}_f v(t, x) = \partial_t v(t, x) + f(t, x, w) \cdot \nabla_x v(t, x). \quad (9)$$

The peak estimation LP in terms of an auxiliary function $v(t, x) \in C^1([0, T] \times X)$ and a parameter $\gamma \in \mathbb{R}$ for a time-varying disturbance $w(t) \in W$ is (Miller et al., 2021),

$$\begin{aligned} d^* = \min_{v(t,x),\gamma} \quad & \gamma & (10a)\\ & \gamma \geq v(0, x) \qquad \forall x \in X_0 & (10b)\\ & \mathcal{L}_f v(t, x) \leq 0 \quad \forall (t, x, w) \in [0, T] \times X \times W & (10c)\\ & v(t, x) \geq p(x) \quad \forall (t, x) \in [0, T] \times X & (10d) \end{aligned}$$

The auxiliary function $v(t, x)$ must decrease along trajectories generated by all possible admissible disturbance processes $w(t)$ (Fantuzzi and Goluskin, 2020). The objectives $P^* = d^*$ between programs (1) and (10) are equal when $[0, T] \times X \times W$ is compact, $p(x)$ is bounded below, and $f(t, x, w)$ is Lipschitz (Lewis and Vinter, 1980).

The infinite-dimensional LP in (10) may be approximated through the moment-SOS hierarchy as discussed in Section 2.2. When the set X satisfies the Archimedean property, the sequence of SOS-tightenings $d_d^* \geq d_{d+1}^* \geq \dots$ will converge as $\lim_{d\to\infty} d_d^* = P^*$. The order-4 moment-SOS peak estimate to program (10) 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 (10c). 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 in (9) may be expanded into,

$$\mathcal{L}_f v(t, x) = \mathcal{L}_{f_0} v(t, x) + \textstyle\sum_{\ell=1}^L w_\ell f_\ell(t, x) \cdot \nabla_x v(t, x), \quad (11)$$

and the Lie derivative constraint in (10c) is,

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

Convex duality may be used to decompose Equation (11) 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 (Ben-Tal et al., 2015), and work applying duality to control constraints includes (Cheng et al., 2015; Dai and Sznaier, 2018).

If the Lie constraint (11) 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):

$$\begin{aligned} \text{find} \quad & (t, x, w) \in [0, T] \times X \times \mathbb{R}^L & (13a)\\ & \mathcal{L}_f v(t, x) > 0 & (13b)\\ & - Aw \geq -b & (13c) \end{aligned}$$

This program may be decomposed through convex alternatives. There is a single strict inequality is Eq. (13b) with $\mathcal{L}_f v(t, x) > 0$, and each linear constraint $-\sum_\ell A_{\ell j} w_\ell \geq -b_j$ in (13c) is a non-strict inequality. Define nonnegative (possibly discontinuous) dual variables $\zeta_j(t, x)$ for each linear constraint $j = 1, \dots, m$ in (13c). A Lagrangian $\mathscr{L}(w; \zeta, t, x)$ for problem (13) with a given auxiliary function $v(t, x)$ is:

$$\begin{aligned}\mathscr{L} &= \mathcal{L}_f v(t,x) + \zeta(t,x)^T(b - Aw) && \text{(14a)}\\ &= \mathcal{L}_{f_0} v(t,x) + b^T\zeta(t,x) && \text{(14b)}\\ &\quad + \textstyle\sum_{\ell=1}^{L} w_\ell \left(f_\ell \cdot \nabla_x v(t,x) - \sum_{j=1}^{m} A_{j\ell}\zeta_j(t,x)\right)\end{aligned}$$

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

$$\begin{cases}\mathcal{L}_{f_0} v + b^T\zeta & \forall \ell : f_\ell \cdot \nabla_x v(t,x) = \sum_{j=1}^m A_{j\ell}\zeta_j \\ \infty & \text{else}\end{cases} \tag{15}$$

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 (Boyd et al., 2004) The dual problem of finding a $\zeta(t, x)$ that infimizes $\mathscr{L}(w; \zeta, t, x)$ is,

$$\begin{aligned}\text{find} \quad & \zeta_j \geq 0 & \forall j && \text{(16a)}\\ & \mathcal{L}_{f_0} v + b^T\zeta(t,x) \leq 0 & && \text{(16b)}\\ & \textstyle\sum_{j=1}^m A_{j\ell}\zeta_j(t,x) = f_\ell(t,x)\cdot\nabla_x v(t,x) & \forall \ell && \text{(16c)}\end{aligned}$$

The nonnegative multiplier functions $\zeta_j(t, x)$ are constant in the input w_ℓ for each $\ell = 1, \ldots, L$. The pair of programs (13) and (16) 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 (11) and (16) are therefore equivalent for a given $v(t, x)$, and (16) 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 of the Arxiv paper

4. POLYHEDRAL UNCERTAINTY PEAK PROBLEM

This section will present peak estimation programs where the Lie constraint (10c) is decomposed into (16).

4.1 Polytopic Decomposition (Function Program)

The polytopic-decomposed peak program only changes the Lie derivative constraint (10c), forming the program,

$$\begin{aligned}d^* = \min_{\gamma\in\mathbb{R}} \; & \gamma && && \text{(17a)}\\ & \gamma \geq v(0,x) && \forall x \in X_0 && \text{(17b)}\\ & \mathcal{L}_{f_0} v(t,x) + b^T\zeta(t,x) \leq 0 && \forall (t,x) \in [0,T]\times X && \text{(17c)}\\ & (A^T)_\ell\zeta(t,x) = f_\ell \cdot \nabla_x v(t,x) && \forall \ell = 1,\ldots,L && \text{(17d)}\\ & v(t,x) \geq p(x) && \forall (t,x) \in [0,T]\times X && \text{(17e)}\\ & v(t,x) \in C^1([0,T]\times X) && && \text{(17f)}\\ & \zeta_j(t,x) \in C_+([0,T]\times X) && \forall j = 1,\ldots,m && \text{(17g)}\end{aligned}$$

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

Theorem 1. The objectives d^* between programs (10) and (17) are equal when A1-5 hold.

Proof. Tightness is assured by application of the Theorem of Alternatives in (16). The Lie constraint (10c) is replaced by constraints (17c)-(17d). The functions ζ may be restricted to be continuous as per Appendix A of the Arxiv version (Miller and Sznaier, 2021b).

Theorem 2. The objectives from (17) and (1) are equal when assumptions A1-5 are satisfied.

Proof. Problems (1) and (10) have equal objectives under assumptions A1-A5 by Theorem 2.1 of (Lewis and Vinter, 1980). It is therefore proven that $P^* = d^*$ from problems (17) and (1) by application of Theorem 1.

4.2 Polytopic Decomposition (SOS Program)

Problem (17) 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,

$$X_0 = \{x \mid g_k^0(x) \; \forall k = 1,\ldots,N_c^0\} \tag{18}$$
$$X = \{x \mid g_k(x) \; \forall k = 1,\ldots,N_c\}. \tag{19}$$

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 $c(x) \in \Sigma[X]_{\leq 2d}$ will denote that the degree $\leq 2d$ polynomial $c(x)$ has an appropriately degree $\leq 2d$ Putinar certificate of nonnegativity (6) over the Archimedian domain X in (19). Define the dynamics degree d' as $d' = \max_\ell\lceil\deg(f_\ell)/2\rceil + d - 1$. The degree-$d$ SOS tightening of program (17) forms the SDP,

$$\begin{aligned}d_d^* = \inf_{\gamma\in\mathbb{R}} \; & \gamma && \text{(20a)}\\ & \gamma - v(0,x) \in \Sigma[X_0]_{\leq 2d} && \text{(20b)}\\ & -\mathcal{L}_{f_0} v(t,x) - b^T\zeta(t,x) \in \Sigma[[0,T]\times X]_{\leq 2d'} && \text{(20c)}\\ & v(t,x) - p(x) \in \Sigma[[0,T]\times X]_{\leq 2d} && \text{(20d)}\\ & \forall \ell : (A^T)_\ell\zeta(t,x) = f_\ell(t,x)\cdot\nabla_x v(t,x) && \text{(20e)}\\ & v(t,x) \in \mathbb{R}[t,x]_{\leq 2d} && \text{(20f)}\\ & \forall j : \zeta_j(t) \in \Sigma[[0,T]\times X]_{\leq 2d'} && \text{(20g)}\end{aligned}$$

Constraints (20b)-(20e) are linear equality constraints involving coefficient vectors of v, ζ and elements of the Gram matrices of each multiplier σ in the Putinar expressions (6). The degree-d SOS tightening of programs (10) and (17) each involve a polynomial auxiliary function $v(t, x)$ of degree $2d$.

Theorem 3. The objective of (20) will approach (17) 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 of the Arxiv version (Miller and Sznaier, 2021b).

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 η satisfying $\|\eta\|_\infty \leq \epsilon$ is $\dot{x}_{observed} = y = F(t, x) + \eta$.

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

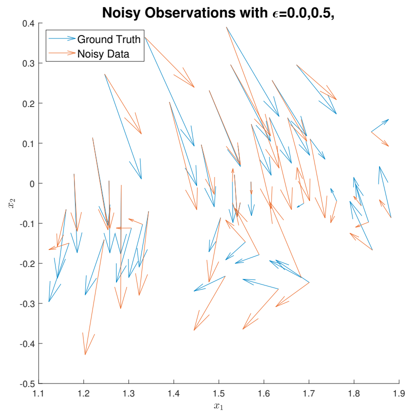


Fig. 2. Observed data of Flow system (8) within a circle

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 $\dot{x} = f(t, x; w) = f_0(t, x) + \sum_{\ell=1}^{L} w_\ell f_\ell(t, x)$ be a continuous ODE depending affinely on a set of parameters $w \in \mathbb{R}^L$. The function f_0 encodes prior knowledge about system dynamics, and the basis functions f_ℓ 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 $k = 1, \ldots, N_s$ noisy observations, which may arise from multiple time sequences and traces of the same system. The set of system parameters w that agree with the data $\mathcal{D}$ forms a set,

$$W = \{w \in \mathbb{R}^L \mid \forall k : \|y_k - f(t_k, x_k; w)\|_\infty \leq \epsilon \quad \}. \quad (21)$$

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

$$\|y_k - f(t_k, x_k; w)\|_\infty = \|y_k - f_0(t_k, x_k) - \textstyle\sum_{\ell=1}^{L} w_\ell f_\ell(t_k, x_k)\|_\infty \quad (22)$$

The L_∞ norm constraint in (22) can be broken up into n absolute value constraints in the states x_i,

$$|y_{ik} - f_{i0}(t_k, x_k) - \textstyle\sum_{\ell=1}^{L} w_\ell f_{i\ell}(t_k, x_k)| \leq \epsilon. \quad (23)$$

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

$$y_{ik} - f_{i0}(t_k, x_k) - \textstyle\sum_{\ell=1}^{L} w_\ell f_{i\ell}(t_k, x_k) \leq \epsilon \quad (24a)$$

$$y_{ik} - f_{i0}(t_k, x_k) - \textstyle\sum_{\ell=1}^{L} w_\ell f_{i\ell}(t_k, x_k) \geq -\epsilon. \quad (24b)$$

With further reformulation by taking the w_ℓ terms to the left side, the constraints are,

$$-\textstyle\sum_{\ell=1}^{L} w_\ell f_{i\ell}(t_k, x_k) \leq \epsilon - y_{ik} + f_{i0}(t_k, x_k) \quad (25a)$$

$$\textstyle\sum_{\ell=1}^{L} w_\ell f_{i\ell}(t_k, x_k) \leq \epsilon + y_{ik} - f_{i0}(t_k, x_k). \quad (25b)$$

New terms that are constant in w_ℓ can be defined,

$$z_{ik\ell} = [-f_{i\ell}(t_k, x_k); f_{i\ell}(t_k, x_k)] \quad (26a)$$

$$d_{ik} = \begin{bmatrix} \epsilon - y_{ik} + f_{i0}(t_k, x_k) \\ \epsilon + y_{ik} - f_{i0}(t_k, x_k) \end{bmatrix}. \quad (26b)$$

The polytope W for use in the peak estimation program (1) is the set, (refined from Equation (25)),

$$W = \left\{ w \in \mathbb{R}^L \middle| \textstyle\sum_{\ell=1}^{L} z_{ik\ell} w_\ell \leq d_{ik} \right\}. \quad (27)$$

Remark 2. Assumption A4 is satisfied when each corrupting noise term η_k is L_∞-norm bounded and sufficiently many samples (t_k, x_k, y_k) are taken.

Data driven peak estimation involves solving the SOS program (20) with respect to the set W in (27). 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 Caron et al. (1989)). reduces the dimensionality of the multiplier term $\zeta(t, x)$.

6. DATA-DRIVEN PEAK EXAMPLE

Code for peak estimation under polytopic uncertainty (SOS program (20)) are available in `https://github.com/Jarmill/data_driven_occ`. All routines were written in MATLAB 2021a, and dependencies include YALMIP (Löfberg, 2004) and MOSEK (ApS, 2020) (or a YALMIP-compatible SDP solver). Redundant constraints in W were identified through the LP method of Caron et al. (1989).

The $\dot{x}_2$ subsystem of the Flow dynamics (8) is modeled by a cubic polynomial in all possible monomials of (x_1, x_2),

$$\dot{x} = [x_2,\ \mathrm{cubic}(x_1, x_2)]. \quad (28)$$

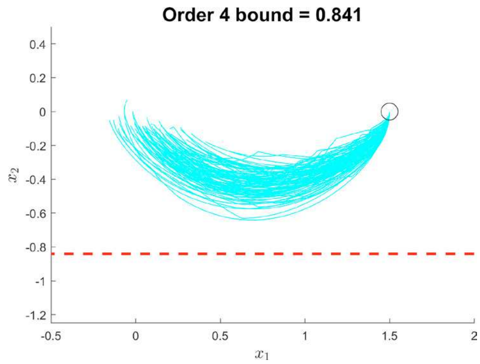
The polynomial model cubic(x_1, x_2) has $L = \binom{2+3}{3} = 10$ free parameters, and the x_1 dynamics of $\dot{x}_1 = x_2$ are perfectly known. The parameters of Table 1 originate from this Flow peak estimation problem. The $N = 40$ observed data points sampled from the circle $\{x \mid (x_1 - 1.5)^2 + x_2^2 \leq 0.4^2\}$ are shown in Figure 2, yielding $2N = 80$ affine constraints. This $L = 10$-dimensional consistency polytope $\Theta_{\mathcal{D}} = W$ has has 33 faces and 7534 vertices ($80 - 33 = 47$ redundant affine constraints) with an ϵ 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 3a. The first 4 SOS bounds of Program (20) to maximize $p(x) = -x_2$ starting at the point X_0 are $d^*_{1:4} = [2.828, 2.448, 1.018, 0.8407]$. Trajectories in Figure (3b) 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 $d^*_{1:4} = [2.828, 2.557, 1.245, 0.894]$.

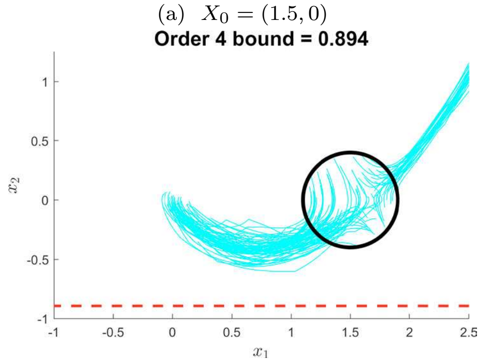
7. CONCLUSION

This paper formulates and applies a polytopic facial decomposition in order to simplify the computational complexity of peak estimation approximation programs. The Lie derivative constraint $\mathcal{L}_f(t, x, w) \leq 0$ 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_∞ 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 (Majumdar et al., 2014), distance-to-unsafe-set estimation (Miller and Sznaier, 2021a), and reachable set approximation (Henrion and Korda, 2013).

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



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



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

Fig. 3. Minimizing x_2 on Flow system (8) at degree-4 SOS tightening

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

Alizadeh, F. (1995). Interior Point Methods in Semidefinite Programming with Applications to Combinatorial Optimization. *SIAM J. Optim.*, 5(1), 13–51.

ApS, M. (2020). *The MOSEK optimization toolbox for MATLAB manual. Version 9.2.* URL `https://docs.mosek.com/9.2/toolbox/index.html`.

Ben-Tal, A., Den Hertog, D., and Vial, J.P. (2015). Deriving robust counterparts of nonlinear uncertain inequalities. *Mathematical programming*, 149(1-2), 265–299.

Blekherman, G. (2006). There are significantly more nonegative polynomials than sums of squares. *Israel Journal of Mathematics*, 153(1), 355–380.

Boyd, S., Boyd, S.P., and Vandenberghe, L. (2004). *Convex Optimization.* Cambridge university press.

Caron, R., McDonald, J., and Ponic, C. (1989). A degenerate extreme point strategy for the classification of linear constraints as redundant or necessary. *Journal of Optimization Theory and Applications*, 62(2), 225–237.

Cheng, Y., Sznaier, M., and Lagoa, C. (2015). Robust Superstabilizing Controller Design from Open-Loop Experimental Input/Output Data. *IFAC-PapersOnLine*, 48(28), 1337–1342.

Cho, M.J. and Stockbridge, R.H. (2002). Linear Programming Formulation for Optimal Stopping Problems. *SIAM Journal on Control and Optimization*, 40(6), 1965–1982.

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

Fantuzzi, G. and Goluskin, D. (2020). Bounding Extreme Events in Nonlinear Dynamics Using Convex Optimization. *SIAM Journal on Applied Dynamical Systems*, 19(3), 1823–1864.

Henrion, D. and Korda, M. (2013). Convex Computation of the Region of Attraction of Polynomial Control Systems. *IEEE TAC*, 59(2), 297–312.

Henrion, D., Lasserre, J.B., and Savorgnan, C. (2008). Nonlinear optimal control synthesis via occupation measures. In *2008 47th IEEE Conference on Decision and Control*, 4749–4754. IEEE.

Korda, M., Henrion, D., and Jones, C.N. (2015). Controller design and value function approximation for nonlinear dynamical systems.

Lasserre, J.B. (2009). *Moments, Positive Polynomials And Their Applications.* Imperial College Press Optimization Series. World Scientific Publishing Company.

Lewis, R. and Vinter, R. (1980). Relaxation of Optimal Control Problems to Equivalent Convex Programs. *Journal of Mathematical Analysis and Applications*, 74(2), 475–493.

Löfberg, J. (2004). YALMIP : A Toolbox for Modeling and Optimization in MATLAB. In *In Proceedings of the CACSD Conference*. Taipei, Taiwan.

Majumdar, A., Vasudevan, R., Tobenkin, M.M., and Tedrake, R. (2014). Convex Optimization of Nonlinear Feedback Controllers via Occupation Measures. *The International Journal of Robotics Research*, 33(9), 1209–1230.

Miller, J., Henrion, D., Sznaier, M., and Korda, M. (2021). Peak Estimation for Uncertain and Switched Systems. In *2021 60th IEEE Conference on Decision and Control (CDC)*, 3222–3228. doi:10.1109/CDC45484.2021.9683778.

Miller, J. and Sznaier, M. (2021a). Bounding the Distance to Unsafe Sets with Convex Optimization. ArXiv: 2110.14047.

Miller, J. and Sznaier, M. (2021b). Facial input decompositions for robust peak and reachable set estimation under polyhedral uncertainty. ArXiv: 2112.14838.

Prajna, S. and Jadbabaie, A. (2004). Safety Verification of Hybrid Systems Using Barrier Certificates. In *International Workshop on Hybrid Systems: Computation and Control*, 477–492. Springer.

Putinar, M. (1993). Positive Polynomials on Compact Semi-algebraic Sets. *Indiana University Mathematics Journal*, 42(3), 969–984.