The Lasserre hierarchy in robotics

Didier Henrion

LAAS-CNRS Univ. Toulouse, FR
Czech Tech. Univ. Prague, CZ

Stockholm, May 2016

What is the Lasserre hierarchy ?

Who is Lasserre ?

The Lasserre hierarchy

Introduced by Jean Bernard Lasserre in 2000 in the context of global continuous and discrete optimization

J. B. Lasserre. Optimisation globale et théorie des moments. Comptes Rendus de l'Académie des Sciences, Paris, Série I, Mathématique, 331(11):929-934, 2000

Main idea: solving **globally nonconvex** optimization problems at the price of solving **convex optimization** problems

Assumption: problem data are **polynomial**

Revisit R. T. Rockafellar's motto (SIAM Review 1993): "the great watershed in optimization isn't between linearity and non-linearity, but convexity and nonconvexity"

The Lasserre hierarchy strategy

1. Reformulate the original non-convex problem as an infinite-dimensional convex linear problem (LP) on **measures**
   (and its dual on continuous functions)

2. Relax this measure LP by truncated LPs on **moments**
   (and their duals on polynomial sums-of-squares, SOS)

3. Solve these moment LPs with **semidefinite programming** (SDP), i.e. via linear matrix inequalities (LMI)

The dual SOS strategy was promoted by Naum Shor, Yurii Nesterov, Pablo Parrilo and many others, see the next talks

Why measures ?

Think of the global optimization problem

$$\min_{x \in X} \; p(x)$$

where a continuous function p is minimized on a compact $X \subset \mathbb{R}^n$

Consider samples x_k in X with relative weights w_k, and the LP

$$\min_w \sum_k p(x_k) w_k \text{ s.t. } \sum_k w_k = 1, \;\; w_k \geq 0$$

Why measures ?

Think of the global optimization problem

$$\min_{x \in X} \; p(x)$$

where a continuous function p is minimized on a compact $X \subset \mathbb{R}^n$

Consider samples x_k in X with relative weights w_k, and the LP

$$\min_w \sum_k p(x_k) w_k \text{ s.t. } \sum_k w_k = 1, \;\; w_k \geq 0$$

Passing to the limit we get an LP

$$\min_\mu \int_X p(x) d\mu(x) \text{ s.t. } \int_X d\mu(x) = 1, \;\; \mu \geq 0$$

where the unknown is a **probability measure** supported on X

What are moments ?

A measure μ is an infinite-dimensional object, just like a signal

It can be characterized by infinitely countably many real numbers, its (algebraic) **moments**

$$y_k = \int_X x^k d\mu(x) \in \mathbb{R}, \;\; k \in \mathbb{N}^n$$

just like a signal can be characterized e.g. by its Fourier coeffs

The integral of a polynomial function p against μ becomes a **linear functional** of y, namely

$$\int_X p(x)d\mu(x) = \int_X \sum_k p_k x^k d\mu(x) = \sum_k p_k \int_X x^k d\mu(x) = \sum_k p_k y_k$$

The moment cone

We can replace measures with moments provided we know the relations satisfied by the sequence elements $(y_k)_{k\in\mathbb{N}^n}$ for them to represent a measure

Said otherwise we need a membership oracle for the cone of moments

Since the cone of moments is dual to the cone of positive polynomials, this turns out to be a hard question dating back to Hilbert's 17th problem

Both cones are convex semialgebraic, yet nobody knows how to minimize efficiently linear functions on them

In practice, we have to **approximate** these cones

Semidefinite approximations of the moment cone

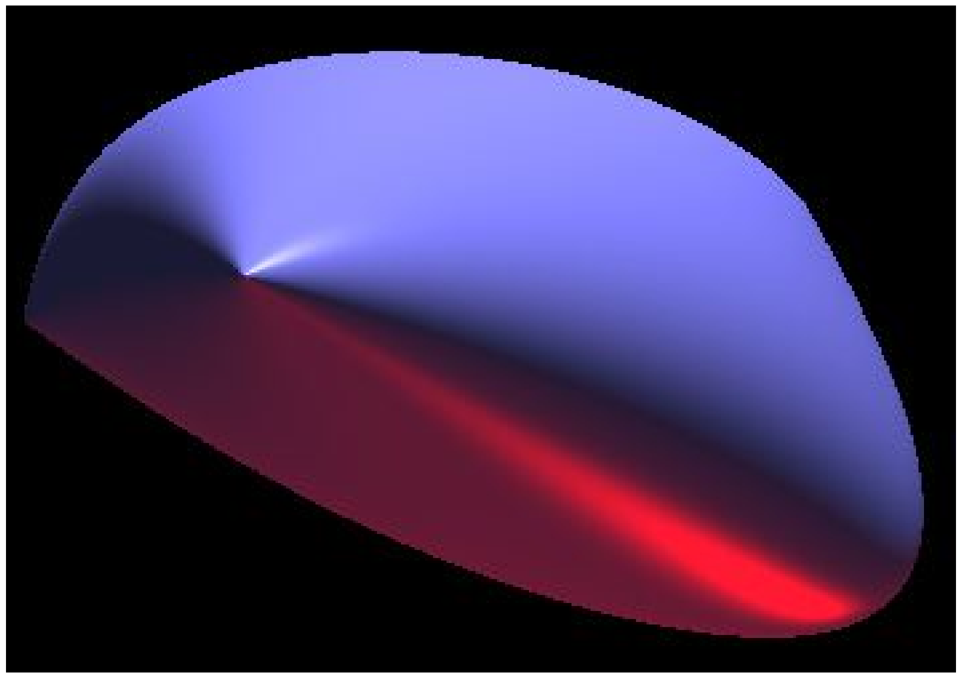
A sufficient condition for a positive polynomial to be non-negative is that it is SOS, i.e. a sum of squares of polynomials

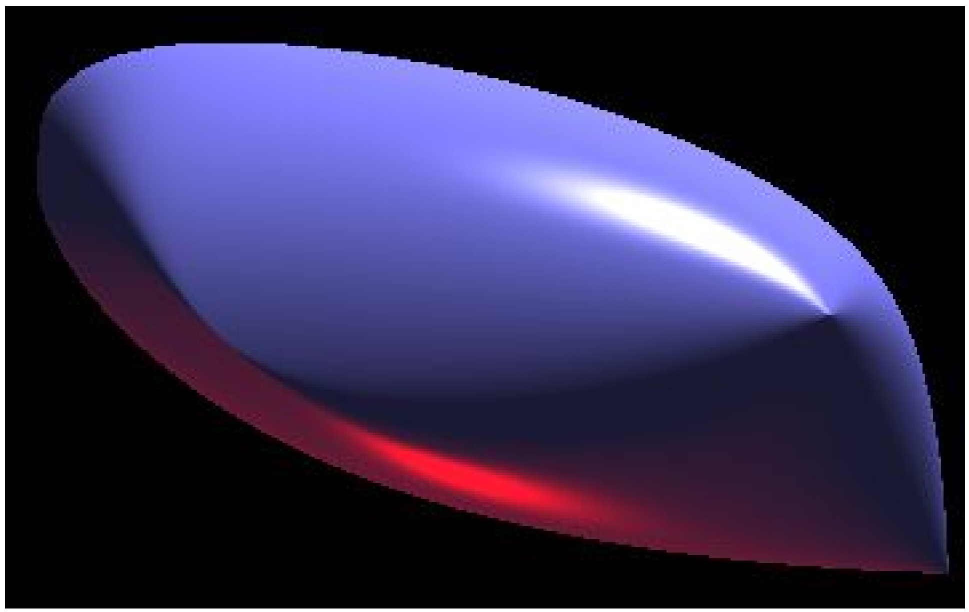
The good news is that the SOS cone is the cone of non-negative quadratic forms, also called the **semidefinite cone**

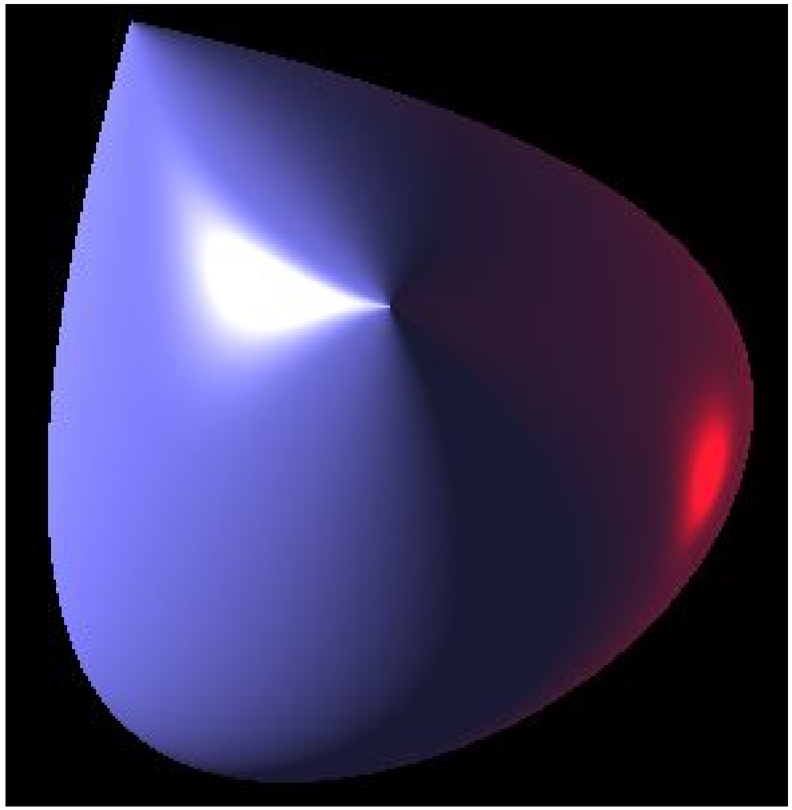
We can minimize linear functions efficiently on the semidefinite cone, this is called semidefinite programming (SDP)

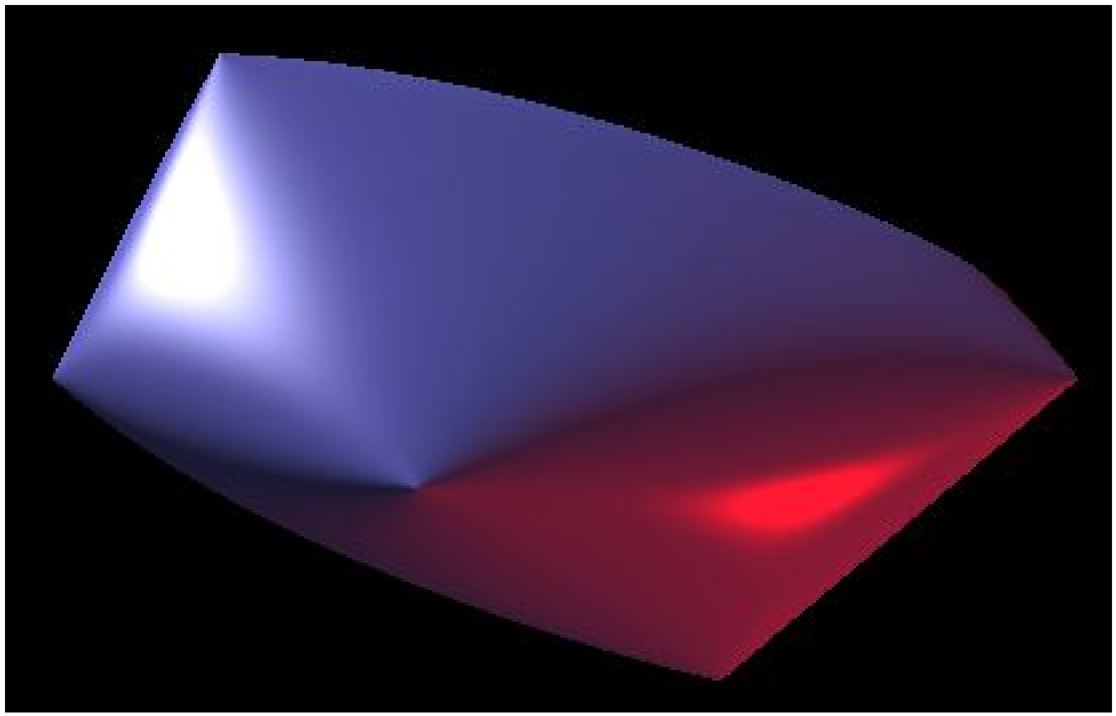
In the engineering literature (control, signals, robotics), linear sections of this cone are called **linear matrix inequalities** (LMI)

In the maths literature (real algebraic geometry, optimization), these are called **spectrahedra**









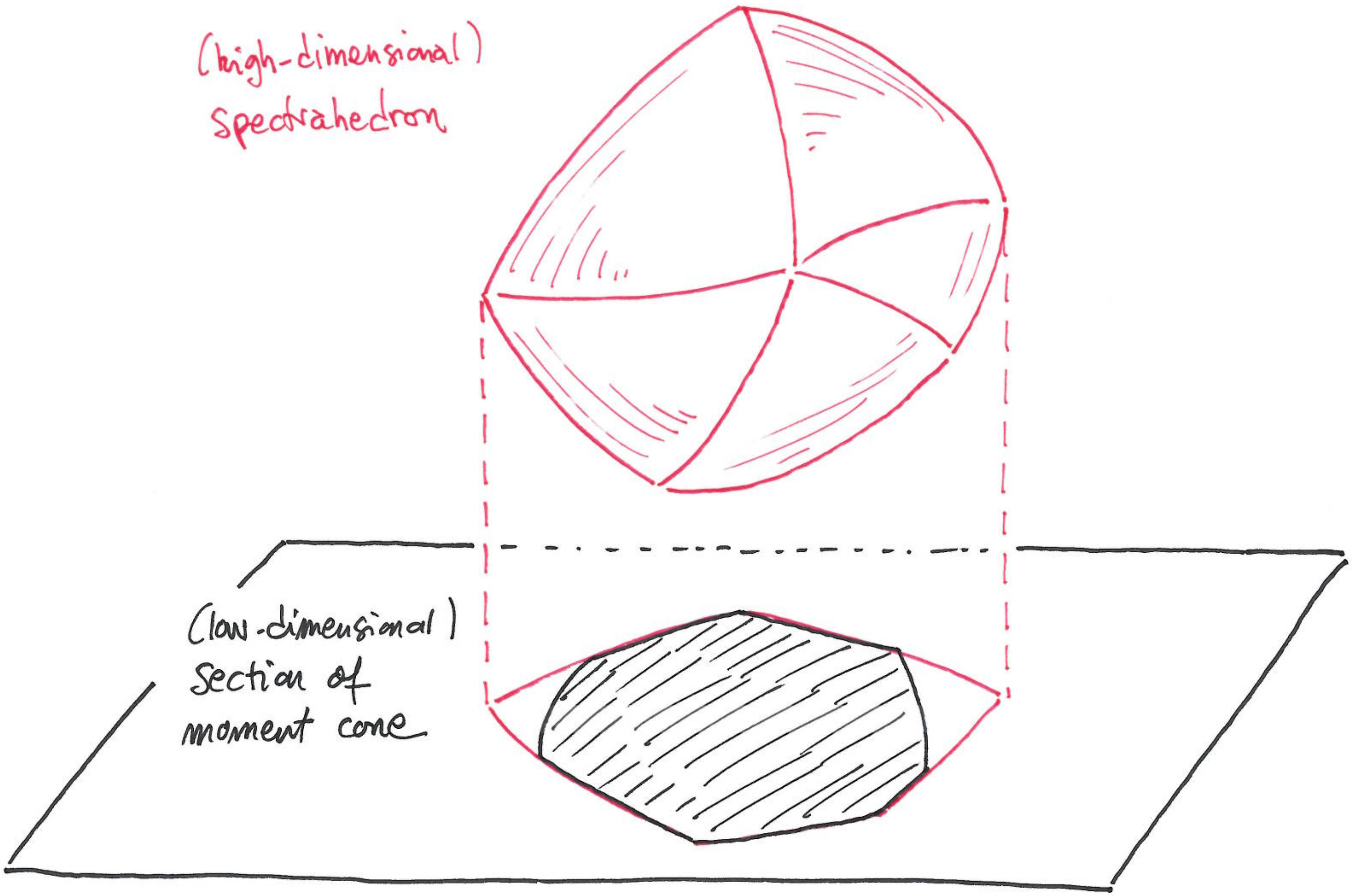
Semidefinite approximations of the moment cone

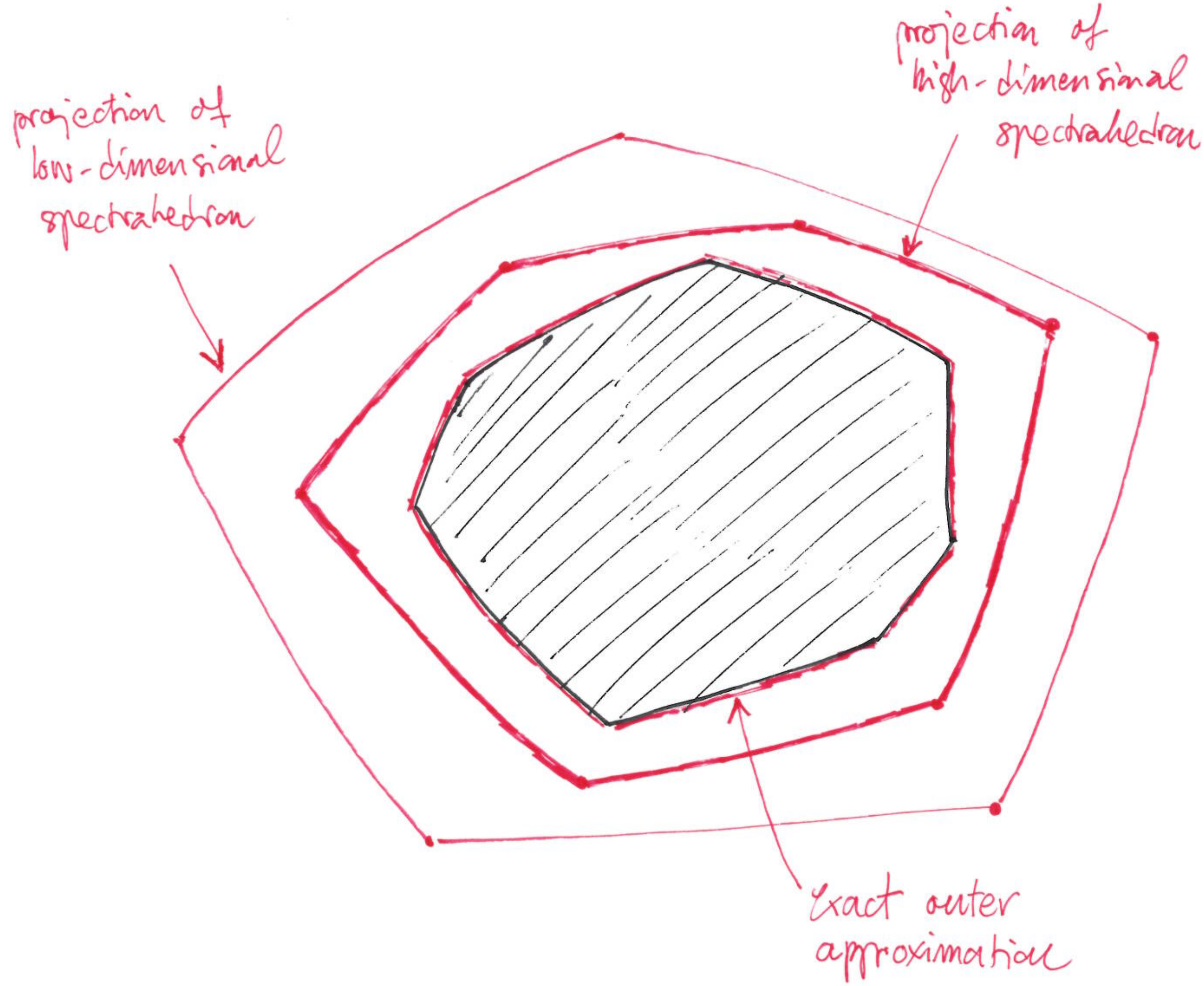
It can be shown that any finite-dimensional moment cone can be approximated or relaxed by projections of spectrahedra of increasing size, **as closely as desired**

This is called the Lasserre hierarchy, or Lasserre's relaxations

We do not know a priori when the approximation is **exact**

In many cases, we can however check exactness **a posteriori**


(high-dimensional)
spectrahedron
(low-dimensional)
section of
moment cone


projection of low-dimensional spectrahedron
projection of high-dimensional spectrahedron
Exact outer approximation

The Lasserre hierarchy strategy (again)

1. Reformulate the original non-convex problem as an infinite-dimensional convex linear problem (LP) on **measures** (and its dual on continuous functions)

2. Relax this measure LP by truncated LPs on **moments** (and their duals on polynomial sums-of-squares, SOS)

3. Solve these moment LPs with **semidefinite programming** (SDP), i.e. via linear matrix inequalities (LMI)

4. Check exactness a posteriori

The Lasserre hierarchy in robotics

In robotics the polynomial optimization problems typically have to be solved in real-time, and the Lasserre hierarchy can be computationally demanding

Experiments reported in [Martin Gurtner. Real-time optimization-based control and estimation for dielectrophoretic micromanipulation. MSc thesis manuscript, Czech Tech Prague, Jan. 2016] show however that small SDP problems can be solved in less than 50ms on a standard PC

Experiments reported in [J. Heller, D. Henrion, T. Pajdla. Hand-eye and robot-world calibration by global polynomial optimization. ICRA 2014] show that many polynomial optimization problems from geometric vision can be solved globally with small SDP relaxations

Coping with dynamics

Let us move

Recall that the polynomial optimization problem $\min_{x \in X} p(x)$ was reformulated as the measure LP

$$\min_{\mu} \int_X p \, d\mu \text{ s.t. } \mu \in \mathscr{P}(X)$$

Similarly, the **polynomial optimal control** problem

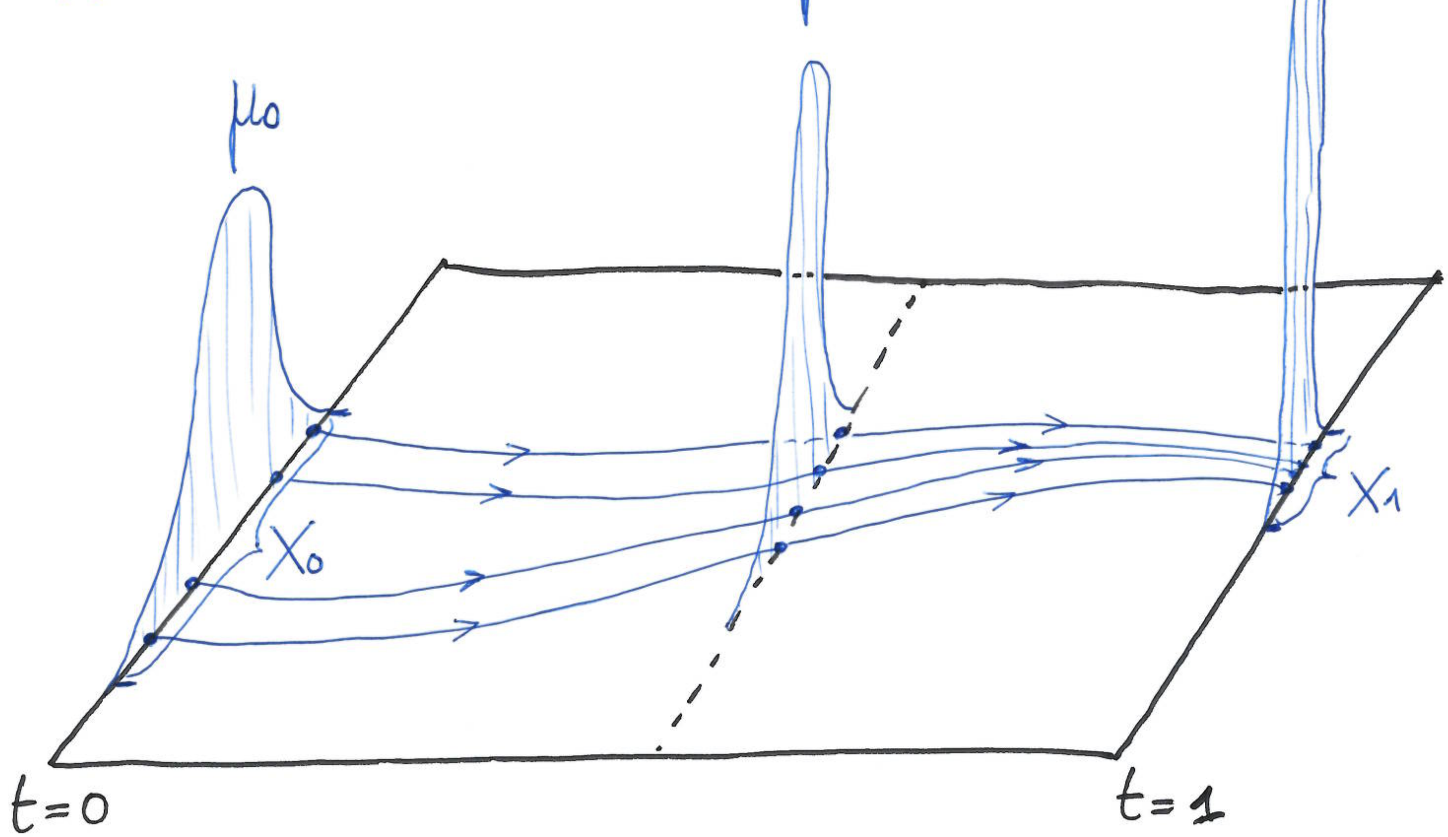
$$\begin{array}{ll} \min & \displaystyle\int_0^1 p(t, x(t), u(t))dt \\ \text{s.t.} & \dot{x}(t) = f(t, x(t), u(t)), \;\; x(0) \in X_0, \;\; x(1) \in X_1 \\ & x \in \mathscr{W}^{1,\infty}([0,1]; X), \;\; u \in \mathscr{L}^{\infty}([0,1]; U) \end{array}$$

is reformulated as the measure LP

$$\begin{array}{ll} \min & \displaystyle\int p \, d\mu \\ \text{s.t.} & \frac{\partial \mu}{\partial t} + \operatorname{div} f\mu + \mu_1 = \mu_0 \\ & \mu_0 \in \mathscr{P}(\{0\} \times X_0), \;\; \mu_1 \in \mathscr{P}(\{1\} \times X_1), \;\; \mu \in \mathscr{P}([0,1] \times X \times U) \end{array}$$

Liouville equation

$$\frac{\partial \mu}{\partial t} + \operatorname{div} f\mu + \mu_1 = \mu_0$$



Trajectory planning

Occupation measures and Lasserre hierarchy in optimal control

J. B. Lasserre, D. Henrion, C. Prieur, E. Trélat. Nonlinear optimal control via occupation measures and LMI relaxations. SIAM J. Control Opt. 47(4):1643-1666, 2008

Region of attraction approximation with Lasserre hierarchy

D. Henrion, M. Korda. Convex computation of the region of attraction of polynomial control systems. IEEE Trans. Autom. Control 59(2):297-312, 2014

Application in robotics

A. Majumdar, R. Vasudevan, M. M. Tobenkin, R. Tedrake. Convex optimization of nonlinear feedback controllers via occupation measures. Intl. J. Robotics Research 33(9):1209:1230, 2014

Take home messages

The Lasserre hierarchy allows to solve **nonconvex** optimization and control problem at the price of solving a family of **convex** semidefinite programs (SDP) on increasing size

Two sides of the same coin: primal SDP on **moments**, dual SDP on polynomial sums of squares (SOS)

Blurs divide between convex=easy and nonconvex=hard

More relevant measure of **complexity** is depth in hierarchy

The Lasserre hierarchy is aimed at small-dimensional instances (up to 10 variables) yet data **sparsity** can be heavily exploited

Can use simpler (e.g. LP, QP) outer or inner approximations of the moment cones or SOS cones