Robust to Dynamics Optimization

Amir Ali Ahmadi

Princeton, ORFE

(PACM, COS, MAE)

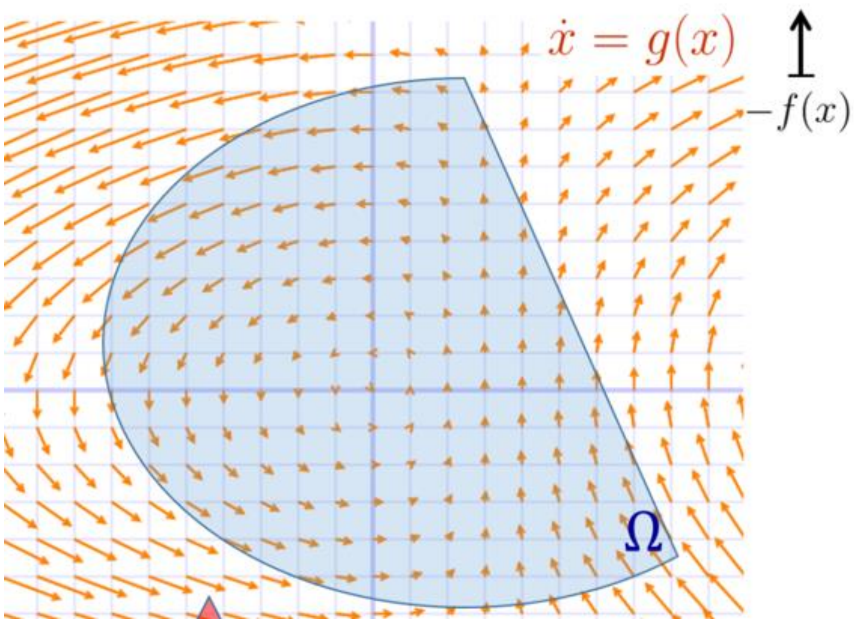
Oktay Gunluk

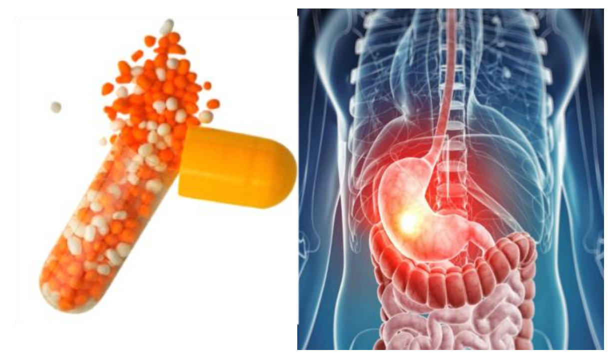
IBM Watson Research Center

Workshop on Design and Analysis of Robust Systems – CPS Week 2018

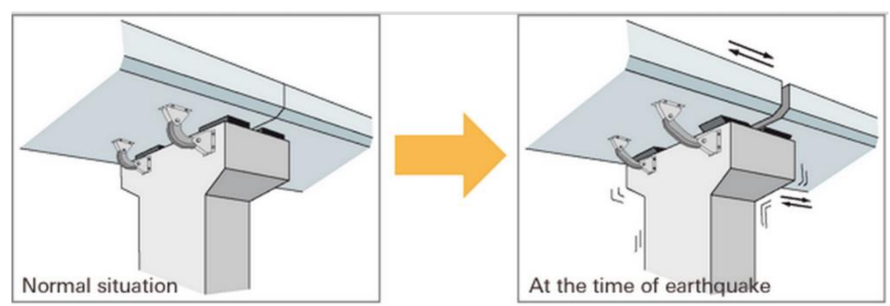
RDO (informally)

- You solve a constrained optimization problem today
- An external dynamical system may move your optimal point in the future and make it infeasible
- You want your initial decision to be "safe enough" to not let this happen





Drug design



Earthquake-resistant structures

Robust to Dynamics Optimization (RDO)

An RDO is describe by two pieces of input:

1) An optimization problem: $\min_{x} \{f(x) : x \in \Omega\}$

2) A dynamical system: $x_{k+1} = g(x_k)$ (discrete time case)

RDO is then the following problem:

$$\min_{x_0} \{f(x_0) : x_k \in \Omega, k = 0, 1, 2, \ldots\}$$

This talk:

Optimization Problem	Dynamics
Linear Program	Linear
Quadratic Program	Nonlinear
Integer Program	Uncertain
Semidefinite Program	Time-varying
Polynomial Program, …	Hybrid, …

Agenda for the rest of the talk

1) Robust to linear dynamics linear programming

2) Stability of uncertain/time-varying systems
   - The joint spectral radius (JSR)
   - SDP-based techniques for bounding the JSR

3) Robust to uncertain dynamics linear programming

R-LD-LP

Robust to linear dynamics linear programming (R-LD-LP)

Classical LP:

$$\min_x \{c^T x : Ax \le b\}$$

Robust LP:

$$\min_x \{c^T x : Ax \le b, \forall A \in \mathbb{A}, b \in \mathbb{B}\}$$

R-LD-LP:

$$\min_{x_0} \{c^T x_0 : Ax_k \le b, k = 0, 1, 2, \ldots ; x_{k+1} = Gx_k\}$$

R-LD-LP

Robust to linear dynamics linear programming (R-LD-LP)

$$\min_{x_0}\{c^T x_0 : Ax_k \leq b, k = 0, 1, 2, \ldots ; x_{k+1} = Gx_k\}$$

Input data: A, b, c, G

Alternative form:

$$\min_{x}\{c^T x: Ax \leq b, AGx \leq b, AG^2x \leq b, AG^3x \leq b, \ldots\}$$

(An infinite LP)

Feasible set of R-LD-LP:

$$S := \bigcap_{k=0}^{\infty}\{x \mid AG^k x \leq b\}$$

An example...

$$\min_{x_0}\{c^T x_0 : Ax_k \leq b, k = 0, 1, 2, \ldots ; x_{k+1} = Gx_k\}$$

$$A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 3 \end{bmatrix}, c = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, G = \begin{bmatrix} 0.6 & -0.4 \\ 0.8 & 0.5 \end{bmatrix}$$

$$\boldsymbol{S} = \cap_{k=0}^{\infty} \{\boldsymbol{A}\boldsymbol{G}^{k}\boldsymbol{x} \leq \boldsymbol{b}\} = \cap_{k=0}^{2} \{\boldsymbol{A}\boldsymbol{G}^{k}\boldsymbol{x} \leq \boldsymbol{b}\}$$

PRINCETON UNIVERSITY ORFE

Obvious way to get lower bounds

$$\min_x \{c^T x : Ax \le b, AGx \le b, AG^2 x \le b, AG^3 x \le b, ..\}$$

Truncate!

(outer approximations to the feasible set)

Natural questions:

- Is the feasible set of R-LD-LP always a polyhedron?
- When it is, how large are the number of facets?
- Does the feasible set have a tractable description?

- How to get **upper bounds**?!
  - (We'll see later: from semidefinite programming)

The feasible set of an R-LD-LP

$$\mathcal{S} := \bigcap_{k=0}^{\infty} \{x \in \mathbb{R}^n |\ AG^k x \leq b\}$$

Theorem.

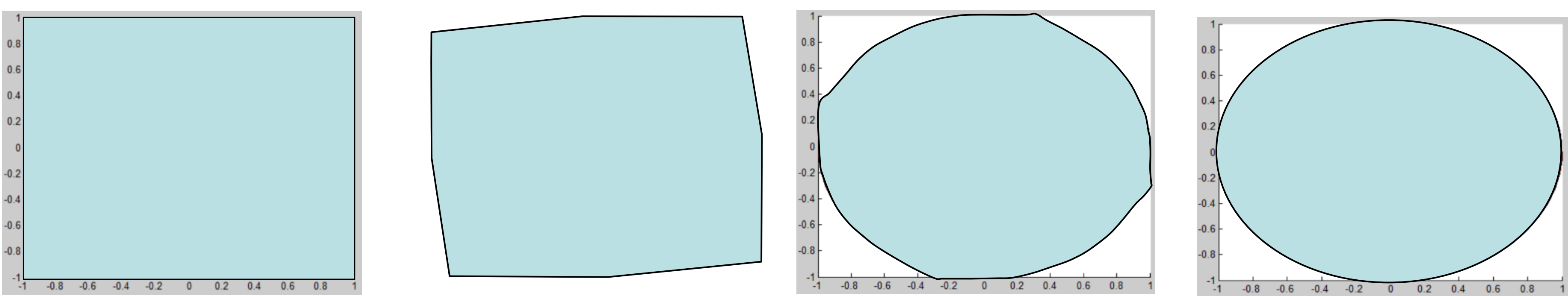
(1) This set is always closed, convex, and invariant.

(2) It is not always polyhedral.

(3) Given A, b, G, and $z \in \mathbb{Q}^n$, it is NP-hard to check whether $z \in S$.

Proof of (2).

$$\{Ax \leq b\} \quad G = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}, \ \theta \text{ irrational}$$



Finite convergence of outer approximations

$$\mathcal{S} := \bigcap_{k=0}^{\infty} \{x \in \mathbb{R}^n | \; AG^k x \leq b\}$$

$$S_r := \bigcap_{k=0}^{r} \{x \in \mathbb{R}^n | \; AG^k x \leq b\}$$

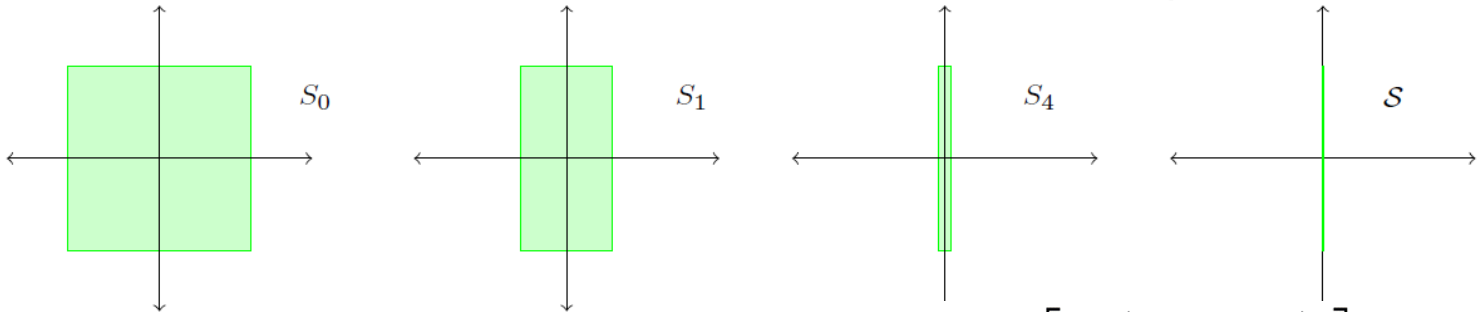
$$\mathcal{S} \subseteq \ldots S_{r+1} \subseteq S_r \subseteq \ldots \subseteq S_2 \subseteq S_1 \subseteq S_0 = P.$$

Lemma. If $S_r = S_{r+1}$, then $S_r = \mathcal{S}$.
(Poly-time checkable condition for fixed r.)

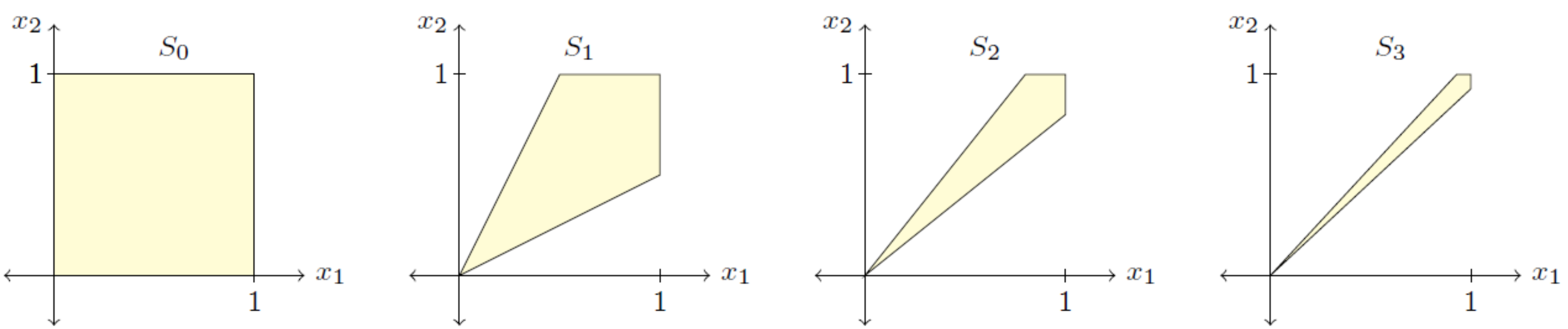
Proposition. There are **three barriers to finite convergence**:
(1) Having $\rho(G) \geq 1$.
(2) Having the origin on the boundary of P.
(3) Having an unbounded polyhedron P.

Barriers to finite convergence

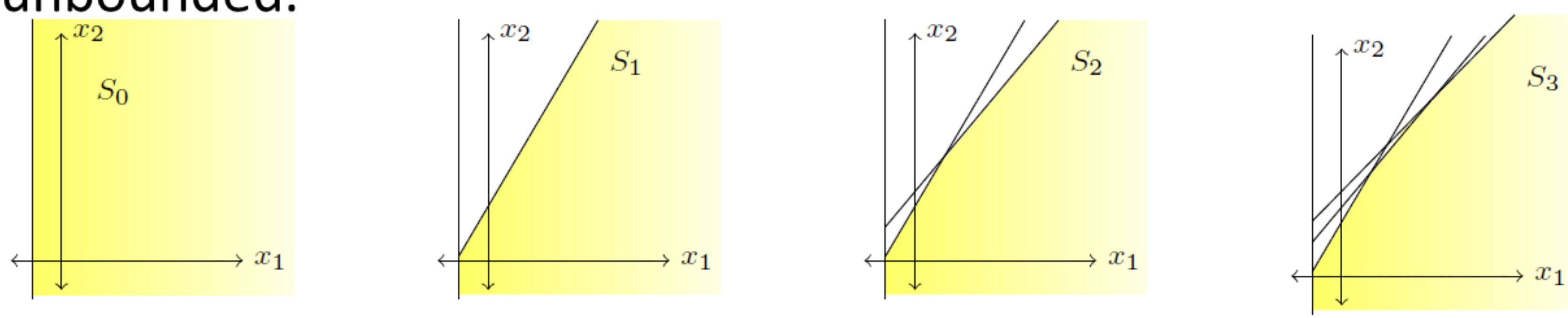
(1) $\rho(G) \geq 1.$ $\qquad G = \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix}$



(2) The origin on the boundary of P. $\qquad G = \begin{bmatrix} 1/3 & -1/6 \\ -1/6 & 1/3 \end{bmatrix} \qquad \rho(G) = \frac{1}{2}$



(3) P unbounded.



Computing time to convergence

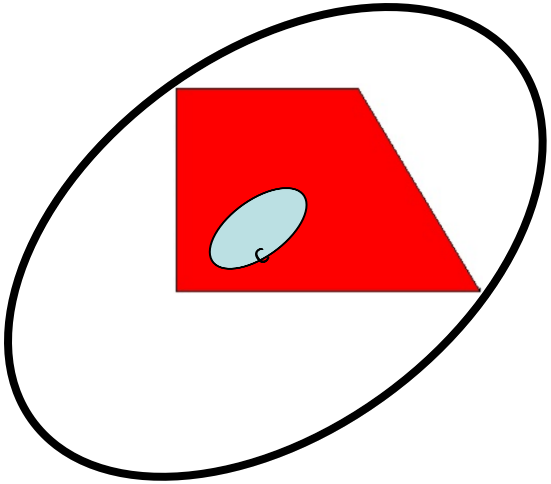
Theorem: If $\rho(G) < 1$, $P = \{Ax \leq b\}$ is bounded and contains the origin in its interior, then

(1) $S = S_r$ for an integer r that can be computed in time $poly\big(\sigma(A, b, G)\big)$.

(2) For any fixed $\rho^* < 1$, all instances of R-LD-LP with $\rho(G) \leq \rho^*$ can be solved in time $poly\big(\sigma(A, b, c, G)\big)$.

Proof idea.

Invariant ellipsoid:
$\{x^T P x \leq 1\}$



Upper bound on the number of iterations

- Find an invariant ellipsoid defined by a positive definite matrix P
- Find a shrinkage factor $\gamma \in (0, 1)$; i.e., a scalar satisfying $G^T PG \preceq \gamma P$
- Find a scalar $\alpha_2 > 0$ such that

$$\{Ax \leq b\} \subseteq \{x^T Px \leq \alpha_2\}$$

- Find a scalar $\alpha_1 > 0$ such that

$$\{x^T Px \leq \alpha_1\} \subseteq \{Ax \leq b\}$$

- Let

$$r = \lceil \frac{\log \frac{\alpha_1}{\alpha_2}}{\log \gamma} \rceil$$

Finding an invariant ellipsoid

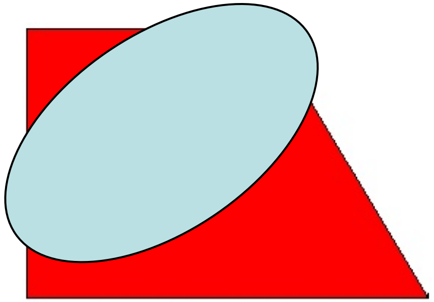
- **Computation of P.**

To find an invariant ellipsoid for G, we solve the linear system

$$G^T PG - P = -I,$$

where I is the $n \times n$ identity matrix. This is called the Lyapunov equation.

The matrix P will automatically turn out to be positive definite.



Finding the shrinkage factor

- **Computation of γ.**

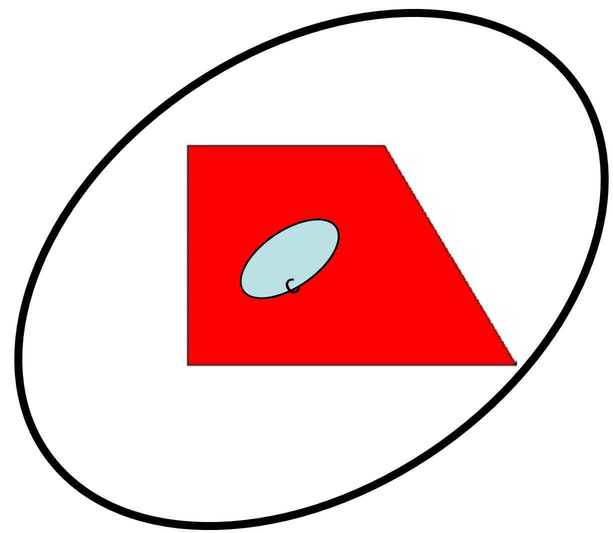
$$\gamma = 1 - \frac{1}{\max_i \{P_{ii} + \sum_{j \neq i} |P_{i,j}|\}}.$$

Proof idea.

$$\begin{aligned} x^T G^T P G x &= x^T P x - x^T x \\ &\leq x^T P x (1 - \eta) \end{aligned}$$

where η is any number such that

$$\eta x^T P x \leq x^T x$$



Shrinkage is at least $1 - \frac{1}{\lambda_{max}(P)}$

$$\lambda_{max}(P) \leq \max_i \{P_{ii} + \sum_{j \neq i} |P_{i,j}|\}.$$

(Bound from Greshgorin's circle theorem)

Finding the outer ellipsoid

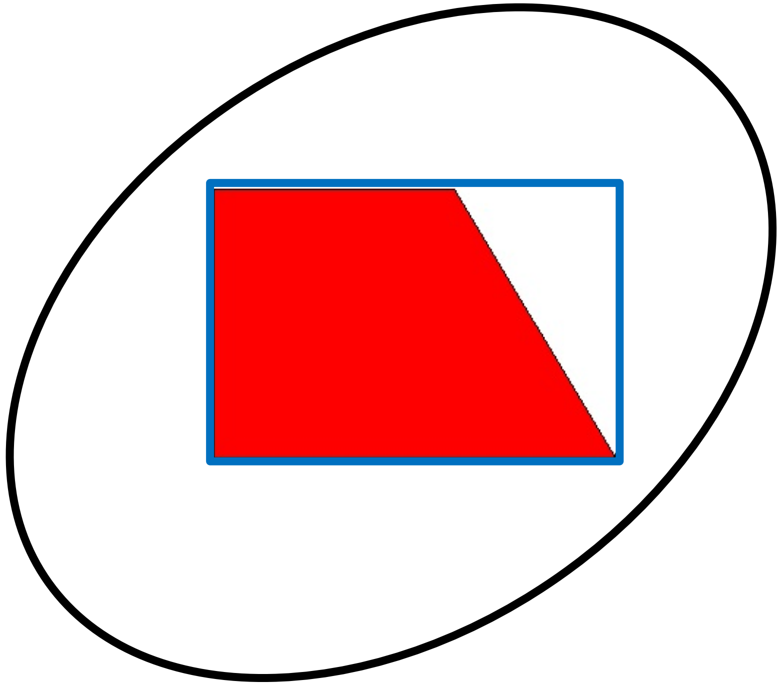
- **Computation of α_2.** By solving, e.g., n LPs, we can place our polytope $\{Ax \leq b\}$ in a box; i.e., compute $2n$ scalars l_i, u_i such that

$$\{Ax \leq b\} \subseteq \{l_i \leq x_i \leq u_i\}.$$

We then bound $x^T P x = \sum_{i,j} P_{i,j} x_i x_j$ term by term to get α_2:

$$\alpha_2 = \sum_{i,j} \max\{P_{i,j} u_i u_j, P_{i,j} l_i l_j, P_{i,j} u_i l_j, P_{i,j} l_i u_j\}.$$

This ensures that $\{l_i \leq x_i \leq u_i\} \subseteq \{x^T P x \leq \alpha_2\}$. Hence, $\{Ax \leq b\} \subseteq \{x^T P x \leq \alpha_2\}$.



Finding the inner ellipsoid

- **Computation of α_1.** For $i = 1, \ldots, m$, we compute a scalar η_i by solving the convex program

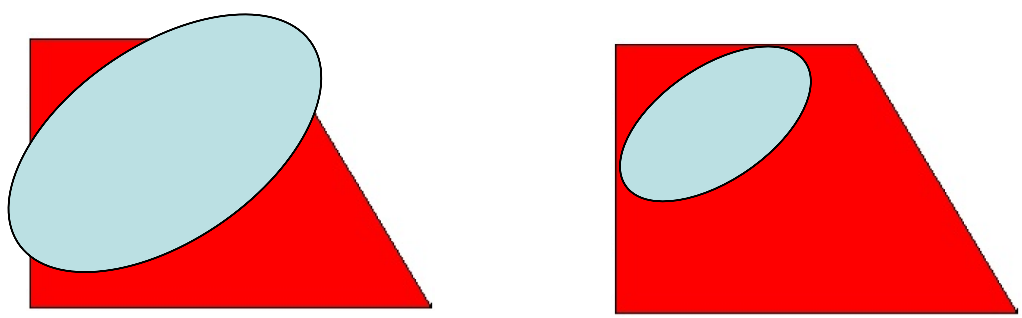
$$\eta_i := \min_x \{a_i^T x : x^T P x \leq 1\},$$

where a_i is the i-th row of the constraint matrix A. This problem has a closed form solution:

$$\eta_i = -\sqrt{a_i^T P^{-1} a_i}.$$

Note that P^{-1} exists since $P \succ 0$. We then let

$$\alpha_1 = \min_i \{\frac{b_i^2}{\eta_i^2}\}.$$



Recap

R-LD-LP:

$$\min_x \{c^T x : Ax \le b, AGx \le b, AG^2 x \le b, AG^3 x \le b, \dots\}$$

$$\mathcal{S} := \bigcap_{k=0}^{\infty} \{x \in \mathbb{R}^n \mid AG^k x \le b\}$$

Outer approximations:

(gives lower bounds on the optimal value)

$$S_r := \bigcap_{k=0}^{r} \{x \in \mathbb{R}^n \mid AG^k x \le b\}$$

$$\mathcal{S} \subseteq \dots S_{r+1} \subseteq S_r \subseteq \dots \subseteq S_2 \subseteq S_1 \subseteq S_0 = P.$$

What about upper bounds? Need inner approximations!

Upper bounds on R-LD-LP via SDP

- Goal: Find the best invariant ellipsoid inside the original polytope and optimize over that.

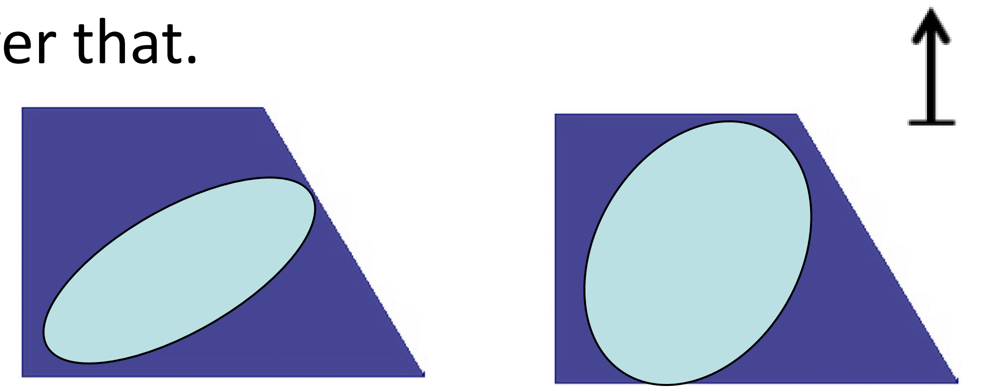
$$\min_{x,P} \quad c^T x$$

$$P \succ 0$$

$$G^T P G \preceq P$$

$$x^T P x \leq 1$$

$$[\forall z, \quad z^T P z \leq 1 \Rightarrow Az \leq b]$$



Non-convex formulation
(even after the application of the S-lemma)

Upper bounds on R-LD-LP via SDP

- If we parameterize in terms of P^{-1} instead, then it becomes convex!

An improving sequence of SDPs

- Goal: Find the best point that lands in an invariant set.

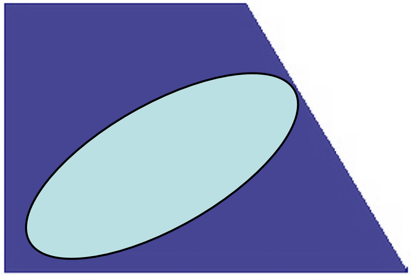
$$\min_{x,P} \quad c^T x$$

$$P \succ 0$$

$$G^T P G \preceq P$$

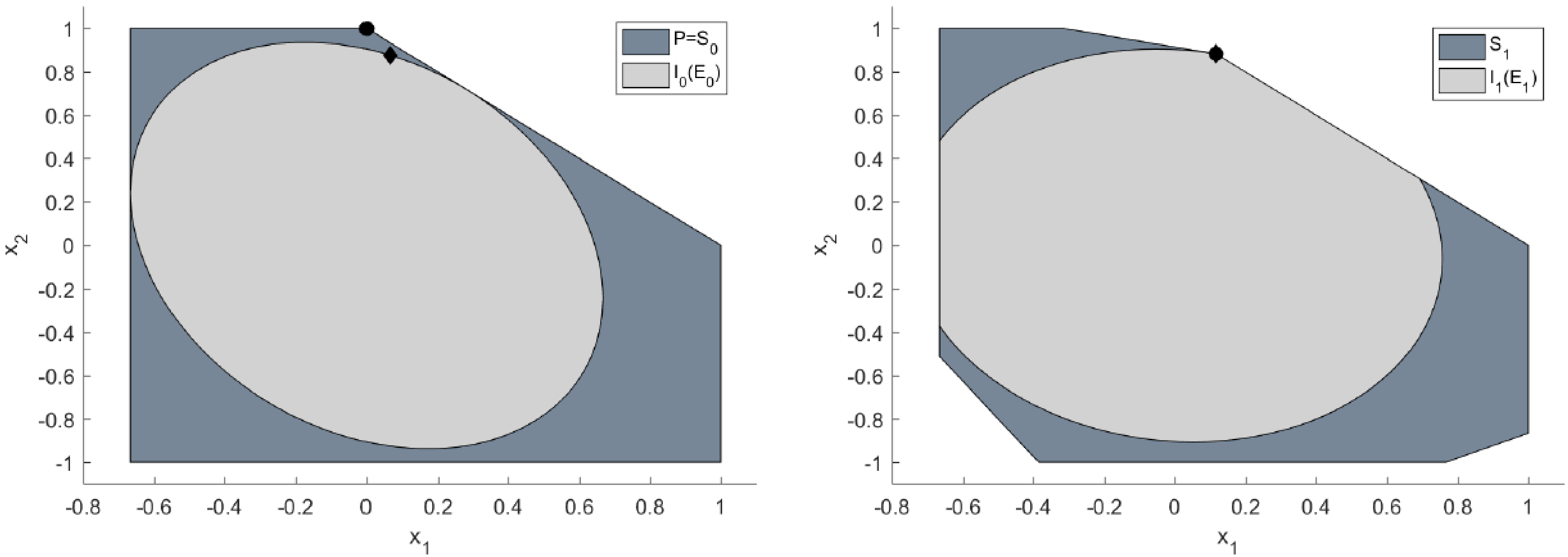
$$x^T P x \leq 1$$

$$\left[\forall z, \quad z^T P z \leq 1 \Rightarrow Az \leq b\right]$$



An example

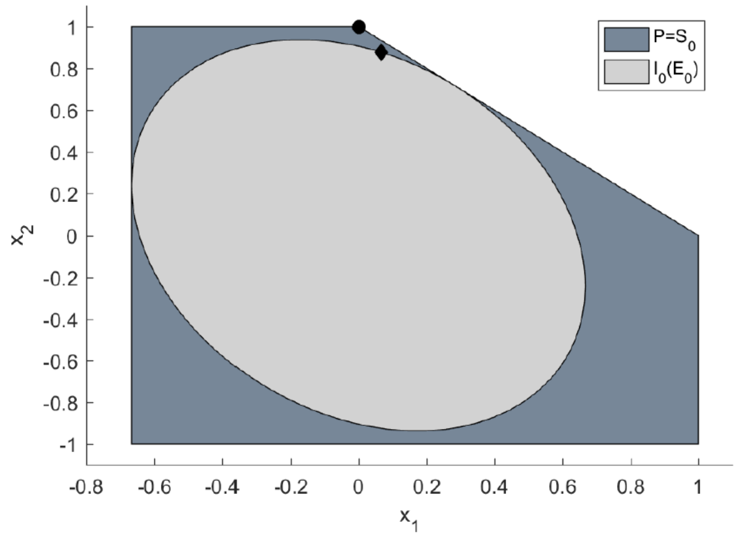
$$A = \begin{pmatrix} 1 & 0 \\ -1.5 & 0 \\ 0 & 1 \\ 0 & -1 \\ 1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad c = -\begin{pmatrix} 0.5 & 1 \end{pmatrix}, \quad G = \frac{4}{5}\begin{pmatrix} \cos(\theta) & \sin(\theta), \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \text{ where } \theta = \frac{\pi}{6}.$$



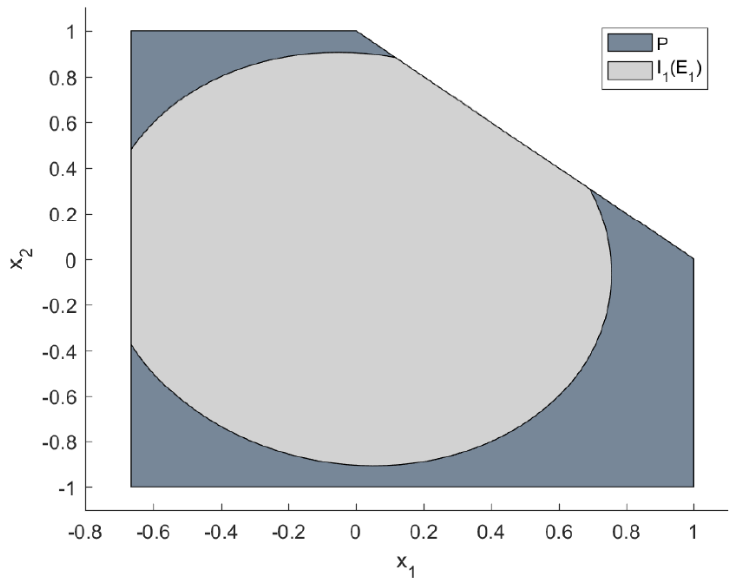
Thm. The SDP upper bound monotonically improves and gives the exact optimal value of R-LD-LP in r^* steps, where r^* is polynomially computable.

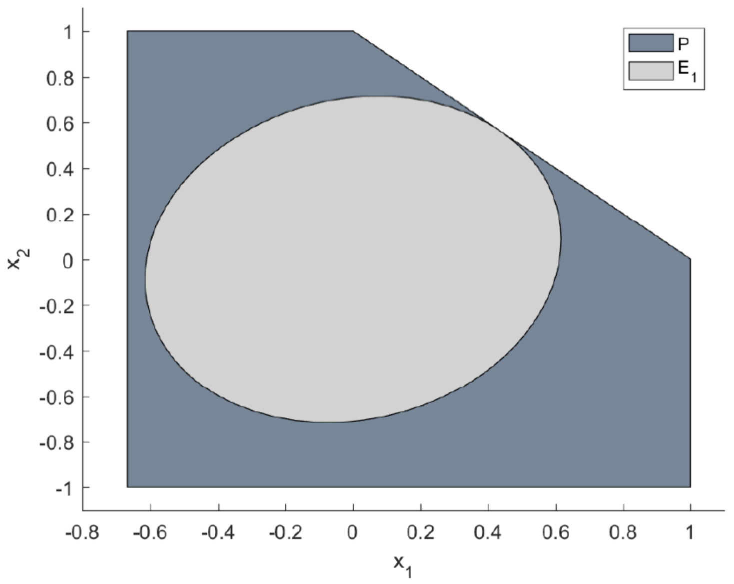
Another interpretation

r=0



r=1





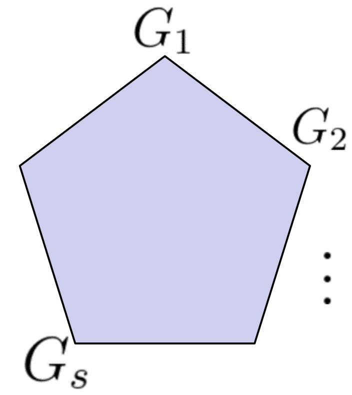
PRINCETON UNIVERSITY ORFE

LP
+
Uncertain & time-varying linear systems

R-ULD-LP

Robust to uncertain linear dynamics linear programming (R-ULD-LP)

$$x_{k+1} \in conv\{G_1, \ldots, G_s\} x_k$$



Models **uncertainty** and **variations with time** in the dynamics

$$\min_x \{c^T x : AGx \le b, \forall G \in \mathbb{G}^*\}$$ (An infinite LP)

$\mathbb{G}^*$: set of all finite products of $G_1, \ldots, G_s$

Input data: $A, b, c, G_1, \ldots, G_s$

The joint spectral radius: Upper bounds via SDP

The Joint Spectral Radius

Given a finite set of $n \times n$ matrices $\mathcal{G} = \{G_1, \ldots, G_s\}$

Joint spectral radius (JSR):

$$\rho(\mathcal{G}) = \lim_{k \to \infty} \max_{\sigma \in \{1,\ldots,s\}^k} ||G_{\sigma_1} \ldots G_{\sigma_k}||^{1/k}$$

If only one matrix:

$$\mathcal{G} = \{G\}$$

Spectral Radius

$$\rho(G) = \lim_{k \to \infty} ||G^k||^{1/k}$$

G. C. Rota and W. G. Strang
A note on the joint spectral radius
Indag. Math., 22:379–381, 1960.

JSR and Uncertain&Time-Varying Linear Systems

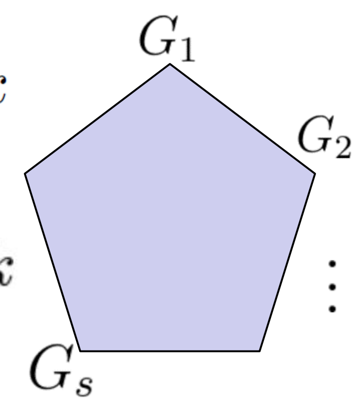
Linear dynamics: $x_{k+1} = Gx_k$

Spectral radius: $\rho(G) = \lim_{k\to\infty} ||G^k||^{1/k}$

"Stable" iff $\rho(G) < 1$

Uncertain and time-varying linear dynamics:

$$x_{k+1} \in conv\{G_1, \ldots, G_s\}x_k$$



Joint spectral radius (JSR):

$$\rho(\mathcal{G}) = \lim_{k\to\infty} \max_{\sigma\in\{1,\ldots,s\}^k} ||G_{\sigma_1} \ldots G_{\sigma_k}||^{1/k}$$

"Uniformly stable" iff $\rho(\mathcal{G}) < 1$

Computation of JSR

If only one matrix: $\mathcal{G} = \{G\}$

Testing if "$\rho(G) < 1$" can be done in poly time.

For more than one matrix:

(even for 2 matrices of size 47x47)

Testing if "$\rho(\mathcal{G}) \leq 1$" is undecidable.

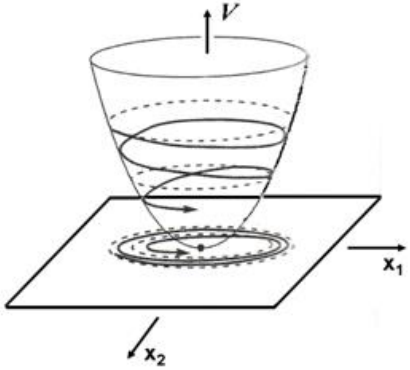
[Blondel, Tsitsiklis]

- **Goal: compute upper bounds on the JSR**

(equivalently, give sufficient conditions for uniform stability)

Common Lyapunov function

$$x_{k+1} \in conv\{G_1, \ldots, G_s\} x_k$$



If we can find a function $V(x) : \mathbb{R}^n \to \mathbb{R}$

such that

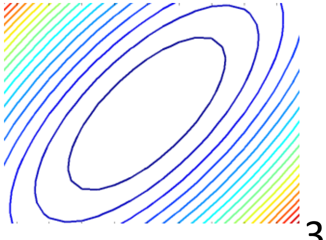
$$V(x) > 0, \forall x \neq 0$$
$$V(G_i x) < V(x), \forall x \neq 0, i = 1, \ldots, s$$

then, $\rho(\mathcal{G}) < 1$

- Such a function always exists!
- May be extremely hard to find. **Can easily fail to be quadratic.**

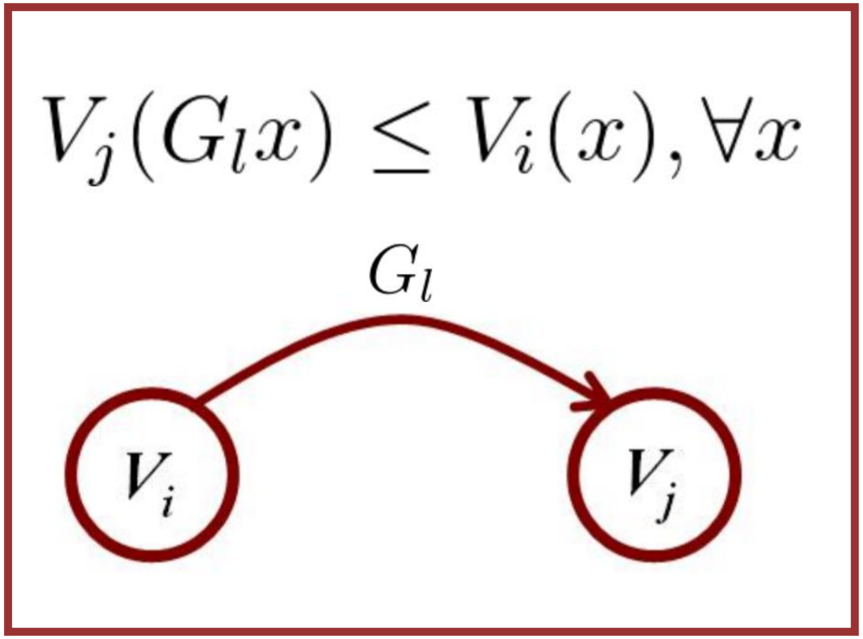
$$P \succ 0, G_i^T P G_i \preceq P, \forall i$$

Can be infeasible even if JSR<1



Multiple Lyapunov functions

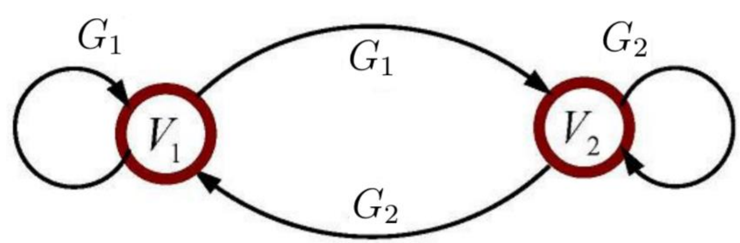
Representation of Lyapunov inequalities via labeled graphs



[AAA, Jungers, Parrilo, Roozbehani
SIAM J. on Control and Opt]

(Best SICON Paper Prize, 2013-2015)

$$
\begin{aligned}
G_1^T P_1 G_1 &\preceq P_1 \\
G_1^T P_2 G_1 &\preceq P_1 \\
G_2^T P_1 G_2 &\preceq P_2 \\
G_2^T P_2 G_2 &\preceq P_2 \\
P_{1,2} &\succ 0
\end{aligned}
$$

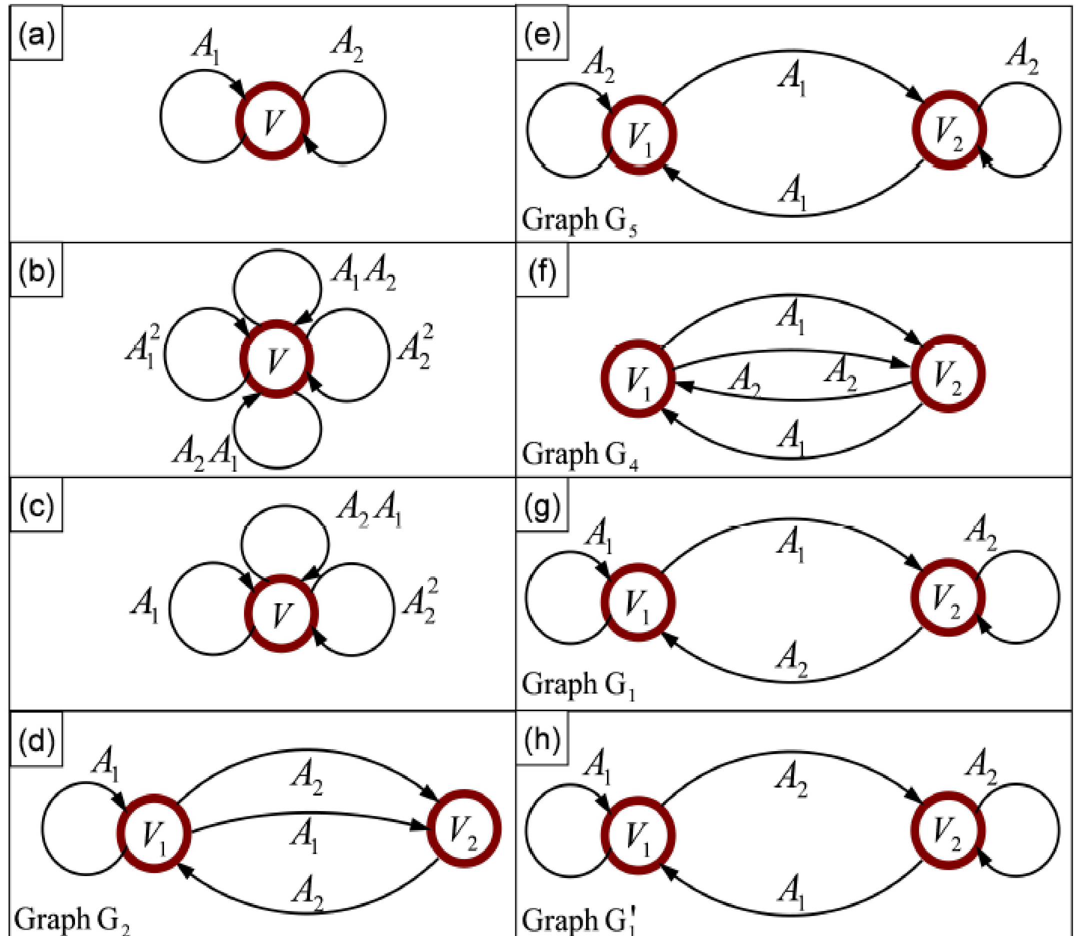


- What property of the graph implies stability?

Path-complete graphs

Defn. A labeled directed graph $G(N,E)$ is path-complete if for every word of finite length there is an associated directed path which reads that word.

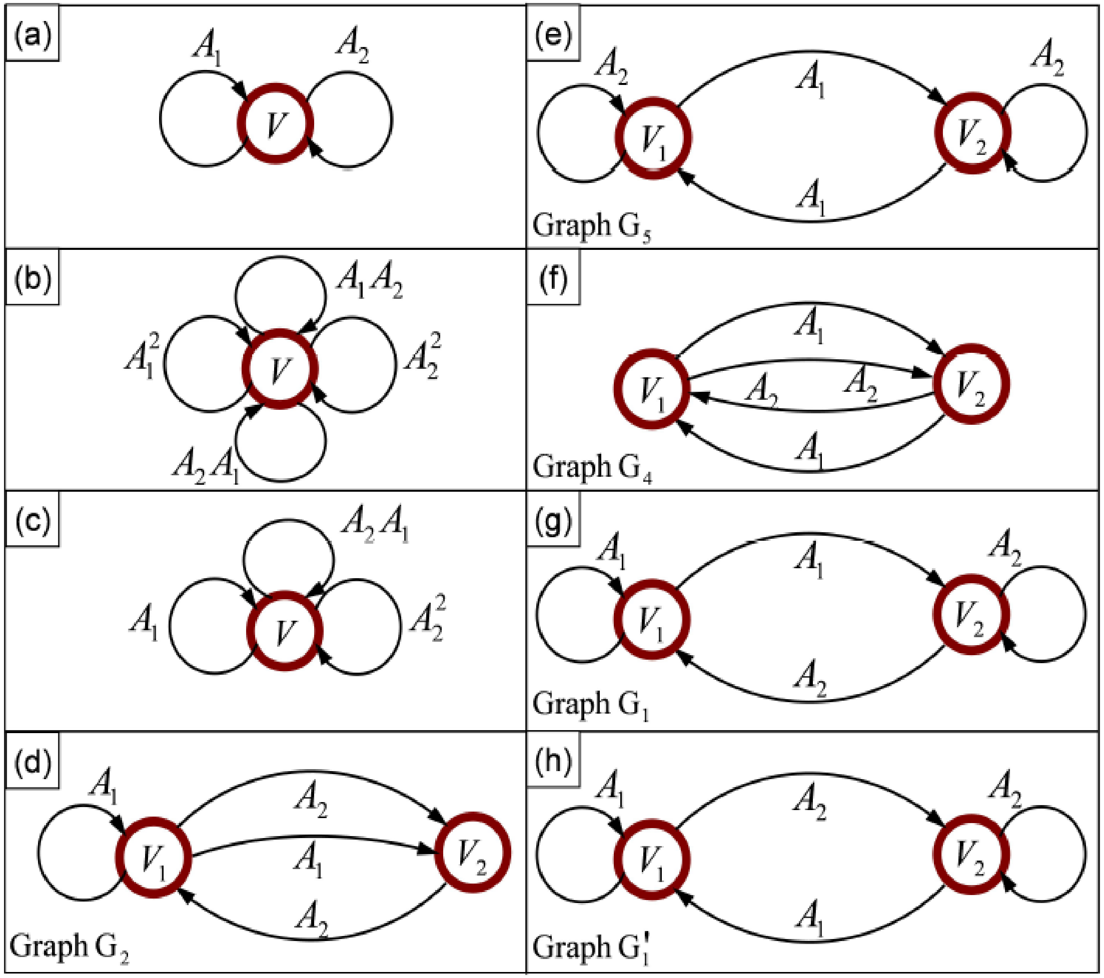
- Path-completeness can be checked with standard algorithms in **automata theory**



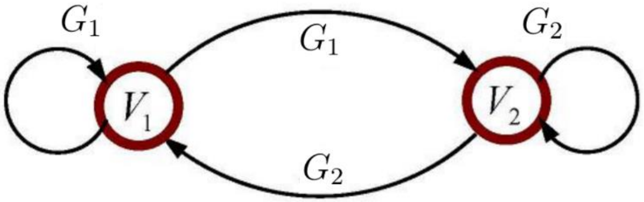
Path-complete graphs and stability

THM. If Lyapunov functions satisfying Lyapunov inequalities associated with **any path-complete graph** are found, then the dynamical system is uniformly stable (i.e., JSR<1).

- Gives immediate proofs for existing methods
- Introduces numerous new methods



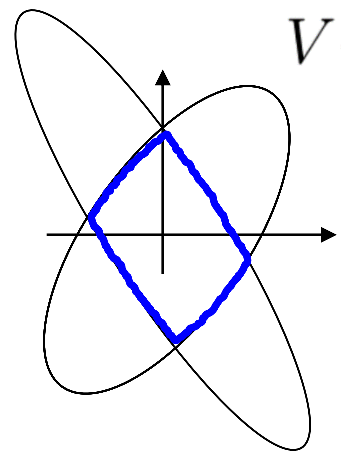
Special cases

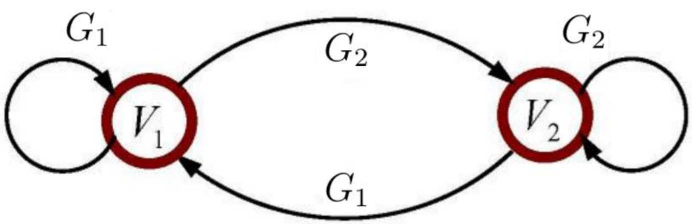


$$G_1^T P_1 G_1 \preceq P_1 \quad P_{1,2} \succ 0$$
$$G_1^T P_2 G_1 \preceq P_1$$
$$G_2^T P_1 G_2 \preceq P_2$$
$$G_2^T P_2 G_2 \preceq P_2$$

max-of-quadratics

$$V(x) = \max\{x^T P_1 x, x^T P_2 x\}$$

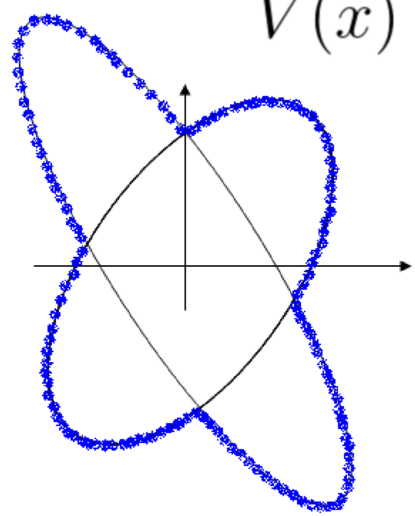




$$G_1^T P_1 G_1 \preceq P_1 \quad P_{1,2} \succ 0$$
$$G_1^T P_1 G_1 \preceq P_2$$
$$G_2^T P_2 G_2 \preceq P_1$$
$$G_2^T P_2 G_2 \preceq P_2$$

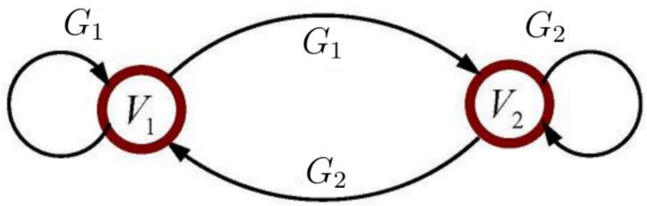
min-of-quadratics

$$V(x) = \min\{x^T P_1 x, x^T P_2 x\}$$

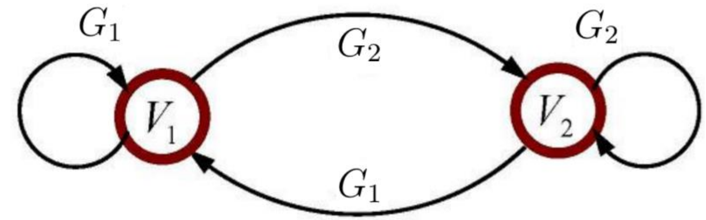


Approximation guarantees

min-of-quadratics



max-of-quadratics



$$\frac{1}{\sqrt[4]{n}}\hat{\rho}(\mathcal{G}) \leq \rho(\mathcal{G}) \leq \hat{\rho}(\mathcal{G})$$

- proof relies on the John's ellipsoid thm

THM. Given any desired accuracy

$$\frac{1}{\sqrt[2l]{n}}\hat{\rho}(\mathcal{G}) \leq \rho(\mathcal{G}) \leq \hat{\rho}(\mathcal{G})$$

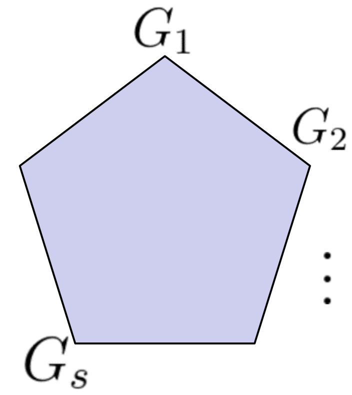
we can explicitly construct a graph G (**with s^{l-1} nodes**) such that the corresponding **SDP** achieves the accuracy.

Back to R-ULD-LP

R-ULD-LP

Robust to uncertain linear dynamics linear programming (R-ULD-LP)

$$x_{k+1} \in conv\{G_1, \ldots, G_s\} x_k$$



Models **uncertainty** and **variations with time** in the dynamics

$$\min_x \{c^T x : AGx \leq b, \forall G \in \mathbb{G}^*\}$$ (An infinite LP)

$\mathbb{G}^*$: set of all finite products of $G_1, \ldots, G_s$

Input data: $A, b, c, G_1, \ldots, G_s$

Finite convergence of outer approximations

$$\mathcal{S} := \bigcap_{k=0}^{\infty} \{x \in \mathbb{R}^n | \; AGx \leq b, \forall G \in \mathcal{G}^*\}$$

$$S_r := \bigcap_{k=0}^{r} \{x \in \mathbb{R}^n | \; AGx \leq b, \forall G \in \mathcal{G}^k\}$$

$$\mathcal{S} \subseteq \ldots S_{r+1} \subseteq S_r \subseteq \ldots \subseteq S_2 \subseteq S_1 \subseteq S_0 = P.$$

Joint spectral radius (JSR):

$$\rho(G_1, \ldots, G_s) = \lim_{k \to \infty} \max_{\sigma \in \{1,\ldots,s\}^m} ||G_{\sigma_1} \cdots G_{\sigma_k}||^{\frac{1}{k}}$$

Theorem. If $\rho(G_1, \ldots, G_s) < 1$, and $P = \{Ax \leq b\}$ is bounded and contains the origin in its interior, then $S = S_r$, for some r.

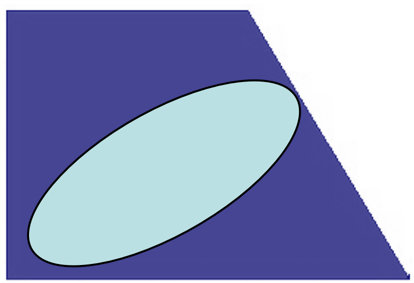
(However, number of facets of S is typically very large.)

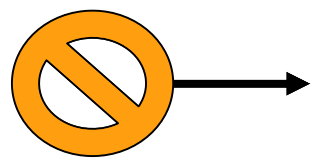
Lower and upper bounds for R-ULD-LP

- To get **lower bounds**, truncate the sequence and solve an LP. For example,

$$\min_{x}\{c^T x: Ax \leq b, AG_1 x \leq b, AG_1 G_2 x \leq b, \dots, AG_1 G_2 G_1 x \leq b\}$$

- What about **upper bounds?**





$$\min_{x,Q} \quad c^T x$$

$$Q \succ 0$$

$$G Q G^T \preceq Q$$

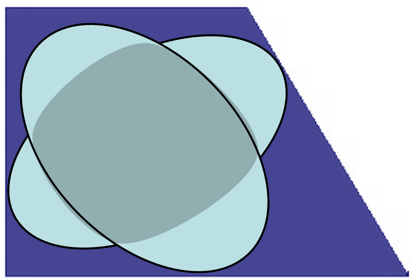
$$\begin{bmatrix} Q & x \\ x^T & 1 \end{bmatrix} \succeq 0$$

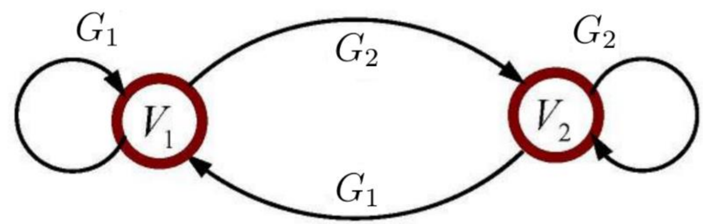
$$a_i^T Q a_i \leq 1$$

Invariant ellipsoid may not exist even when JSR<1

Idea: search for intersection of ellipsoids instead!

$$\begin{aligned}
\underset{x\in\mathbb{R}^n, Q_{1,2}\in S^{n\times n}}{\text{minimize}} \quad & c^T x \\
\text{s.t.} \quad & Q_1 \succ 0, Q_2 \succ 0 \\
& G_1 Q_1 G_1^T \preceq Q_1 \\
& G_2 Q_1 G_2^T \preceq Q_2 \\
& G_1 Q_2 G_1^T \preceq Q_1 \\
& G_2 Q_2 G_2^T \preceq Q_2 \\
& \begin{bmatrix} Q_1 & \tilde{G}x \\ (\tilde{G}x)^T & 1 \end{bmatrix} \succeq 0, \ \forall \tilde{G} \in \mathcal{G}^r \\
& \begin{bmatrix} Q_2 & \tilde{G}x \\ (\tilde{G}x)^T & 1 \end{bmatrix} \succeq 0, \ \forall \tilde{G} \in \mathcal{G}^r \\
& a_i^T Q_1 a_i \leq 1 \\
& a_i^T Q_2 a_i \leq 1 \\
& A\tilde{G}x \leq 1, \ \forall \tilde{G} \in \mathcal{G}^k, k = 0, \ldots, r-1
\end{aligned}$$

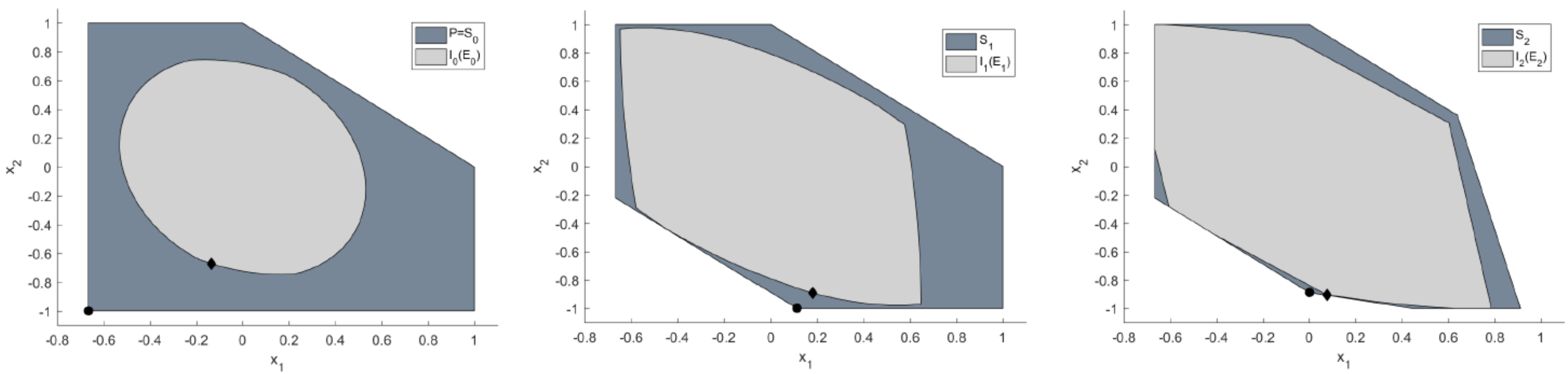




- The convexification tricks go through!
- Finite convergence of upper bounds is guaranteed.

A numerical example of an R-ULD-LP

$$A = \begin{pmatrix} 1 & 0 \\ -1.5 & 0 \\ 0 & 1 \\ 0 & -1 \\ 1 & 1 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, c = \begin{pmatrix} 0.5 \\ 1 \end{pmatrix}, G_1 = \begin{pmatrix} -1/4 & -1/4 \\ -1 & 0 \end{pmatrix}, \text{ and } G_2 = \begin{pmatrix} 3/4 & 3/4 \\ -1/2 & 1/4 \end{pmatrix}.$$



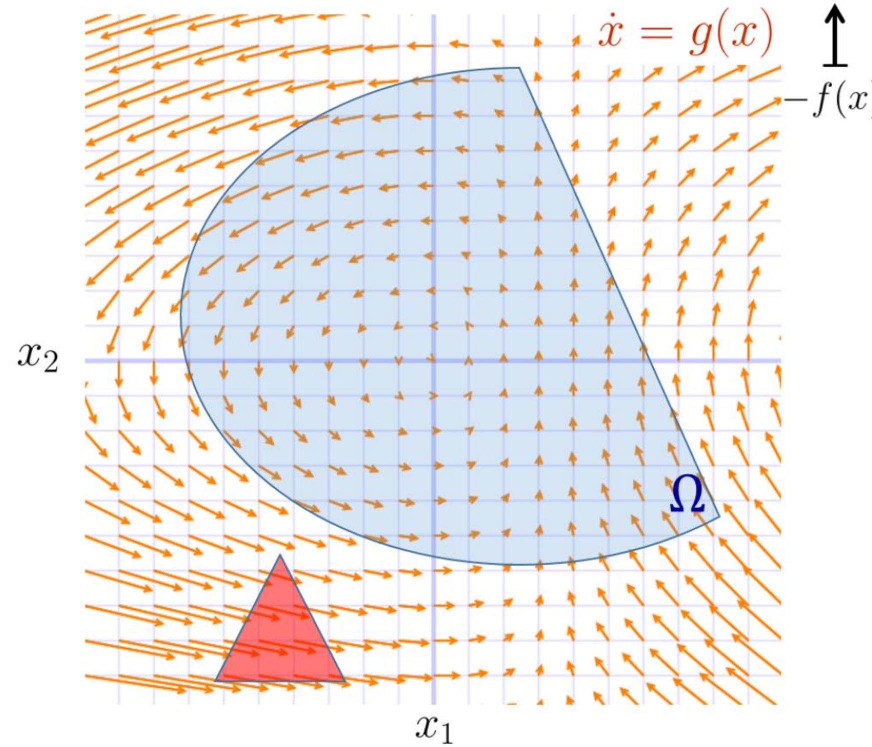
(a) $r = 0$ (b) $r = 1$ (c) $r = 2$

	$r = 0$	$r = 1$	r=2
Lower bounds	-1.3333	-0.9444	-0.8889
Upper bounds	-0.7395	-0.8029	-0.8669

The broader perspective

Optimization problems with dynamical systems (DS) constraints

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in \Omega \cap \Omega_{DS}. \end{array}$$



Optimization Problem "f, Ω"	Type of Dynamical System "g"	DS Constraint "Ω_{DS}"
Linear program*	Linear*	Invariance*
Convex quadratic program*	Linear and uncertain/stochastic	Inclusion in region of attraction
Semidefinite program	Linear and time-varying*	Collision avoidance
Robust linear program	Nonlinear (polynomial)	Reachability
Polynomial program	Nonlinear and time-varying	Orbital stability
Integer program	Discrete/continuous/hybrid of both	Stochastic stability
⋮	⋮	⋮

CDC'15 + Journal preprint in preparation: http://aaa.princeton.edu