Facial Input Decompositions for Robust Peak Estimation under Polyhedral Uncertainty

Jared Miller, Mario Sznaier

ROCOND, Convex Optimization (WeTS2.6), Aug 31, 2022

Northeastern University

Main Ideas

Peak estimation problems posed under polyhedral uncertainty

Use polytopic structure to simplify Lie constraints

Apply to L_∞ bounded noise setting

Peak Estimation Background

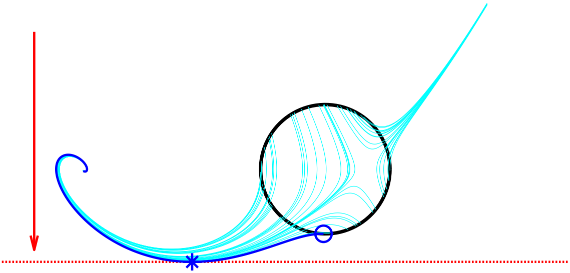
Peak Estimation Problem

Find maximum value of $p(x)$ along trajectories

$$P^* = \max_{t,\ x_0 \in X_0} p(x(t))$$

$$\dot{x}(t) = f(t, x(t)) \qquad t \in [0, T]$$

$$x(0) = x_0 \in X_0$$



$$\dot{x} = [x_2, -x_1 - x_2 + x_1^3/3]$$

Peak Function Program

Infinite dimensional linear program (Fantuzzi, Goluskin, 2020)

Uses auxiliary function $v(t, x)$

$$d^* = \min_{\gamma \in \mathbb{R}} \quad \gamma \tag{1}$$

$$\gamma \geq v(0, x) \qquad \forall x \in X_0 \tag{2}$$

$$\mathcal{L}_f v(t, x) \leq 0 \qquad \forall (t, x) \in [0, T] \times X \tag{3}$$

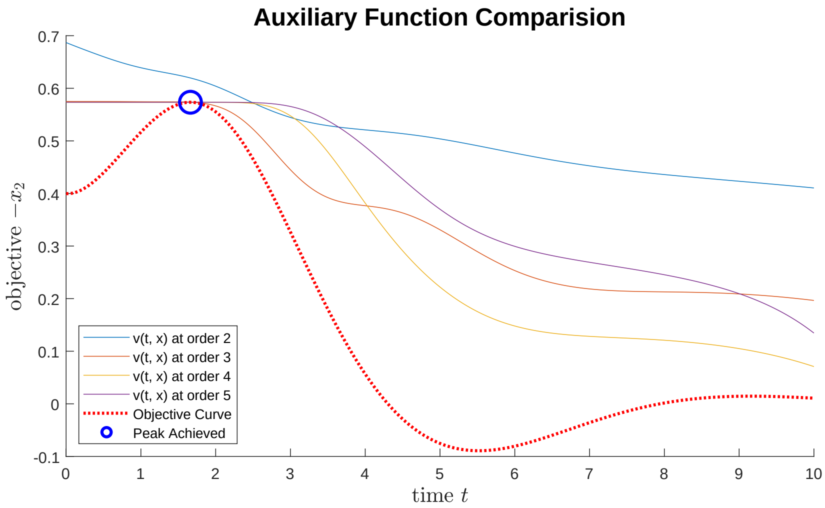
$$v(t, x) \geq p(x) \qquad \forall (t, x) \in [0, T] \times X \tag{4}$$

$$v \in C^1([0, T] \times X) \tag{5}$$

Lie Derivative $\mathcal{L}_f v(t, x) = \partial_t v + f(t, x) \cdot \nabla_x v$

$P^* = d^*$ holds if $[0, T] \times X$ is compact, f Lipschitz

Auxiliary Evaluation along Optimal Trajectory



Optimal $v(t, x)$ should be constant until peak is achieved

Peak Estimation with Uncertainty

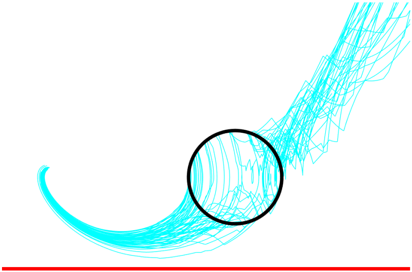
Dynamics $\dot{x} = f(t, x(t), w(t))$

Uncertain process $w(t) \in W, \ \forall t \in [0, T]$

Time-dependent w

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

System with Uncertainty Example



$$\dot{x}(t) = [x_2(t), -x_1 w(t) - x_2(t) + x_1(t)^3/3]$$

$$w(t) \in [0.5, 1.5]$$

Assumptions

Set $[0, T] \times X$ is compact

Uncertainty W is a compact polytope $\{w \mid Aw \leq b\}$

Nonempty interior: $\exists w \in \mathbb{R}^L \mid Aw < b$

Dynamics $f(t, x, w)$ are Lipschitz

Input-affine $f(t, x, w) = f_0(t, x) + \sum_{\ell=1}^{L} w_\ell f_\ell(t, x)$

Constraint Decomposition

Feasibility Pair

Polyhedral-constrained input $w(t) \in W$

Problem is Feasible

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

Problem is Infeasible

$$\mathcal{L}_{f(t,x,w)} v(t,x) > 0 \qquad \exists (t,x,w) \in [0,T] \times X \times W$$

Pair of Strong Alternatives

Theorem of Alternatives

Lagrange dual function g

$$g(\zeta) = \sup_{w \in \mathbb{R}^L} \ S(w;\zeta) = \sup_{w \in \mathbb{R}^L} \mathcal{L}_f v + \sum_{k=1}^{m} \zeta_k(t,x)(b_k - A_k w)$$

Certificate ζ that $\mathcal{L}_f v > 0$ is empty:

$$\forall (t,x) \in [0,T] \times X : \qquad g(\zeta) \leq 0, \quad \zeta(t,x) \geq 0$$

Weak alternative $g(\zeta) \leq 0$ also strong:

- w-affine functions $\mathcal{L}_f v$ and $b - Aw$ are w-concave
- Exists a point $w : Aw < b$ (Slater)

Lie Polytopic Decomposition

Original

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

Decomposed

$$\begin{aligned} &\mathcal{L}_{f_0} v(t,x) + b^T \zeta(t,x) \leq 0 && \forall (t,x) \in [0,T] \times X \\ &(A^T)_\ell \zeta(t,x) = f_\ell \cdot \nabla_x v(t,x) && \forall \ell = 1,\ldots,L \\ &\zeta_k(t,x) \in C_+([0,T] \times X) && \forall k = 1,\ldots,m \end{aligned}$$

Strong equivalence (given affine structure in w)

Peak Estimation (revisited)

Peak Original Program

Include time-varying uncertainty $w(t) \in W$

$$
\begin{aligned}
d^* = \min_{\gamma \in \mathbb{R}} \quad & \gamma \\
& \gamma \geq v(0, x) && \forall x \in X_0 \\
& \mathcal{L}_f v(t, x) \leq 0 && \forall (t, x, w) \in [0, T] \times X \times W \\
& v(t, x) \geq p(x) && \forall (t, x) \in [0, T] \times X \\
& v \in C^1([0, T] \times X)
\end{aligned}
$$

Peak Decomposed Program

Only the Lie Derivative constraint changes

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

Data Driven Setting

Noise Model

Ground truth $\dot{x} = F(t, x)$

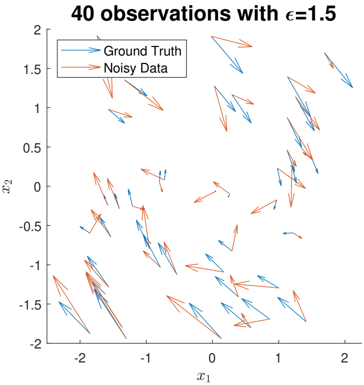
Corrupted observations of system F in $t \in [0, T]$

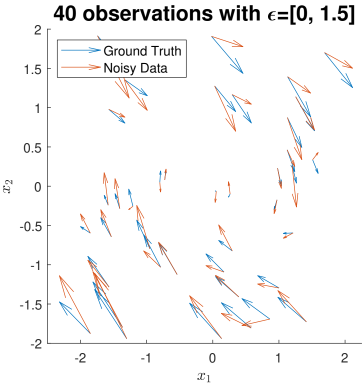
$$(t_j, x_j, \dot{x}_j) \qquad \forall j = 1, \ldots, N_s$$

Assumption of L_∞ bounded noise

$$\|F(t_j, x_j) - \dot{x}_j\|_\infty \leq \epsilon \qquad \forall j = 1, \ldots, N_s$$

Sampling: Flow System





$$\dot{x} = [x_2, -x_1 - x_2 + x_1^3/3]$$

Dynamics Model

Parameterize unknown F by functions in dictionary

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

Affine in uncertainties w

Bounded noise constraint ϵ

$$\begin{aligned} \|F(t_j, x_j) - \dot{x}_j\|_\infty &\leq \epsilon & \forall j = 1, \ldots, N_s \\ \|f(t_j, x_j, w) - \dot{x}_j\|_\infty &\leq \epsilon & \forall j = 1, \ldots, N_s \end{aligned}$$

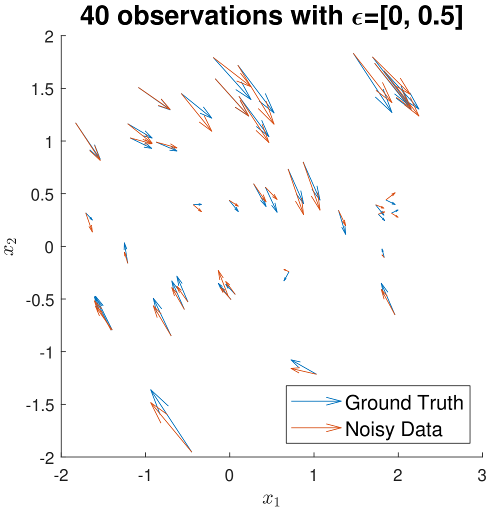
Noise Constraints

2 linear constraints for each coordinate i, sample j

$$-\epsilon \leq f_0(t_j, x_j)_i + \sum_{\ell=1}^{L} w_\ell f_\ell(t_j, x_j)_i - (\dot{x}_j)_i \leq \epsilon$$

Polytopic region $W = \{w \in \mathbb{R}^L \mid Aw \leq b\}$ with $b \in \mathbb{R}^{2N_x N_s}$

Peak Estimation Example (Flow)



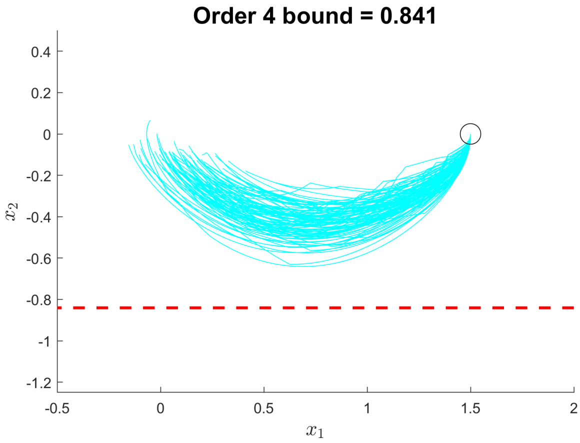
$$\dot{x} = [x_2, \; -x_1 - x_2 + x_1^3/3], \;\; T = 5$$

Peak Estimation Example (Flow)

$$\dot{x} = [x_2,\ -wx_1 - x_2 + x_1^3/3]$$

$$L = 1,\ m = 80 \text{ (2 nonredundant)}$$

Peak Estimation Example (Flow)



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

$$L = 10,\ \ m = 80\ (33\ \text{nonredundant})$$

Computational Considerations (Flow)

Size of largest PSD matrix in SDP (without symmetries):

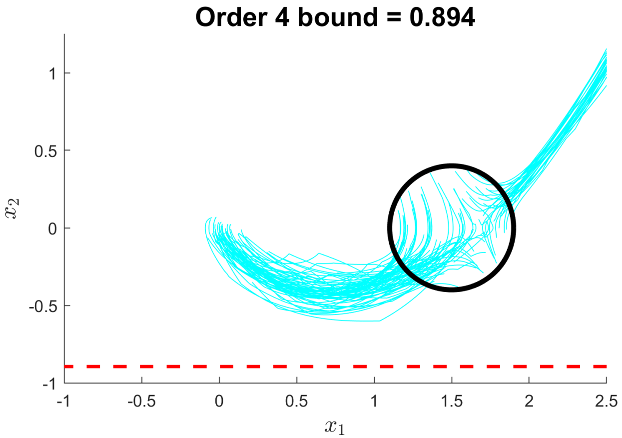
$$\text{Original} \quad \binom{1+N_x+L+d+\lceil \deg(f)/2\rceil-1}{1+N_x+L} = \binom{18}{13} = 8568$$

$$\text{Decomposed} \quad \binom{1+N_x+d+\lceil \deg(f)/2\rceil-1}{1+N_x} = \binom{8}{3} = 56$$

Order $d = 4$, $L = 10$, $N_x = 2$

W has 33 faces, 7534 vertices

Peak Estimation Example (Flow)



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

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

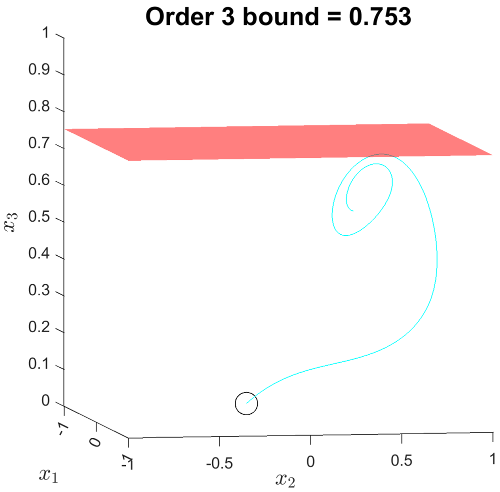
Peak Estimation Example (Twist)

Dynamics model:

$$\dot{x}_i = A_{ij}x_j - B_{ij}(4x_j^3 - 3x_j)$$

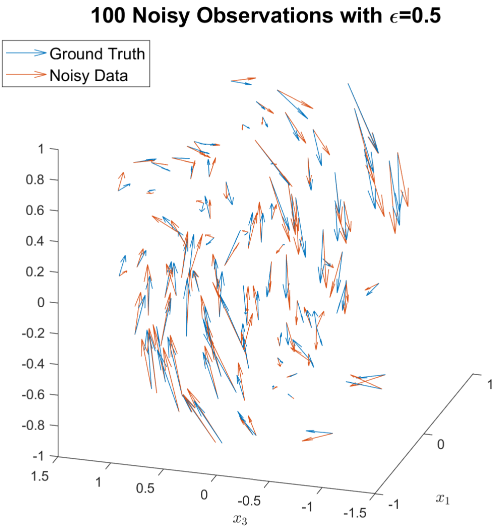
$$A = \begin{bmatrix} -1 & 1 & 1 \\ -1 & 0 & -1 \\ 0 & 1 & -2 \end{bmatrix}$$

$$B = \begin{bmatrix} -\frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$



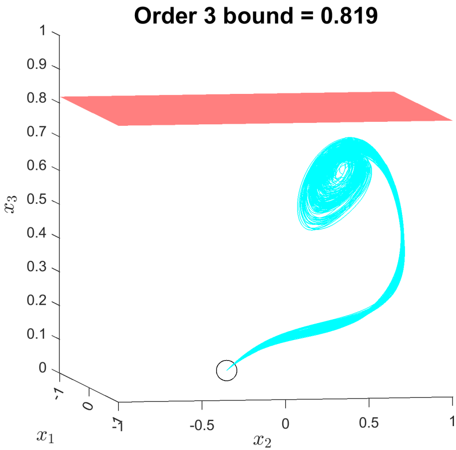
$$X_0 = [-1, 0, 0], \quad T = 8$$

Peak Estimation Example (Twist)



$m = 2N_sN_x = 600$ constraints

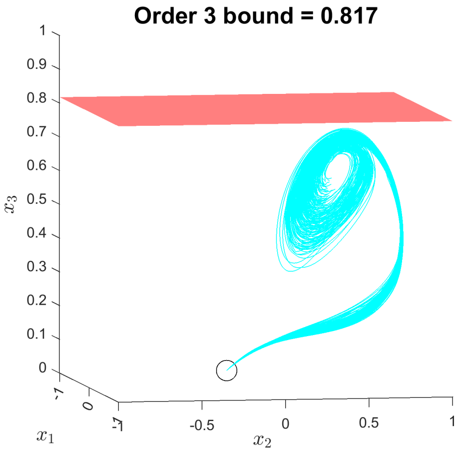
Peak Estimation Example (Twist)



Unknown A, Known B

$$L = 9, \ m = 600 \ (34 \text{ nonredundant})$$

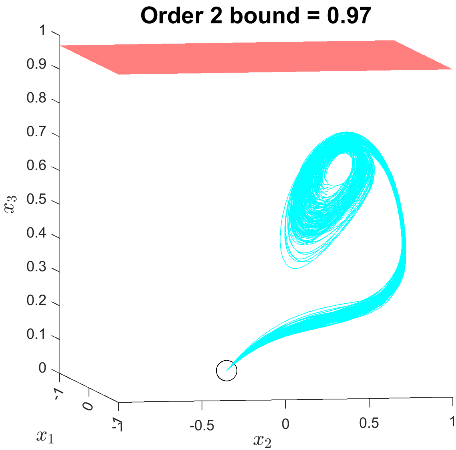
Peak Estimation Example (Twist)



Known A, Unknown B

$$L = 9, \;\; m = 600 \text{ (30 nonredundant)}$$

Peak Estimation Example (Twist)



Unknown A, Unknown B

$L = 18, \; m = 600$ (70 nonredundant)

Take-aways

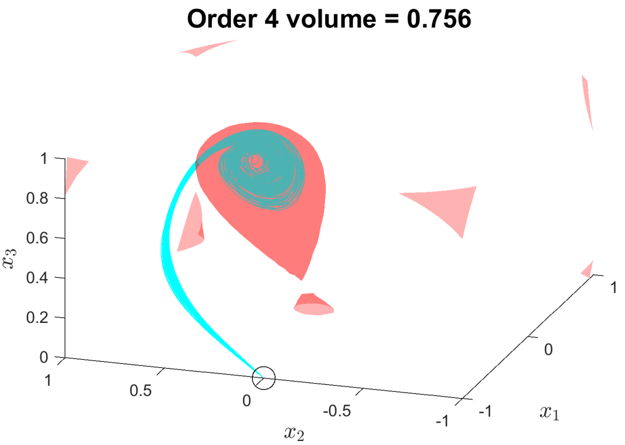
Conclusion

Tractable peak estimation problems (after preprocessing)

More SOS constraints in fewer variables

Data-driven estimates given L_∞-bounded noise

Reachable Set Estimation Example (Twist)



Unknown A, Known B

$L = 9, \; m = 600$ (34 nonredundant)

Future Work

- Streaming data and warm starts
- Maximum positively invariant sets
- Hybrid systems
- Compatibility with structure (e.g. sparsity)

Acknowledgements

LAAS-CNRS: Didier Henrion, POP group

Robust Systems Lab: Tianyu Dai

Chateaubriand Fellowship of the Office for Science Technology of the Embassy of France in the United States.

National Science Foundation

Air Force Office of Scientific Research

Thank you for your attention

arxiv:2112.14838

github.com/jarmill/data_driven_occ

Extra Material

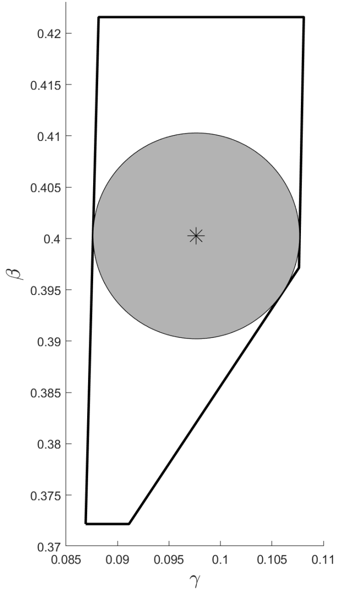
Preprocessing: Centering

Chebyshev center c: center of sphere with largest radius in W

Find through linear programming

$$\begin{aligned} &\max r \\ &A_k c + r\|A_k\|_2 \leq b_k \qquad \forall k \\ &r \geq 0, c \in \mathbb{R}^L \end{aligned}$$

Shifted dynamics $f_0 \leftarrow f_0 + \sum_{\ell=1}^{L} c_\ell f_\ell$



Preprocessing: Redundancy

Majority of $m = 2N_x N_s$ constraints are often redundant

Convex hull of dual polytope:
Time: $\Omega(m \log m + m^{\lfloor L/2 \rfloor})$

Linear program per constraint:
Time: $m \times \tilde{O}(mL + L^3)$
(Jan van den Brand *et. al.* 2020)