Bonus: Data-Driven Program
Auxiliary Evaluation along Optimal Trajectory

**Auxiliary Function Comparison**

Optimal $v(t, x)$ should be constant until peak is achieved.
Polytopic region for $L_{\infty}$-bounded noise

2 linear constraints for each coordinate $i$, sample $j$

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

Intersection of ellipsoids for $L_2$-bounded noise

$$\|f_0(t_j, x_j) + \sum_{\ell=1}^L w_{\ell} f_{\ell}(t_j, x_j) - (\dot{x}_j)\|_2 \leq \epsilon$$
Robust Counterpart Theory

Semidefinite-representable uncertainty set

\[ W = \cap_s \{ \exists \lambda_s \in \mathbb{R}^{q_s} : A_s w + G_s \lambda_s + e_s \in K_s \} \]

Lie constraint (based on Ben-Tal, Nemirovskii, 2009)

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

Nonconservative robust counterpart with multipliers \( \zeta \)

\[ \mathcal{L}_{f_0} v(t, x) + \sum_{s=1}^{N_s} e_s^T \zeta_s(t, x) \leq 0 \quad \forall [0, T] \times X \]

\[ G_s^T \zeta_s(t, x) = 0 \quad \forall s = 1..N_s \]

\[ \sum_{s=1}^{N_s} (A_s^T \zeta_s(t, x))_\ell + f_\ell(t, x) \cdot \nabla_x v(t, x) = 0 \quad \forall \ell = 1..L \]

\[ \zeta_s(t, x) \in K_s^* \quad \forall s = 1..N_s \]
Example: Polytopic uncertainty $W = \{ w \mid Aw \leq b \}$

Only the Lie Derivative constraint changes

\[
d^* = \min_{\gamma \in \mathbb{R}} \gamma
\]

\[
\begin{align*}
\gamma & \geq v(0, x) & \forall x \in X_0 \\
\mathcal{L}_{f_0} v(t, x) + b^T \zeta(t, x) & \leq 0 & \forall (t, x) \in [0, T] \times X \\
(A^T)_{\ell} \zeta(t, x) & = (f_\ell \cdot \nabla_x) v(t, x) & \forall \ell = 1..L \\
v(t, x) & \geq p(x) & \forall (t, x) \in [0, T] \times X \\
v(t, x) & \in C^1([0, T] \times X) \\
\zeta_k(t, x) & \in C_+([0, T] \times X) & \forall k = 1..m
\end{align*}
\]
Peak Estimation Example (Flow)

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

\[ L = 1, \ m = 80 \ (2 \text{ nonredundant}) \]
Peak Estimation Example (Flow)

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

\[ L = 10, \ m = 80 \ (33 \text{ nonredundant}) \]
Consistency sets

\[ Z = [0, J_{\text{max}}] \quad \Omega = \{ (w, z) \in W \times Z : J(w) \leq z \} \].

Optimal Control Problem with auxiliary \( v(t, x, z) \in C^1 \)

\[ d^* = \sup_{\gamma \in \mathbb{R}, \nu} \gamma \]

\[ v(0, x, z) \geq \gamma \quad \forall (x, z) \in X_0 \times Z \]

\[ v(t, x, z) \leq z \quad \forall (t, x, z) \in [0, T] \times X_u \times Z \]

\[ \mathcal{L}_f v(t, x, z, w) \geq 0 \quad \forall (t, x, z, w) \in [0, T] \times X \times \Omega \]
Crash Lie-decomposition

Exploit affine structure of $J(w) = \|\Gamma w - h\|_\infty$

Nonconservatively robustified Lie constraint

$$d^* = \sup_{\gamma \in \mathbb{R}, v} \gamma$$

$$v(0, x, z) \geq \gamma \quad \forall (x, z) \in X_0 \times Z$$

$$v(t, x, z) \leq z \quad \forall (t, x, z) \in [0, T] \times X_u \times Z$$

$$\mathcal{L}_{f_0} v - (z \mathbf{1} + h)^T \zeta \geq 0 \quad \forall (t, x, z) \in [0, T] \times X \times [0, J_{\text{max}}]$$

$$(\Gamma^T)_{\ell} \zeta + f_{\ell} \cdot \nabla_x v = 0 \quad \forall \ell = 1..L$$

$$\zeta_j \in C_+([0, T] \times X \times Z) \quad \forall j = 1..2nT.$$
Sum-of-Squares Method

Every \( c \in \mathbb{R} \) satisfies \( c^2 \geq 0 \)

Sufficient: \( q(x) \in \mathbb{R}[x] \) nonnegative if \( q(x) = \sum_i q_i^2(x) \)

Exists \( v(x) \in \mathbb{R}[x]^s \), Gram matrix \( Z \in S^s_+ \) with \( q = v^T Z v \)

Sum-of-Squares (SOS) cone \( \Sigma[x] \)

\[
x^2y^4 - 6x^2y^2 + 10x^2 + 2xy^2 + 4xy - 6x + 4y^2 + 1 = (x + 2y)^2 + (3x - 1 - xy^2)^2
\]

Motzkin Counterexample (nonnegative but not SOS)

\[
x^2y^4 + x^4y^2 - x^2y^2 + 1
\]
Putinar Positivstellensatz (Psatz) nonnegativity certificate over set $\mathbb{K} = \{ x \mid g_i(x) \geq 0, h_j(x) = 0 \}$:

$$q(x) = \sigma_0(x) + \sum_i \sigma_i(x) g_i(x) + \sum_j \phi_j(x) h_j(x) \quad (1a)$$

$$\exists \sigma_0(x) \in \Sigma[x], \quad \sigma_i(x) \in \Sigma[x], \quad \phi_j \in \mathbb{R}[x]. \quad (1b)$$

Psatz at degree $2d$ is an SDP, monomial basis: $s = \binom{n+d}{d}$

Archimedean: $\exists R \geq 0$ where $R - \| x \|_2^2$ has Psatz over $\mathbb{K}$
Optimal trajectories described by \((x^*_p, y^*, x^*_0, t^*_p)\):

- \(x^*_p\): location on trajectory of closest approach
- \(y^*\): location on unsafe set of closest approach
- \(x^*_0\): initial condition to produce \(x^*_p\)
- \(t^*_p\): time to reach \(x^*_p\) from \(x^*_0\)
Measures from Optimal Trajectories

Form measures from each \((x^*_p, x^*_0, t^*_p, y^*)\)

Atomic Measures (rank-1)

\[\begin{align*}
\mu^*_0 : & \quad \delta_{x=x^*_0} \\
\mu^*_p : & \quad \delta_{t=t^*_p} \otimes \delta_{x=x^*_p} \\
\eta^* : & \quad \delta_{x=x^*_p} \otimes \delta_{y=y^*}
\end{align*}\]

Occupation Measure \(\forall \nu(t, x) \in C([0, T] \times X)\)

\[\mu^* : \quad \langle \nu(t, x), \mu \rangle = \int_0^{t^*_p} \nu(t, x^*(t | x^*_0)) dt\]
Hybrid Systems

State guards and transitions

$L_2$ bound 0.0891: uncontrolled to boundary, controlled to sphere
Bonus: Chance-Peak
Reformulate as infinite-dimensional second-order cone program

SOC set $Q^3 = \{(s, \kappa) \in \mathbb{R}^3 \times \mathbb{R}_{\geq 0} \mid \|s\|_2 \leq \kappa\}$

\begin{align*}
p_r^* &= \sup_{z \in \mathbb{R}} rz + \langle p, \mu_\tau \rangle \quad (2a) \\
\mu_\tau &= \delta_0 \otimes \mu_0 + \mathcal{L}^\dagger \mu \quad (2b) \\
s &= [1 - \langle p^2, \mu_\tau \rangle, 2z, 2\langle p, \mu_\tau \rangle] \quad (2c) \\
(s, 1 + \langle p^2, \mu_\tau \rangle) &\in Q^3 \quad (2d) \\
\mu, \mu_\tau &\in \mathcal{M}_+([0, T] \times X). \quad (2e)
\end{align*}

Moment-SOS: $p_d^* \geq p_{d+1}^* \geq \ldots \geq p_r^* = P_r^* \geq P^*$
Bonus: Time Delay
Use moment-SOS hierarchy (Archimedean assumption)

Degree $d$, dynamics degree $\tilde{d} = d + \max(\lfloor \deg f / 2 \rfloor, \deg g - 1)$

Bounds: $p_d^* \geq p_{d+1}^* \geq \ldots \geq p_r^* = P_r^* \geq P^*$

Measure $\mu_p(t, x)$, $\mu(t, x)$

PSD Size $\binom{1+n+d}{d}$, $\binom{1+n+\tilde{d}}{\tilde{d}}$

Timing scales approximately as $(1 + n)^{6\tilde{d}}$ or $\tilde{d}^{4(n+1)}$
Propagation of Continuity

Increasing Continuity

\[ x'(t) = -2x(t) - 2x(t - 1) \]

Continuity increases every \( \tau_r \) time steps
Use moment-SOS hierarchy (Archimedean assumption)

Degree $d$, dynamics degree $\tilde{d} = d + \lceil \deg f / 2 \rceil$

Bounds: $p_d^* \geq p_{d+1}^* \geq \cdots = p^* \geq P^*$

Size of Moment Matrices Peak Estimation

Measure: $\mu_0$  $\mu^p$  $\mu_h$
Size: $\binom{n+d}{d}$  $\binom{n+1+d}{d}$  $\binom{n+1+\tilde{d}}{\tilde{d}}$

Measure: $\bar{\mu}_0$  $\bar{\mu}_1$  $\nu$
Size: $\binom{2n+1+\tilde{d}}{\tilde{d}}$  $\binom{2n+1+\tilde{d}}{\tilde{d}}$  $\binom{n+1+\tilde{d}}{\tilde{d}}$

Timing scales approximately as $(2n + 1)^{6\tilde{d}}$ or $\tilde{d}^4(2n+1)$
Upper bound $I_{max} \geq 56.9\%$ with order 3 LMI

Recovery: $t^* = 15.6$ days, $(S^*, I^*) = (56.9\%, 5.61\%)$
Maximize $x_1$ on $\dot{x}(t) = \begin{bmatrix} x_2(t)t - 0.1x_1(t) - x_1(t - \tau)x_2(t - \tau) \\ -x_1(t)t - x_2(t) + x_1(t)x_1(t - \tau) \end{bmatrix}$
Time-Varying Histories

History restrictions and trajectories of system
Joint Component Consistency

\( (t, x_0) \) marginal of \( \bar{\mu} \)

For all test functions \( \phi_0 \in C([0, T] \times X) \)

\[
\langle \phi_0(t, x_0), \bar{\mu} \rangle = \int_0^T \phi_0(t, x(t | x_h)) dt \\
= \left( \int_0^{T-\tau} + \int_{T-\tau}^T \right) \phi_0(t, x(t | x_h)) dt \\
= \langle \phi_0(t, x), \nu_0 + \nu_1 \rangle
\]
For all test functions $\phi_1 \in C([0, T] \times \mathcal{X})$

$$\langle \phi_1(t, x_1), \bar{\mu} \rangle = \int_0^T \phi_1(t, x(t - \tau | x_h))dt$$

$$= \int_{-\tau}^{T-\tau} \phi_1(t + \tau, x(t | x_h))dt$$

$$= \int_{-\tau}^0 \phi_1(t + \tau, x_h(t))dt + \langle \phi_1(t + \tau, x), \nu_0 \rangle$$
Table 1: Objective values for Flow experiment

<table>
<thead>
<tr>
<th>degree $d$</th>
<th>Joint+Component</th>
<th>Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.25</td>
<td>1.25</td>
</tr>
<tr>
<td>2</td>
<td>1.223</td>
<td>1.2183</td>
</tr>
<tr>
<td>3</td>
<td>1.1937</td>
<td>1.1913</td>
</tr>
<tr>
<td>4</td>
<td>1.1751</td>
<td>1.1727</td>
</tr>
<tr>
<td>5</td>
<td>1.1636</td>
<td>1.1630</td>
</tr>
</tbody>
</table>

Table 2: Time (seconds) to obtain SDP bounds in Table 1

<table>
<thead>
<tr>
<th>degree $d$</th>
<th>Joint+Component</th>
<th>Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.782</td>
<td>0.937</td>
</tr>
<tr>
<td>2</td>
<td>0.991</td>
<td>1.190</td>
</tr>
<tr>
<td>3</td>
<td>5.271</td>
<td>9.508</td>
</tr>
<tr>
<td>4</td>
<td>31.885</td>
<td>105.777</td>
</tr>
<tr>
<td>5</td>
<td>336.509</td>
<td>552.496</td>
</tr>
</tbody>
</table>
Bonus: Measure Background
Measures

Nonnegative Borel Measure $\mu$
Assigns each set $A \subseteq X$ a ‘size’ $\mu(A) \geq 0$ (Measure)
Mass $\mu(X) = \langle 1, \mu \rangle = 1$: Probability distribution

$\mu \in M_+(X)$: space of measures on $X$
$f \in C(X)$: continuous function on $X$
Pairing by Lebesgue integration $\langle f, \mu \rangle = \int_X f(x) d\mu(x)$
Dirac Delta Measure

Dirac delta \( \delta_{x'}(A) = \begin{cases} 
1 & x' \in A \\
0 & x' \notin A 
\end{cases} \)

Probability:
\( \delta_{x'}(X) = 1, \langle f(x), \delta_{x'} \rangle = f(x') \)

\( \mu(A) = 1: \) Solid Box

\( \mu(A) = 0: \) Dashed Box
Atomic Measure

Rank-1 atomic measure

\[ \mu = c \delta_{x'} \quad c > 0 \]

Rank-2 atomic measure

\[ \mu = c_1 \delta_{x'_1} + c_2 \delta_{x'_2} \quad c > 0, \quad x'_1 \neq x'_2 \]

Rank-r atomic measure

\[ \mu = \sum_{i=1}^{r} c_i \delta_{x'_i} \quad c > 0, \quad \{x'_i\}_{i=1}^{r} \text{distinct} \]
Example of Measure Optimization

Expected Values of Probability Distributions

- $E[f] = 1.46$ for Uniform(-2, -1)
- $E[f] = 1.46$ for Normal(-1, 1)
- $E[f] = 2.78$ for Normal(1, 0.25)
- $E[f] = 1.26$ for Delta(-1.525)

Optimum $\mathbb{E}_\mu[f] = \langle f, \mu \rangle$ at $\mu = \delta_{x^*}$
Nonconvex problems could be convex in measures

$$\min_{x \in K} p(x) \rightarrow \min_{\mu \in \mathcal{M}_+(K)} \langle p, \mu \rangle, \quad \langle 1, \mu \rangle = 1$$

$$f\left(\frac{1}{2}(1 + (-1))\right) = 1, \text{ but } \frac{1}{2}(f(1) + f(-1)) = 0$$
Bonus: Approximating Measure LPs
Need for Approximation

Measure LPs are infinite-dimensional

Linear Matrix Inequality: convex problem

\[
\max_y b^T y \quad C + \sum_{i=1}^m A_i y_i \geq 0
\]

Solve LMIs through (interior point, ADMM, etc.)

Approximate infinite LPs by finite-dimensional LMIs
Moments

Monomial $x^\alpha = \prod_i x_i^{\alpha_i}$ for power $\alpha \in \mathbb{N}^n$

Degree $|\alpha| = \sum_i \alpha_i$

$\alpha$-moment of measure $y_\alpha = \langle y_\alpha, \mu \rangle$

Measure uniquely described by infinite set $\{y_\alpha\}_{\alpha \in \mathbb{N}^n}$

When does a sequence $\{y_\alpha\}_{\alpha \in \mathcal{A}}$ correspond to a measure $\mu$?
Linear Functional polynomial → moments

\[ f(x) \to \int_X f(x) d\mu = \int_X \sum_{\alpha} f_\alpha x^\alpha d\mu = \sum_{\alpha} f_\alpha y_\alpha \]

Bivariate Example

\[ 2 + x_1 x_2 - 3x_1^2 + x_1 x_2^3 \to 2 + y_{11} - 3y_{20} + y_{13} \]
Squares $f(x)^2$ are nonnegative (real)

$f(x)^2 \geq 0$ implies that $\langle f(x)^2, \mu \rangle \geq 0 \quad \forall f \in \mathbb{R}[x]$

$$\langle f(x)^2, \mu \rangle = \int_X \sum_{\alpha,\beta} (f_\alpha x^\alpha)(f_\beta x^\beta) d\mu = \int_X \sum_{\alpha,\beta} (f_\alpha f_\beta x^{\alpha+\beta}) d\mu \geq 0$$

Moment matrix $M[y] \succeq 0$ has $M[y]_{\alpha,\beta} = y_{\alpha+\beta}$

$$\langle f(x)^2, \mu \rangle = f^T M[y] f \geq 0$$
Moments up to degree $2 \times 2 = 4$

$$\mathbf{M}_2[y] = \begin{bmatrix}
    y_{00} & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} \\
    y_{10} & y_{20} & y_{11} & y_{00} & y_{30} & y_{21} \\
    y_{01} & y_{11} & y_{02} & y_{21} & y_{12} & y_{00} \\
    y_{20} & y_{30} & y_{21} & y_{40} & y_{31} & y_{20} \\
    y_{11} & y_{21} & y_{12} & y_{31} & y_{22} & y_{11} \\
    y_{02} & y_{12} & y_{03} & y_{22} & y_{13} & y_{02}
\end{bmatrix}$$
Localizing Matrices

$\mu$ supported on set $K = \{x \mid g_i(x) \geq 0, i = 1 \ldots N\}$

$g_i(x)f(x)^2 \geq 0$ implies that $\langle g_i(x)f(x)^2, \mu \rangle \geq 0$

$$\langle g_i(x)f(x)^2, \mu \rangle = \int_X \sum_{\alpha, \beta, \gamma} (f_\alpha f_\beta g_\gamma x^{\alpha+\beta+\gamma}) d\mu \geq 0$$

Localizing matrix $M[g_i m] \succeq 0$ has $M[g_i m]_{\alpha, \beta} = \sum_\gamma g_\gamma m_{\alpha+\beta+\gamma}$

$$\langle g_i(x)f(x)^2, \mu \rangle = f^T M[g_i y] f \geq 0$$
Polynomial optimization problem example:

\[ p^* = \max_{x \in K} p(x) = \max_{\mu \in \mathcal{M}_+(K)} \langle p(x), \mu \rangle, \quad \mu(K) = 1 \]

Keep moments up to degree \( d \):

\[ p_d^* = \max_y \sum_{|\alpha| \leq 2d} p_\alpha m_\alpha \]

\[ \mathbb{M}_d[y], \quad \mathbb{M}_{d - \deg(g_i)}[g_i y] \succeq 0 \]

Finite-dimensional SDP: \( \mathbb{M}_d[y] \) has size \( \binom{n+d}{d} \)

Bounds \( p_d^* \geq p_{d+1}^* \geq p_{d+2}^* \ldots \) converge to \( p^* \) as \( d \to \infty \)
1. Trajectory Program
2. Measure LP
3. Moment LMI

Increase degree $d$ of LMI to get better bounds

Prove conditions under which $\lim_{d \to \infty} p_d^* \to p^* = P^*$