Noncommutative aspects of dynamic programming

Stephane.Gaubert@inria.fr

INRIA and CMAP, École polytechnique, CNRS, Université Paris-Saclay

INDAM Workshop Roma June 19–23, 2017

Based on work with Stott (Phd thesis), Allamigeon, and Goubault and Putot from LIX, École polytechnique (ACM Trans Embedded Computing Systems, 2016) followup in arXiv:1706.04471, and on earlier work with McEneaney and Qu (CDC2011, JDE2014).

Three different methods

- max-plus method of McEneaney (SICON2007) for optimal control of switched LQ systems $\to$ approximation of the value function by suprema of quadratic forms, curse of dimensionality attenuation
- Ahmadi, Jungers, Parrilo, and Roozbehani: path complete LMI automata method (SICON2014) to approximate the joint spectral radius, approximation of the Barabanov ball by an intersection of ellipsoids
- static analysis of program by abstract interpretation, template method / discretized support function Sankaranarayanan and Sipma and Manna (VMCAI'05), nonlinear templates Adjé, SG, Goubault ESOP'10, Seidl, Gawlitza, program invariant as a conjunction of quadratic inequalities.

This talk

- Relations between three classes of methods. (Common bottleneck: large scale semidefinite optimization.)
- Geometry of the space of positive semi-definite matrices
- Noncommutative Bellman operators = tropicalization of the Kraus maps in quantum information
- new (almost) LMI-free methods to compute piecewise quadratic invariants
- methods of non-linear Perron-Frobenius theory

Part I.

McEneaney's max-plus method

curse of dimensionality attenuation

Max-plus or tropical algebra

In an exotic country, children are taught that:

$$\text{“}a + b\text{”} = \max(a, b) \qquad \text{“}a \times b\text{”} = a + b$$

So

- “$2 + 3$” $=$

Max-plus or tropical algebra

In an exotic country, children are taught that:

$$\text{“}a + b\text{”} = \max(a, b) \qquad \text{“}a \times b\text{”} = a + b$$

So

- “$2 + 3$” $= 3$

Max-plus or tropical algebra

In an exotic country, children are taught that:

$$\text{“}a + b\text{”} = \max(a, b) \qquad \text{“}a \times b\text{”} = a + b$$

So

- “$2 + 3$” $= 3$
- “2×3” $=$

Max-plus or tropical algebra

In an exotic country, children are taught that:

$$\text{“}a + b\text{”} = \max(a, b) \qquad \text{“}a \times b\text{”} = a + b$$

So

- “$2 + 3$” $= 3$
- “2×3” $=5$

Lagrange problem / Lax-Oleinik semigroup

$$v(t,x) = \sup_{\mathbf{x}(0)=x,\ \mathbf{x}(\cdot)} \int_0^t L(\mathbf{x}(s),\dot{\mathbf{x}}(s))ds + \phi(\mathbf{x}(t))$$

Lax-Oleinik semigroup: $(S_t)_{t\geqslant 0}$, $S_t\phi := v(t,\cdot)$.

Superposition principle: $\forall \lambda \in \mathbb{R}$, $\forall \phi,\psi$,

$$\begin{aligned} S_t(\sup(\phi,\psi)) &= \sup(S_t\phi, S_t\psi) \\ S_t(\lambda+\phi) &= \lambda + S_t\phi \end{aligned}$$

So S_t is max-plus linear.

Lagrange problem / Lax-Oleinik semigroup

$$v(t,x) = \sup_{\mathbf{x}(0)=x,\ \mathbf{x}(\cdot)} \int_0^t L(\mathbf{x}(s),\dot{\mathbf{x}}(s))ds + \phi(\mathbf{x}(t))$$

Lax-Oleinik semigroup: $(S_t)_{t\geqslant 0}$, $S_t\phi := v(t,\cdot)$.

Superposition principle: $\forall\lambda\in\mathbb{R}$, $\forall\phi,\psi$,

$$\begin{aligned} S_t(\text{“}\phi+\psi\text{''}) &= \text{“}S_t\phi + S_t\psi\text{''} \\ S_t(\text{“}\lambda\phi\text{''}) &= \text{“}\lambda S_t\phi\text{''} \end{aligned}$$

So S_t is max-plus linear.

The function v is solution of the Hamilton-Jacobi equation

$$\frac{\partial v}{\partial t} = H(x, \frac{\partial v}{\partial x}) \qquad v(0, \cdot) = \phi$$

Max-plus linearity $\Leftarrow$ Hamiltonian convex in p

$$H(x, p) = \sup_u (L(x, u) + p \cdot u)$$

Hopf formula, when $L = L(u)$ concave:

$$v(t, x) = \sup_{y \in \mathbb{R}^n} tL(\frac{x - y}{t}) + \phi(y) \ .$$

The function v is solution of the Hamilton-Jacobi equation

$$\frac{\partial v}{\partial t} = H(x, \frac{\partial v}{\partial x}) \qquad v(0, \cdot) = \phi$$

Max-plus linearity $\Leftarrow$ Hamiltonian convex in p

$$H(x,p) = \sup_u (L(x,u) + p \cdot u)$$

Hopf formula, when $L = L(u)$ concave:

$$v(t,x) = \text{``}\int G(x-y)\phi(y)dy\text{''} \ .$$

Max-plus basis methods

Fleming, McEneaney 00;
Akian, Lakhoua, SG 04, Dower, Kaise, Sridharan, ... Approximate the value function by a "linear comb." of "basis" functions with coeffs. $\lambda_i(t) \in \mathbb{R}$:

$$v(t,\cdot) \simeq \text{``}\sum_{i\in[p]} \lambda_i(t) w_i\text{''}$$

The w_i are given max-plus basis functions, to be chosen depending on the regularity of $v(t,\cdot)$

Max-plus basis methods

Fleming, McEneaney 00;
Akian, Lakhoua, SG 04, Dower, Kaise, Sridharan, ... Approximate the value function by a "linear comb." of "basis" functions with coeffs. $\lambda_i(t) \in \mathbb{R}$:

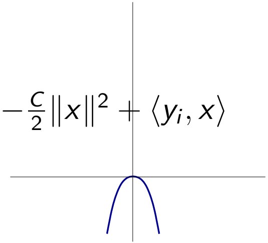
$$v(t, \cdot) \simeq \sup_{i \in [p]} \lambda_i(t) + w_i$$

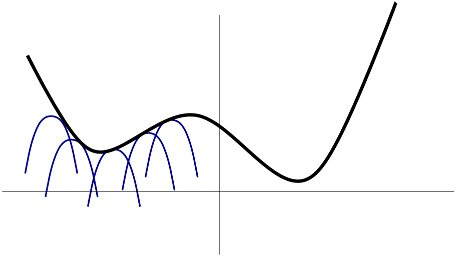
The w_i are given max-plus basis functions, to be chosen depending on the regularity of $v(t, \cdot)$

Best max-plus approximation

$$P(f) := \max\{g \leqslant f \mid g \text{ “linear comb.” of } w_i\}$$

$w_i : x \mapsto -\frac{C}{2}\|x\|^2 + \langle y_i, x\rangle$





adapted if v is C-semi-convex, i.e. $v + C\|x\|^2/2$ convex, or $v'' \geqslant -CI$

Example: switched optimal control problem

$$V(x) = \sup_{\mu} \sup_{\mathbf{u}} \int_0^{\infty} \frac{1}{2}\mathbf{x}(t)' D^{\mu(t)}\mathbf{x}(t) - \frac{\gamma^2}{2}|\mathbf{u}(t)|^2 dt,$$

$$\dot{\mathbf{x}}(t) = A^{\mu(t)}\mathbf{x}(t) + \sigma^{\mu(t)}\mathbf{u}(t), \;\; \mathbf{x}(0) = x \in \mathbb{R}^d \;\;,$$

$\mu : [0,\infty) \to \{1,\dots,M\}$ discrete valued control, $\mathbf{u} \in L_2^{\mathrm{loc}}([0,\infty);\mathbb{R}^k)$

Problem is nonconvex ($D^1,\dots,D^M \not\succcurlyeq 0$).

(McEneaney SICON 07)

McEneaney's curse of dimensionality attenuation method

- Semigroup approximation, τ small time:

$$S_\tau \simeq \tilde{S}_\tau = \sup_{m \in \{1,\dots,M\}} S_\tau^m$$

McEneaney's curse of dimensionality attenuation method

- Semigroup approximation, τ small time:

$$S_\tau \simeq \tilde{S}_\tau = \sup_{m \in \{1,\dots,M\}} S_\tau^m$$

- S_t^m semigroup associated to the unswitched control problem in which $\mu \equiv m \in \{1, \dots, M\}$:

$$S_t^m[\phi](x) = \sup_{\mathbf{u}} \int_0^t \frac{1}{2}\mathbf{x}(t)' D^m \mathbf{x}(t) - \frac{\gamma^2}{2}|\mathbf{u}(t)|^2 dt + \phi(\mathbf{x}(t)).$$

$$\dot{\mathbf{x}}(s) = A^m \mathbf{x}(s) + \sigma^m \mathbf{u}(s); \;\; \mathbf{x}(0) = x \in \mathbb{R}^d \;\; .$$

McEneaney's curse of dimensionality attenuation method

- Semigroup approximation, τ small time:

$$S_\tau \simeq \tilde{S}_\tau = \sup_{m \in \{1,\ldots,M\}} S^m_\tau$$

- S^m_t semigroup associated to the unswitched control problem in which $\mu \equiv m \in \{1,\ldots,M\}$:

$$S^m_t[\phi](x) = \sup_{\mathbf{u}} \int_0^t \frac{1}{2}\mathbf{x}(t)' D^m \mathbf{x}(t) - \frac{\gamma^2}{2}|\mathbf{u}(t)|^2 dt + \phi(\mathbf{x}(t)).$$

$$\dot{\mathbf{x}}(s) = A^m \mathbf{x}(s) + \sigma^m \mathbf{u}(s);\;\; \mathbf{x}(0) = x \in \mathbb{R}^d \;\; .$$

- $S^m_t[\phi]$ is a quadratic function if ϕ is. (Riccati)

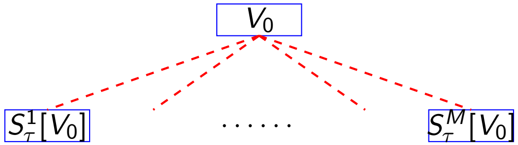
Arborescent propagation

$$V \simeq V_T = \{S_\tau\}^N[V_0] \simeq \{\tilde{S}_\tau\}^N[V_0] = \sup_{i_N,\dots,i_1} S_\tau^{i_N} \circ \cdots \circ S_\tau^{i_1}[V_0] \ .$$

V_0

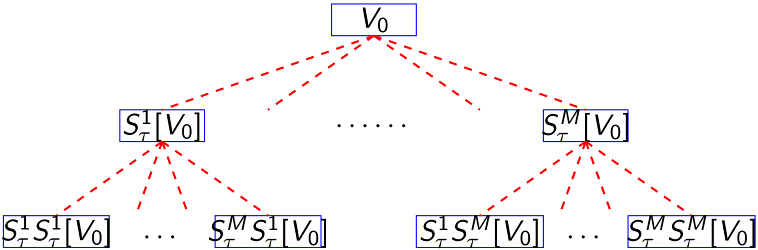
Arborescent propagation

$$V \simeq V_T = \{S_\tau\}^N[V_0] \simeq \{\tilde{S}_\tau\}^N[V_0] = \sup_{i_N,\dots,i_1} S_\tau^{i_N} \circ \cdots \circ S_\tau^{i_1}[V_0] \ .$$



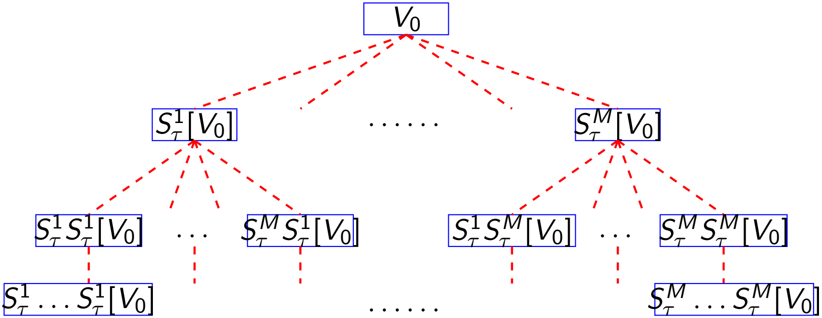
Arborescent propagation

$$V \simeq V_T = \{S_\tau\}^N[V_0] \simeq \{\tilde{S}_\tau\}^N[V_0] = \sup_{i_N,\dots,i_1} S_\tau^{i_N} \circ \cdots \circ S_\tau^{i_1}[V_0] \ .$$



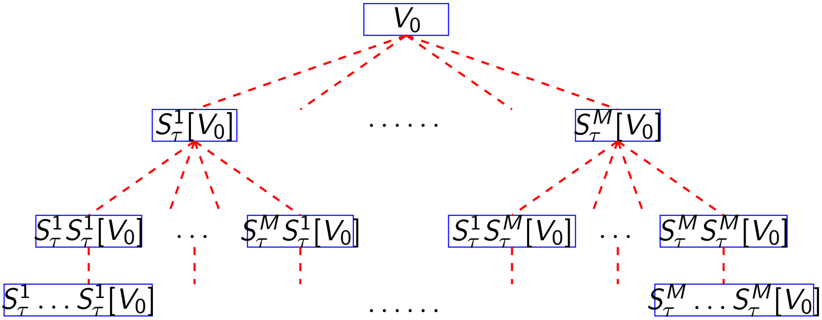
Arborescent propagation

$$V \simeq V_T = \{S_\tau\}^N[V_0] \simeq \{\tilde{S}_\tau\}^N[V_0] = \sup_{i_N,\dots,i_1} S_\tau^{i_N} \circ \cdots \circ S_\tau^{i_1}[V_0] \ .$$



Arborescent propagation

$$V \simeq V_T = \{S_\tau\}^N[V_0] \simeq \{\tilde{S}_\tau\}^N[V_0] = \sup_{i_N,\dots,i_1} S_\tau^{i_N} \circ \cdots \circ S_\tau^{i_1}[V_0] \ .$$



Comput. complexity: $O(M^N d^3)$ $\Rightarrow$ polynomial in the dimension

Kluberg, McEneaney (SICON 2010) obtained the first error bound, refined in:

Theorem (Zheng Qu, SICON 2014)

The computational complexity to reach an error of order ϵ is

$$O(M^{-\log(\epsilon)/\epsilon} d^3)$$

and this is tight (matched by numerical experiments).

Compare with $O(1/\epsilon^{d/r})$ for a grid scheme with an error of order $(\Delta x)^r$.

The complexity of the max-plus method is dramatically improved if the method is combined with pruning:

the value function V is approximated by a supremum of an exponential number of quadratic forms, many of which are redundant.

These can be dynamically eliminated

The Pruning Problem

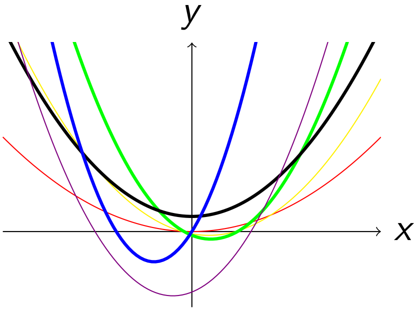
Given

$$f = \sup_{i \in [p]} \phi_i, \qquad \phi_i \text{ quadratic } \mathbb{R}^d \to \mathbb{R}$$

and $k \ll p$, find $I \subset [p]$, $|I| = k$, with a best approximation of f by

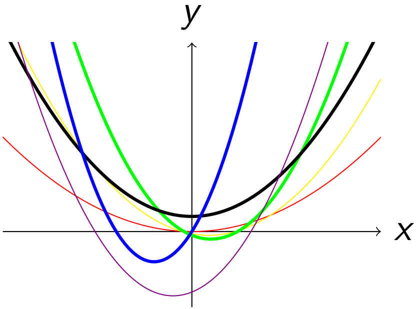
$$\sup_{i \in I} \phi_i \quad .$$

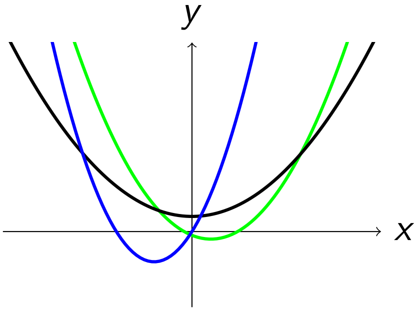
Pruning redundant quadratic forms



$$\phi = \sup(\phi_{green}, \phi_{red}, \phi_{violet}, \phi_{yellow}, \phi_{black}, \phi_{blue})$$

Pruning redundant quadratic forms





$$\phi = \sup(\phi_{green}, \phi_{red}, \phi_{violet}, \phi_{yellow}, \phi_{black}, \phi_{blue})$$

$$\phi = \sup(\phi_{green}, \phi_{black}, \phi_{blue})$$

Pruning algorithm

$$\phi_i(x) = (x^T \;\; 1) Q_i \begin{pmatrix} x \\ 1 \end{pmatrix}, \quad i = 1, \dots, p.$$

Importance metric

$$\sup_{i \neq j} \phi_i \geqslant \phi_j \Leftrightarrow \nu_j := \sup_x \frac{\phi_j(x) - \sup_{i \neq j} \phi_i(x)}{1 + |x|^2} \leqslant 0$$

$\nu_j \leqslant 0 \implies \phi_j$ can be pruned.

Optimisation problem

$$\begin{aligned} \nu_j = \max \;\; & \nu \\ & \nu \leqslant z^\top (Q_j - Q_i) z \\ & z^T z = 1. \end{aligned}$$

SDP relaxation

$$\begin{aligned} \overline{\nu}_j = \max \;\; & \nu \\ & \nu \leqslant \mathrm{tr}((Q_j - Q_i) Z) \\ & Z \geqslant 0, \;\; \mathrm{tr}(Z) = 1. \end{aligned}$$

Tree approximation + pruning

The method has been applied to solve approximately problems of

- dimension $d = 4$, number of switches $M = 3$, in McEneaney 07
- dimension $d = 6$, number of switches $M = 6$, in McEneaney, Deshpande, SG 08, SG, McEneney, Qu 2011 Shor relaxation + improved pruning / facility location for Bregman distances.
- dimension $d = 15$, number of switches $M = 6$, in Sridharan, James, McEneaney 10 (quantum optimal gate synthesis, $SU(4)$)

SDP is the bottleneck

Table : HJ dim 6, CPU time Matlab/cvx/sdpt3

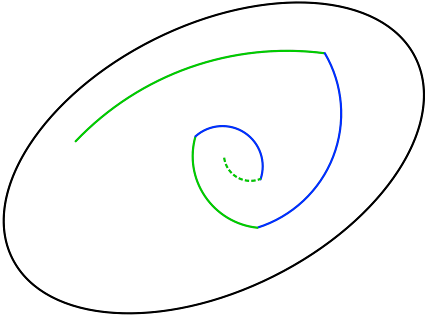
τ=0.2, K=25	Total time	Propagation	SDP	Pruning
sort lower	1.04h	1.85%	98.15%	0.00%
sort upper	1.34h	1.52%	98.43%	0.05%
Facility location	1.38h	1.45%	89.47%	9.08%
greedy	1.43h	1.63%	97.84%	0.53%

SG, McEneney, Qu 2011

Part II.
Linear switched systems and LMI automata

Switching system

```
x <- init();
while (rand_bool)
  switch (rand_bool)
    case 0:
      x <- A*x;
    case 1:
      x <- B*x;
  end
end
```



Problem

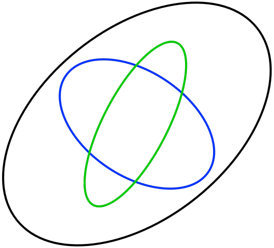
Given a switched discrete dynamical system

$$x_{k+1} = A_{\phi(k)} x_k \in \mathbb{R}^n \, , \; k \in \mathbb{N} \, , \; A_{\phi(k)} \in \{A_1, \cdots, A_m\} \, ,$$

Bound the set of reachable values

Look for a bounded invariant set $\mathcal{S}$:

$$\forall i, \; A_i(\mathcal{S}) \subseteq \mathcal{S}$$



We will look for piecewise quadratic invariants (union or intersection of ellipsoids).

Joint spectral radius / discrete HJ equation

$$\rho := \lim_{k\to\infty} \sup_{i_1,i_2\dots i_k} \|A_{i_1}\dots A_{i_k}\|^{1/k}$$

ν is an α-approximate extremal norm if

$$\nu(A_i x) \leqslant \alpha\nu(x), \quad \forall i$$

The existence of such a ν implies $\rho \leqslant \alpha$.

Barabanov norm / sol. of max-plus eigenvalue problem

$$\max_{1\leqslant i\leqslant M} \nu(A_i x) = \rho\nu(x)$$

$\rho \leqslant 1 \implies$ Barabanov ball is invariant

path complete LMI automata

$[m] := \{1, \dots, m\}$ switching modes, $[m]^N$ switching sequences of length N.
$[m]$ acts on $[m]^N$: concatenate and forget, eg, $728 \cdot 3 = 283$ for $N = 3$.

Theorem (Ahmadi, Jungers, Parrilo, and Roozbehani SICON 2014)

Suppose we can find positive definite matrices $(Q_w)_{w \in [m]^N}$ and $\alpha > 0$ such that

$$A_i^\top Q_w A_i \preccurlyeq \alpha^2 Q_{w \cdot i}, \qquad \forall w \in [m]^N, i \in [m]$$

Then, $x \mapsto \sup_w (x^\top Q_w x)^{1/2}$ is an α-approximate extremal norm, in particular,

$$\rho \leqslant \alpha$$

α_N^{opt} converges to ρ as $N \to \infty$.

The path complete LMI automata method requires solving LMI of size $\Omega(m^N)$.

hardly feasible for large N.

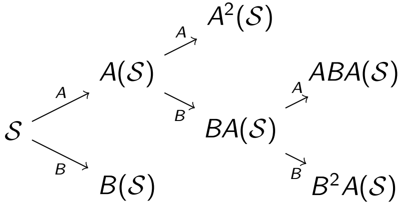
difficulty similar to pruning in McEneaney's method.

Part III.

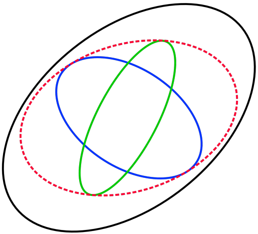
Avoiding LMIs: noncommutative Bellman operators

work with X. Allamigeon, E. Goubault, S. Putot, N. Stott

Basic tool for quadratic invariants



Avoid exploration of disjunctions !



Over-approximation by a single quadric

Basic tool

Given 2 centered ellipsoids, find a tight overapproximating quadric.

Avoid recourse to LMI (semidefinite programming) – too slow since the primitive will be called repeatedly.

Löwner order

A symmetric matrix A is positive semidefinite (PSD) if

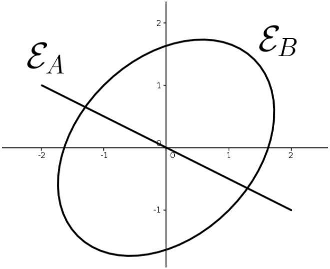
$$\lambda_{min}(A) \geqslant 0 \equiv \forall x,\ x^T A x \geqslant 0 \equiv A \succcurlyeq 0$$

The order $\preccurlyeq$ is called the Löwner order for symmetric matrices:

$$A \preccurlyeq B \iff B - A \succcurlyeq 0$$

Given a PSD matrix A, we define the ellipsoid as $\mathcal{E}_A$:

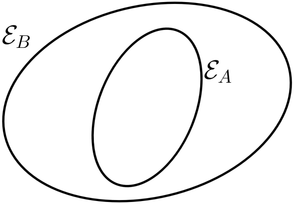
$$\mathcal{E}_A = \left\{ x \in \mathbb{R}^n \mid x^T A^{-1} x \leqslant 1 \right\} = \left\{ x \in \mathbb{R}^n \mid x x^T \preccurlyeq A \right\}$$



The ordered set of ellipsoids

Orderings on ellipsoids and matrices are equivalent:

$$\mathcal{E}_A \subseteq \mathcal{E}_B \iff A \preccurlyeq B$$



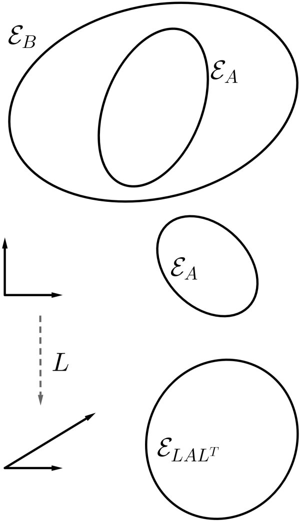
The ordered set of ellipsoids

Orderings on ellipsoids and matrices are equivalent:

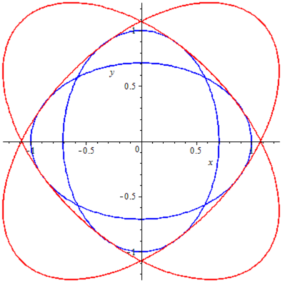
$$\mathcal{E}_A \subseteq \mathcal{E}_B \iff A \preccurlyeq B$$

Invertible linear transformations yield order automorphisms on the space of ellispoids :

$$L \bullet \mathcal{E}_A = \mathcal{E}_{LAL^T}$$



Krein-Rutman (1950): A cone defines a lattice order iff it is simplicial. So, neither quadratic sets nor ellipsoids constitutes a lattice



Kadison (1951): Symmetric matrices under Löwner order constitute an **anti-lattice**: A, B have a least upper bound only if $A \preccurlyeq B$ or $A \succcurlyeq B$.

What is the structure of the set of least upper bounds?

The inertia of the symmetric matrix M is the tuple (p, q, r), where

- p: number of positive eigenvalues of M,
- q: number of negative eigenvalues of M,
- r: number of zero eigenvalues of M.

Definition (Indefinite orthogonal group)

$\mathcal{O}(p, q)$ is the group of matrices S preserving the quadratic form $x_1^1 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2$:

$$S \begin{pmatrix} I_p & \\ & -I_q \end{pmatrix} S^T = \begin{pmatrix} I_p & \\ & -I_q \end{pmatrix} =: J_{p,q}$$

The inertia of the symmetric matrix M is the tuple (p, q, r), where

- p: number of positive eigenvalues of M,
- q: number of negative eigenvalues of M,
- r: number of zero eigenvalues of M.

Definition (Indefinite orthogonal group)

$\mathcal{O}(p, q)$ is the group of matrices S preserving the quadratic form $x_1^1 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2$:

$$S \begin{pmatrix} I_p & \\ & -I_q \end{pmatrix} S^T = \begin{pmatrix} I_p & \\ & -I_q \end{pmatrix} =: J_{p,q}$$

$\mathcal{O}(1, 1)$ is the group of hyperbolic isometries $\begin{pmatrix} \epsilon_1 ch\, t & \epsilon_2 sh\, t \\ \epsilon_1 sh\, t & \epsilon_2 ch\, t \end{pmatrix}$, where $\epsilon_1, \epsilon_2 \in \{-1, 1\}$

The inertia of the symmetric matrix M is the tuple (p, q, r), where

- p: number of positive eigenvalues of M,
- q: number of negative eigenvalues of M,
- r: number of zero eigenvalues of M.

Definition (Indefinite orthogonal group)

$\mathcal{O}(p, q)$ is the group of matrices S preserving the quadratic form $x_1^1 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2$:

$$S \begin{pmatrix} I_p & \\ & -I_q \end{pmatrix} S^T = \begin{pmatrix} I_p & \\ & -I_q \end{pmatrix} =: J_{p,q}$$

$\mathcal{O}(1,1)$ is the group of hyperbolic isometries $\begin{pmatrix} \epsilon_1 ch\, t & \epsilon_2 sh\, t \\ \epsilon_1 sh\, t & \epsilon_2 ch\, t \end{pmatrix}$,
where $\epsilon_1, \epsilon_2 \in \{-1, 1\}$
$\mathcal{O}(p) \times \mathcal{O}(q)$ is a maximal compact subgroup of $\mathcal{O}(p, q)$.

Quantitative version of Kadison theorem

Theorem (Stott, arXiv:1612.05664, Proc. AMS)

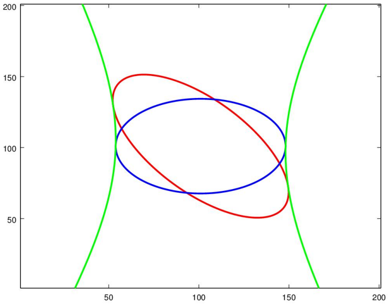
Given symmetric matrices A, B and (p, q, r) the inertia of $A - B$, the set of maximal lower bounds of $\{A, B\}$ can be identified to

$$\mathcal{O}(p,q) \Big/ \mathcal{O}(p) \times \mathcal{O}(q) \cong \mathbb{R}^{pq}$$

Given $P \in \mathcal{M}_{n,p+q}$ s.t. $A - B = PJ_{p,q}P^T$, the parametrization is:

$$S \mapsto A + PS\begin{pmatrix} 0_p & \\ & I_q \end{pmatrix} S^T P^T$$

Illustration of the theorem



- The inertia of $A - B$ is $(1, 1, 0)$.
- $\mathcal{O}(1,1) \Big/ \mathcal{O}(1) \times \mathcal{O}(1)$ is the group of hyperbolic rotations:

$$\left\{ \begin{pmatrix} ch(t) & sh(t) \\ sh(t) & ch(t) \end{pmatrix} \mid t \in \mathbb{R} \right\}$$

Minimal upper bound selection requirements

```
while (rand_bool)
  y <- x;
  x <- 0.5*x + y;
end
```

$$A = \begin{pmatrix} 0.5 & 1 \\ 1 & 0 \end{pmatrix}$$

```
while (rand_bool)
  z' <- x';
  x' <- 0.5*x' + 1.5*y';
  y' <- z' - y';
end
```

$$B = \begin{pmatrix} 0.5 & 1.5 \\ 1 & -1 \end{pmatrix}$$

The programs are similar up to a linear change of variable:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

$\Rightarrow$ Invariants should only differ by a linear change of variables.

Natural requirement for a minimal upper bound selection:

Invariance under linear change of variables.

Minimal upper bound selection

A minimal upper bound selection $\sqcup$ should satisfy:

- $\mathcal{E}_A \sqcup \mathcal{E}_B$ is a minimal upper bound of $(\mathcal{E}_A, \mathcal{E}_B)$,
- $(L \bullet \mathcal{E}_A) \sqcup (L \bullet \mathcal{E}_B) = L \bullet (\mathcal{E}_A \sqcup \mathcal{E}_B)$ for any linear operator L.

Minimal upper bound selection

A minimal upper bound selection $\sqcup$ should satisfy:

- $\mathcal{E}_A \sqcup \mathcal{E}_B$ is a minimal upper bound of $(\mathcal{E}_A, \mathcal{E}_B)$,
- $(L \bullet \mathcal{E}_A) \sqcup (L \bullet \mathcal{E}_B) = L \bullet (\mathcal{E}_A \sqcup \mathcal{E}_B)$ for any linear operator L.

Several selections have been used:

- $\frac{A+B}{2} + \frac{1}{2}[(A-B)(A-B)]^{1/2}$, compare with

$$\max(a,b) = (a+b)/2 + |a-b|/2$$

- $\lim_{n \to +\infty} \left(A^n + B^n\right)^{-1/2} \left(A^{n+1} + B^{n+1}\right) \left(A^n + B^n\right)^{-1/2}$,

Minimal upper bound selection

The minimum volume covering ellipsoid has been vastly studied:

- Optimal design and statistics (Titterington 1973-80),
- Computational geometry (Khachiyan & Todd 1993)
- Pattern recognition (Glineur 1998),
- Computer graphics (Bouville 1985, Eberly '01),
- Convex Optimization (Nesterov, Boyd & Vanderberghe '04),

and applied to abstract interpretation (Roux '11).

Minimal upper bound selection

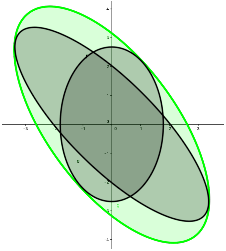
The minimum volume covering ellipsoid has been vastly studied:

- Optimal design and statistics (Titterington 1973-80),
- Computational geometry (Khachiyan & Todd 1993)
- Pattern recognition (Glineur 1998),
- Computer graphics (Bouville 1985, Eberly '01),
- Convex Optimization (Nesterov, Boyd & Vanderberghe '04),

and applied to abstract interpretation (Roux '11).

Usually computed using SDP:

$$\left[A \sqcup B\right]^{-1} = \arg \max_{\begin{cases} X \succcurlyeq 0 \\ X \preccurlyeq A^{-1} \\ X \preccurlyeq B^{-1} \end{cases}} \log\det X$$



Theorem (Allamigeon et al., EMSOFT 2015 + ACM Emb. Software)

The mimimum volume ellipsoid is the unique selection of a minimal upper bound which commutes with the action of the linear group. It is given by

$$A \sqcup B = \frac{1}{2} A^{1/2} \Big[I + A^{-1/2} B A^{-1/2} + |I - A^{-1/2} B A^{-1/2}| \Big] A^{1/2}$$

with $|X| = (XX^T)^{1/2}$.
It can be computed using $\mathcal{O}(n^3)$ *arithmetic operations, involving*

- *2 Cholesky decompositions,*
- *1 inversion (triangular),*
- *5 matrix multiplications.*

The computation is much more tractable than SDP.

Structure of analyzed programs

We shall use the invariant join to compute invariants of some switched programs written in a toy language:

```
while (rand_bool)
  switch (rand_int)
    case 0:
      I0(x);
    ...
    case p:
      Ip(x);
  end
end
```

Ik is a statement of the form:

- a variable declaration or deletion
- a linear variable assignment `xi <- L(x);`
- a nested sub-program of the form on the left

This is a very small language, but already very difficult to analyze.

Reduction to a non-linear fix-point problem

A simple example

```
init(x);
while (rand_bool)
  switch (rand_bool)
    case 0:
      x <- A0(x);
    case 1:
      x <- A1(x);
  end
end
```

A0 and A1 are invertible linear operators.

Our stability problem is

$$(A_0 \bullet \mathcal{E}_S) \cup (A_1 \bullet \mathcal{E}_S) \subseteq \mathcal{E}_S \,.$$

This is relaxed into

$$(A_0 S A_0^T) \sqcup (A_1 S A_1^T) \preccurlyeq S \,.$$

We refine this into:

$$T(S) = (A_0 S A_0^T) \sqcup (A_1 S A_1^T) = \lambda S,\ \lambda < 1 \,.$$

This is our post-fix-point/eigenvector problem.

What we saw

$$T(S) = (A_0 S A_0^T) \sqcup (A_1 S A_1^T)$$

is the simplest instance of “noncommutative” Bellman operator.

Classical Bellman operators

$T : \mathbb{R}^n \to \mathbb{R}^n$,

$$T_i(x) = \max_{a \in A} \left(r_i^a + \sum_j P_{ij}^a x \right)$$

$P_{ij} \geqslant 0$, $\sum_j P_{ij} = 1$.

Passing to noncommutative: replace $\mathbb{R}^n$ by S_n^+, space of PSD matrices.

Zoo of "noncommutative" dynamic programming operators

- 0-player, stochastic, quantum channels = completely positive operators = Kraus maps,

$$T(X) = \sum_i A_i X A_i^\dagger , \qquad \sum_i A_i^\dagger A_I = I$$

- 1-player deterministic

$$T(X) = (A_0 X A_0^T) \sqcup (A_1 X A_1^T)$$

- 1-player stochastic, generalized Riccati flows (Qu, SG, JDE 2014)

$$\dot{P} + A'P + PA + C'PC + Q = (PB + C'PD + L')(R + D'PD)^{-1}(B'P + D'PC + L)$$

Tropical Kraus map

Definition

$$T(X) := \bigvee_{1 \leqslant i \leqslant m} A_i^\top X A_i$$

where $\bigvee$ denotes the set of least upper bounds in Löwner order (multivalued map).

Proposition (SG, Stott arXiv:1706.04471)

Suppose that $\alpha^2 X \in T(X)$ with $\alpha > 0$ and X positive definite. Then,

$$\rho(A_1, \ldots, A_m) \leqslant \alpha$$

Moreover, if $\alpha \leqslant 1$, the ellipsoid $\{y \mid y^\top X y \leqslant 1\}$ is invariant.

Tropical Kraus map associated to a path complete automaton

$p := m^N$ number of switching sequences.
$T : (S_n^+)^p \to (S_n^+)^p$

$$T_z(X) := \bigvee_{\substack{w \in [m]^N, i \in [m] \\ w \cdot i = z}} A_i^\top X_w A_i$$

Theorem (SG, Stott arXiv:1706.04471)

Suppose that $\alpha^2 X \in T(X)$ with $\alpha > 0$ and X positive definite. Then,

$$\rho(A_1, \dots, A_m) \leqslant \alpha$$

Reduces to the earlier $p = 1$ case by a block diagonal construction.

Perron-Frobenius theorem for Tropical Kraus maps

Let $\sqcup$ denote the minimal volume selection,

$$T_z(X) \ni T_z(X)^{\mathrm{sel}} := \sqcup_{\substack{w\in[m]^N, i\in[m] \\ w\cdot i = z}} A_i^\top X_w A_i$$

Theorem (SG, Stott ibid.)

Suppose $\langle A_1, \dots, A_m\rangle$ is irreducible. Then, there exists $\alpha > 0$ and X positive definite such that $\alpha^2 X = T^{sel}(X) \in T(X)$.

Sketch of proof

Apply Brouwer fixed point theorem to $X \mapsto T(X)/\operatorname{trace}(X)$ which preserves the noncommutative simplex (trace one PSD matrices).
QED

Sketch of proof

Apply Brouwer fixed point theorem to $X \mapsto T(X)/\operatorname{trace}(X)$ which preserves the noncommutative simplex (trace one PSD matrices).
QED

Caveat. Incorrect proof. $\sqcup$ is continuous on the open cone $\operatorname{int} S_n^+$ but not on the closed cone S_n^+. We must construct an invariant set in the interior of the cone by exploiting the irreducibility condition.

Same idea works for McEneaney's switched LQ problem, consider the following noncommutative Bellman operator

$$T_z(X) := \bigvee_{\substack{w\in[m]^N, i\in[m] \\ w\cdot i=z}} \mathrm{RiccFlow}_i^{\tau}(X_w)$$

Lemma

$X \in T(X) \Rightarrow$

$$V(y) := \sup_{w\in[m]^N} y^\top X_w y$$

satisfies $S^t V \leqslant V$, $t \geqslant 0$.

$S^t V \simeq V$ if the selection of X is well chosen.

How to solve the nonlinear eigenproblem?

$$T_z^{\mathrm{sel}}(X) := \bigsqcup_{\substack{w\in[m]^N, i\in[m] \\ w\cdot i=z}} A_i^\top X_w A_i = \alpha^2 X_z, \qquad \forall z \in [m]^N$$

or the fixed point problem

$$T_z^{\mathrm{sel}}(X) := \bigsqcup_{\substack{w\in[m]^N, i\in[m] \\ w\cdot i=z}} \mathrm{RiccFlow}_i^\tau(X_w) = X_z \qquad \forall z \in [m]^N$$

$\rightarrow$ evaluating every coordinate $T_z^{\mathrm{sel}}(X)$ is tractable, SDP or Cholesky in small dimension ($n \ll m^N$)

Idea: iterative fixed point methods (large scale, SDP-free)

Take home message :

invariant Finsler metrics on the cone of positive definite matrices

the standard Riccati flow is a contraction in these metrics

$$\dot{P} = A'P + PA + D - P\Sigma P \ , \qquad D, \Sigma > 0 \ .$$

Thompson's part metric

C closed convex pointed cone in a Banach space C,
$x \leqslant y$ iff $y - x \in C$

$$d_T(x,y) = \inf\{\log \alpha \mid \alpha^{-1} x \leqslant y \leqslant \alpha x\}$$

$C = \mathbb{R}^n_+$, $d_T(x,y) = \| \log x - \log y \|_\infty$.

$C = S_n^+$, $d_T(A,B) = \| \operatorname{spec} \log A^{-1/2} B A^{-1/2} \|_\infty = \| \log \operatorname{spec} A^{-1} B \|_\infty$.

Invariant metrics on the cone of positive matrices

$$d(UAU', UBU') = d(A, B) \ , U \in GL(n)$$

- Thompson's part metric:

$$d_T(A, B) = \| \log \operatorname{spec} A^{-1} B \|_\infty, \quad A, B \succ 0$$

$$d_T(A, B) = \inf_\gamma \int_0^1 \|\dot\gamma(t)\gamma(t)^{-1}\|_\infty \, dt.$$

- Riemannian metric:

$$d_2(A, B) = \inf_\gamma \int_0^1 \|\dot\gamma(t)\gamma(t)^{-1}\|_2 dt = \| \log \operatorname{spec} A^{-1} B \|_2$$

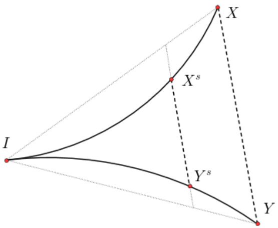
- Invariant Finsler metric, ν convex positively homogeneous

$$d_\nu(A, B) = \inf_\gamma \int_0^1 \nu(\dot\gamma(t)\gamma(t)^{-1}) dt = \nu(\log \operatorname{spec} A^{-1} B)$$

The map $X \mapsto X^s$ is a geodesic in any of the invariant Finsler metrics (including Thompson d_T and Riemann d_2).

All these metrics have nonpositive curvature in the sense of Busemann (Corach, Bhatia,...).

$$d\big(X^s, Y^s\big) \leqslant s\, d(X, Y)$$



Theorem (Bougerol 93)

The standard Riccati operator (flow) is a strict contraction mapping in the invariant Riemannian metric.

Theorem (Liverani, Wojtkowski 94)

The standard Riccati operator (flow) is a strict contraction mapping in Thompson's part metric.

Theorem (Lawson and Lim 08)

The standard Riccati operator (flow) is a strict contraction mapping in any invariant Finsler metric.

The proof relies on the symplectic structure of the standard Riccati flow

contraction of Riccati flow / McEneaney's problem

For all $m \in \{1, \ldots, M\}$, the semigroup $\{S_t^m\}_t$ corresponds to the flow of an indefinite Riccati equation:

$$\dot{P} = (A^m)'P + PA^m + D^m + P\Sigma^m P \quad . \tag{1}$$

sign changed, $-P\Sigma^m P$ in Bougerol, Liverani, Wojtwoski...!

Theorem (SG, Qu JDE 12)

Under assumptions (a bit more than well posedness), there is $P_0 \succ 0$ and $\alpha > 0$ such that for all solutions $P_1(\cdot), P_2(\cdot) : [0, T] \to (0, P_0)$ of the indefinite Riccati flow (1) we have:

$$d_T(P_1(t), P_2(t)) \leqslant e^{-\alpha t} d_T(P_1(0), P_2(0)), \;\; \forall t \in [0, T] \;\; .$$

Contraction rate in Thompson's part metric

Denote by $M_s^t(\cdot)$ the flow associated to the ODE

$$\dot{x}(t) = \phi(t, x(t))$$

ϕ is a continuous function, $\mathscr{C}^1$ in the second variable with bounded differential. Let $\mathcal{U} \subseteq \operatorname{int} C$ be an open set satisfying $\lambda\mathcal{U} \subseteq \mathcal{U}$ for all $\lambda \in (0, 1]$.

Theorem (SG, Qu JDE 12)

If the flow is order-preserving, then the best constant α such that

$$d_T(M_s^t(x_1), M_s^t(x_2)) \leqslant e^{-\alpha(t-s)} d_T(x_1, x_2),\ s \leqslant t < t_{\mathcal{U}}(s, x_i)$$

holds for all $x_i \in \mathcal{U}$, $i = 1, 2$, is given by

$$\alpha := -\sup_{s\in J,\ x\in\mathcal{U}} \lambda_{\max}\Big(\big(D\phi_s(x)x - \phi(s,x)\big)x^{-1}\Big) \ .$$

Lipschitz properties of the invariant join

$$\operatorname{Lip} \sqcup := \sup_{X_1, X_2, Y_1, Y_2} \frac{d(X_1 \sqcup X_2,\, Y_1 \sqcup Y_2)}{d_m(X_1 \oplus X_2,\, Y_1 \oplus Y_2)}$$

where $X_1 \oplus X_2 := \operatorname{blockdiag}(X_1, X_2)$.

Theorem (Stott, PhD)

The Lipschitz constant of the invariant join in the Thompson metric satisfies:

$$\frac{1}{\pi} \log \frac{n}{4} \leqslant \operatorname{Lip}_T \sqcup \leqslant 2 + \frac{4}{\pi} \log n + o(\log n)\,.$$

The invariant join is nonexpansive in the Riemann metric:

$$\operatorname{Lip}_R \sqcup = 1\,.$$

The classical power algorithm

The power algorithm is used to find the leading eigenvector of large matrices, in the PageRank algorithm for instance:

PageRank algorithm

The PageRank iteration is

$$x_{k+1}^T = x_k^T \left[(1-\gamma)P + \gamma E\right]$$

with $0 < \gamma < 1$, E a rank 1 matrix and P denotes the transition matrix of a random walker on the web-graph.

This algorithm introduces a perturbation through a damping process with γ.

The modified power algorithm

From our example, we look for an eigenvector of

$$T : \mathcal{S}_n^+ \to \mathcal{S}_n^+$$
$$X \mapsto (A_0 X A_0^T) \sqcup (A_1 X A_1^T) \,.$$

The modified iteration is:

$$X_{k+1} = \frac{(1-\epsilon) T(X_k) + \epsilon I}{\operatorname{tr}\left[(1-\epsilon) T(X_k) + \epsilon I\right]}$$

with a damping parameter $0 < \epsilon < 1$.

The modified power algorithm

From our example, we look for an eigenvector of

$$T : \mathcal{S}_n^+ \to \mathcal{S}_n^+$$
$$X \mapsto (A_0 X A_0^T) \sqcup (A_1 X A_1^T)\,.$$

The modified iteration is:

$$X_{k+1} = \frac{(1-\epsilon)T(X_k) + \epsilon I}{\operatorname{tr}\left[(1-\epsilon)T(X_k) + \epsilon I\right]}$$

with a damping parameter $0 < \epsilon < 1$.

Upon convergence, X_∞ is an invariant if

$$\frac{1}{1-\epsilon}\operatorname{tr}\left[(1-\epsilon)T(X_\infty) + \epsilon I\right] < 1$$

Allamigeon et al., EMSOFT15 + ACM Emb. Soft.

The multiplicative power algorithm

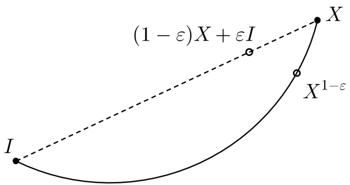
Multiplicative damping:

Allamigeon et al., ibid.

The modified iteration is:

$$X_{k+1} = \frac{T(X_k)^{1-\epsilon}}{\operatorname{tr}\left[T(X_k)^{1-\epsilon}\right]}$$

with a damping parameter $0 < \epsilon < 1$.



The multiplicative power algorithm

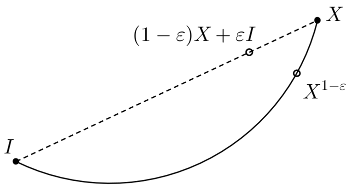
Multiplicative damping:

Allamigeon et al., ibid.

The modified iteration is:

$$X_{k+1} = \frac{T(X_k)^{1-\epsilon}}{\operatorname{tr}\left[T(X_k)^{1-\epsilon}\right]}$$

with a damping parameter $0 < \epsilon < 1$.



Upon convergence,

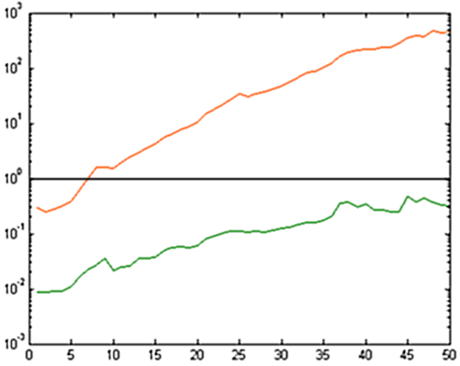
$$T(X_\infty) \approx \lambda X_\infty$$

The eigenvalue λ is fundamental in absorbing numerical errors, ensuring robustness of the invariant if $\lambda < 1$.

Benchmarks

We simulate a highly correlated damped oscillator, with a switch mechanism, yielding non-sparse matrices:

$$\begin{pmatrix} x_{k+1} \\ v_{k+1} \end{pmatrix} = \begin{pmatrix} I_n & hI_n \\ -hM & I_n-hD \end{pmatrix}_{\phi(k)} \begin{pmatrix} x_k \\ v_k \end{pmatrix},$$

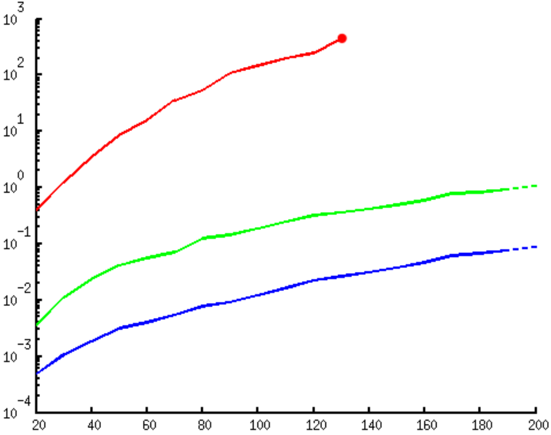


Comparison between computation times:

- SDP
- Power

Usually, less than 100 Power Iterations are needed.

Benchmarks



Comparison between computation times:

- SDP
- Power (mult)
- Power (add)

Again, less than 100 Power Iterations were needed.

Convergence analysis of the power algorithm relies on the nonpositive curvature properties of the Thompson and Riemann metrics.

Convergence if damping is large enough

See Allamigeon, SG, Goubault, Putot, Stott, EMSOFT 2015 and ACM Trans on Embedded Software

Current refinements, using Krasnoselski-Mann iteration in spaces of nonpositive curvature, $x_{k+1} = \text{midpoint}(T(x_k), x_k)$ with T nonexpansive.

LMI automata vs tropical Kraus

depth 3 LMI versus depth 6 tropical Kraus / joint spectral radius

Dimension n	5	10	20	30	40	45	50	100	500
CPU time (tropical)	0.9 s	1.5 s	3.5 s	7.9 s	13.7 s	18.1 s	25.2 s	1min	8min
CPU time (LMI)	3.1 s	4.2 s	31 s	3min	18min	—	—	—	—
Upper bound ρ (tropical)	2.767	3.797	5.4093	6.2038	7.3402	7.687	8.1591	11.487	25.44
Upper bound ρ (LMI)	2.7627	3.7426	5.3891	6.1942	7.3363	—	—	—	—

LMI versus tropical Kraus when the depth d varies

Depth d	2	4	6	8	10
Size of $[m]^N$	8	32	128	512	2048
CPU time (tropical)	0.03s	0.07s	0.4s	2.0s	9.0s
CPU time (LMI)	1.9s	4.0s	24s	1min	10min
Upper bound ρ (tropical)	1.842	1.821	1.804	1.800	1.801
Upper bound ρ (LMI)	1.8216	1.7974	1.7957	1.7922	1.7905

Tropical Kraus method applied to McEneaney's switched linear quadratic problem

Dimension	2	6	6	20	20
$\sharp$ switches	3	6	6	2	4
τ	0.05 s	0.2	0.1	0.1	0.1
$(\sharp \text{ switches})^d$	81	216	1296	128	256
Initial error	0.78	1.12	1.12	4.2	4.79
Final error	0.047	0.071	0.090	0.0006	0.17
Iterations	194	115	200	55	288
CPU time	8 s	41 s	5 min	5 s	2.5 min

Concluding remarks

- Tropical Kraus operators = analogues of completely positive maps in quantum information = noncommutative 1-player deterministic Bellman operators
- Generalized Riccati flows are other incarnations of noncommutative Bellman operators
- Non-linear fixed point schemes for tropical Kraus operators yield new, scalable methods, avoiding interior points.
- Potential of major speedup for switched control problems (Barabanov norms, HJ PDE with piecewise quadratic Hamiltonian).
- Convergence analysis: too conservative estimates so far!
- The selection $\sqcup$ of a least upper bound creates a relaxation gap. We would need to change the selection dynamically to reduce this gap and get that the error does tend to zero.

References

- X. Allamigeon, S. Gaubert, E. Goubault, S. Putot ad N. Stott. A scalable algebraic method to infer quadratic invariants of switched systems ACM Transactions on Embedded Computing Systems (TECS), Volume 15 Issue 4, August 2016, abridged version in EMSOFT2015.
- S. Gaubert and N. Stott, Tropical Kraus maps for optimal control of switched systems, arXiv:1706.04471
- N. Sott, Maximal lower bounds in the Löwner order, arXiv:1612.05664, Proc. AMS., published on line.

Thank you !