Computing invariant measures with the Lasserre hierarchy

Didier Henrion

with Marcelo Forets, Milan Korda, Victor Magron, Igor Mezić

IPAM UCLA 2019

homepages.laas.fr/henrion/papers/invsdp.pdf
homepages.laas.fr/henrion/papers/invmeas.pdf

Partly funded by the ERC Advanced Grant TAMING

Consider a dynamical system

$$x_{t+1} = f(x_t), \quad x_t \in X, \quad t = 0, 1, \ldots$$

on a compact set $X \subset \mathbb{R}^n$

An **invariant measure** $\mu \in \mathscr{M}(X)$ satisfies

$$\mu(B) = \mu(f^{-1}(B)) = \mu(\{x \in X : f(x) \in B\})$$

for all Borel sets $B \subset X$

The set of all invariant measures is a convex cone of $\mathscr{M}(X)$

Let $f_{\#}\mu := \mu \circ f^{-1}$ denote the push-forward or image measure

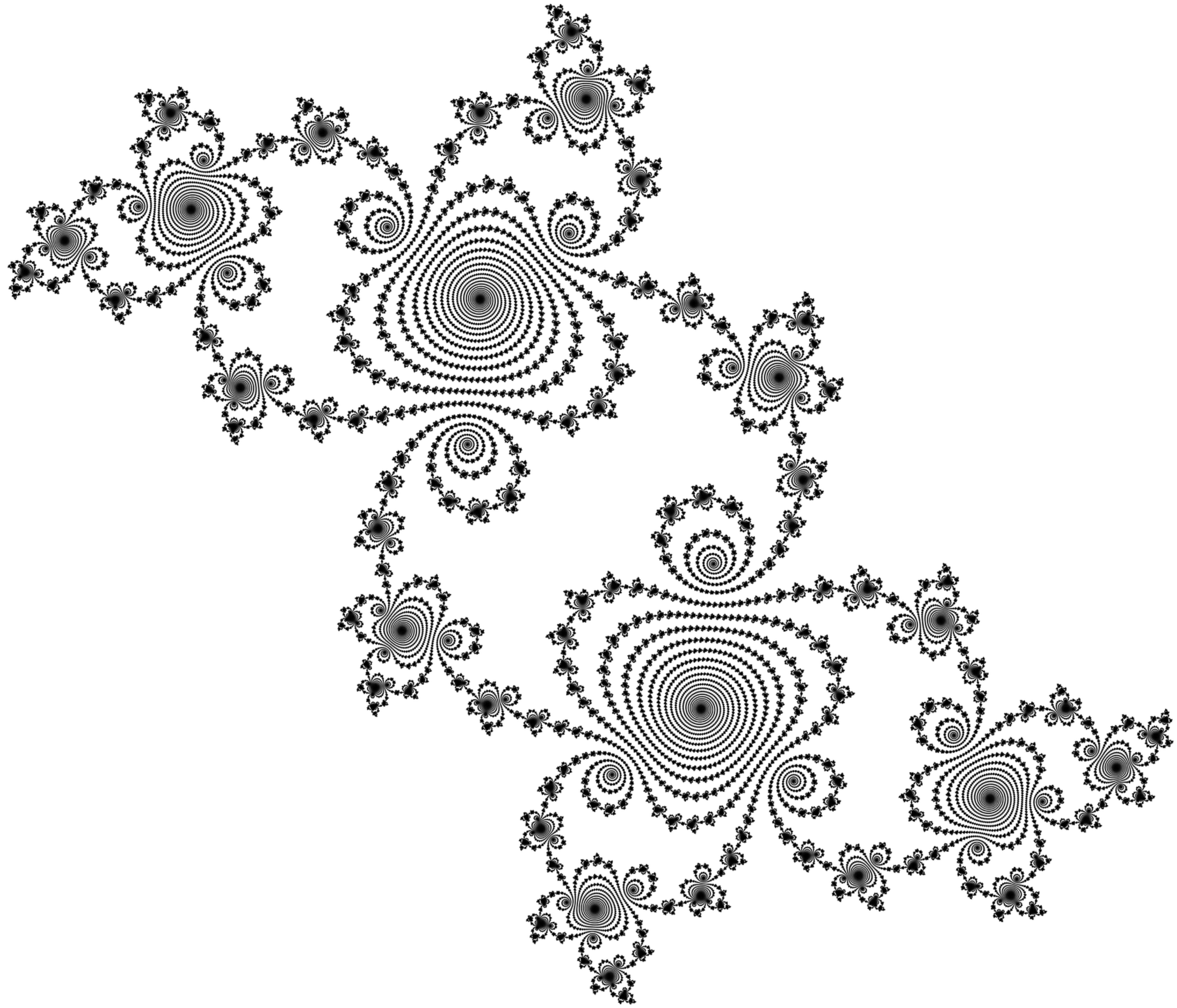
This includes fixed points ($x = f(x)$) with

$$\operatorname{spt}\mu = \{x\}$$

and finite orbits ($x_0, x_1 = f(x_0), x_2 = f(x_1) .., x_0 = f(x_k)$) with

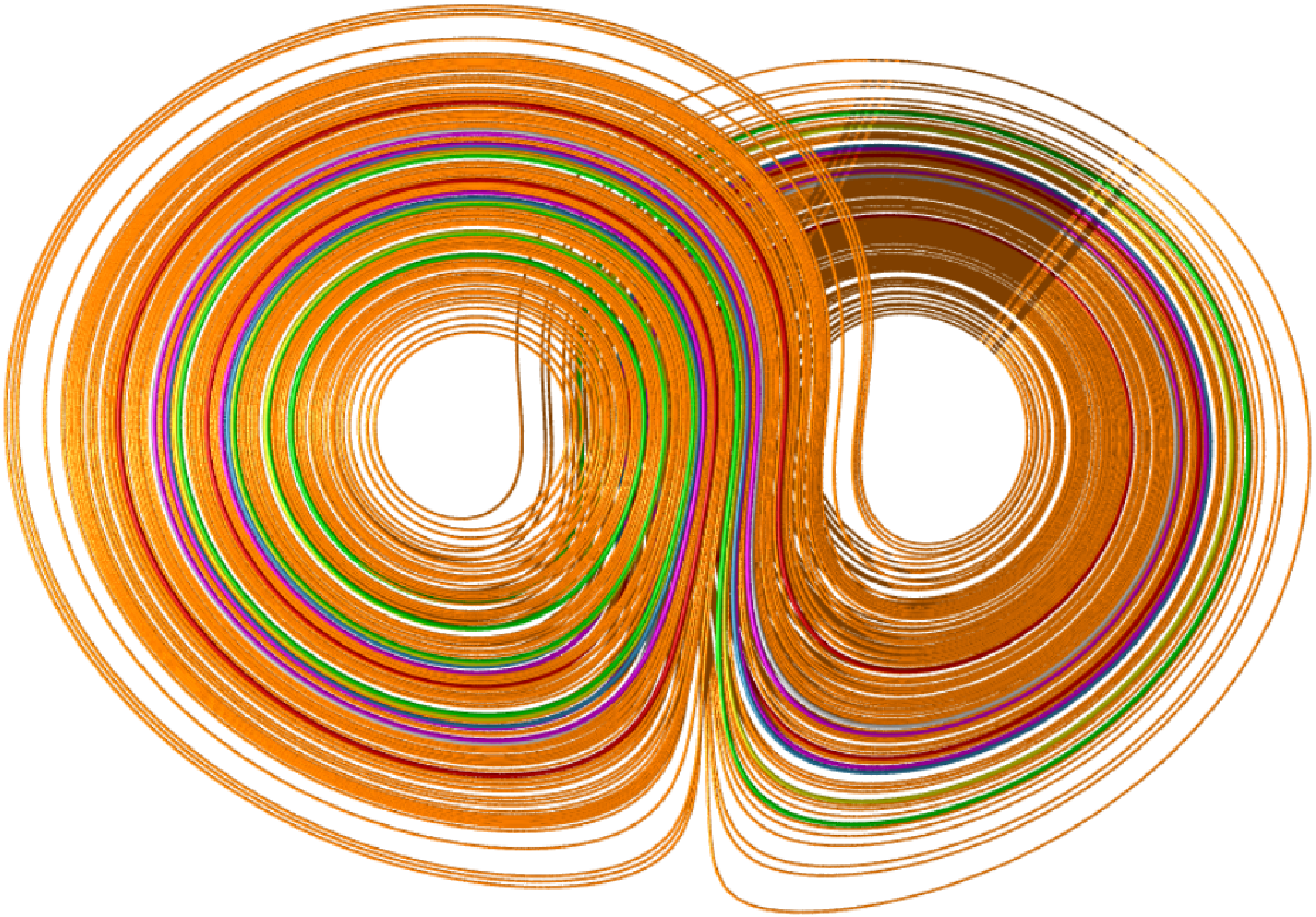
$$\operatorname{spt}\mu = \{x_0, x_1, x_2, \ldots, x_k\}$$

but also infinite orbits, countable or uncountable .. attractors ..



Invariant measures can be also defined for

- continuous time systems
- piecewise dynamical systems
- iterative functional systems
- Markov stochastic systems



www.chaos-math.org/en/chaos-viii-statistics

We would like to approximate the support of invariant measures

We propose two approaches:

1. by regularity: absolutely continuous, singular continuous, singular discrete

2. by convex optimization: ergodic measures, physical measures

For both approaches, the approximation is carried with the **Lasserre hierarchy** (moments - sums of squares) and convex semidefinite programming, provided the data are **polynomial**

Part 1 - invariant measures by regularity

Part 1.1 - absolutely continuous measures

Consider the **infinite-dimensional** conic problem

$$\begin{array}{rll} \rho^*_{\mathrm{ac}} = & \mathrm{sup}_\mu & \int \mu \\ & \mathrm{s.t.} & f_\# \mu = \mu \\ & & \|\mu\|_{\mathscr{L}^p(X)} \leq 1 \end{array}$$

Theorem: this problem has an optimal solution, and if $\rho^*_{\mathrm{ac}} > 0$ then the solution is an invariant measure in $\mathscr{L}^p(X)$

Theorem: if there is a unique invariant probability measure $\mu_{\mathrm{ac}} \in \mathscr{L}^p(X)$ then this problem has a unique optimal solution

$$\mu^*_{\mathrm{ac}} = \rho^*_{\mathrm{ac}} \, \mu_{\mathrm{ac}}$$

How do we compute with measures ?

Since $X \subset \mathbb{R}^n$ is compact, a measure $\mu \in \mathscr{M}(X)$ is uniquely characterized by its **moments**

$$\mathbf{y}_a := \int_X b_a(x) d\mu(x) \in \mathbb{R}, \quad a \in \mathbb{N}^n$$

wrt a dense family $(b_a)_{a \in \mathbb{N}^n} \subset \mathbb{R}[x]$ (e.g. monomials, Chebyshev polynomials)

An invariant measure μ satisfies $f_{\#}\mu = \mu$ i.e.

$$\int_X b_a(f(x)) d\mu(x) = \int b_a(x) d\mu(x), \quad a \in \mathbb{R}^n$$

which is a **linear system** of equations in the moments

$$\mathbf{A}(\mathbf{y}) = 0$$

Given a sequence $(\mathbf{y}_a)_a \subset \mathbb{R}$, define the Riesz linear functional

$$\ell_{\mathbf{y}} : \mathbb{R}[x] \to \mathbb{R}, \quad p(x) = \sum_a \mathbf{p}_a b_a(x) \mapsto \sum_a \mathbf{p}_a \mathbf{y}_a$$

Define the **moment cone**

$$M(X) := \{\mathbf{y} : \mathbf{y}_a = \int_X b_a(x)\, d\mu(x),\ \mu \in \mathscr{M}(X)\}$$

This cone can be approximated from outside by semidefinite cones of increasing size: this is the **Lasserre hierarchy**

Given a compact basic semialgebraic sets

$$X := \{x \in \mathbb{R}^n : g_1(x) = N^2 - \sum_{i=1}^{n} x_i^2 \geq 0,\ g_2(x) \geq 0, \ldots, g_m(x) \geq 0\}$$

with $g_0 := 1$, define the **semidefinite relaxations**

$$\begin{aligned} M_d(X) &:= \{\mathbf{y} : \ell_{\mathbf{y}} \underbrace{(g_k h^2)}_{\text{deg } 2d} \geq 0,\ \forall h \in \mathbb{R}[x],\ k = 0, 1, \ldots, m\} \\ &= \{\mathbf{y} : \underbrace{\mathbf{M}_{d_k}(g_k \mathbf{y})}_{\text{moment matrices}} \succeq 0,\ k = 0, 1, \ldots, m\} \end{aligned}$$

Theorem [Putinar]:

$M_d(X) \supset M_{d+1}(X) \supset \cdots \supset M_\infty(X) = M(X)$

Theorem: if $1 \leq q < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$, the sequence $\mathbf{y}$ has a representing measure μ such that $\|\mu\|_{\mathscr{L}^p(X)} \leq 1$ if and only if

$$\ell_{\mathbf{y}}(g_k h^2) \geq 0,\ \forall h \in \mathbb{R}[x],\ k = 0, 1, \ldots, m$$

and

$$|\ell_{\mathbf{y}}(h)| \leq \ell_{\mathbf{z}}(|h|^q)^{\frac{1}{q}},\ \forall h \in \mathbb{R}[x]$$

where $\mathbf{z}$ are moments of λ_X, the Lebesgue measure on X

Build the Lasserre hierarchy of moment relaxations for $p = 2$

$$
\begin{array}{rcll}
\rho_{\mathsf{ac}}^{d} & := & \sup & \mathbf{y}_0 \\
 & & \text{s.t.} & \mathbf{A}_d(\mathbf{y}) = 0 \\
 & & & \mathbf{M}_{d_k}(g_k \mathbf{y}) \succeq 0,\ k = 0, 1, \ldots, m \\
 & & & \begin{pmatrix} \mathbf{M}_d(\mathbf{z}) & (\mathbf{y}_a)_{|a| \le d} \\ \star & 1 \end{pmatrix} \succeq 0
\end{array}
$$

and for $p = \infty$

$$
\begin{array}{rcll}
\rho_{\mathsf{ac}}^{d} & := & \sup & \mathbf{y}_0 \\
 & & \text{s.t.} & \mathbf{A}_d(\mathbf{y}) = 0 \\
 & & & \mathbf{M}_{d_k}(g_k \mathbf{y}) \succeq 0,\ k = 0, 1, \ldots, m \\
 & & & \mathbf{M}_d(\mathbf{z}) - \mathbf{M}_d(\mathbf{y}) \succeq 0
\end{array}
$$

which are finite-dimensional semidefinite programming problems

Theorem: if there is a unique invariant probability measure $\mu_{\text{ac}} \in \mathscr{L}^p(X)$ with moments $(\mathbf{y}_a)_a$, then the moment relaxations have a sequence $\mathbf{y}^d$ of solutions **converging pointwise**

$$\lim_{d\to\infty} \mathbf{y}_a^d = \rho_{\text{ac}}^* \, \mathbf{y}_a$$

Moreover, the polynomial $h^d \in \mathbb{R}[x]$ of degree $2d$ with coefficients

$$\mathbf{M}_d(g_0\mathbf{y})^{-1}\mathbf{y}^d$$

converges weakly to the density of the invariant measure:

$$\lim_{d\to\infty} \int_X g(x)h^d(x)dx = \rho_{\text{ac}}^* \int_X g(x)\mu_{\text{ac}}(dx), \quad \forall g \in \mathbb{R}[x]$$

Part 1.2 - singular measures

Consider the infinite-dimensional conic problem

$$
\begin{array}{rll}
\rho^*_{\mathsf{sing}} = & \sup & \displaystyle\int \nu \\
& \text{s.t.} & \displaystyle\int \mu = 1 \\
& & f_{\#}\mu = \mu \\
& & \nu + \psi = \mu \\
& & \nu + \widehat{\nu} = \lambda_X \\
& & \mu, \nu, \widehat{\nu}, \psi \in \mathscr{M}(X)
\end{array}
$$

Theorem: if there is a unique invariant probability measure $\mu^* \in \mathscr{M}(X)$, then this problem has a unique optimal solution $(\mu^*, \nu_1^*, \lambda_X - \nu_1^*, \mu^* - \nu_1^*)$ where $\nu_1^* := \max\{1, \nu^*\}$ and $(\nu^*, \mu^* - \nu^*)$ is the **Lebesgue decomposition** of μ^* wrt λ_X

How can we visualize the support of a measure ?

Given the moments of a probability measure μ, perform an eigenvalue decomposition of its moment matrix

$$\mathbf{M}_d(g_0\mathbf{y}) = \mathbf{P}\mathbf{E}\mathbf{P}^T$$

with diagonal $\mathbf{E}$ with entries $e_k \geq 0$, and orthonormal $\mathbf{P}$ with columns $\mathbf{p}_k$ coefficients of polynomials $p_k(x)$

Construct the **Christoffel polynomial**

$$p_{\text{sos}}(x) := \sum_k p_k^2(x)$$

Lemma: the measure μ concentrates on the sublevel sets of the Christoffel polynomial, i.e. for all $\beta \in (0,1)$

$$\mu(\{x : p_{\text{sos}}(x) \leq \frac{\sum_k e_k}{\beta}\}) \geq 1 - \beta$$

Part 2 - invariant measures by convex optimization

Amongst all invariant probability measures, we may want to single out a specific one by solving the **moment** problem

$$\begin{array}{rll} \rho^* = & \min & F(\mathbf{y}) \\ & \text{s.t.} & \mathbf{y}_0 = 1 \\ & & \mathbf{A}(\mathbf{y}) = 0 \\ & & \mathbf{y} \in M(X) \end{array}$$

with its Lasserre hierarchy of **semidefinite relaxations**

$$\begin{array}{rll} \rho^d = & \min & F(\mathbf{y}) \\ & \text{s.t.} & \mathbf{y}_0 = 1 \\ & & \mathbf{A}_d(\mathbf{y}) = 0 \\ & & \mathbf{M}_{d_k}(g_k\mathbf{y}) \succeq 0,\ k = 0, 1, \ldots, m \end{array}$$

Theorem: if F is lower semi-continuous, then $\lim_{d\to\infty} \rho^d = \rho^*$. If $\mathbf{y}^d$ denote an optimal solution of the relaxation, then there is a subsequence converging pointwise to the moments of an optimal invariant measure

Typical choices of objective functions include

$$F(\mathbf{y}) = \sum_a (\mathbf{y}_a - \mathbf{z}_a)^2$$

where $(\mathbf{z}_a)_a$ is a finite vector of given reference moments, or

$$F(\mathbf{y}) = \sum_a \mathbf{f}_a \, \mathbf{y}_a$$

where $(\mathbf{f}_a)_a$ is given for ergodic optimization

An **ergodic measure** μ is an invariant probability measure such that for any Borel set $B \subset X$ such that $f^{-1}(B) = B$, its measure $\mu(B)$ is either 0 or 1

Ergodic measures are extreme points of the convex set of invariant probability measures

Examples

Logistic map

$$f(x) = 2x^2 - 1$$

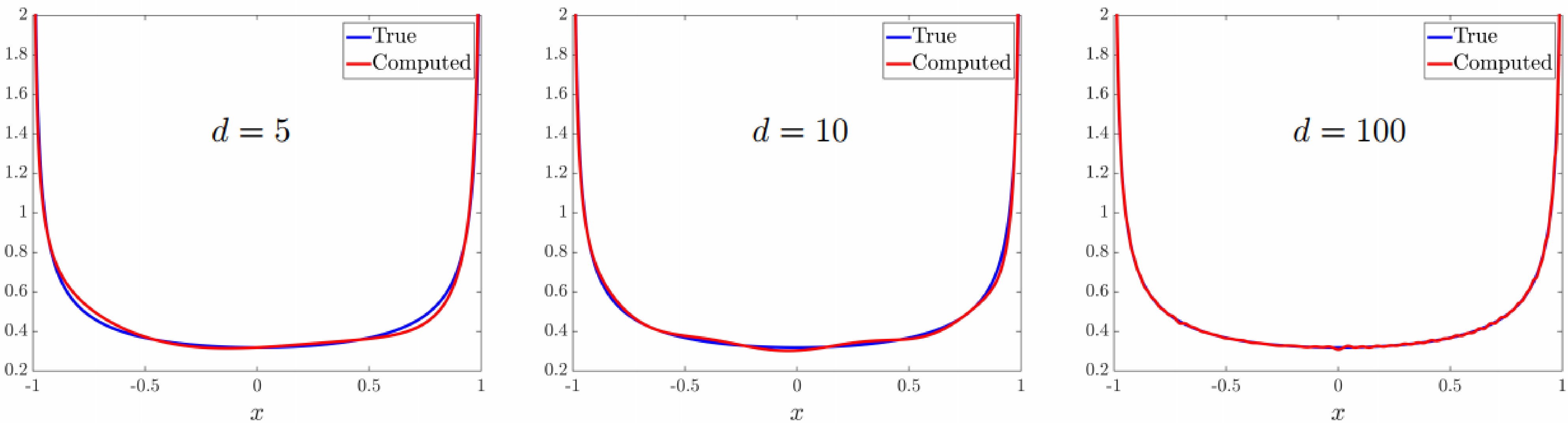
on $X = [-1, 1] = \{x \in \mathbb{R} : (1 + x)(1 - x) \geq 0\}$ for which there is a unique absolutely continuous invariant probability measure

$$\mu(dx) = \frac{1}{\pi}\frac{dx}{\sqrt{1 - x^2}}$$

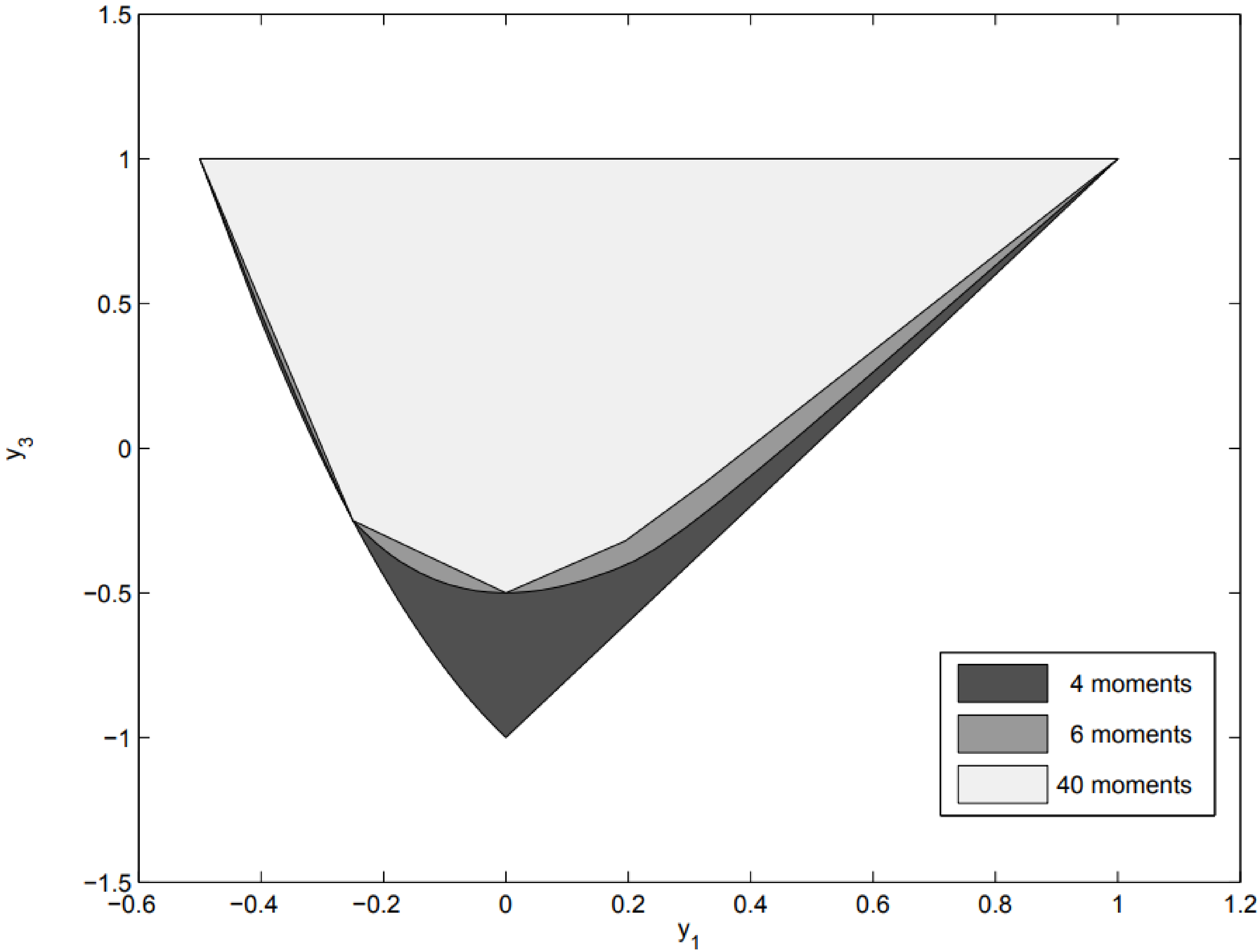
We use the first moment $\mathbf{z}_1 = 0$ for the regression, i.e. $F(\mathbf{y}) = \mathbf{y}_1^2$

We use Chebyshev polynomials for the semidefinite relaxations

homepages.laas.fr/henrion/papers/odds.pdf



True
Computed
$d = 5$
$d = 10$
$d = 100$
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
-1
-0.5
0
0.5
1
x


1.5
1
0.5
0
−0.5
−1
−1.5
−0.6
−0.4
−0.2
0
0.2
0.4
0.6
0.8
1
1.2
y3
y1
4 moments
6 moments
40 moments

For the Hénon map

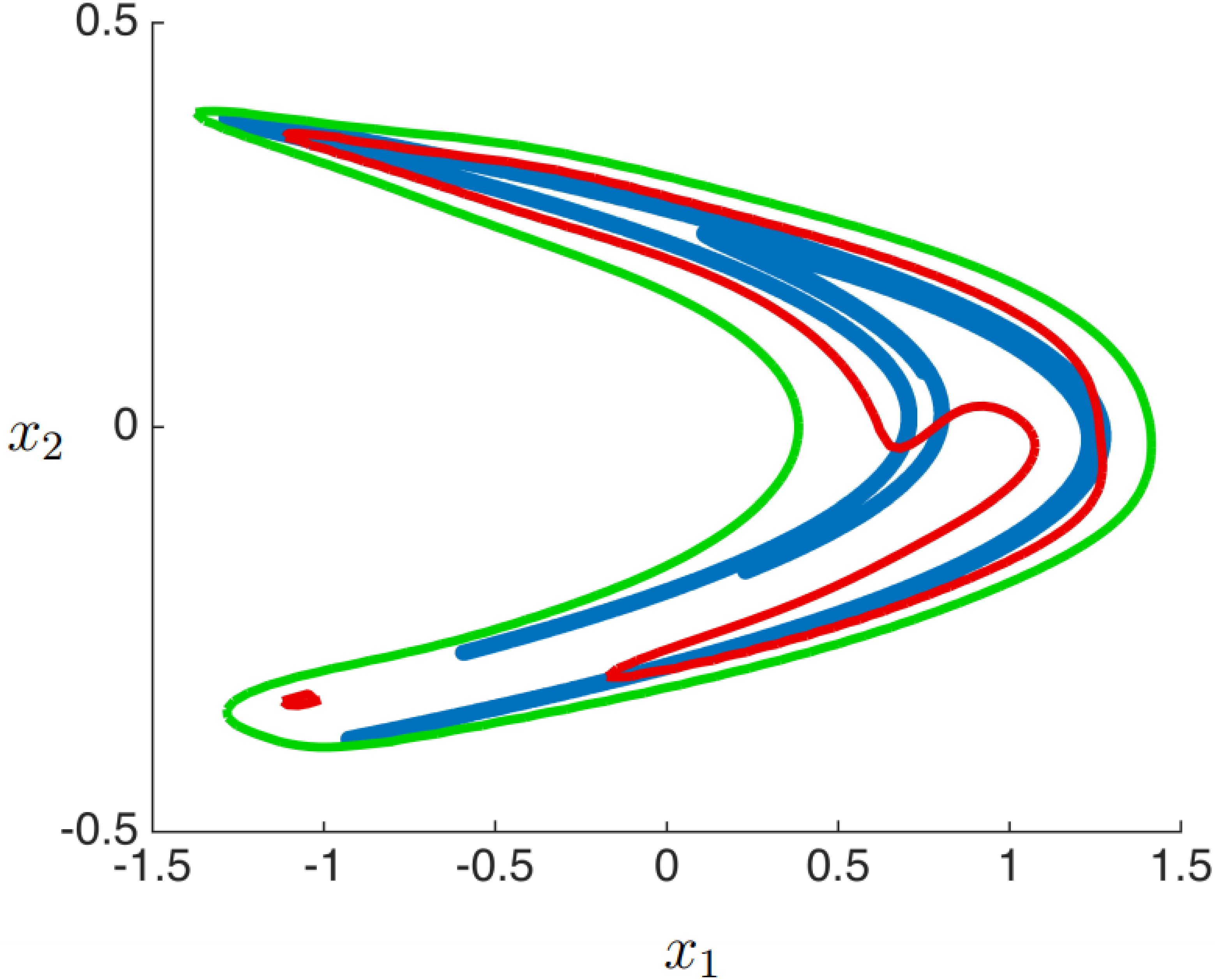
$$f(x_1,\, x_2) = (1 - 1.4x_1^2 + x_2,\, 0.3x_1)$$

on the box $X = [-1.5, 1.5] \times [-0.4, 0.4]$, we use only the first moment $\mathbf{z}_{1,0} = 0.2570$ for the regression

We compare the moments obtained with the relaxation of order $d = 10$ with the simulated moments $\int x^a d\mu(x) \approx \frac{1}{N}\sum_{i=1}^{N}(f(x))^a$

moments	x_1	x_2	x_1^2	x_1x_2	x_2^2
relax	0.2570	0.0771	0.5858	-0.0379	0.0527
simu	0.2570	0.0771	0.5858	-0.0291	0.0527

moments	x_1^3	$x_1^2x_2$	$x_1x_2^2$	x_2^3
relax	0.2468	0.0131	-0.0140	0.0067
simu	0.2320	0.0510	-0.0174	0.0063


0.5
0
-0.5
x_2
-1.5
-1
-0.5
0
0.5
1
1.5
x_1

Computing invariant measures
with the Lasserre hierarchy

Didier Henrion

with Marcelo Forets, Milan Korda, Victor Magron, Igor Mezić

IPAM UCLA 2019

homepages.laas.fr/henrion/papers/invsdp.pdf
homepages.laas.fr/henrion/papers/invmeas.pdf

Partly funded by the ERC Advanced Grant TAMING