

Computation of the Joint Spectral Radius with Optimization Techniques

Amir Ali Ahmadi

Goldstine Fellow, IBM Watson Research Center

Joint work with:

- **Raphaël Jungers** (UC Louvain)
- **Pablo A. Parrilo** (MIT)
- **R. Jungers, P.A. Parrilo, Mardavij Roozbehani** (MIT)

The Joint Spectral Radius

Given a finite set of $n \times n$ matrices

$$\mathcal{A} := \{A_1, \dots, A_m\}$$

Joint spectral radius (JSR):

$$\rho(\mathcal{A}) = \lim_{k \rightarrow \infty} \max_{\sigma \in \{1, \dots, m\}^k} \|A_{\sigma_k} \dots A_{\sigma_2} A_{\sigma_1}\|^{1/k}$$

If only one matrix:

$$\mathcal{A} = \{A\}$$

Spectral Radius

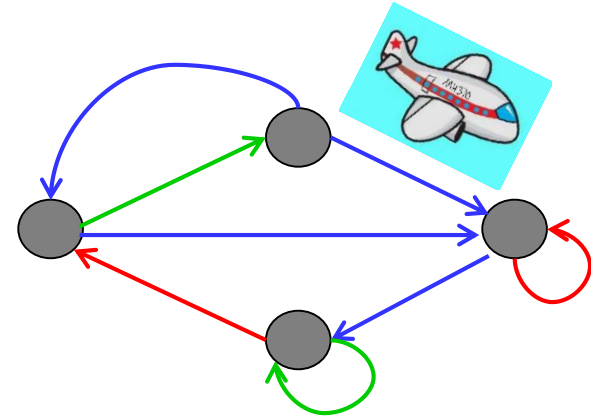
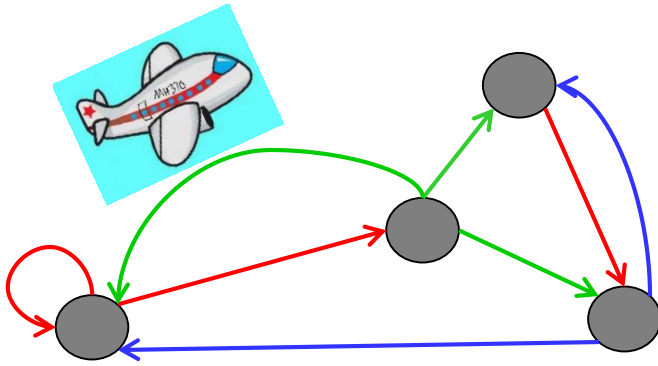
$$\rho(A) = \lim_{k \rightarrow \infty} \|A^k\|^{1/k}$$



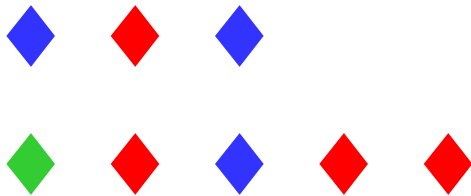
G. C. Rota and W. G. Strang
A note on the joint spectral radius
Indag. Math., 22:379–381, 1960.2



Trackability of Graphs



Noisy observations:



How does the number of possible paths grow with length of observation?

$N(t)$: max. number of possible paths over all observations of length t

Graph is called **trackable** if $N(t)$ is bounded by a polynomial in t

$$\mathcal{A} = \{A_1, A_2, A_3\}$$

$$\rho(\mathcal{A}) = \lim_{t \rightarrow \infty} N(t)^{1/t}$$

Graph **trackable** iff $\rho(\mathcal{A}) \leq 1$



JSR and Switched/Uncertain Linear Systems

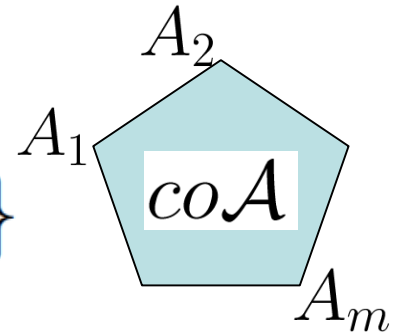
Linear dynamics: $x_{k+1} = Ax_k$

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

“Stable” iff $\rho(A) < 1$

Switched linear dynamics: $x_{k+1} = A_i x_k$

$$\mathcal{A} := \{A_1, \dots, A_m\}$$

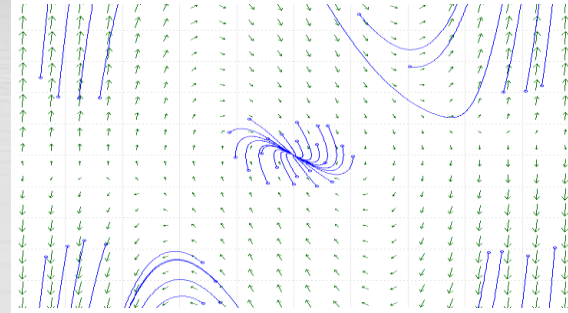
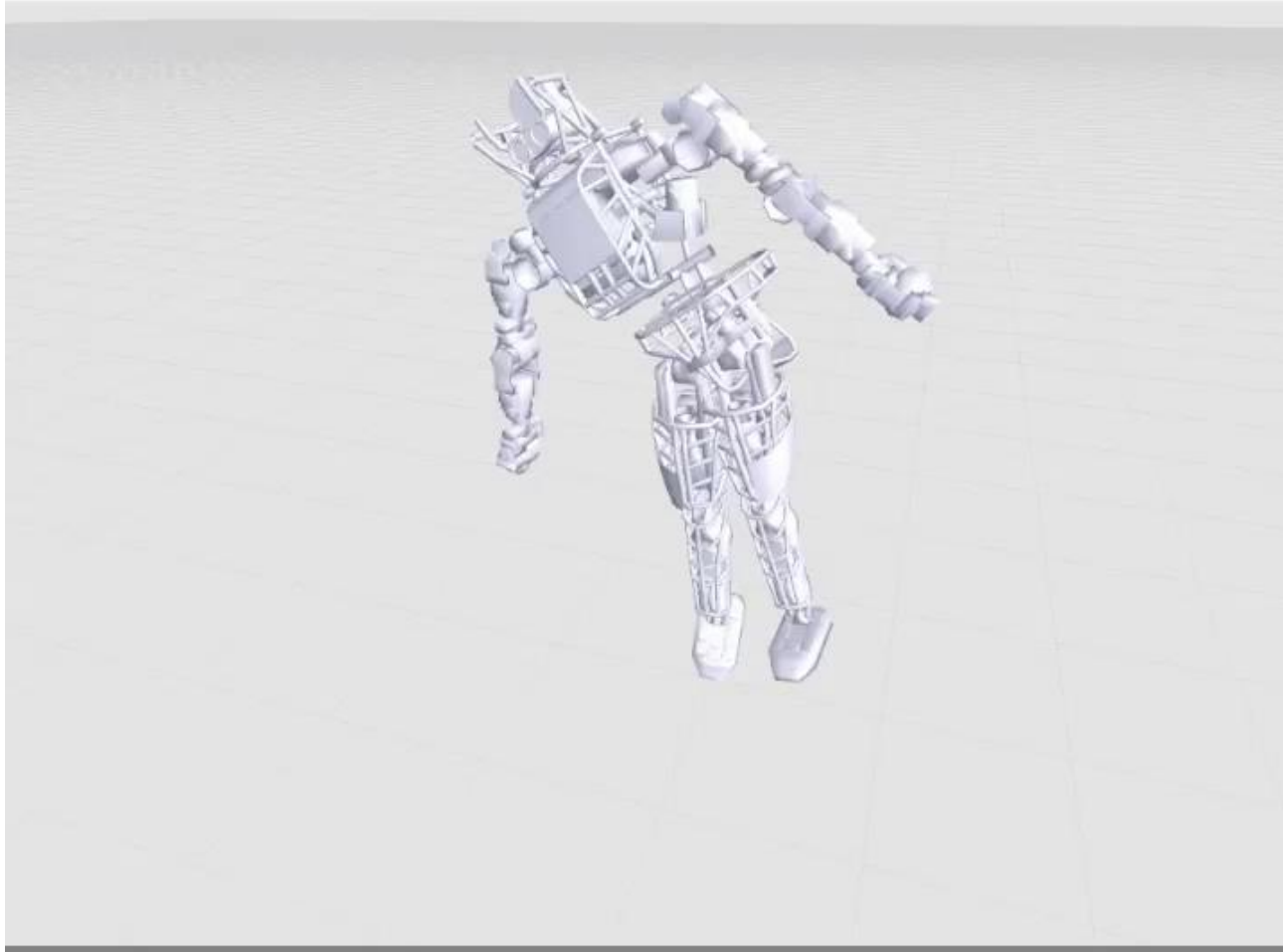


Joint spectral radius (JSR):

$$\rho(\mathcal{A}) = \lim_{k \rightarrow \infty} \max_{\sigma \in \{1, \dots, m\}^k} \|A_{\sigma_k} \dots A_{\sigma_2} A_{\sigma_1}\|^{1/k}$$

“Uniformly stable” iff $\rho(\mathcal{A}) < 1$

Stability around an equilibrium point



Controller design for this humanoid presented in:

[Majumdar, AAA, Tedrake, CDC'14 – submitted]

Done by **SDSOS Optimization**
[AAA, Majumdar,'13]



Computation of JSR

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

Testing “ $\rho(A) < 1$?” decidable in polynomial time

For more than one matrix:

Testing “ $\rho(\mathcal{A}) \leq 1$?” undecidable [Blondel, Tsitsiklis]

(even for 2 matrices of size 47x47 !!)

- **Open problem:** decidability of testing “ $\rho(\mathcal{A}) < 1$?”
- Would become decidable if *rational finiteness conjecture* is true
- **Lower bounds on JSR:** $\rho(A_{\sigma_k} \cdots A_{\sigma_1})^{1/k} \leq \rho(\mathcal{A})$
 - **Finiteness conjecture:** equality achieved at finite k
- **Upper bounds on JSR:** from Lyapunov theory



This Talk

1. A meta-SDP-algorithm for computing upper bounds

Lyapunov theory + basic automata theory

2. Exact JSR of rank-one matrices

via dynamic programming

3. Uncertain nonlinear systems

SOS-convex Lyapunov functions

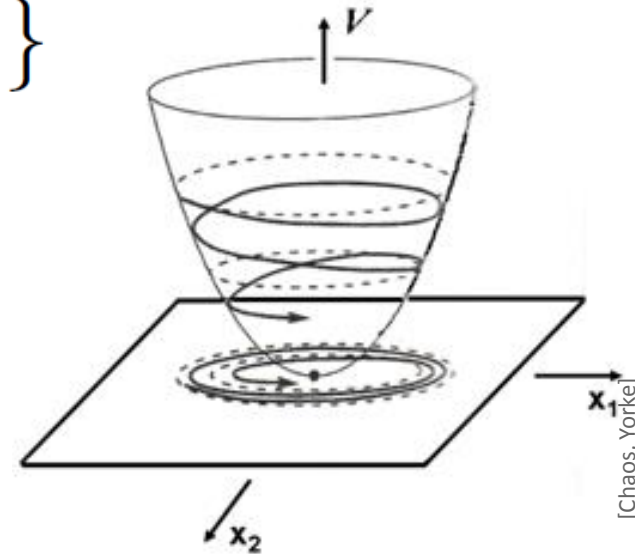


Common Lyapunov function

$$x_{k+1} = A_i x_k$$

$$\mathcal{A} := \{A_1, \dots, A_m\}$$

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



[Chaos, Yorke]

such that

$$V(x) > 0,$$
$$V(A_i x) < V(x), \quad \forall i = 1, \dots, m$$

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

Such a function always exists! But may be extremely difficult to find!!



Computationally-friendly common Lyapunov functions

$$x_{k+1} = A_i x_k \quad \mathcal{A} := \{A_1, \dots, A_m\}$$

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

such that

$$V(x) > 0,$$

$$V(A_i x) < V(x), \quad \forall i = 1, \dots, m$$

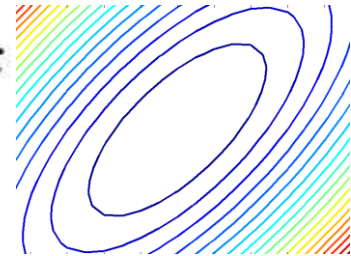
then, $\rho(\mathcal{A}) < 1$

■ Common quadratic Lyapunov function: $V(x) = x^T P x$

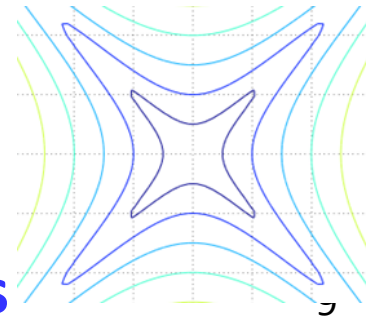
$$A_i^T P A_i \preceq P \quad \forall i = 1, \dots, m$$

[Ando, Shih]

[Blondel, Nesterov, Theys]



■ Common SOS Lyapunov function [Parrilo, Jadbabaie]



Our approach: use multiple Lyapunov functions



Multiple Lyapunov functions

■ Can we do better with more than one Lyapunov function?

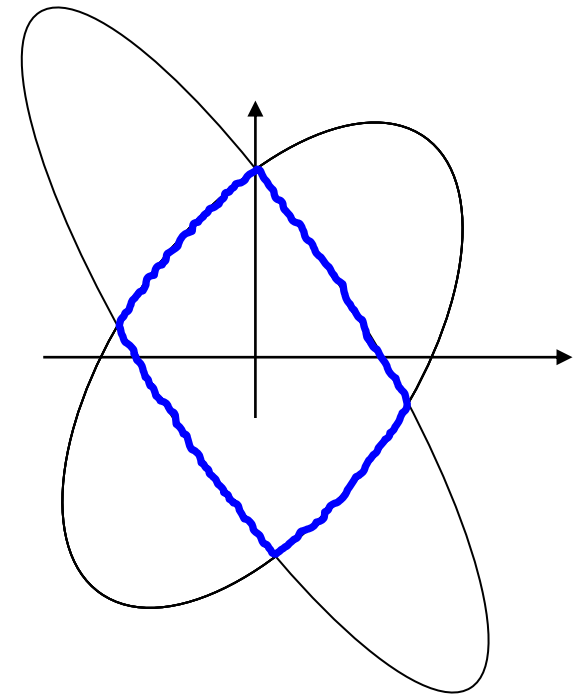
■ How?

■ Consider the SDP:

$$\begin{array}{l} A_1^T P_1 A_1 \preceq P_1, \\ A_2^T P_1 A_2 \preceq P_2, \\ A_1^T P_2 A_1 \preceq P_1, \\ A_2^T P_2 A_2 \preceq P_2, \\ P_i \succeq 0. \end{array}$$

→ $\rho(\mathcal{A}) \leq 1$

max-of-quadratics



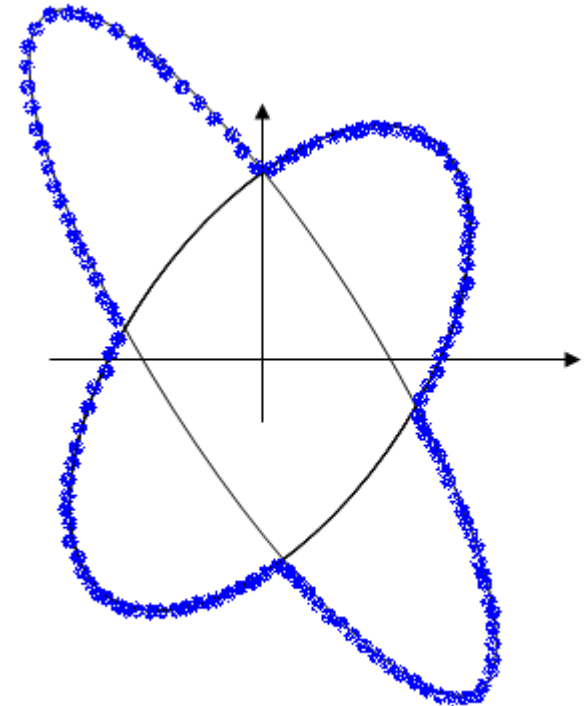
Multiple Lyapunov functions

- Consider another SDP:

$$\begin{array}{lcl} A_1^T P_1 A_1 & \preceq & P_1, \\ A_2^T P_2 A_2 & \preceq & P_1, \\ A_1^T P_1 A_1 & \preceq & P_2, \\ A_2^T P_2 A_2 & \preceq & P_2, \\ P_i & \succeq & 0. \end{array}$$

 $\rho(\mathcal{A}) \leq 1$

min-of-quadratics



Even stranger SDPs...

- Feasibility of the following SDP also implies stability:

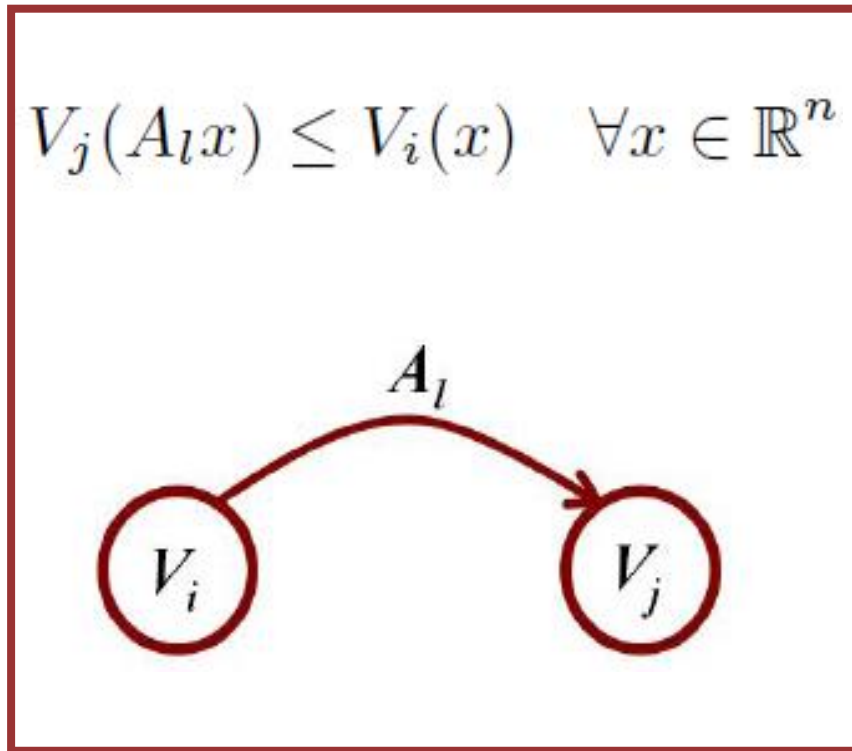
$$\begin{array}{l} A_1^T P A_1 \\ (A_2 A_1)^T P (A_2 A_1) \\ (A_2^2)^T P (A_2^2) \\ P \end{array} \preceq \begin{array}{l} P, \\ P, \\ P, \\ 0. \end{array}$$

 $\rho(\mathcal{A}) \leq 1$

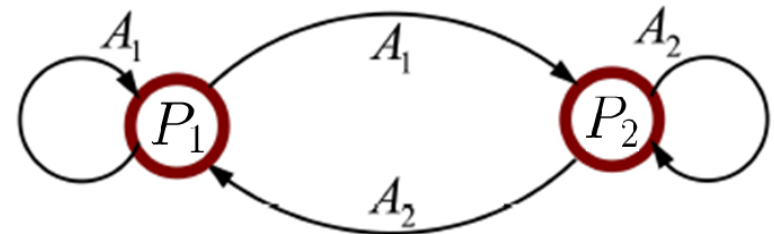
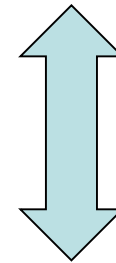
- Where do these conditions come from?
- Can we give a unifying framework?



Representation of Lyapunov inequalities via labeled graphs



$$\begin{aligned} A_1^T P_1 A_1 &\preceq P_1 \\ A_1^T P_2 A_1 &\preceq P_1 \\ A_2^T P_1 A_2 &\preceq P_2 \\ A_2^T P_2 A_2 &\preceq P_2 \\ P_{1,2} &\succ 0 \end{aligned}$$



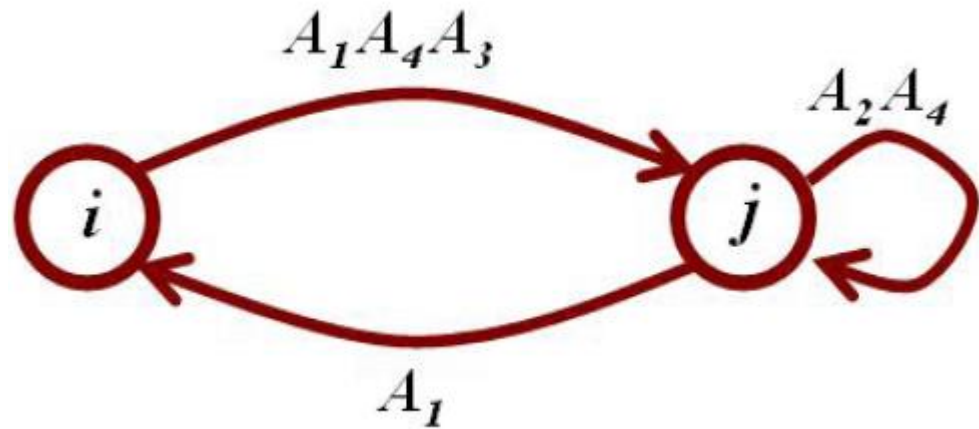
[AAA, Jungers, Parrilo, Roazbehani
SIAM J. on Control and Opt., '13]

- What property of the graph implies stability?

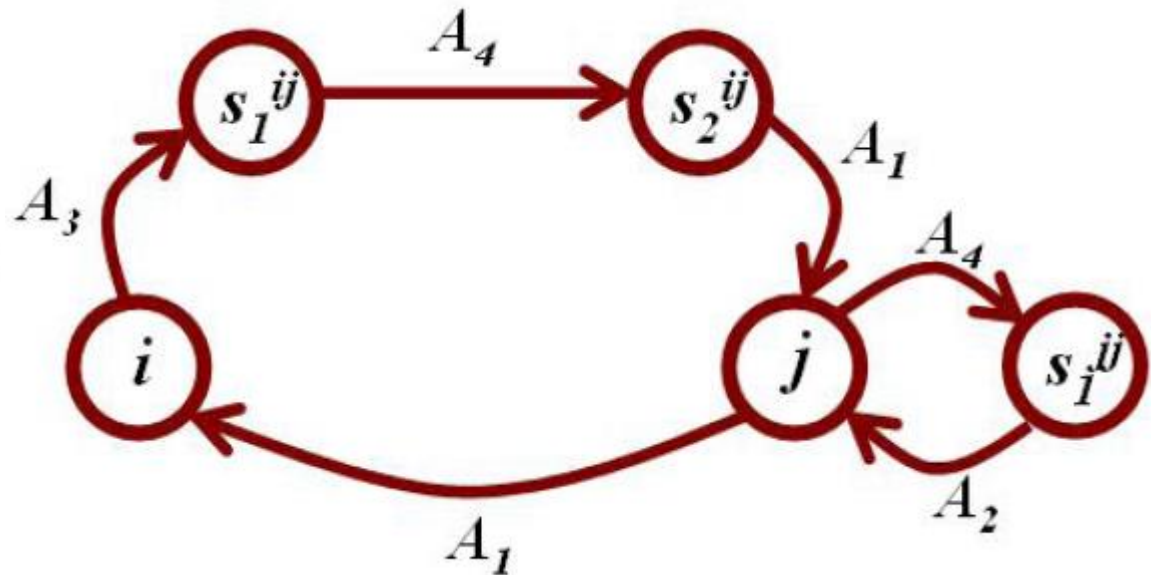


Graph expansion

Graph $G(N,E)$



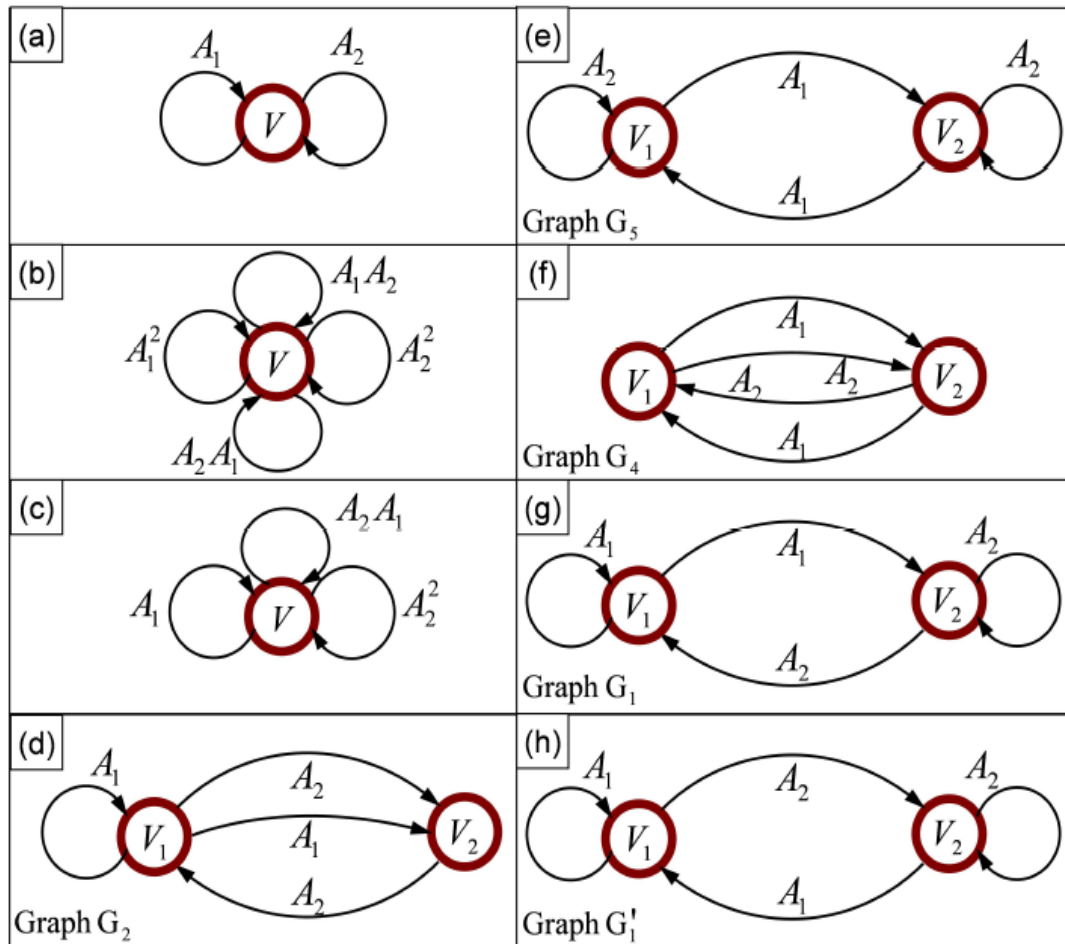
Expanded Graph $G^e(N^e, E^e)$



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 in its expanded graph $G^e(N^e,E^e)$.

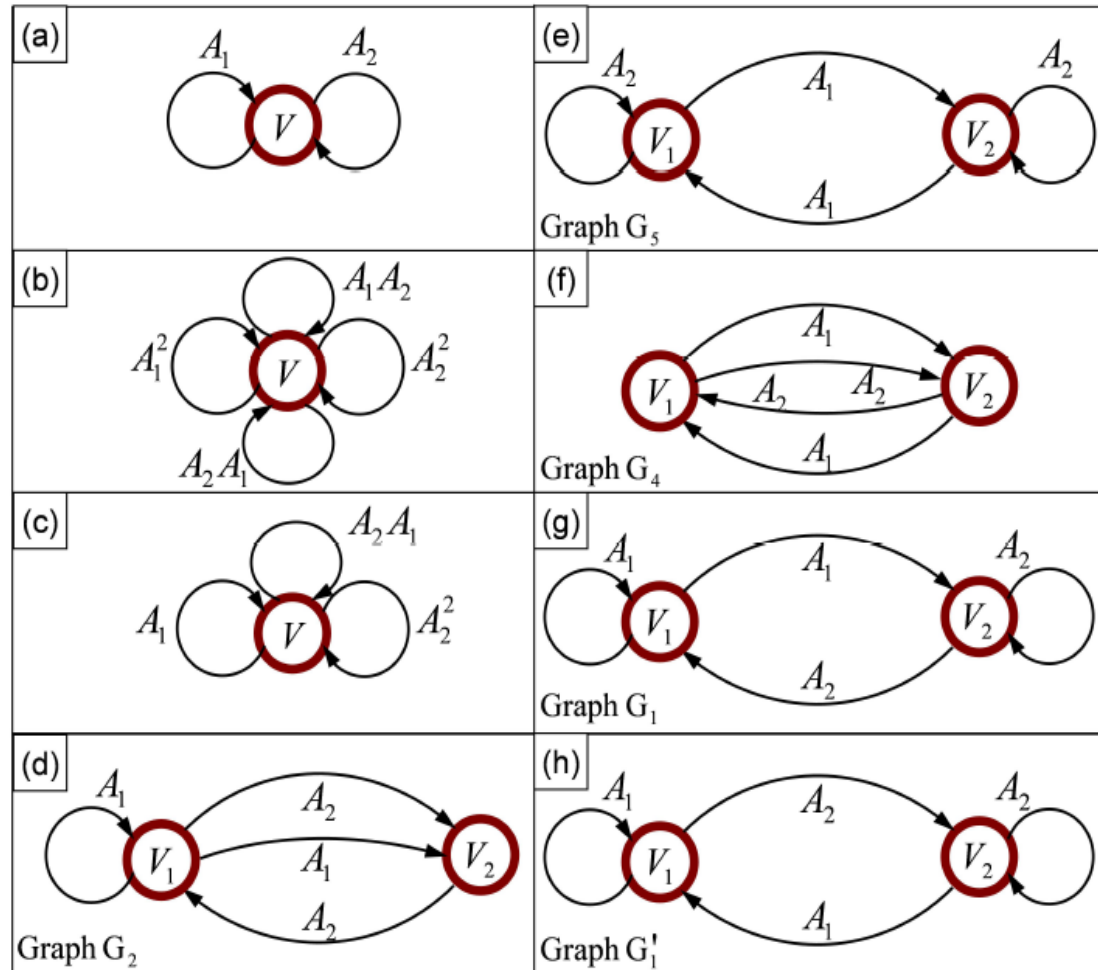
■ Path-completeness can be checked with standard algorithms in **automata theory**



Path-complete graphs and stability

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

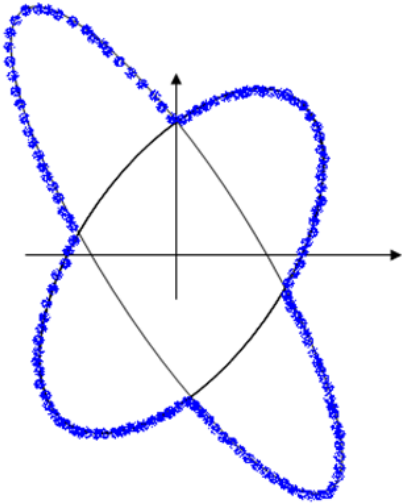
- Gives immediate proofs for existing methods
- Introduces numerous new methods



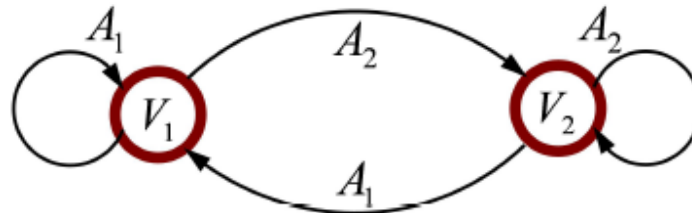
Quick proofs

For example:

min-of-quadratics

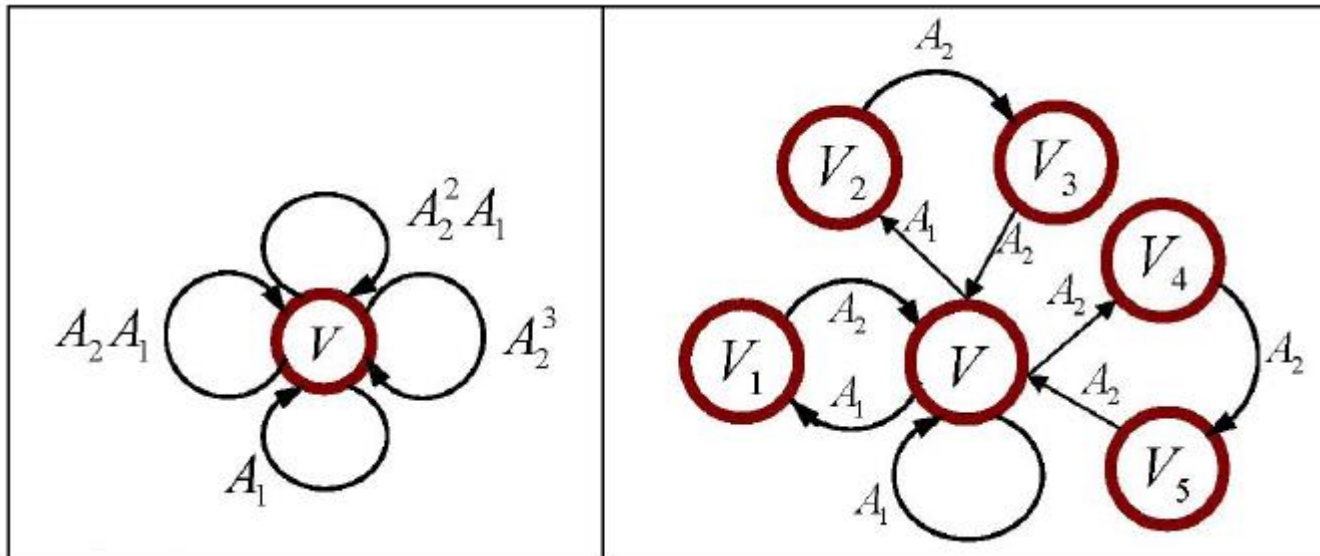


$$\begin{array}{lcl} A_1^T P_1 A_1 & \preceq & P_1 \\ A_2^T P_2 A_2 & \preceq & P_1 \\ A_1^T P_1 A_1 & \preceq & P_2 \\ A_2^T P_2 A_2 & \preceq & P_2 \end{array}$$



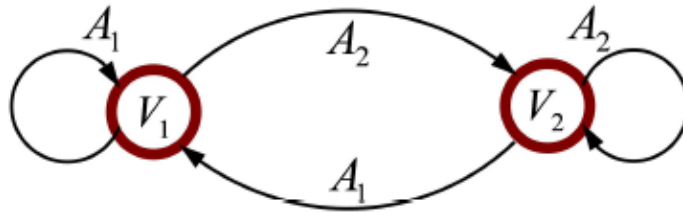
Let's revisit our strange SDP

$$\begin{aligned}
 A_1^T P A_1 &\preceq P \\
 (A_2 A_1)^T P (A_2 A_1) &\preceq P \\
 (A_2^2 A_1)^T P (A_2^2 A_1) &\preceq P \\
 (A_2^3)^T P (A_2^3) &\preceq P
 \end{aligned}$$

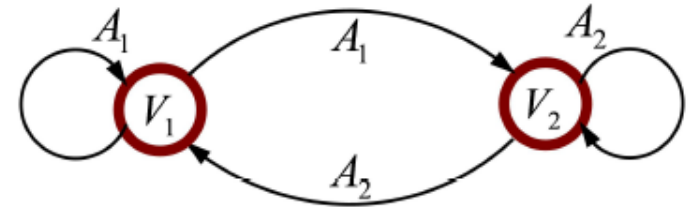


Approximation guarantees

min-of-quadratics



max-of-quadratics



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

- tighter than known SOS bounds
- proof relies on the John's ellipsoid thm

THM. Given any desired accuracy

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

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



No bound on size of SDP

THM. [AAA, Jungers, IFAC'14]

Given any positive integer d , there are families of switched systems that are uniformly stable (i.e., have $\text{JSR} < 1$), but yet this fact **cannot be proven with**

- a polynomial Lyapunov function of degree $\leq d$
- a max-of-quadratics Lyapunov function with $\leq d$ pieces
- a min-of-quadratics Lyapunov function with $\leq d$ pieces
- a polytopic Lyapunov function with $\leq d$ facets.

Kozyakin'90:

$$A_1 = \frac{(1 - t^4)}{(1 - 3\pi t^3/2)} \begin{bmatrix} \sqrt{1 - t^2} & -t \\ 0 & 0 \end{bmatrix}, \quad A_2 = (1 - t^4) \begin{bmatrix} \sqrt{1 - t^2} & -t \\ t & \sqrt{1 - t^2} \end{bmatrix}$$

$$t = \sin \frac{2\pi}{2k+1} \quad \text{JSR} < 1$$

$$t = \sin \frac{2\pi}{2k} \quad \text{JSR} > 1$$



JSR of Rank One Matrices and the Maximum Cycle Mean Problem

[AAA, Parrilo, *IEEE Conf. on Decision and Control*, '12]

Basic facts about rank one matrices

A rank one iff $A = xy^T$

spectral radius: $|y^T x|$

Products of rank one matrices have rank at most one:

$$\begin{aligned} A_i A_j &= x_i y_i^T x_j y_j^T \\ &= (y_i^T x_j) x_i y_j^T \end{aligned}$$



Cycles and cycle gains

Easy definitions:

- Cycle
- Simple cycle
- Cycle gain

$$\rho_c = \prod_{i=1}^k w(e_i)$$

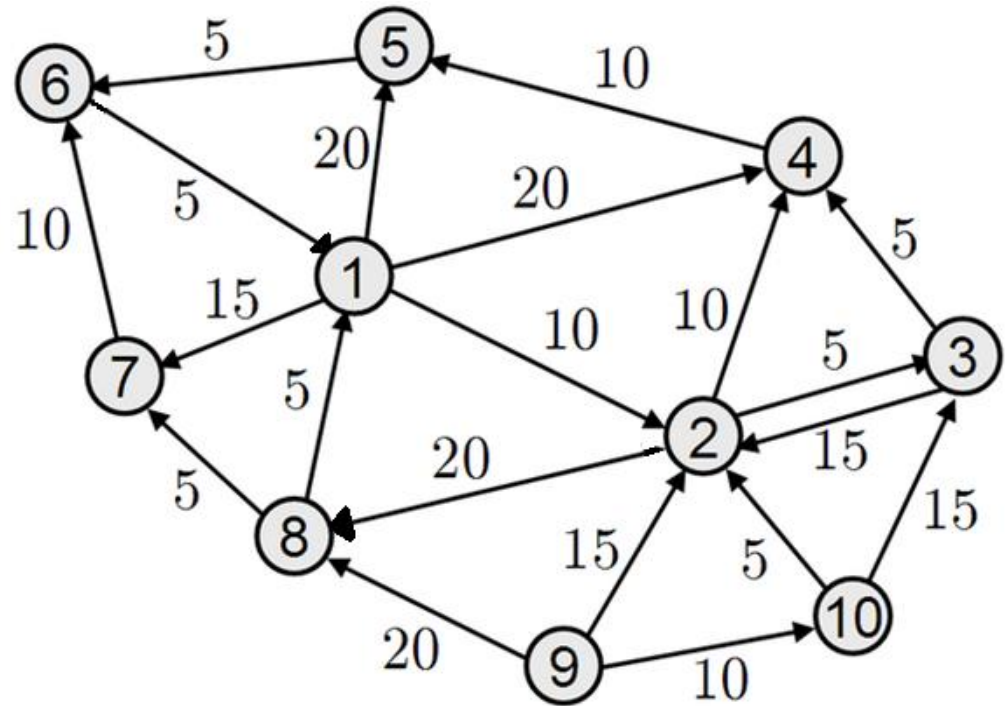
$$g(c) = \rho_c^{1/k}$$

- Maximum cycle gain

$$\max_c g(c)$$

- Gain-maximizing cycle

$$c_{max}$$



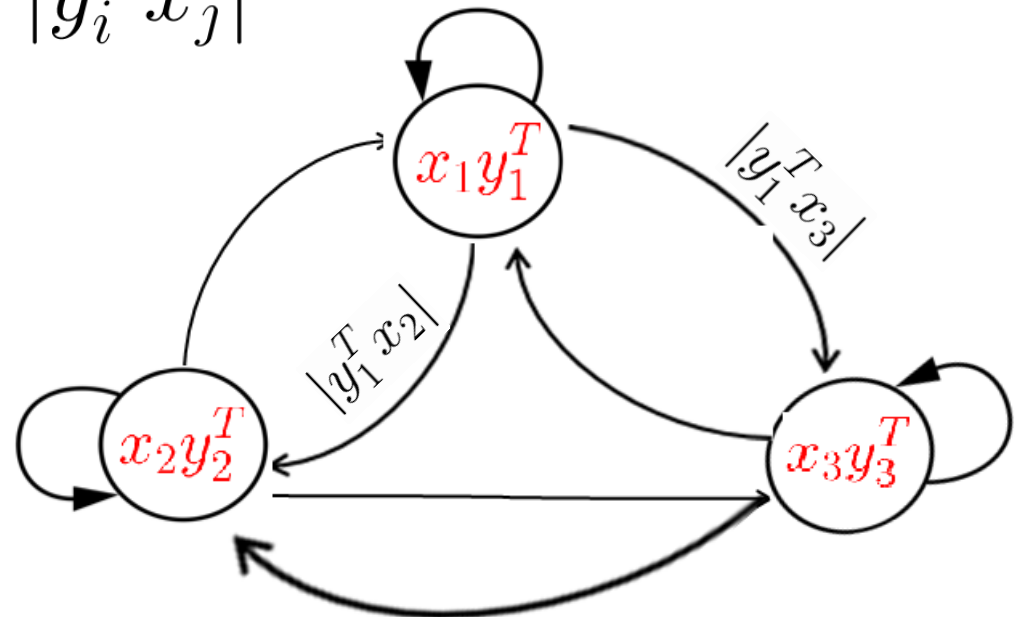
From matrix products to cycles in graphs

$$\mathcal{A} = \{A_1, \dots, A_m\} \quad A_i = x_i y_i^T$$

$G_{\mathcal{A}}$ Complete directed graph on m nodes:

Nodes: matrices A_i

Edge weights: $w(e_{ij}) = |y_i^T x_j|$



Maximum cycle gain gives the JSR

Thm: Let c_{max} be a gain-maximizing cycle, with l_{max} and $\rho_{c_{max}}$ denoting its length and the product of the weights on its edges, respectively.

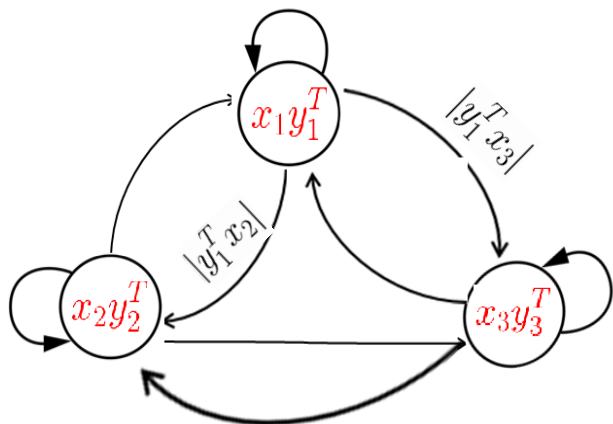
Then, the joint spectral radius is given by:

$$\rho(\mathcal{A}) = \rho_{c_{max}}^{1/l_{max}}$$

Proof sketch:

$$\rho(A_{\sigma_k} \cdots A_{\sigma_1}) = \rho_c$$

$$\begin{aligned} \rho(A_{\sigma_1} \cdots A_{\sigma_k})^{1/k} &= \left(\prod_{i=1}^s \rho_{c_i}^{m_i} \right)^{1/k} \\ &= \prod_{i=1}^s (\rho_{c_i}^{1/l_i})^{m_i l_i / k} \\ &\leq \rho_{c_{max}}^{1/l_{max}}, \end{aligned}$$



$$\rho(\mathcal{A}) = \limsup_{k \rightarrow \infty} \max_{A \in \mathcal{A}^k} \rho^{1/k}(A)$$



Finiteness property and the optimal product

Corollary:

- The JSR is achieved by the spectral radius of a **finite matrix product, of length at most m** .
(In particular, the finiteness property holds. – independently shown by Gurvits et al.)
- There always exists an optimal product where no matrix appears more than once.

Proof: A simple cycle does not visit a node twice. ■

Bound of m is tight.



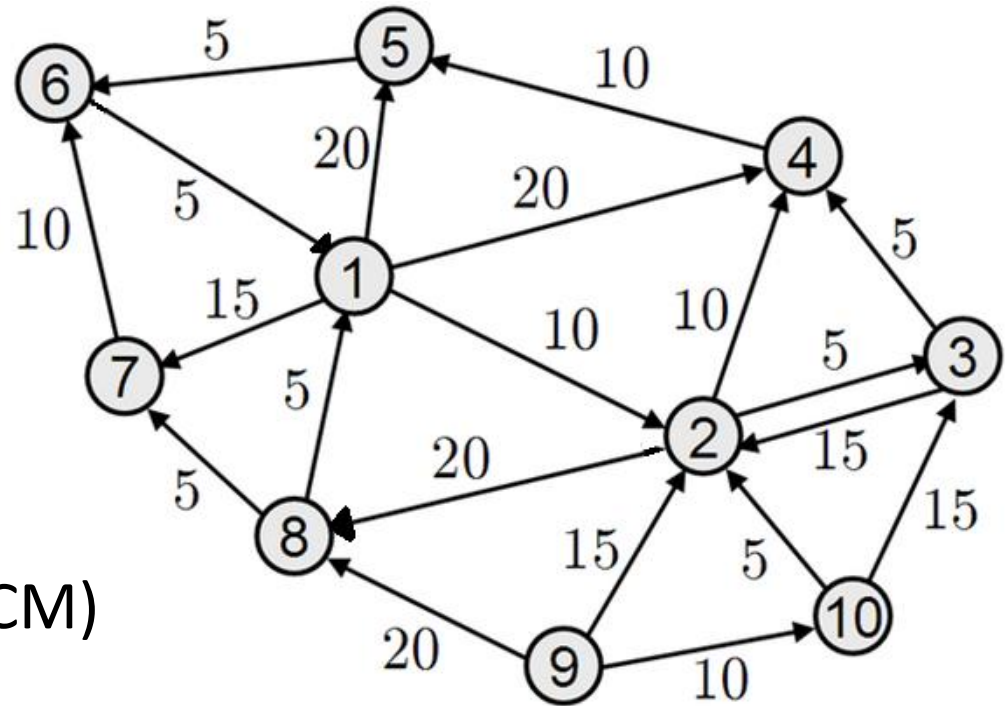
Maximum Cycle Mean Problem (MCMP)

- Cycle mean

$$m(c) = \sum_{i=1}^k \frac{w(e_i)}{k}$$

- Maximum cycle mean (MCM)

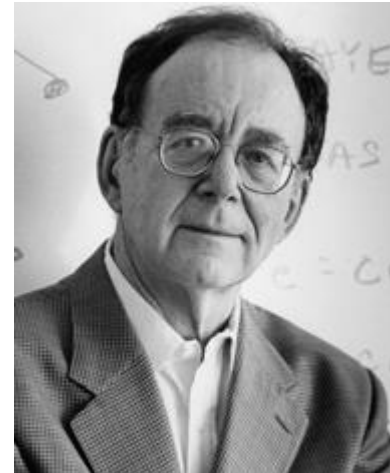
$$\lambda^* = \max_c m(c)$$



Karp's algorithm for MCMP

- Proposed by Karp in 1978, based on *dynamic programming*
- Let s be an arbitrary vertex
- For every vertex v and integer k , define $F_k(v)$ as the minimum weight of an edge progression of length k from s to v
- $F_k(v)$ can be computed via a simple DP recursion
- From this, the MCM can be computed as:

$$\min_v \max_{0 \leq k \leq n-1} \left[\frac{F_n(v) - F_k(v)}{n - k} \right]$$



- Run time $O(|N| |E|)$

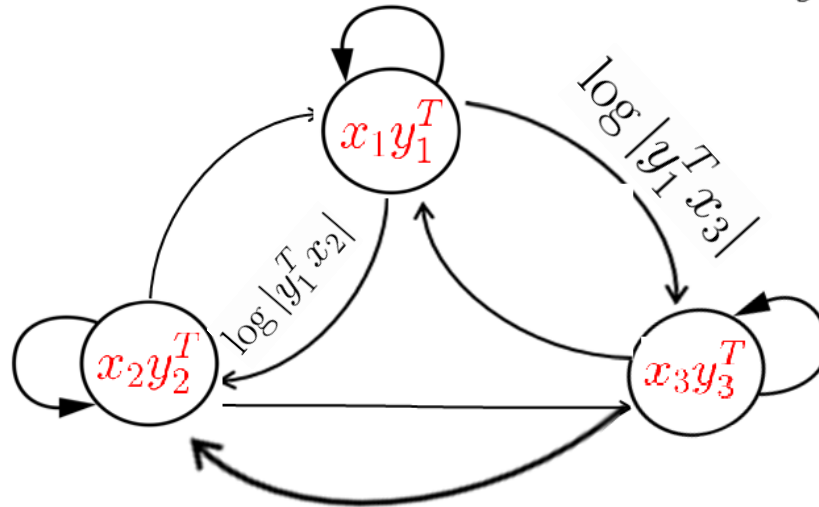


Take logs and apply Karp

$$\mathcal{A} = \{A_1, \dots, A_m\} \quad A_i = x_i y_i^T$$

$\tilde{\mathbf{G}}_{\mathcal{A}}$ Complete directed graph on m nodes:

Nodes: matrices A_i Edge weights: $w(e_{ij}) = \log |y_i^T x_j|$



$$\rho(\mathcal{A}) = e^{\lambda^* / k^*}$$

Run time: $O(m^3 + m^2n)$



Common quadratic Lyapunov function can fail

$$A_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}$$

- $\rho(\mathcal{A}) = 1$ (can be proven e.g. using our algorithm)
- An LMI searching for a common quadratic Lyapunov function can only prove

$$\rho(\mathcal{A}) \leq \sqrt{2}$$



Nonlinear Switched Systems & SOS-Convex Lyapunov Functions

[AAA, Jungers, *IEEE Conf. on Decision and Control*, '13]

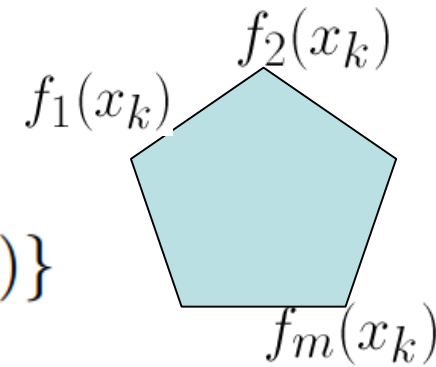


Nonlinear switched systems

$$x_{k+1} = \tilde{f}(x_k)$$

$$\tilde{f}(x_k) \in \text{conv}\{f_1(x_k), \dots, f_m(x_k)\}$$

$$f_1, \dots, f_m : \mathbb{R}^n \rightarrow \mathbb{R}^n$$



Lemma:

Unlike the linear case, a common Lyapunov function for the corners **does not imply** stability of the convex hull.

Ex.

Common Lyapunov function:

$$f_1(x) = (x_1 x_2, 0)^T$$

$$V(x) = x_1^2 x_2^2 + (x_1^2 + x_2^2)$$

$$f_2(x) = (0, x_1 x_2)^T$$

$$V(f_i(x)) = x_1^2 x_2^2 < V(x) = x_1^2 x_2^2 + (x_1^2 + x_2^2)$$

But unstable:

$$f(x) = \left(\frac{x_1 x_2}{2}, \frac{x_1 x_2}{2} \right) \in \text{conv}\{f_1(x_k), f_2(x_k)\}$$

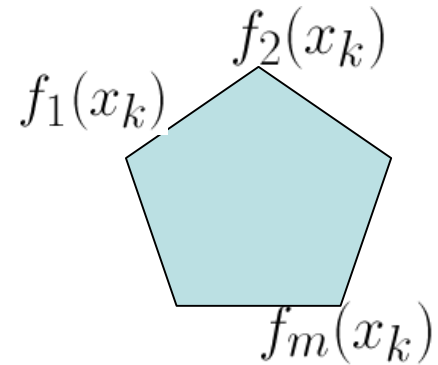


But a convex Lyapunov function implies stability

$$x_{k+1} = \tilde{f}(x_k)$$

$$\tilde{f}(x_k) \in \text{conv}\{f_1(x_k), \dots, f_m(x_k)\}$$

$$f_1, \dots, f_m : \mathbb{R}^n \rightarrow \mathbb{R}^n$$



Suppose we can find a **convex** common Lyapunov function:

$$V(x) > 0, \quad V(f_i(x)) < V(x) \text{ for } i = 1, \dots, m$$

Then, then we have stability of the convex hull.

Proof:
$$V(\tilde{f}(x)) = V\left(\sum_{i=1}^m \alpha_i f_i(x)\right) \leq \sum_{i=1}^m \alpha_i V(f_i(x)) < V(x)$$



SOS-Convexity

sos-convex
polynomial: $y^T H(x)y$ sos (Helton & Nie)

sos-convex
Lyapunov function: $V(x)$ SOS-CONVEX
 $V(x) - V(f_i(x))$ SOS

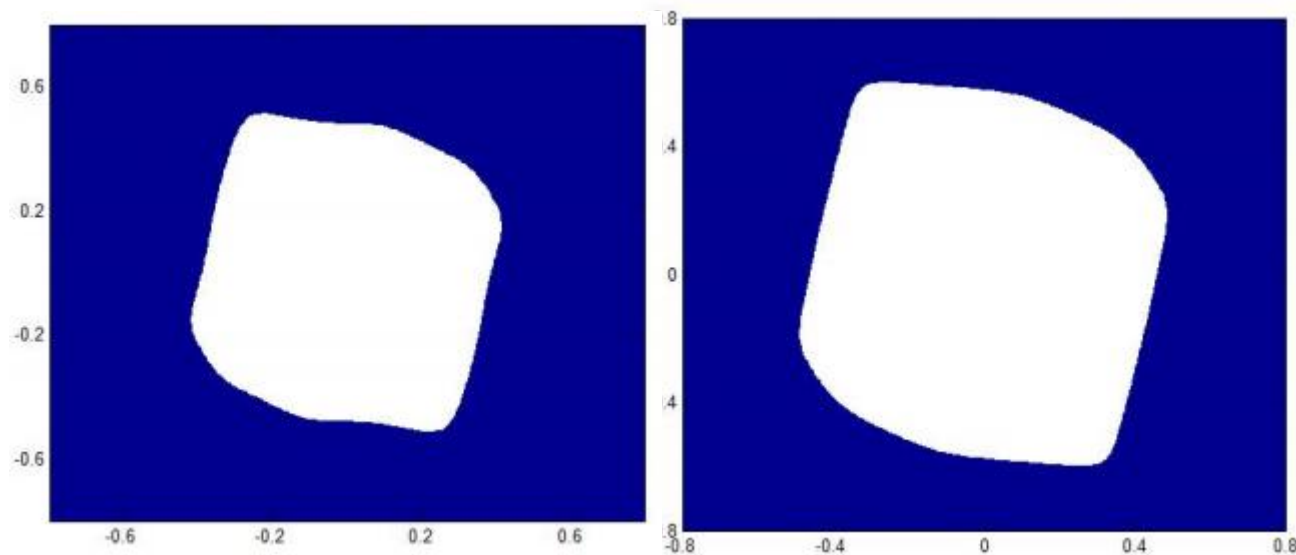
- Search for an sos-convex Lyapunov function is an SDP!
- But except for some specific degrees and dimensions, there are convex polynomials that are not sos-convex:

$$\begin{aligned} p(x) = & 32x_1^8 + 118x_1^6x_2^2 + 40x_1^6x_3^2 + 25x_1^4x_2^4 - 43x_1^4x_2^2x_3^2 \\ & - 35x_1^4x_3^4 + 3x_1^2x_2^4x_3^2 - 16x_1^2x_2^2x_3^4 + 24x_1^2x_3^6 + 16x_2^8 \\ & + 44x_2^6x_3^2 + 70x_2^4x_3^4 + 60x_2^2x_3^6 + 30x_3^8 \end{aligned}$$



ROA Computation via SDP

$$f_1(x) = \begin{pmatrix} 0.687x_1 + 0.558x_2 - .0001x_1x_2 \\ -0.292x_1 + 0.773x_2 \end{pmatrix}$$
$$f_2(x) = \begin{pmatrix} 0.369x_1 + 0.532x_2 - .0001x_1^2 \\ -1.27x_1 + 0.12x_2 - .0001x_1x_2 \end{pmatrix}$$



non-convex, deg=12

sos-convex, deg=14

- **Left:**
Cannot make any statements about ROA
- **Right:**
Level set is part of ROA under arbitrary switching



A converse Lyapunov theorem

$$x_{k+1} \in \text{conv}\{A_i x_k\}, \quad i = 1, \dots, m$$

Thm: SOS-convex Lyapunov functions are *universal* (i.e., *necessary and sufficient*) for stability.

Proof idea:

- Approximate original Lyapunov function with convex polynomials
- In a second step, go from convex to sos-convex
 - Uses a [Positivstellensatz result of Claus Scheiderer](#):

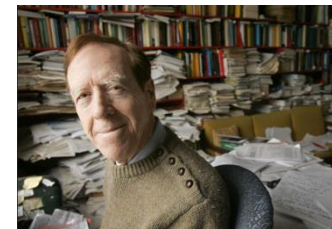
Given any two positive definite forms g and h , there exists an integer k such that $g \cdot h^k$ is sos.



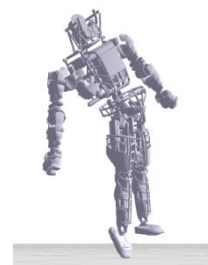
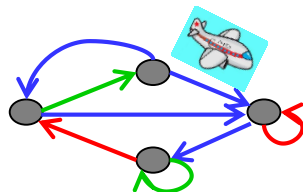
Joint Spectral Radius



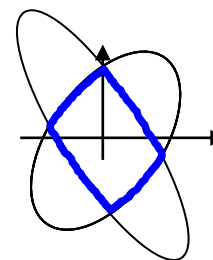
$$\rho(\mathcal{A}) = \lim_{k \rightarrow \infty} \max_{\sigma \in \{1, \dots, m\}^k} \|A_{\sigma_k} \dots A_{\sigma_2} A_{\sigma_1}\|^{1/k}$$



Has lots of applications...



Powerful approximation algorithms based on Lyapunov theory + optimization...



$$\begin{array}{l} A_1^T P_1 A_1 \quad \preceq \quad P_1 \\ A_2^T P_1 A_2 \quad \preceq \quad P_2 \\ A_1^T P_2 A_1 \quad \preceq \quad P_1 \\ A_2^T P_2 A_2 \quad \preceq \quad P_2 \end{array}$$

A lot less understood for nonlinear switched systems...

Want to know more? <http://aaa.princeton.edu/>

