

Semidefinite Approximations of Invariant Measures for Polynomial Systems

Victor Magron¹ Marcelo Forets¹ Didier Henrion^{2,3}

July 2, 2018

Abstract

We consider the problem of approximating numerically the moments and the supports of measures which are invariant with respect to the dynamics of continuous- and discrete-time polynomial systems, under semialgebraic set constraints. First, we address the problem of approximating the density and hence the support of an invariant measure which is absolutely continuous with respect to the Lebesgue measure. Then, we focus on the approximation of the support of an invariant measure which is singular with respect to the Lebesgue measure.

Each problem is handled through an appropriate reformulation into a linear optimization problem over measures, solved in practice with two hierarchies of finite-dimensional semidefinite moment-sum-of-square relaxations, also called Lasserre hierarchies.

Under specific assumptions, the first Lasserre hierarchy allows to approximate the moments of an absolutely continuous invariant measure as close as desired and to extract a sequence of polynomials converging weakly to the density of this measure.

The second Lasserre hierarchy allows to approximate as close as desired in the Hausdorff metric the support of a singular invariant measure with the level sets of the Christoffel polynomials associated to the moment matrices of this measure.

We also present some application examples together with numerical results for several dynamical systems admitting either absolutely continuous or singular invariant measures.

Keywords: invariant measures, dynamical systems, polynomial optimization, semidefinite programming, moment-sum-of-square relaxations

1 Introduction

Given a polynomial system described by a discrete-time (difference) or continuous-time (differential) equation under general semialgebraic constraints, we propose numerical

¹CNRS; VERIMAG; 700 av Centrale, 38401 Saint-Martin d'Hères; France

²CNRS; LAAS; Université de Toulouse; 7 avenue du colonel Roche, F-31400 Toulouse; France

³Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 4, CZ-16206 Prague, Czechia

methods to approximate the moments and the supports of the measures which are invariant under the system dynamics. The characterization of invariant measures allows to determine important features of long term dynamical behaviors [21].

One classical way to approximate such features is to perform numerical integration of the equation satisfied by the system state after choosing some initial conditions. However, the resulting trajectory could exhibit some chaotic behaviors or great sensitivity with respect to the initial conditions.

Numerical computation of invariant sets and measures of dynamical systems have previously been studied using domain subdivision techniques, where the density of an invariant measure is recovered as the solution of fixed point equations of the discretized Perron-Frobenius operator [8, 5]. The underlying method, integrated in the software package GAIO [7], consists of covering the invariant set by boxes and then approximating the dynamical behaviour by a Markov chain based on transition probabilities between elements of this covering. More recently, in [9] the authors have developed a multilevel subdivision scheme that can handle uncertain ordinary differential equations as well.

By contrast with most of the existing work in the literature, our method does not rely neither on time nor on space discretization. In our approach, the invariant measures are modeled with finitely many moments, leading to approximate recovery of densities in the absolutely continuous case or supports in the singular case. Our contribution is in the research trend aiming at characterizing the behavior of dynamical nonlinear systems through linear programs (LP), whose unknown are measures supported on the system constraints. This methodology was introduced in the static case by Lasserre [22] and consists of reformulating a polynomial optimization problem as an infinite-dimensional LP over probability measures. To handle practically such an LP, one can then rely on a hierarchy of semidefinite programming (SDP) problems, called moment-sum-of-square or Lasserre hierarchy, see [23] for a global view.

In the context of polynomial dynamical systems, extensions to this hierarchy have been studied, for example, to obtain converging approximations for regions of attraction [12], maximum controlled invariants [18] and reachable sets [31].

In our case, we first focus on the characterization of densities of absolutely continuous invariant measures with respect to some reference measure (for instance the Lebesgue measure). For this first problem, our method is inspired by previous contributions looking for moment conditions ensuring that the underlying unknown measure is absolutely continuous [24] with a bounded density in a Lebesgue space, as a follow-up of the volume approximation results of [13]. When the density function is assumed to be square-integrable, one can rely on [16] to build a sequence of polynomial approximations converging to this density in the L^2 -norm.

We focus later on the characterization of supports of singular invariant measures. For this second problem, we rely on previous works [13, 25, 27] aiming at extracting as much information as possible on the support of a measure from the knowledge of its moments. The numerical scheme proposed in [13] allows to approximate as close as desired the moments of a measure uniformly supported on a given semialgebraic set. The framework from [25] uses similar techniques to compute effectively the Lebesgue decomposition of

a given measure, while [27] relies on the Christoffel function associated to the moment matrix of this measure. When the measure is uniform or when the support of the measure satisfies certain conditions, the sequence of level sets of the Christoffel function converges to the measure support with respect to the Hausdorff distance.

Previous work by the third author [11] shows how to use the Lasserre hierarchy to characterize invariant measures for one-dimensional discrete polynomial dynamical systems. We extend significantly this work in the sense that we now characterize invariant measures on more general multidimensional semialgebraic sets, in both discrete and continuous settings, and we establish convergence guarantees under certain assumptions. In the concurrent work [19], the authors are also using the Lasserre hierarchy for approximately computing extremal measures, i.e. invariant measures optimal w.r.t. a convex criterion. They have weaker convergence guarantees than ours, but the problem is formulated in a more general set-up including physical measures, ergodic measures or atomic measures.

Our contribution is twofold:

- A first Lasserre hierarchy allowing to approximate the moments of an invariant measure which is absolutely continuous with respect to the Lebesgue measure. For an invariant measure with either square integrable or essentially bounded density, one has convergence guarantees of the hierarchy and one can compute asymptotically the exact moment sequence of this measure. At each step of the hierarchy, one can recover an approximate polynomial density from the solution of the SDP problem. The resulting sequence of polynomial approximations converges weakly to the exact density when the degree of the polynomial goes to infinity.
- A second Lasserre hierarchy allowing to approximate as close as desired the support of a singular invariant measure. At each step of this hierarchy, one can recover an approximation of the support with the superlevel set of the Christoffel polynomial associated to the moment matrix of the singular measure. Under certain assumptions, the maximal distance between the exact support and the points among this superlevel set converges to zero when the size of the moment matrix goes to infinity, i.e. we can ensure convergence in the Hausdorff metric.

In both cases, our results apply for both discrete-time and continuous-time polynomial systems. Each problem is handled through an adequate reformulation into an LP problem over probability measures. We show how to solve in practice this infinite-dimensional problem with moment relaxations. In both cases, this boils down to solving a Lasserre hierarchy of finite-dimensional SDP problems of increasing size.

In Section 2, we describe preliminary materials about discrete/continuous-time polynomial systems, invariant measures as well as polynomial sums of squares and moment matrices. The first problem of density approximation for absolutely continuous invariant measures is handled in Section 3 with our first hierarchy. The second problem of support approximation for singular invariant measures is investigated in Section 4 with our second hierarchy. Finally, we illustrate the method with several numerical experiments in Section 5.

2 Invariant Measures and Polynomial Systems

2.1 Discrete-Time and Continuous-Time Polynomial Systems

Given $r, n \in \mathbb{N}$, let $\mathbb{R}[\mathbf{x}]$ (resp. $\mathbb{R}_{2r}[\mathbf{x}]$) stand for the vector space of real-valued n -variate polynomials (resp. of degree at most $2r$) in the variable $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$. Let $\mathbb{C}[\mathbf{x}]$ be the vector space of complex-valued n -variate polynomials. We are interested in the polynomial system defined by

- a polynomial transition map

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad \mathbf{x} \mapsto f(\mathbf{x}) := (f_1(\mathbf{x}), \dots, f_n(\mathbf{x})) \in \mathbb{R}^n[\mathbf{x}] \quad (1)$$

with $f_1, \dots, f_n \in \mathbb{R}[\mathbf{x}]$. The degree of f is given by $d_f := \max\{\deg f_1, \dots, \deg f_n\}$;

- a set of constraints assumed to be compact basic semialgebraic:

$$\mathbf{X} := \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geq 0, \dots, g_m(\mathbf{x}) \geq 0\}, \quad (2)$$

defined by given polynomials $g_1, \dots, g_m \in \mathbb{R}[\mathbf{x}]$.

We develop our approach in parallel for discrete-time and continuous-time systems. With f being a polynomial transition as in (1) and \mathbf{X} being a set of semialgebraic state constraints as in (2), we consider either the discrete-time system:

$$\mathbf{x}_{t+1} = f(\mathbf{x}_t), \quad \mathbf{x}_t \in \mathbf{X}, \quad t \in \mathbb{N}, \quad (3)$$

or the continuous-time system:

$$\dot{\mathbf{x}}(t) = \frac{d\mathbf{x}(t)}{dt} = f(\mathbf{x}(t)), \quad \mathbf{x}(t) \in \mathbf{X}, \quad t \in [0, \infty). \quad (4)$$

2.2 Invariant Measures

Given a compact set $\mathbf{A} \subset \mathbb{R}^n$, we denote by $\mathcal{M}(\mathbf{A})$ the vector space of finite signed Borel measures supported on \mathbf{A} , namely real-valued functions on the Borel sigma algebra $\mathcal{B}(\mathbf{A})$. The support of a measure $\mu \in \mathcal{M}(\mathbf{A})$ is defined as the closure of the set of all points \mathbf{x} such that $\mu(\mathbf{B}) \neq 0$ for an open neighborhood \mathbf{B} of \mathbf{x} . We note $\mathcal{C}(\mathbf{A})$ (resp. $\mathcal{C}^1(\mathbf{A})$) the Banach space of continuous (resp. continuously differentiable) functions on \mathbf{A} equipped with the sup-norm. Let $\mathcal{C}(\mathbf{A})'$ be the topological dual of $\mathcal{C}(\mathbf{A})$ (equipped with the sup-norm), i.e. the set of continuous linear functionals of $\mathcal{C}(\mathbf{A})$. By a Riesz representation theorem, $\mathcal{C}(\mathbf{A})'$ is isomorphically identified with $\mathcal{M}(\mathbf{A})$ equipped with the total variation norm denoted by $\|\cdot\|_{\text{TV}}$. Let $\mathcal{C}_+(\mathbf{A})$ (resp. $\mathcal{M}_+(\mathbf{A})$) be the cone of non-negative elements of $\mathcal{C}(\mathbf{A})$ (resp. $\mathcal{M}(\mathbf{A})$). A probability measure on \mathbf{A} is an element $\mu \in \mathcal{M}_+(\mathbf{A})$ such that $\mu(\mathbf{A}) = 1$. The topology in $\mathcal{C}_+(\mathbf{A})$ is the strong topology of uniform convergence in contrast with the weak-star topology that can be defined in $\mathcal{M}_+(\mathbf{A})$.

The restriction of the Lebesgue measure on a subset $\mathbf{A} \subseteq \mathbf{X}$ is $\lambda_{\mathbf{A}}(d\mathbf{x}) := \mathbf{1}_{\mathbf{A}}(\mathbf{x}) d\mathbf{x}$, where $\mathbf{1}_{\mathbf{A}} : \mathbf{X} \rightarrow \{0, 1\}$ stands for the indicator function of \mathbf{A} , namely $\mathbf{1}_{\mathbf{A}}(\mathbf{x}) = 1$ if $\mathbf{x} \in \mathbf{A}$ and $\mathbf{1}_{\mathbf{A}}(\mathbf{x}) = 0$ otherwise.

The moments of the Lebesgue measure on \mathbf{A} are denoted by

$$z_{\beta}^{\mathbf{A}} := \int \mathbf{x}^{\beta} \lambda_{\mathbf{A}}(d\mathbf{x}) \in \mathbb{R}, \quad \beta \in \mathbb{N}^n \quad (5)$$

where we use the multinomial notation $\mathbf{x}^{\beta} := x_1^{\beta_1} x_2^{\beta_2} \dots x_n^{\beta_n}$. The sequence of Lebesgue moments on \mathbf{A} is denoted by $\mathbf{z}^{\mathbf{A}} := (z_{\beta}^{\mathbf{A}})_{\beta \in \mathbb{N}^n}$. The Lebesgue volume of \mathbf{A} is $\text{vol } \mathbf{A} := z_0^{\mathbf{A}} = \int \lambda_{\mathbf{A}}(d\mathbf{x})$. When $\mathbf{A} = \mathbf{X}$, we define $\mathbf{z} := \mathbf{z}^{\mathbf{X}}$ by omitting the superscript notation. A sequence $\mathbf{y} := (y_{\beta})_{\beta \in \mathbb{N}^n} \in \mathbb{R}^{\mathbb{N}^n}$ is said to have a representing measure on \mathbf{X} if there exists $\mu \in \mathcal{M}(\mathbf{X})$ such that $y_{\beta} = \int \mathbf{x}^{\beta} \mu(d\mathbf{x})$ for all $\beta \in \mathbb{N}^n$.

Given $\mu, \nu \in \mathcal{M}(\mathbf{A})$, the notation

$$\mu \leq \nu$$

means that $\nu - \mu \in \mathcal{M}_+(\mathbf{A})$, namely that μ is dominated by ν .

Given $\mu \in \mathcal{M}_+(\mathbf{A})$, there exists a unique Lebesgue decomposition $\mu = \nu + \psi$ with $\nu, \psi \in \mathcal{M}_+(\mathbf{A})$, $\nu \ll \lambda$ and $\psi \perp \lambda$. Here, the notation $\nu \ll \lambda$ means that ν is absolutely continuous with respect to (w.r.t.) λ , that is, for every $\mathbf{A} \in \mathcal{B}(\mathbf{X})$, $\lambda(\mathbf{A}) = 0$ implies $\nu(\mathbf{A}) = 0$. The notation $\psi \perp \lambda$ means that ψ is singular w.r.t. λ , that is, there exist disjoint sets $\mathbf{A}, \mathbf{B} \in \mathcal{B}(\mathbf{X})$ such that $\mathbf{A} \cup \mathbf{B} = \mathbf{X}$ and $\psi(\mathbf{A}) = \lambda(\mathbf{B}) = 0$.

The so-called pushforward measure or image measure of a given $\mu \in \mathcal{M}_+(\mathbf{X})$ under f is defined as follows:

$$f_{\#}\mu(\mathbf{A}) := \mu(f^{-1}(\mathbf{A})) = \mu(\{\mathbf{x} \in \mathbf{X} : f(\mathbf{x}) \in \mathbf{A}\}), \quad (6)$$

for every set $\mathbf{A} \in \mathcal{B}(\mathbf{X})$.

See [32, Section 21.7] and [6, Chapter IV] or [29, Section 5.10] for additional background on functional analysis, measure theory and applications in convex optimization. For more details on image measures, see e.g. [2, Section 1.5].

Let us define the linear operator $\mathcal{L}_f^{\text{disc}} : \mathcal{C}(\mathbf{X})' \rightarrow \mathcal{C}(\mathbf{X})'$ by:

$$\mathcal{L}_f^{\text{disc}}(\mu) := f_{\#}\mu - \mu$$

and the linear operator $\mathcal{L}_f^{\text{cont}} : \mathcal{C}^1(\mathbf{X})' \rightarrow \mathcal{C}(\mathbf{X})'$ by:

$$\mathcal{L}_f^{\text{cont}}(\mu) := \text{div}(f\mu) = \sum_{i=1}^n \frac{\partial(f_i \mu)}{\partial x_i}$$

where the derivatives of measures are understood in the sense of distributions, that is, through their action on test functions of $\mathcal{C}^1(\mathbf{X})$. In the sequel, we use the more concise notation \mathcal{L}_f to refer to $\mathcal{L}_f^{\text{disc}}$ (resp. $\mathcal{L}_f^{\text{cont}}$) in the context of discrete-time (resp. continuous-time) systems.

Definition 2.1. (Invariant measure) We say that a measure μ is invariant w.r.t. f when $\mathcal{L}_f(\mu) = 0$ and refer to such a measure as an invariant measure. We also omit the reference to the map f when it is obvious from the context and write $\mathcal{L}(\mu) = 0$.

When considering discrete-time systems as in (3), a measure μ is called invariant w.r.t. f when it satisfies $\mathcal{L}_f^{\text{disc}}(\mu) = 0$. When considering continuous-time systems as in (4), a measure is called invariant w.r.t. f when it satisfies $\mathcal{L}_f^{\text{cont}}(\mu) = 0$.

It was proved in [20] that a continuous map of a compact metric space into itself has at least one invariant probability measure. A probability measure μ is ergodic w.r.t. f if for all $\mathbf{A} \in \mathcal{B}(\mathbf{X})$ with $f^{-1}(\mathbf{A}) = \mathbf{A}$, one has either $\mu(\mathbf{A}) = 0$ or $\mu(\mathbf{A}) = 1$. The set of invariant probability measures is a convex set and the extreme points of this set consist of the so-called ergodic measures. For more material on dynamical systems and invariant measures, we refer the interested reader to [21].

2.3 Sums of Squares and Moment Matrices

With \mathbf{X} a basic compact semialgebraic set as in (2), we set $r_j := \lceil (\deg g_j)/2 \rceil, j = 1, \dots, m$. For the ease of further notation, we set $g_0(\mathbf{x}) := 1$. Let $\Sigma[\mathbf{x}]$ stand for the cone of polynomials that can be expressed as sums of squares (SOS) of some polynomials, and let us note $\Sigma_r[\mathbf{x}]$ the cone of SOS polynomials of degree at most $2r$, namely $\Sigma_r[\mathbf{x}] := \Sigma[\mathbf{x}] \cap \mathbb{R}_{2r}[\mathbf{x}]$.

For each $r \geq r_{\min} := \max\{1, r_1, \dots, r_m\}$, let $\mathcal{Q}_r(\mathbf{X})$ be the r -truncated quadratic module generated by g_0, \dots, g_m :

$$\mathcal{Q}_r(\mathbf{X}) := \left\{ \sum_{j=0}^m s_j(\mathbf{x})g_j(\mathbf{x}) : s_j \in \Sigma_{r-r_j}[\mathbf{x}], j = 0, \dots, m \right\},$$

a convex cone of $\mathbb{R}_{2r}[\mathbf{x}]$. Let $\mathcal{P}(\mathbf{X}) := \mathbb{R}[\mathbf{x}] \cap \mathcal{C}_+(\mathbf{X})$ denote the cone of polynomials that are nonnegative on \mathbf{X} . To guarantee the convergence behavior of the relaxations presented in the sequel, we need to ensure that polynomials which are in the interior of $\mathcal{P}(\mathbf{X})$ lie in $\mathcal{Q}_r(\mathbf{X})$ for some $r \in \mathbb{N}$. The existence of such SOS-based representations is guaranteed by Putinar's Positivstellensatz (see e.g. [13, Theorem 2.2]), when the following algebraic compactness condition holds:

Assumption 2.2. There exists a large enough integer N such that one of the polynomials describing the set \mathbf{X} is equal to $N - \|\mathbf{x}\|_2^2$.

In addition, the semialgebraic set \mathbf{X} should fulfill the following condition:

Assumption 2.3. The moments (5) of the Lebesgue measure on \mathbf{X} are available analytically.

This is the case e.g. if \mathbf{X} is a ball. From now on we assume that \mathbf{X} is a compact basic semialgebraic set as in (2) and that it satisfies both Assumptions 2.2 and 2.3.

For all $r \in \mathbb{N}$, we set $\mathbb{N}_r^n := \{\beta \in \mathbb{N}^n : \sum_{j=1}^n \beta_j \leq r\}$, whose cardinality is $\binom{n+r}{r}$. Then a polynomial $p \in \mathbb{R}_r[\mathbf{x}]$ is written as follows:

$$\mathbf{x} \mapsto p(\mathbf{x}) = \sum_{\beta \in \mathbb{N}_r^n} p_\beta \mathbf{x}^\beta,$$

and p is identified with its vector of coefficients $\mathbf{p} = (p_\beta)_{\beta \in \mathbb{N}_r^n}$ in the canonical basis $(\mathbf{x}^\beta)_{\beta \in \mathbb{N}_r^n}$.

Given a real sequence $\mathbf{y} = (y_\beta)_{\beta \in \mathbb{N}^n} \in \mathbb{R}^{\mathbb{N}^n}$, let us define the linear functional $\ell_{\mathbf{y}} : \mathbb{R}[\mathbf{x}] \rightarrow \mathbb{R}$ by $\ell_{\mathbf{y}}(p) := \sum_{\beta} p_\beta y_\beta$, for every polynomial p . For each $j = 0, 1, \dots, m$, we associate to \mathbf{y} a localizing matrix, that is a real symmetric matrix $\mathbf{M}_r(g_j \mathbf{y})$ with rows and columns indexed by $\mathbb{N}_{r-r_j}^n$ and the following entrywise definition:

$$(\mathbf{M}_r(g_j \mathbf{y}))_{\beta, \gamma} := \ell_{\mathbf{y}}(g_j(\mathbf{x}) \mathbf{x}^{\beta+\gamma}), \quad \forall \beta, \gamma \in \mathbb{N}_{r-r_j}^n.$$

When $j = 0$ the localizing matrix is called the moment matrix $\mathbf{M}_r(\mathbf{y}) := \mathbf{M}_{r-r_0}(g_0 \mathbf{y})$.

For a given invariant measure $\mu \in \mathcal{M}(\mathbf{X})$, one has $\mathcal{L}(\mu) = 0$. It follows from the Stone-Weierstrass Theorem that monomials are dense in continuous functions on the compact set \mathbf{X} . The equation $\mathcal{L}(\mu) = 0$ is then equivalent to

$$\ell_{\mathbf{y}}(f(\mathbf{x})^\alpha) - \ell_{\mathbf{y}}(\mathbf{x}^\alpha) = 0, \quad \forall \alpha \in \mathbb{N}^n,$$

in the context of a discrete-time system (3) and

$$\sum_{i=1}^n \ell_{\mathbf{y}} \left(\frac{\partial(\mathbf{x}^\alpha)}{\partial x_i} f_i(\mathbf{x}) \right) = 0, \quad \forall \alpha \in \mathbb{N}^n,$$

in the context of a continuous-time system (4).

Hence, we introduce the linear functionals $\mathcal{I}_{\mathbf{y}}^{\text{disc}} : \mathbb{R}[\mathbf{x}] \rightarrow \mathbb{R}$ defined by

$$\mathcal{I}_{\mathbf{y}}^{\text{disc}}(p) := \ell_{\mathbf{y}}(p \circ f) - \ell_{\mathbf{y}}(p)$$

and $\mathcal{I}_{\mathbf{y}}^{\text{cont}} : \mathbb{R}[\mathbf{x}] \rightarrow \mathbb{R}$ defined by

$$\mathcal{I}_{\mathbf{y}}^{\text{cont}}(p) := \ell_{\mathbf{y}}(\text{grad } p \cdot f)$$

for every polynomial p and where $\text{grad } p := (\frac{\partial p}{\partial x_i})_{i=1, \dots, n}$. In the sequel, we use the more concise notation $\mathcal{I}_{\mathbf{y}}$ to refer to $\mathcal{I}_{\mathbf{y}}^{\text{disc}}$ (resp. $\mathcal{I}_{\mathbf{y}}^{\text{cont}}$) in the context of discrete-time (resp. continuous-time) systems.

3 Absolutely Continuous Invariant Measures

For $p = 1, 2, \dots$, let $L^p(\mathbf{X})$ (resp. $L_+^p(\mathbf{X})$) be the space of (resp. nonnegative) Lebesgue integrable functions f on \mathbf{X} , i.e. such that $\|f\|_p := (\int_{\mathbf{X}} |f(\mathbf{x})|^p \lambda(d\mathbf{x}))^{1/p} < \infty$. Let $L^\infty(\mathbf{X})$ (resp. $L_+^\infty(\mathbf{X})$) be the space of (resp. nonnegative) Lebesgue integrable functions f on \mathbf{X} which are essentially bounded on \mathbf{X} , i.e. such that $\|f\|_\infty := \text{ess sup}_{x \in \mathbf{X}} |f(x)| < \infty$. Two integers p and q are said to be conjugate if $1/p + 1/q = 1$, and by Riesz's representation theorem (see e.g. [28, Theorem 2.14]), the dual space of $L^q(\mathbf{X})$ for $1 \leq q < \infty$ (i.e. the set of continuous linear functionals on $L^q(\mathbf{X})$) is isometrically isomorphic to $L^p(\mathbf{X})$.

For $\mu \in \mathcal{M}(\mathbf{X})$, if $\mu \ll \lambda$ then there exists a measurable function h on \mathbf{X} such that $d\mu = h d\lambda$ and the function h is called the density of μ . If $h \in L^p(\mathbf{X})$, by a slight abuse

of notation, we write $\mu \in L^p(\mathbf{X})$ and $\|\mu\|_p := \|h\|_p$. If in addition μ is invariant w.r.t. f then we say that f has an invariant density in $L^p(\mathbf{X})$.

In Section 3.1, we state some conditions fulfilled by the moments of an absolutely continuous measure with a density in $L^p(\mathbf{X})$. In the case of invariant measures, we rely on these conditions to provide an infinite-dimensional LP characterization in Section 3.2. In Section 3.3, we show how to approximate the solution of this LP by using a hierarchy of finite-dimensional SDP relaxations. Section 3.4 is dedicated to the approximation of the invariant density. In Section 3.5, we extend the approach to piecewise-polynomial systems.

3.1 Invariant Densities in Lebesgue Spaces

Theorem 3.1. *Let p and q be conjugate with $1 \leq q < \infty$. Consider a sequence $\mathbf{y} \subset \mathbb{R}$. The following statements are equivalent:*

- (i) \mathbf{y} has a representing measure $\mu \in L^p_+(\mathbf{X})$ with $\|\mu\|_p \leq \gamma < \infty$ for some $\gamma \geq 0$;
- (ii) there exists $\gamma \geq 0$ such that for all $g \in \mathbb{R}[x]$ it holds

$$|\ell_{\mathbf{y}}(g)| \leq \gamma \ell_{\mathbf{z}}(|g|^q)^{1/q}. \quad (7)$$

and for all $g \in \mathcal{P}(\mathbf{X})$, it holds $\ell_{\mathbf{y}}(g) \geq 0$.

Proof. (i) \implies (ii). Let $h \in L^p_+(\mathbf{X})$ be the density of μ . By Hölder's inequality, it follows that for all $g \in \mathbb{R}[x]$,

$$|\ell_{\mathbf{y}}(g)| \leq \int_{\mathbf{X}} |g| d\mu = \int_{\mathbf{X}} |g| h d\lambda = \|gh\|_1 \leq \gamma \|g\|_q = \gamma \ell_{\mathbf{z}}(|g|^q)^{1/q},$$

which proves (7). Moreover, from the Riesz-Haviland Theorem (see e.g. [23, Theorem 3.1]) one has $\ell_{\mathbf{y}}(g) \geq 0$ for all $g \in \mathcal{P}(\mathbf{X})$.

(i) \longleftarrow (ii). Since $\ell_{\mathbf{y}}(g) \geq 0$ for all $g \in \mathcal{P}(\mathbf{X})$, we rely again on the Riesz-Haviland Theorem to show that \mathbf{y} has a representing measure $\mu \in \mathcal{M}_+(\mathbf{X})$. It remains to prove that $\mu \in L^p_+(\mathbf{X})$. For this, we use a modified version of [37, Theorem 2, p.106], originally stated for complex linear spaces. This theorem relies on the Hahn-Banach extension Theorem in complex linear spaces stated in [37, p.105]. One obtains a similar version for real linear spaces by using the Hahn-Banach extension Theorem in real linear spaces [37, p.102].

By the real version of [37, Theorem 2, p.106] with the notations $i \leftarrow \alpha$, $X \leftarrow L^q_+(\mathbf{X})$, $\beta_i \leftarrow g_\alpha$, $\alpha_i \leftarrow y_\alpha$, $x_i \leftarrow x^\alpha$, $f \leftarrow \tilde{\ell}_{\mathbf{y}}$, there exists a linear functional $\tilde{\ell}_{\mathbf{y}}$ in $L^q(\mathbf{X})$ such that $\tilde{\ell}_{\mathbf{y}}(\mathbf{x}^\alpha) = y_\alpha$ for all $\alpha \in \mathbb{N}^n$, its operator norm $\|\cdot\|$ being bounded by a nonnegative real constant γ . Moreover, the restriction of $\tilde{\ell}_{\mathbf{y}}$ to $\mathbb{R}[\mathbf{x}]$ is $\ell_{\mathbf{y}}$ hence $\|\ell_{\mathbf{y}}\| \leq \|\tilde{\ell}_{\mathbf{y}}\| \leq \gamma$. Since the dual space of $L^q_+(\mathbf{X})$ is isometrically isomorphic to $L^p_+(\mathbf{X})$, there exists $\mu \in L^p_+(\mathbf{X})$ such that $\|\tilde{\ell}_{\mathbf{y}}\| = \|\mu\|_p \leq \gamma$ and satisfying $\tilde{\ell}_{\mathbf{y}}(g) = \ell_{\mathbf{y}}(g) = \int_{\mathbf{X}} g d\mu$, for all $g \in \mathbb{R}[\mathbf{x}]$. \square

Theorem 3.1 provides necessary and sufficient conditions satisfied by the moments of an absolutely continuous Borel measure with a density in $L_+^p(\mathbf{X})$. We now state further characterizations when $p = q = 2$ in Theorem 3.2 and when $p = \infty$ and $q = 1$ in Theorem 3.3. For a given sequence $\mathbf{y} = (y_\alpha)_\alpha$ and $r \in \mathbb{N}$, the notation \mathbf{y}^r stands for the truncated sequence $(y_\alpha)_{|\alpha| \leq 2r}$. The notation $\succeq 0$ means positive semidefinite.

Theorem 3.2. *Consider a sequence $\mathbf{y} \subset \mathbb{R}$. The following statements are equivalent:*

- (i) \mathbf{y} has a representing measure $\mu \in L_+^2(\mathbf{X})$ with $\|\mu\|_2 \leq \gamma < \infty$ for some $\gamma \geq 0$;
- (ii) there exists $\gamma \geq 0$ such that for all $r \in \mathbb{N}$:

$$\begin{pmatrix} \mathbf{M}_r(\mathbf{z}) & \mathbf{y}^r \\ (\mathbf{y}^r)^T & \gamma^2 \end{pmatrix} \succeq 0 \quad (8)$$

and

$$\mathbf{M}_{r-r_j}(g_j \mathbf{y}) \succeq 0, \quad j = 0, 1, \dots, m. \quad (9)$$

Proof. (i) \implies (ii). By using the first necessary condition of Theorem 3.1, there exists $\gamma \geq 0$ such that $\ell_{\mathbf{y}}(g) \leq \gamma \ell_{\mathbf{z}}(g^2)^{1/2}$ for all $g \in \mathbb{R}_r[\mathbf{x}]$. Let $r \in \mathbb{N}$ and choose an arbitrary $g \in \mathbb{R}_r[\mathbf{x}]$ with vector of coefficients \mathbf{g} . Thus, one has $\ell_{\mathbf{y}}(g)^2 \leq \gamma^2 \ell_{\mathbf{z}}(g^2)$, yielding $(\mathbf{g}^T \mathbf{y})^2 \leq \gamma^2 \int_{\mathbf{X}} g^2(\mathbf{x}) d\lambda$. Therefore this implies $\mathbf{g}^T \mathbf{y}^r (\mathbf{y}^r)^T \mathbf{g} \leq \gamma^2 \mathbf{g}^T \mathbf{M}_r(\mathbf{z}) \mathbf{g}$. We obtain (8) by using a Schur complement. We prove that (9) holds in a similar way by using the second necessary condition of Theorem 3.1.

(i) \longleftarrow (ii). Since Assumption 2.2 holds, one can apply Putinar's Positivstellensatz [13, Theorem 2.2] to prove that (9) implies that \mathbf{y} has a representing measure $\mu \in \mathcal{M}_+(\mathbf{X})$. As above, we show that (8) implies that $|\ell_{\mathbf{y}}(g)| \leq \gamma \ell_{\mathbf{z}}(g^2)^{1/2}$ for all $g \in \mathbb{R}_r[\mathbf{x}]$. We conclude the proof by using the sufficient condition of Theorem 3.1. \square

Theorem 3.3. *Consider a sequence $\mathbf{y} \in \mathbb{R}$. The following statements are equivalent:*

- (i) \mathbf{y} has a representing measure $\mu \in L_+^\infty(\mathbf{X})$ with $\|\mu\|_\infty \leq \gamma$ for some $\gamma \geq 0$;
- (ii) there exists $\gamma \geq 0$ such that for all $r \in \mathbb{N}$:

$$\gamma \mathbf{M}_r(\mathbf{z}) \succeq \mathbf{M}_r(\mathbf{y}) \quad (10)$$

$$\mathbf{M}_{r-r_j}(g_j \mathbf{y}) \succeq 0, \quad j = 0, 1, \dots, m. \quad (11)$$

Proof. (i) \implies (ii). By using the first necessary condition of Theorem 3.1, there exists a real $\gamma \geq 0$ such that (i) implies that $|\ell_{\mathbf{y}}(g^2)| \leq \gamma \ell_{\mathbf{z}}(g^2)$, for all $g \in \mathbb{R}_r[\mathbf{x}]$ with vector of coefficients \mathbf{g} . This implies that $\int_{\mathbf{X}} g(\mathbf{x})^2 d\mu \leq \gamma \int_{\mathbf{X}} g(\mathbf{x})^2 d\lambda$. Since $\int_{\mathbf{X}} g^2 d\mu = \mathbf{g}^T \mathbf{M}_r(\mathbf{y}) \mathbf{g}$ and $\int_{\mathbf{X}} g^2 d\lambda = \mathbf{g}^T \mathbf{M}_r(\mathbf{z}) \mathbf{g}$, this shows that $\gamma \mathbf{M}_r(\mathbf{z}) \succeq \mathbf{M}_r(\mathbf{y})$. The remaining inequalities are proved as in Theorem 3.2.

(i) \longleftarrow (ii). As in Theorem 3.2, \mathbf{y} has a representing measure $\mu \in \mathcal{M}_+(\mathbf{X})$. Since (10) holds, we prove as in [13, Lemma 2.4] that $\mu \leq \gamma \lambda_{\mathbf{X}}$ which implies that $\mu \in L^\infty(\mathbf{X})$ with $\|\mu\|_\infty \leq \gamma$. \square

In Section 3.3, we will restrict to the case where $p = 2$ or $p = \infty$ while relying on the characterizations stated in the two previous theorems.

3.2 Infinite-dimensional Conic Formulation

Let us consider the following infinite-dimensional conic program:

$$\begin{aligned} \rho_{ac}^* := \sup_{\mu} \quad & \int_{\mathbf{X}} \mu \\ \text{s.t.} \quad & \mathcal{L}(\mu) = 0, \\ & \|\mu\|_p \leq 1, \\ & \mu \in L_+^p(\mathbf{X}). \end{aligned} \tag{12}$$

Theorem 3.4. *Problem (12) admits an optimal solution. If the optimal value ρ_{ac}^* is positive, then the optimal solution is a nonzero invariant measure.*

Proof. First, let us prove that the feasible set of problem (12) is nonempty and weak-star compact. Nonemptiness follows from the fact the zero measure is admissible. Now, let us consider an admissible sequence $(\mu_n)_{n \in \mathbb{N}}$. One has $\|\mu_n\|_{\text{TV}} = \|\mu_n\|_p \leq 1 < \infty$, which shows that the feasible set of problem (12) is bounded for the weak-star topology. Now, let us assume that the sequence converges weakly-star to μ . For all $\mathbf{A} \in \mathcal{B}(\mathbf{X})$, one has $0 = \mathcal{L}(\mu_n)(\mathbf{A}) \rightarrow \mathcal{L}(\mu)(\mathbf{A})$ as n tends to infinity. In addition, one has $\|\mu_n\|_p \leq 1$ which yields $\|\mu\|_p \leq 1$ as n tends to infinity. Thus μ is feasible for problem (12), which proves that the feasible set of problem (12) is closed in the metric induced by the weak-star topology. This proves that this feasible set is weak-star compact. Problem (12) has a linear cost function and a weak-star compact feasible set, which implies the existence of an optimal solution. The proof of the second statement is straightforward. \square

Assumption 3.5. There exists a unique invariant probability measure $\mu_{ac} \in L^p(\mathbf{X})$ for some $p \geq 1$.

Note that Assumption 3.5 is equivalent to supposing that there exists a unique ergodic probability measure.

Theorem 3.6. *If Assumption 3.5 holds, then problem (12) admits a unique optimal solution $\mu_{ac}^* := \rho_{ac}^* \mu_{ac}$.*

Proof. If Assumption 3.5 holds, then the nonzero invariant measure $\mu_{ac}/\|\mu_{ac}\|_p$ is feasible for problem (12), which proves that $\rho_{ac}^* \geq \int_{\mathbf{X}} \mu_{ac}/\|\mu_{ac}\|_p = 1/\|\mu_{ac}\|_p > 0$. Finally, let μ_{ac}^* be an optimal solution of problem (12), yielding the optimal value $\rho_{ac}^* = \int_{\mathbf{X}} \mu_{ac}^*$. Then, the measure $(\rho_{ac}^*)^{-1} \mu_{ac}^*$ is an invariant probability measure, ensured to be unique from Assumption 3.5, which concludes the proof. \square

The choice of maximizing the mass of the invariant measure in problem (12) is motivated by the following reasons:

- If we consider to solve only the feasibility constraints associated to problem (12), one could end up with a solution being the zero measure, even under Assumption 3.5.

- Enforcing the feasibility constraints by adding the condition for μ to be a probability measure (i.e. $\int_{\mathbf{X}} \mu = \|\mu\|_1 = 1$) would not provide any guarantee to obtain a feasible solution as the inequality constraints $\|\mu\|_p \leq 1$ may not be fulfilled since $\|\mu\|_1 \leq \text{vol } \mathbf{X} \|\mu\|_p \leq \|\mu\|_p$ when $\mu \in L^p(\mathbf{X})$ for some $p \geq 1$.

3.3 A Hierarchy of SDP Relaxations

Let

$$\mathbf{C}_r^2(\mathbf{y}) := \begin{pmatrix} \mathbf{M}_r(\mathbf{z}) & \mathbf{y}^r \\ (\mathbf{y}^r)^T & 1 \end{pmatrix}, \quad \mathbf{C}_r^\infty(\mathbf{y}) := \mathbf{M}_r(\mathbf{z}) - \mathbf{M}_r(\mathbf{y}),$$

and from now on, let $p = 2$ or $p = \infty$ and $r \in \mathbb{N}$ be fixed, with $r \geq r_{\min}$. We build the following hierarchy of finite-dimensional semidefinite programming (SDP) relaxations for problem (12):

$$\begin{aligned} \rho_{ac}^r &:= \sup_{\mathbf{y}} y_0 \\ \text{s.t. } & \mathcal{J}_{\mathbf{y}}(\mathbf{x}^\alpha) = 0, \quad \forall \alpha \in \mathbb{N}_{2r}^n, \\ & \mathbf{C}_r^p(\mathbf{y}) \succeq 0, \\ & \mathbf{M}_{r-r_j}(g_j \mathbf{y}) \succeq 0, \quad j = 0, 1, \dots, m. \end{aligned} \tag{13}$$

Lemma 3.7. *Problem (13) has a compact feasible set and an optimal solution \mathbf{y}^r .*

Proof. First, let us note that the zero sequence is feasible for SDP (13). Each diagonal element of $\mathbf{M}_r(\mathbf{y})$ has to be nonnegative, thus $y_\alpha \geq 0$ for $\alpha \in \mathbb{N}_{2r}^n$. Let $\tau := \max_{\alpha \in \mathbb{N}_{2r}^n} |z_\alpha|$. For all $\alpha \in \mathbb{N}_{2r}^n$, from the constraints $\mathbf{C}_r^p(\mathbf{y}) \succeq 0$, we consider the two following cases.

- For $p = 2$, one has $\mathbf{g}^T \mathbf{y} \mathbf{y}^T \mathbf{g} \leq \mathbf{g}^T \mathbf{M}_r(\mathbf{z}) \mathbf{g}$, for each $g \in \mathbb{R}_r[\mathbf{x}]$ with vector of coefficients \mathbf{g} . In particular for $g(\mathbf{x}) = \mathbf{x}^\alpha$, one obtains $y_\alpha^2 \leq z_{2\alpha} \leq \tau$.
- For $p = \infty$, each diagonal element of $\mathbf{M}_r(\mathbf{z}) - \mathbf{M}_r(\mathbf{y})$ has to be nonnegative, yielding $0 \leq y_\alpha \leq z_\alpha \leq \tau$.

Therefore, this implies that the feasible set of problem (13) is nonempty and compact, ensuring the existence of an optimal solution \mathbf{y}^r . \square

Let us denote by $\mathbb{R}[\mathbf{x}]'$ the dual set of $\mathbb{R}[\mathbf{x}]$, i.e., the linear functionals acting on $\mathbb{R}[\mathbf{x}]$.

Lemma 3.8. *Let Assumption 3.5 hold and let μ_{ac}^* be the unique optimal solution of problem (12). For every $r \geq r_{\min}$, let \mathbf{y}^r be an arbitrary optimal solution of problem (13) and by completing with zeros, consider \mathbf{y}^r as an element of $\mathbb{R}[\mathbf{x}]'$. Then the sequence $(\mathbf{y}^r)_{r \geq r_{\min}} \subset \mathbb{R}[\mathbf{x}]'$ converges pointwise to $\mathbf{y}^* \in \mathbb{R}[\mathbf{x}]'$, that is, for any fixed $\alpha \in \mathbb{N}^n$:*

$$\lim_{r \rightarrow +\infty} y_\alpha^r = y_\alpha^*. \tag{14}$$

Moreover, \mathbf{y}^* has representing measure μ_{ac}^* . In addition, one has:

$$\lim_{r \rightarrow +\infty} \rho_{ac}^r = \rho_{ac}^* = \|\mu_{ac}^*\|_1. \tag{15}$$

Proof. As in the proof of Theorem 3.4 in [25], we show that $(\mathbf{y}^r)_{r \geq r_{\min}}$ converges to (\mathbf{y}^*) since there exists a subsequence of integers (r_k) with $r_k \geq r_{\min}$ such that:

$$\lim_{k \rightarrow \infty} y_\alpha^{r_k} = y_\alpha^*. \quad (16)$$

For an arbitrary integer $r \geq r_{\min}$, it follows from (16) that $0 \preceq \mathbf{M}_r(\mathbf{y}^*)$. As a consequence of [23, Proposition 3.5], \mathbf{y}^* has a representing measure $\mu \in \mathcal{M}_+(\mathbf{X})$. Using the fact that \mathbf{y}^{r_k} is feasible for SDP (13) together with (16), one has $\mathcal{J}_{\mathbf{y}^*}(\mathbf{x}^\alpha) = 0$, for all $\alpha \in \mathbb{N}_{2r}^n$, $\mathbf{C}_r^p(\mathbf{y}) \succeq 0$ and $\mathbf{M}_{r-r_j}(g_j \mathbf{y}) \succeq 0$, for all $j = 0, 1, \dots, m$. Thus this determinate representing measure satisfies $\mathcal{L}(\mu) = 0$ and has a density in $L^p(\mathbf{X})$ by using Theorem 3.2 for the case $p = 2$ and Theorem 3.3 for the case $p = \infty$. This proves that μ is feasible for problem (12) and ensures that $\rho_{\text{ac}}^* \geq \int_{\mathbf{X}} \mu$.

Since problem (13) is a relaxation of problem (12), one has $\rho_{\text{ac}}^* \leq \rho_{\text{ac}}^{r_k}$ for all $k \in \mathbb{N}$. Hence, one has $\rho_{\text{ac}}^* \leq \lim_{k \rightarrow \infty} \rho_{\text{ac}}^{r_k} = \lim_{k \rightarrow \infty} y_0^{r_k} = \int_{\mathbf{X}} \mu$. This shows that μ is an optimal solution of problem (12), which is unique from Theorem 3.6. The accumulation point of \mathbf{y}^r is unique as it is the moment sequence of μ_{ac}^* , yielding (14) and (15), the desired results. \square

Remark 1. Note that without the uniqueness hypothesis made in Assumption 3.5, we are not able to guarantee the pointwise convergence of the sequence of optimal solutions $(\mathbf{y}^r)_{r \geq r_{\min}}$ to \mathbf{y}^* .

Remark 2. One could consider the dual of SDP (13), which is an optimization problem over polynomial sums of squares (SOS). One way to prove the non-existence of invariant densities in $L^p(\mathbf{X})$ for $p \in \{2, \infty\}$ is to use the output of this dual program, yielding SOS certificates of infeasibility.

3.4 Approximations of Invariant Densities

Recall that $p = 2$ or ∞ . Given a solution \mathbf{y}^r of the SDP (13) at finite order $r \geq r_{\min}$, let $h^r \in \mathbb{R}_{2r}[\mathbf{x}]$ be the polynomial with vector of coefficients \mathbf{h}^r given by:

$$\mathbf{h}^r := \mathbf{M}_r(\mathbf{z})^{-1} \mathbf{y}^r \quad (17)$$

where the moment matrix $\mathbf{M}_r(\mathbf{z})$ is positive definite hence invertible for all $r \in \mathbb{N}$. Note that the degree of the extracted invariant density depends on the SDP relaxation order r , and higher relaxation orders lead to higher degree approximations.

Lemma 3.9. *Let Assumption 3.5 hold. For every $r \geq r_{\min}$, let \mathbf{y}^r be an optimal solution of SDP (13), let h^r be the corresponding polynomial obtained as in (17) and let μ_{ac}^* be the unique optimal solution of problem (12) with density h_{ac}^* . Then, the following convergence holds:*

$$\lim_{r \rightarrow +\infty} \int_{\mathbf{X}} g(\mathbf{x}) h^r(\mathbf{x}) d\lambda = \int_{\mathbf{X}} g(\mathbf{x}) h_{\text{ac}}^*(\mathbf{x}) d\lambda,$$

for all $g \in \mathbb{R}[\mathbf{x}]$.

Proof. By definition (17), one has:

$$\int_{\mathbf{X}} \mathbf{x}^\alpha h^r(\mathbf{x}) d\lambda = [\mathbf{M}_r(\mathbf{z}) \mathbf{h}^r]_\alpha = y_\alpha^r,$$

for all $\alpha \in \mathbb{N}_{2r}^n$. As a consequence of (14) from Lemma 3.8, this yields:

$$\int_{\mathbf{X}} \mathbf{x}^\alpha h^r(\mathbf{x}) d\lambda = y_\alpha^r \rightarrow y_\alpha^* = \int_{\mathbf{X}} \mathbf{x}^\alpha h_{ac}^*(\mathbf{x}) d\lambda,$$

as r tends to infinity. Thus, for all $g \in \mathbb{R}[\mathbf{x}]$, $\int_{\mathbf{X}} g(\mathbf{x}) h^r(\mathbf{x}) d\lambda \rightarrow \int_{\mathbf{X}} g(\mathbf{x}) h_{ac}^*(\mathbf{x}) d\lambda$, yielding the desired result. \square

3.5 Extension to Piecewise Polynomial Systems

Now we explain how to extend the current methodology to piecewise polynomial systems. The idea, inspired from [1], consists in using the piecewise structure of the dynamics and the state-space partition to decompose the invariant measure into a sum of local invariant measures supported on each partition cell while being invariant w.r.t. the local dynamics.

Let us consider a set of cell indices I and a union of semialgebraic cells $\mathbf{X} = \bigcup_{i \in I} \mathbf{X}_i \subset \mathbb{R}^n$ partitioning the state-space. For each $i \in I$, the state-space cell \mathbf{X}_i is assumed to be a compact basic semialgebraic set:

$$\mathbf{X}_i := \{\mathbf{x} \in \mathbb{R}^n : g_{i,1}(\mathbf{x}) \geq 0, \dots, g_{i,m_i}(\mathbf{x}) \geq 0\}, \quad (18)$$

defined by given polynomials $g_{i,1}, \dots, g_{i,m_i} \in \mathbb{R}[\mathbf{x}]$, $m_i \in \mathbb{N}$ and fulfilling Assumption 2.2 as well as Assumption 2.3. We set $r_{i,j} := \lceil \deg g_{i,j} / 2 \rceil$, for all $j = 0, 1, \dots, m_i$ and $i \in I$. Then, one considers either the following discrete-time piecewise polynomial system:

$$\mathbf{x}_{t+1} = f_i(\mathbf{x}_t) \quad \text{for } \mathbf{x} \in \mathbf{X}_i, \quad i \in I, \quad t \in \mathbb{N}, \quad (19)$$

or the following continuous-time piecewise polynomial system:

$$\dot{\mathbf{x}} = f_i(\mathbf{x}) \quad \text{for } \mathbf{x} \in \mathbf{X}_i, \quad i \in I, \quad t \in [0, \infty). \quad (20)$$

For $p \geq 1$, this leads to the following infinite-dimensional conic program:

$$\begin{aligned} \sup_{\mu_i} \quad & \sum_{i \in I} \int_{\mathbf{X}_i} \mu_i \\ \text{s.t.} \quad & \sum_{i \in I} \mathcal{L}_{f_i}(\mu_i) = 0, \\ & \sum_{i \in I} \|\mu_i\|_p \leq 1, \\ & \mu_i \in L_+^p(\mathbf{X}_i), \quad i \in I. \end{aligned} \quad (21)$$

Given $\mu_i \in \mathcal{M}_+(\mathbf{X}_i)$ let us define $\mu := \sum_{i \in I} \mu_i \in \mathcal{M}_+(\mathbf{X})$ as well as the linear mapping $\mathcal{L} : \mathcal{M}_+(\mathbf{X}) \rightarrow \mathcal{M}_+(\mathbf{X})$ by $\mathcal{L}(\mu) := \sum_{i \in I} \mathcal{L}_{f_i}(\mu_i)$. Since $\|\mu\|_p^p = \sum_{i \in I} \|\mu_i\|_p^p$, one can rewrite problem (21) as problem (12).

Next, we associate to problem (21) the hierarchy of SDP relaxations indexed by $r \geq \max_{i \in I} \{\max_{0 \leq j \leq m_i} \{r_{i,j}\}\}$:

$$\begin{aligned}
& \max_{\mathbf{y}_i} y_{i,0} \\
& \text{s.t.} \quad \sum_{i \in I} \mathcal{J}_{\mathbf{y}_i}(\mathbf{x}^\alpha) = 0, \quad \forall \alpha \in \mathbb{N}_{2r}^n, \\
& \quad \sum_{i \in I} \mathbf{C}_r^p(\mathbf{y}_i) \succeq 0, \\
& \quad \mathbf{M}_r(g_{i,j} \mathbf{y}_i) \succeq 0, \quad j = 0, 1, \dots, m_i, \quad i \in I.
\end{aligned} \tag{22}$$

As for Lemma 3.8 in Section 3.3, one proves that the sequence of optimal values of SDP (22) converges to the optimal value of problem (21). The extraction of approximate invariant densities can be performed in a way similar to the procedure described in Section 3.4.

4 Singular Invariant Measures

In the sequel, we focus on computing the support of singular measures for either discrete-time or continuous-time polynomial systems. Our approach is inspired from the framework presented in [25], yielding a numerical scheme to obtain the Lebesgue decomposition of a measure μ w.r.t. λ , for instance when λ is the Lebesgue measure. By contrast with [25] where all moments of μ and λ are *a priori* known, we only know the moments of the Lebesgue measure λ in our case but we impose an additional constraint on μ to be an invariant probability measure.

4.1 Infinite-Dimensional LP Formulation

We start by considering the infinite-dimensional linear optimization problem:

$$\begin{aligned}
\rho_{\text{sing}}^* &= \sup_{\mu, \nu, \hat{\nu}, \psi} \int_{\mathbf{X}} \nu \\
& \text{s.t.} \quad \int_{\mathbf{X}} \mu = 1, \quad \mathcal{L}(\mu) = 0, \\
& \quad \nu + \psi = \mu, \quad \nu + \hat{\nu} = \lambda_{\mathbf{X}}, \\
& \quad \mu, \nu, \hat{\nu}, \psi \in \mathcal{M}_+(\mathbf{X}).
\end{aligned} \tag{23}$$

Assumption 4.1. There exists a unique invariant probability measure $\mu^* \in \mathcal{M}_+(\mathbf{X})$.

For a measure ν with density $h \in L_+^\infty(\mathbf{X})$, let us denote by $\max\{1, \nu\}$ the measure with density $x \mapsto \max\{1, h(x)\} \in L_+^\infty(\mathbf{X})$.

Theorem 4.2. *Let Assumption 4.1 hold. Then LP (23) has a unique optimal solution $(\mu^*, \nu_1^*, \lambda_{\mathbf{X}} - \nu_1^*, \mu^* - \nu_1^*)$, where $(\nu^*, \mu^* - \nu^*)$ is the Lebesgue decomposition of μ^* w.r.t. $\lambda_{\mathbf{X}}$ and $\nu_1^* := \max\{1, \nu^*\} \in L_+^\infty(\mathbf{X})$.*

Proof. We first prove that the feasible set of LP (23) is nonempty and weak-star compact. Let us denote by μ^* the unique invariant probability measure for f . Then, nonemptiness follows from the fact that $(\mu^*, 0, \lambda_{\mathbf{X}}, \mu^*)$ is feasible for LP (23). Now, let us consider the sequences of measures $((\mu_n)_n, (\nu_n)_n, (\hat{\nu}_n)_n, (\psi_n)_n)$ such that for all $n \in \mathbb{N}$, $(\mu_n, \nu_n, \hat{\nu}_n, \psi_n)$ is feasible for LP (23). One has $\|\mu_n\|_{\text{TV}} = 1 = \|\nu_n\|_{\text{TV}} + \|\psi_n\|_{\text{TV}} < \infty$ and $(\hat{\nu}_n)_n \leq \text{vol } \mathbf{X} < \infty$ (as \mathbf{X} is bounded). This shows that the feasible set of LP (23) is bounded for the weak-star topology. Now, let us assume that the sequences of measures respectively converge weakly-star to $\mu, \nu, \hat{\nu}$ and ψ . For all $A \in \mathcal{B}(\mathbf{X})$, one has $0 = \mathcal{L}(\mu_n)(A) \rightarrow \mathcal{L}(\mu)(A)$ and we easily see that $(\mu, \nu, \hat{\nu}, \psi)$ is feasible for LP (23), which proves that the feasible set is closed in the metric including the weak-star topology. LP (23) has a linear cost function and a weak-star compact feasible set, thus admits an optimal solution.

The proof that $(\mu^*, \nu_1^*, \lambda_{\mathbf{X}} - \nu_1^*, \mu^* - \nu_1^*)$ is the unique optimal solution of LP (23) is similar to the one of Theorem 3.1 in [25] with the notations $f^* \leftarrow h^*, \gamma \leftarrow 1$. \square

Now we explain the rationale behind LP (23). When there is no absolutely continuous invariant probability measure supported on \mathbf{X} , then LP (23) has an optimal solution $(\mu^*, 0, \lambda_{\mathbf{X}}, \mu^*)$ with μ^* being the unique singular invariant probability measure. In this case, the value of LP (23) is $\rho_{\text{sing}}^* = 0$. Note that in the general case where Assumption 4.1 does not hold, there may be several invariant probability measures. In this case, LP (23) still admits an optimal solution and the optimal value is the maximal mass of the ν -component among all invariant probability measures.

By contrast with problem (12) from Section 3, we enforce the feasibility constraints by adding the condition for μ to be a probability measure. The reason is that if we remove this condition, the value $\rho_{\text{sing}}^* = 0$ could still be obtained with another optimal solution $(0, 0, \lambda_{\mathbf{X}}, 0)$, and we could not retrieve the unique invariant probability measure μ^* .

4.2 A Hierarchy of Semidefinite Programs

For every $r \geq r_{\min}$, we consider the following optimization problem:

$$\begin{aligned} \rho_{\text{sing}}^r &:= \sup_{\mathbf{u}, \mathbf{v}, \hat{\mathbf{v}}, \mathbf{y}} v_0 \\ \text{s.t. } & u_0 = 1, \quad \mathcal{I}_{\mathbf{u}}(\mathbf{x}^\alpha) = 0, \quad \forall \alpha \in \mathbb{N}_{2r}^n, \\ & v_\alpha + y_\alpha = u_\alpha, \quad v_\alpha + \hat{v}_\alpha = z_\alpha, \quad \forall \alpha \in \mathbb{N}_{2r}^n, \\ & \mathbf{M}_{r-r_j}(g_j \mathbf{u}), \mathbf{M}_{r-r_j}(g_j \mathbf{v}) \succeq 0, \quad j = 0, \dots, m, \\ & \mathbf{M}_{r-r_j}(g_j \hat{\mathbf{v}}), \mathbf{M}_{r-r_j}(g_j \mathbf{y}) \succeq 0, \quad j = 0, \dots, m. \end{aligned} \tag{24}$$

Problem (24) is a finite-dimensional SDP relaxation of LP (23), implying that $\rho_{\text{sing}}^r \geq \rho_{\text{sing}}^*$ for every $r \geq r_{\min}$.

Theorem 4.3. *Problem (24) has a compact feasible set and an optimal solution $(\mathbf{u}^*, \mathbf{v}^*, \hat{\mathbf{v}}^*, \mathbf{y}^*)$.*

Proof. First, let us denote by μ^* the invariant probability measure with associated moment sequence \mathbf{u} . Since \mathbf{z} is the moment sequence associated with $\lambda_{\mathbf{X}}$, it follows that SDP (24) has the trivial feasible solution $(\mathbf{u}, 0, \mathbf{z}, \mathbf{u})$.

Then, Assumption 2.2 implies that the semidefinite constraint $\mathbf{M}_{r-1}(g^{\mathbf{X}}\mathbf{u}) \succeq 0$ holds. Thus, the first diagonal element of $\mathbf{M}_{r-1}(g^{\mathbf{X}}\mathbf{u})$ is nonnegative, and since $\ell_{\mathbf{u}}(1) = u_0 = 1$, it follows that $\ell_{\mathbf{u}}(x_i^{2r}) \leq N^r$, $i = 1, \dots, n$. We deduce from [26, Lemma 4.3, p. 111] that $|u_\alpha|$ is bounded for all $\alpha \in \mathbb{N}_{2r}^n$. In addition, from the moment constraints of SDP (24), one has for every $i = 1, \dots, n$:

$$\ell_{\mathbf{v}}(x_i^{2r}), \ell_{\hat{\mathbf{v}}}(x_i^{2r}) \leq \int_{\mathbf{X}} x_i^{2r} d\lambda_{\mathbf{X}}, \quad \ell_{\mathbf{y}}(x_i^{2r}) \leq \ell_{\mathbf{u}}(x_i^{2r}) \leq N^r.$$

which similarly yields that $|v_\alpha|$, $|\hat{v}_\alpha|$ and $|y_\alpha|$ are all bounded for all $\alpha \in \mathbb{N}_{2r}^n$. Therefore, we conclude that the feasible set of SDP (24) is compact and that there exists an optimal solution $(\mathbf{u}^*, \mathbf{v}^*, \hat{\mathbf{v}}^*, \mathbf{y}^*)$. \square

Theorem 4.4. *Let Assumption 4.1 hold. For every $r \geq r_{\min}$, let $(\mathbf{u}^r, \mathbf{v}^r, \hat{\mathbf{v}}^r, \mathbf{y}^r)$ be an arbitrary optimal solution of SDP (24) and by completing with zeros, consider $\mathbf{u}^r, \mathbf{v}^r, \hat{\mathbf{v}}^r, \mathbf{y}^r$ as elements of $\mathbb{R}[\mathbf{x}]'$. The sequence $(\mathbf{u}^r, \mathbf{v}^r, \hat{\mathbf{v}}^r, \mathbf{y}^r)_{r \geq r_{\min}} \subset (\mathbb{R}[\mathbf{x}]')^4$ converges pointwise to $(\mathbf{u}^*, \mathbf{v}^*, \hat{\mathbf{v}}^*, \mathbf{y}^*) \subset (\mathbb{R}[\mathbf{x}]')^4$, that is, for any fixed $\alpha \in \mathbb{N}^n$:*

$$\lim_{r \rightarrow \infty} u_\alpha^r = u_\alpha^*, \quad \lim_{r \rightarrow \infty} v_\alpha^r = v_\alpha^*, \quad \lim_{r \rightarrow \infty} \hat{v}_\alpha^r = z_\alpha^{\mathbf{X}} - v_\alpha^*, \quad \lim_{r \rightarrow \infty} y_\alpha^r = u_\alpha^* - v_\alpha^*. \quad (25)$$

Moreover, with $(\mu^*, \nu_1^*, \lambda_{\mathbf{X}} - \nu_1^*, \mu^* - \nu_1^*)$ being the unique optimal solution of LP (23), \mathbf{u}^* is the moment sequence of the unique invariant probability measure μ^* , \mathbf{v}^* and \mathbf{y}^* are the respective moment sequences of ν_1^* , $\hat{\nu}^* = \lambda_{\mathbf{X}} - \nu_1^*$, $\mu^* - \nu_1^*$.

In addition, one has:

$$\lim_{r \rightarrow \infty} \rho_{\text{sing}}^r = \rho_{\text{sing}}^*.$$

Proof. As in the proof of Theorem 3.4 in [25], we show that $(\mathbf{u}^r, \mathbf{v}^r, \hat{\mathbf{v}}^r, \mathbf{y}^r)_{r \geq r_{\min}}$ converges pointwise to $(\mathbf{u}^*, \mathbf{v}^*, \hat{\mathbf{v}}^*, \mathbf{y}^*)$ since there exists a subsequence of integers (r_k) with $r_k \geq r_{\min}$ such that:

$$\lim_{k \rightarrow \infty} u_\alpha^{r_k} = u_\alpha^*, \quad \lim_{k \rightarrow \infty} v_\alpha^{r_k} = v_\alpha^*, \quad \lim_{k \rightarrow \infty} \hat{v}_\alpha^{r_k} = \hat{v}_\alpha^*, \quad \lim_{k \rightarrow \infty} y_\alpha^{r_k} = y_\alpha^*. \quad (26)$$

By fixing an arbitrary integer $r \geq r_{\min}$, it follows from (26) that $0 \preceq \mathbf{M}_r(\mathbf{u}^*)$, $0 \preceq \mathbf{M}_r(\mathbf{v}^*)$, $0 \preceq \mathbf{M}_r(\hat{\mathbf{v}}^*)$ and $0 \preceq \mathbf{M}_r(\mathbf{y}^*)$. Therefore, by using [23, Proposition 3.5], $\mathbf{u}^*, \mathbf{v}^*, \hat{\mathbf{v}}^*$ and \mathbf{y}^* are the respective moment sequences of determinate representing measures $\mu, \nu, \hat{\nu}$ and ψ , supported on \mathbf{X} . Using the fact that $(\mathbf{u}^{r_k}, \mathbf{v}^{r_k}, \hat{\mathbf{v}}^{r_k}, \mathbf{y}^{r_k})$ is feasible for SDP (24) together with (26), these determinate representing measures must satisfy $\mathcal{L}(\mu) = 0$, $\int_{\mathbf{X}} \mu = 1$, $\nu + \psi = \mu$ and $\nu + \hat{\nu} = \lambda_{\mathbf{X}}$. This proves that $(\mu, \nu, \hat{\nu}, \psi)$ is feasible for LP (23), thus $\rho_{\text{sing}}^* \geq \int_{\mathbf{X}} \nu$. In addition, one has $\rho_{\text{sing}}^* \leq \lim_{k \rightarrow \infty} \rho_{\text{sing}}^{r_k} = \lim_{k \rightarrow \infty} v_0^{r_k} = \int_{\mathbf{X}} \nu$. This shows that $(\mu, \nu, \hat{\nu}, \psi)$ is an optimal solution of LP (23), which is unique from Theorem 4.2. All accumulation points of $(\mathbf{u}^r, \mathbf{v}^r, \hat{\mathbf{v}}^r, \mathbf{y}^r)$ are unique as they are the respective moment sequences of $\mu^*, \nu_1^*, \hat{\nu}^* = \lambda_{\mathbf{X}} - \nu_1^*, \mu^* - \nu_1^*$ and yields (25), the desired result. \square

The meaning of Theorem 4.4 is similar to the one of Theorem 3.4 in [25]. By noting (ν^*, ψ^*) the Lebesgue decomposition of the unique invariant probability measure μ^* , we have the two following cases:

1. If $\nu^* \in L_+^\infty(\mathbf{X})$ with $\|\nu^*\|_\infty \leq 1$ then we can obtain all the moment sequences associated to ν^* and ψ^* by computing \mathbf{v}^r and \mathbf{y}^r through solving SDP (24) as $r \rightarrow \infty$. In [25], the sup-norm must be less than an arbitrary fixed $\gamma > 0$ while in the present study we select $\gamma = 1$ as we consider an invariant probability measure μ^* . In particular, when there is no invariant measure which is absolutely continuous w.r.t. λ , one has $\nu^* = \nu_1^* = 0$, $\psi^* = \mu^*$ and we obtain in the limit the moment sequence \mathbf{y}^* of the singular measure μ^* .
2. If $\nu^* \notin L_+^\infty(\mathbf{X})$ or $\nu^* \in L_+^\infty(\mathbf{X})$ with $\|\nu^*\|_\infty > 1$, then the invariant probability measure μ^* is equal to $\nu' + \psi'$, with $\nu' = \max\{1, \nu^*\} \in L_+^\infty(\mathbf{X})$ and $\psi' = \mu^* - \nu'$ is not singular w.r.t. λ .

4.3 Support Approximations for Singular Invariant Measures

Definition 4.5. (Christoffel polynomial) Let $\mu \in \mathcal{M}_+(\mathbf{X})$ be such that its moments are all finite and that for all $r \in \mathbb{N}$, the moment matrix $\mathbf{M}_r(\mathbf{u})$ is positive definite. With $\mathbf{v}_r(\mathbf{x})$ denoting the vector of monomials of degree less or equal than r , sorted by graded lexicographic order, the Christoffel polynomial is the function $p_{\mu,r} : \mathbf{X} \rightarrow \mathbb{R}$ such that

$$\mathbf{x} \mapsto p_{\mu,r}(\mathbf{x}) := \mathbf{v}_r(\mathbf{x})^T \mathbf{M}_r(\mathbf{u})^{-1} \mathbf{v}_r(\mathbf{x}).$$

The following assumption is similar to [27, Assumption 3.6 (b)]. It provides the existence of a sequence of thresholds $(\alpha_r)_{r \in \mathbb{N}}$ for the Christoffel function associated to a given measure μ in order to approximate the support \mathbf{S} of this measure. Here, we do not assume as in [27, Assumption 3.6 (a)] that the closure of the interior of \mathbf{S} is equal to \mathbf{S} .

Assumption 4.6. Given a measure $\mu \in \mathcal{M}_+(\mathbf{X})$ with support $\mathbf{S} \subseteq \mathbf{X}$, \mathbf{S} has nonempty interior and there exist three sequences $(\delta_r)_{r \in \mathbb{N}}$, $(\alpha_r)_{r \in \mathbb{N}}$, $(d_r)_{r \in \mathbb{N}}$ such that:

- $(\delta_r)_{r \in \mathbb{N}}$ is a decreasing sequence of positive numbers converging to 0.
- For every $r \in \mathbb{N}$, d_r is the smallest integer such that:

$$2^{3 - \frac{\delta_r d_r}{\delta_r + \text{diam } \mathbf{S}}} d_r^n \left(\frac{e}{n}\right)^n \exp\left(\frac{n^2}{d_r}\right) \leq \alpha_r. \quad (27)$$

- For every $r \in \mathbb{N}$, α_r is defined as follows:

$$\alpha_r := \frac{\delta_r^n \omega_n}{\text{vol } \mathbf{S}} \frac{(d_r + 1)(d_r + 2)(d_r + 3)}{(d_r + n + 1)(d_r + n + 2)(2d_r + n + 6)},$$

where $\text{diam } \mathbf{S}$ denotes the diameter of the set \mathbf{S} and $\omega_n := \frac{2\pi^{\frac{n+1}{2}}}{\Gamma(\frac{n+1}{2})}$ is the surface of the n -dimensional unit sphere in \mathbb{R}^{n+1} .

Remark 3. Regarding Assumption 4.6, as mentioned in [27, Remark 3.7], d_r is well defined for all $r \in \mathbb{N}$ and the sequence $(d_r)_{r \in \mathbb{N}}$ is nondecreasing. Since $\text{diam } \mathbf{S} \leq \text{diam } \mathbf{X} \leq 1$ and $\text{vol } \mathbf{S} \leq \text{vol } \mathbf{X} \leq 1$, replacing $\text{diam } \mathbf{S}$ and $\text{vol } \mathbf{S}$ by 1 in (27) yields a result similar to Theorem 4.7. For a given sequence $(\delta_r)_{r \in \mathbb{N}}$, one can compute recursively d_r as well as the threshold α_r for the Christoffel polynomial.

Theorem 4.7. *Let Assumption 4.1 hold and let $\mathbf{S} \subseteq \mathbf{X}$ be the support of the invariant probability measure μ^* . Suppose that there exist sequences $(\delta_r)_{r \in \mathbb{N}}$, $(\alpha_r)_{r \in \mathbb{N}}$ and $(d_r)_{r \in \mathbb{N}}$ such that μ^* , \mathbf{S} , $(\delta_r)_{r \in \mathbb{N}}$, $(\alpha_r)_{r \in \mathbb{N}}$ and $(d_r)_{r \in \mathbb{N}}$ fulfill Assumption 4.6. For every $r \in \mathbb{N}$, let:*

$$\mathbf{S}^r := \{\mathbf{x} \in \mathbf{X} : p_{\mu^*, d_r}(\mathbf{x}) \leq \frac{\binom{d_r+n}{n}}{\alpha_r}\}. \quad (28)$$

Then $\lim_{r \rightarrow \infty} \sup_{\mathbf{x} \in \mathbf{S}^r} \text{dist}(\mathbf{x}, \mathbf{S}) = 0$.

Proof. The proof is similar to the first part of the proof of Theorem 3.8 in [27], which relies on Assumption 4.6 and [27, Lemma 6.6]. We only need to show that the Christoffel polynomial associated to the invariant measure μ^* is affine invariant. Given an invertible affine mapping $g : \mathbf{x} \mapsto \mathbf{A}\mathbf{x} + \mathbf{b}$, with $\mathbf{A} \in \mathbb{R}^{n \times n}$ and $\mathbf{b} \in \mathbb{R}^n$, we prove that the pushforward measure $\tilde{\mu}^* := g_{\#} \mu^*$ has all its moments finite, that the associated moment matrix is positive definite and that $p_{\mu^*, d_r}(\mathbf{x}) = p_{\tilde{\mu}^*, d_r}(\mathbf{A}\mathbf{x} + \mathbf{b})$ for all $\mathbf{x} \in \mathbf{X}$. Since \mathbf{X} is compact, $\tilde{\mu}^*$ has all its moments finite. Since g is invertible, there exists an invertible matrix \mathbf{G} such that $v_r(g(\mathbf{x})) = \mathbf{G}v_r(\mathbf{x})$ for all $\mathbf{x} \in \mathbf{X}$. Denoting by $\tilde{\mathbf{u}}$ the sequence of moments associated to $\tilde{\mu}^*$, one has by definition of the pushforward measure:

$$\begin{aligned} \mathbf{M}_r(\tilde{\mathbf{u}}) &= \int_{\mathbf{X}} \mathbf{v}_r(\mathbf{x})\mathbf{v}_r(\mathbf{x})^T d\tilde{\mu}^*(\mathbf{x}) = \int_{\mathbf{X}} \mathbf{v}_r(g(\mathbf{x}))\mathbf{v}_r(g(\mathbf{x}))^T d\mu^*(\mathbf{x}) \\ &= \mathbf{G} \int_{\mathbf{X}} \mathbf{v}_r(\mathbf{x})\mathbf{v}_r(\mathbf{x})^T d\mu^*(\mathbf{x}) \mathbf{G}^T = \mathbf{G}\mathbf{M}_r(\mathbf{u})\mathbf{G}^T. \end{aligned}$$

This proves that the moment matrix $\mathbf{M}_r(\tilde{\mathbf{u}})$ is positive definite.

In addition, one has for all $\mathbf{x} \in \mathbf{X}$, $p_{\tilde{\mu}^*, r}(g(\mathbf{x})) := \mathbf{v}_r(g(\mathbf{x}))^T \mathbf{M}_r(\tilde{\mathbf{u}})^{-1} \mathbf{v}_r(g(\mathbf{x})) = \mathbf{v}_r(\mathbf{x})^T \mathbf{G}^T \mathbf{G} \mathbf{M}_r(\mathbf{u})^{-1} \mathbf{G}^T \mathbf{G} \mathbf{v}_r(\mathbf{x}) = p_{\mu^*, r}(\mathbf{x})$. \square

Remark 4. In the case when the invariant measure is discrete singular, its moment matrix may not be invertible and we can not approximate the support of this measure with the level sets of the Christoffel polynomial. In this case, one way to recover the support of the measure is to rely on the numerical linear algebra algorithm proposed in [14] for detecting global optimality and extracting solutions of moment problems. This algorithm is implemented in the GloptiPoly [15] software and has been already used in previous work [11] to recover finite cycles in the context of discrete-time systems.

5 Numerical Experiments

Here, we present experimental benchmarks that illustrate our method.

In Section 5.1, we compute the optimal solution \mathbf{y}^r of the primal SDP program (13) for a given positive integer r and $p = 2$ or ∞ , as well as the approximate polynomial density h_p^r defined in (17).

In Section 5.2, we compute the optimal solution \mathbf{u}^r of the primal SDP program (24) for a given positive integer r as well as \mathbf{S}^r , the sublevel set of the Christoffel polynomial defined in (28). In practice, the computation of α_r in (28) relies on the following iterative

procedure: we select $d_r = r$, $\delta_r = 1$ and increment the value of δ_r until the inequality (27) from Assumption 4.6 is satisfied.

SDPs (13) and (24) are both modeled through GloptiPoly [15] via YALMIP toolbox [30] available within MATLAB and interfaced with the SDP solver MOSEK [3]. Performance results are obtained with an Intel Core i7-5600U CPU (2.60 GHz) with 16Gb of RAM running on Debian 8. For each problem, we apply a preprocessing step which consists in scaling data (dynamics, general state constraints) so that the constraint sets become unit boxes. Note that our theoretical framework (including convergence of the SDP relaxations) only works after assuming that there exists a unique invariant probability measure (Assumption 3.5 and Assumption 4.1). Even though this may not hold for some of the considered systems, numerical experiments show that satisfying results can be obtained when approximating invariant densities and supports of singular measures.

5.1 Absolutely Continuous Measures

5.1.1 Square Integrable Invariant density

First, let us consider the one-dimensional discrete-time polynomial system defined by

$$t^+ = T(t) := t + w \pmod{1},$$

with t being constrained in the interval $\mathbf{T} := [0, 1]$ and $w \in \mathbb{R} \setminus \mathbb{Q}$ be an arbitrary irrational number. This dynamics corresponds to the circle rotation with an irrational angle w and thus it has a unique invariant measure equal to the restriction of the Lebesgue measure on \mathbf{T} [36], i.e. Assumption 3.5 is fulfilled here.

Let us consider the square integrable probability density $h^*(t) := \frac{3}{4}t^{-1/4}$ and let $F(t) := \int_0^t h^*(s)ds = t^{3/4}$ be its cumulative distribution function. Since F is invertible, then $F^{-1}(t) = t^{4/3}$ is distributed according to h^* and hence the following dynamical system

$$x^+ = F^{-1} \circ T \circ F(x), \quad (29)$$

with x being constrained in the interval $\mathbf{X} := [0, 1]$, has the invariant measure with density $h^* \in L^2(\mathbf{X})$.

Now, let us take $w \in (0, 1)$. In this case one has $F^{-1} \circ T \circ F(x) = (x^{3/4} + w)^{4/3}$ if $x^{3/4} + w \leq 1$ and $F^{-1} \circ T \circ F(x) = (x^{3/4} + w - 1)^{4/3}$ otherwise. In order to cast the dynamical system from (29) as a piecewise polynomial system, we introduce two additional (so-called *lifting*) variables y, z and consider the system defined as follows:

$$x^+ = y \quad \text{for } (x, y, z) \in \mathbf{X}_1 \cup \mathbf{X}_2 \quad (30)$$

where \mathbf{X}_1 and \mathbf{X}_2 are defined by

$$\begin{aligned} \mathbf{X}_1 &:= \{(x, y, z) \in \mathbb{R}^3 : z(1 - w - z) \geq 0, z^4 = x^3, (z + w)^4 = y^3\}, \\ \mathbf{X}_2 &:= \{(x, y, z) \in \mathbb{R}^3 : (1 - z)(z + w - 1) \geq 0, z^4 = x^3, (z + w - 1)^4 = y^3\}. \end{aligned}$$

Note that the collection $\{\mathbf{X}_1, \mathbf{X}_2\}$ is a partition of $[0, 1]^3$. The variable z represents $x^{3/4}$, for all $x \in [0, 1]$. The variable y represents either $(x^{3/4} + w)^{4/3}$ on \mathbf{X}_1 or $(x^{3/4} + w - 1)^{4/3}$

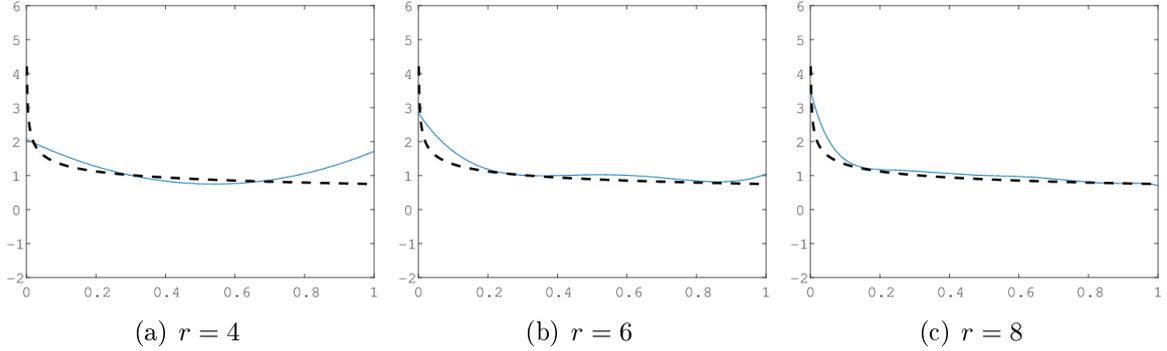


Figure 1: Approximate invariant density for the dynamics from (29) with corresponding approximations h_2^r (solid curve) of the exact density h^* (dashed curve) for $r \in \{4, 6, 8\}$ and $w = \frac{\sqrt{99}}{10}$.

on \mathbf{X}_2 . Using the results from Section 3.5 with $I = \{1, 2\}$, we performed numerical experiments with the irrational number $w = \frac{\sqrt{99}}{10}$. The approximate density h_2^r obtained in (17) from the r first moments (for $r = 4, 6, 8$) and the exact density h^* are displayed on Figure 1. These numerical results indicate that the density approximations become tighter when the value of r increases.

5.1.2 Piecewise Systems

Next, we consider three discrete-time piecewise systems coming respectively from Example 3, Example 4 and Example 5 in [17]. These three systems are known to have unique invariant densities [17, Section 4], thus Assumption 3.5 is fulfilled here.

$$x^+ := \begin{cases} \frac{2x}{1-x^2} & \text{if } x \in \mathbf{X}_1 := [0, \sqrt{2} - 1], \\ \frac{1-x^2}{2x} & \text{if } x \in \mathbf{X}_2 := [\sqrt{2} - 1, 1], \end{cases} \quad (31)$$

$$x^+ := \begin{cases} \frac{2x}{1-x} & \text{if } x \in \mathbf{X}_1 := [0, \frac{1}{3}], \\ \frac{1-x}{2x} & \text{if } x \in \mathbf{X}_2 := [\frac{1}{3}, 1], \end{cases} \quad (32)$$

$$x^+ := \begin{cases} (\frac{1}{8} + 2(x - \frac{1}{2})^3)^{1/3} + \frac{1}{2} & \text{if } x \in \mathbf{X}_1 := [0, \frac{1}{2}], \\ (\frac{1}{8} - 2(x - \frac{1}{2})^3)^{1/3} + \frac{1}{2} & \text{if } x \in \mathbf{X}_2 := [\frac{1}{2}, 1]. \end{cases} \quad (33)$$

As in Section 5.1.1, we introduce an additional lifting variable to handle either the division or the cube root operator. We compute the approximate density h_∞^r from (17) by solving SDP (22). For the system from (31), SDP (22) yields at $r = 6$ (to four significant digits):

$$y_0 = 1, \quad y_1 = 0.3924, \quad y_2 = 0.2463, \quad y_3 = 0.1827, \quad y_4 = 0.1476, \quad y_5 = 0.1254, \quad y_6 = 0.1101.$$

These approximate values are close to the moments associated to the exact invariant density $h^* = \frac{4}{\pi} \cdot \frac{1}{1+x^2}$:

$$y_0^* = 1, \quad y_1^* = 0.4412, \quad y_2^* = 0.2732, \quad y_3^* = 0.1954, \quad y_4^* = 0.1512, \quad y_5^* = 0.1230, \quad y_6^* = 0.1035.$$

The approximate density h_∞^6 and the exact density h^* are displayed on Figure 2(a). Similar results are displayed on Figure 2(b) for the system from (32) and SDP (22) yields:

$$y_0 = 1, y_1 = 0.3628, y_2 = 0.2181, y_3 = 0.1582, y_4 = 0.1262, y_5 = 0.1063, y_6 = 0.0929,$$

with the moments of the exact density $h^*(x) = \frac{2}{(1+x)^2}$ being:

$$y_0^* = 1, y_1^* = 0.3863, y_2^* = 0.2274, y_3^* = 0.1589, y_4^* = 0.1215, y_5^* = 0.0981, y_6^* = 0.0822.$$

For the system from (33), the second order SDP relaxation already provides a very accurate approximation of the exact density $h^*(x) = 12(x - \frac{1}{2})^2$, as shown in Figure 2(c).

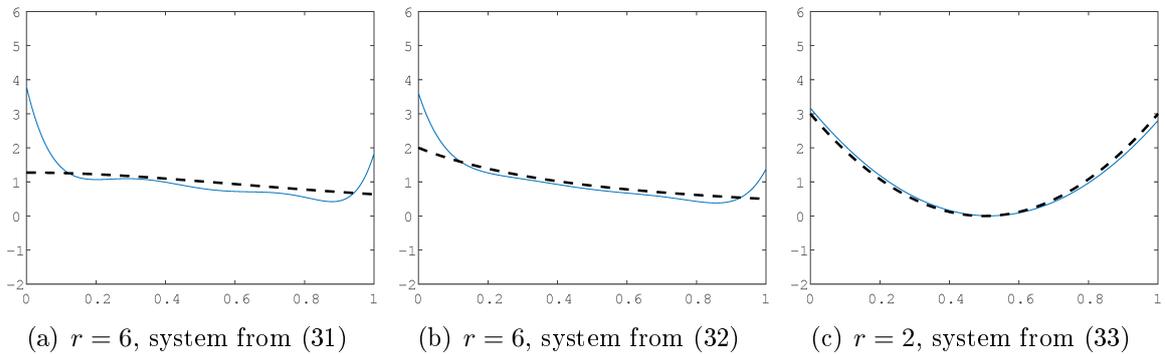


Figure 2: Approximate invariant density for the piecewise systems defined respectively in (31), (32) and (33) with corresponding approximations h_∞^r (solid curve) of the exact density h^* (dashed curve).

5.1.3 Rotational Flow Map

We consider the rotational flow system, i.e. the two-dimensional continuous-time system defined by

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= -x_1. \end{aligned}$$

with general state constraints within the unit disk $\mathbf{X} := \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\|_2 \leq 1\}$. The restriction of the Lebesgue measure to \mathbf{X} is invariant for the rotational flow map $f(\mathbf{x}) := (x_2, -x_1)$. Indeed, denoting by $\mathbf{n}(\mathbf{x})$ the outer normal vector to the unit circle $\partial\mathbf{X}$ at \mathbf{x} , the Green-Ostrogradski formula yields the following, for all $v \in C^1(\mathbf{X})$:

$$\int_{\mathbf{X}} \text{grad } v \cdot f(\mathbf{x}) d\mathbf{x} = \int_{\partial\mathbf{X}} v(\mathbf{x}) \underbrace{(f(\mathbf{x}) \cdot d\mathbf{n}(\mathbf{x}))}_{=0} - \int_{\mathbf{X}} v(\mathbf{x}) \underbrace{\text{div } f(\mathbf{x})}_{=0} d\mathbf{x} = 0,$$

which shows the invariance of $\lambda_{\mathbf{X}}$.

For $p = \infty$, SDP (13) yields at $r = 2$ (to four significant digits): $y_0 = 1, y_1 = y_2 = 0$ and we obtain the approximate polynomial density $h_\infty^r(\mathbf{x}) = 1$, matching the exact invariant density h^* .

5.2 Singular Measures

5.2.1 Hénon Map

The Hénon map is a famous example of two-dimensional discrete-time systems that exhibit a chaotic behavior. The system is defined as follows:

$$\begin{aligned}x_1^+ &= 1 - ax_1^2 + x_2, \\x_2^+ &= bx_1.\end{aligned}$$

with general state constraints within the box $\mathbf{X} := [-3, 1.5] \times [-0.6, 0.4]$. For $a = 1.4$

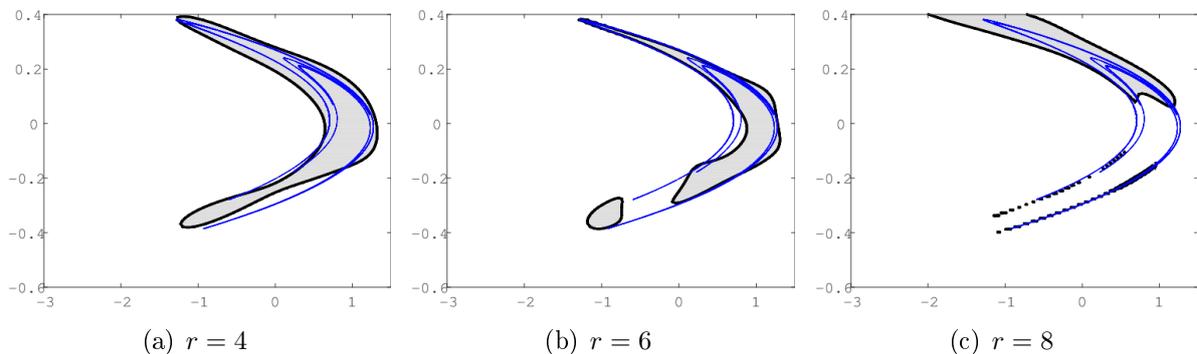


Figure 3: Hénon attractor (blue) and approximations \mathbf{S}^r (light gray) for the support of the invariant measure w.r.t. the map from Example 5.2.1 for $r \in \{4, 6, 8\}$, $a = 1.4$ and $b = 0.3$.

and $b = 0.3$, Hénon proves in [10] that the sequence of points obtained by iteration of this map from an initial point can either diverge or tend to a strange attractor being the product of a one-dimensional manifold by a Cantor set.

Figure 3 displays set approximations \mathbf{S}^r obtained from (28) of the support of the measure invariant w.r.t. the Hénon map, for $r \in \{4, 6, 8\}$. For comparison purpose, we also represented the “true” Hénon attractor by displaying the sequence of points obtained after hundred iterations of the map while starting from random sampled initial conditions within the disk of radius 0.1 and center $[-1, 0.4]$. These numerical experiments show that the level sets of the Christoffel polynomial provide fairly tight approximations of the attractor for modest values of the relaxation order.

5.2.2 Van der Pol Oscillator

The Van der Pol oscillator is an example of an oscillating system with nonlinear damping [35]. The dynamics are given by the following second-order ordinary differential equation:

$$\ddot{x}_1 - a(1 - x_1^2)\dot{x}_1 + x_1 = 0.$$

By setting $x_2 = \dot{x}_1$, one can reformulate this one-dimensional system into a two-dimensional continuous-time system:

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= a(1 - x_1^2)x_2 - x_1. \end{aligned}$$

When $a > 0$, there exists a limit cycle for the system. Here, we consider $a = 0.5$ and general state constraints within the box $\mathbf{X} := [-3, 3] \times [-4, 4]$.

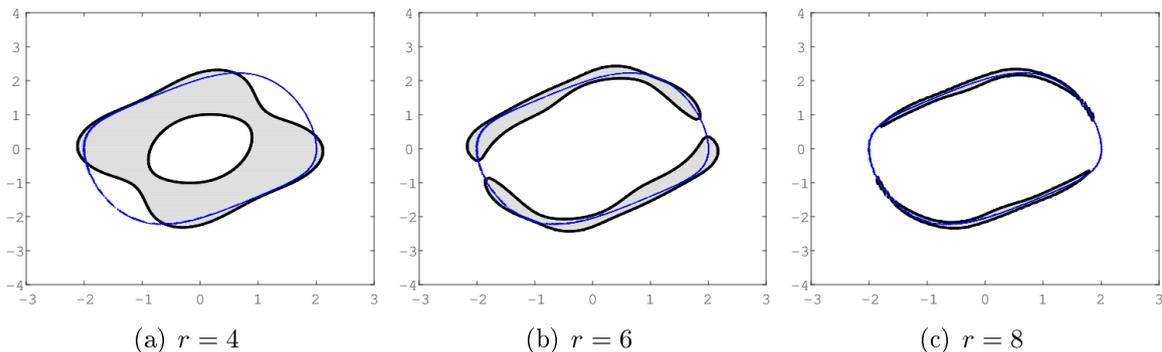


Figure 4: Van der Pol attractor (blue) and approximations \mathbf{S}^r (light gray) for the support of the invariant measure w.r.t. the map from Example 5.2.2 for $r \in \{4, 6, 8\}$ and $a = 0.5$.

Figure 4 shows set approximations \mathbf{S}^r obtained from (28) of the support of the measure invariant w.r.t. the Van der Pol map. As for Example 5.2.1, we also represent the “true” limit cycle after performing a numerical integration of the Van der Pol system from $t_0 = 0$ to $T = 20$ with random sampled initial conditions within the disk of radius 0.1 and center $[1, -1]$. This numerical approximation is done with the `ode45` procedure available inside MATLAB. Once again, the plots exhibit a quite fast convergence behavior of the approximations \mathbf{S}^r of the invariant measure support to the limit cycle when r increases.

5.2.3 Arneodo-Coulet System

Finally, we investigate the Arneodo-Coulet system [4], representing the dynamics of a forced oscillator. This oscillator can be described by the following three-dimensional time-continuous system:

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= x_3, \\ \dot{x}_3 &= -ax_1 - bx_2 - x_3 + cx_1^3. \end{aligned}$$

with general state constraints within the box $\mathbf{X} := [-4, 4] \times [-8, 8] \times [-12, 12]$.

This system exhibits a chaotic behavior for certain parameter values of a , b and c , for instance with $a = -5.5$, $b = 3.5$ and $c = -1$.

As for Example 5.2.2, we represent the “true” attractor after performing numerical integration with the `ode45` procedure, from $t_0 = 0$ to $T = 1000$ with random sampled initial

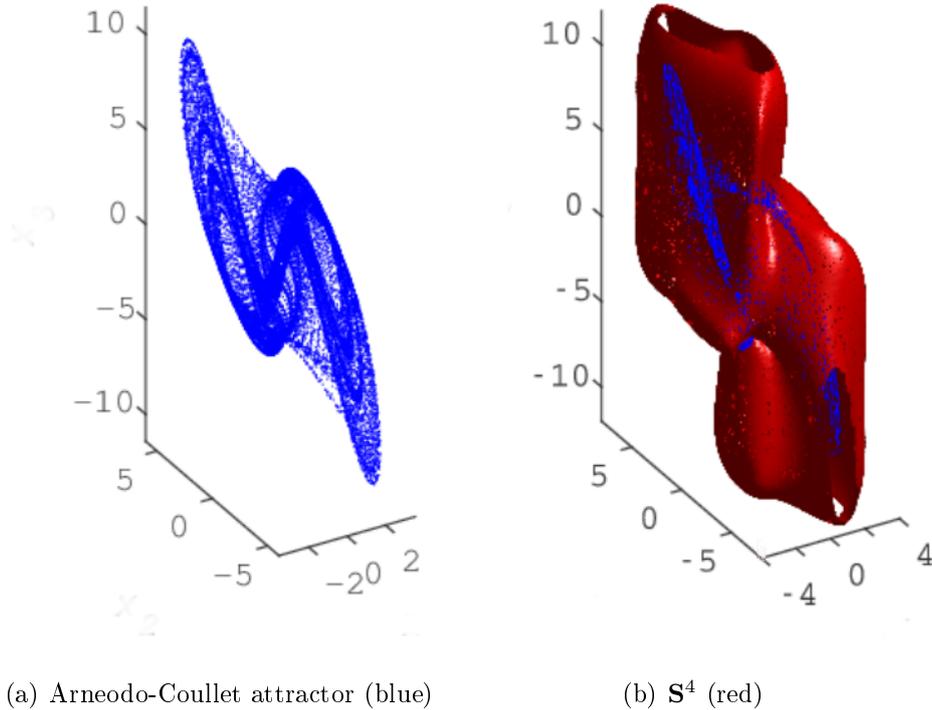


Figure 5: Arneodo-Coulet attractor (blue) and approximations \mathbf{S}^4 (red) for the support of the invariant measure w.r.t. the map from Example 5.2.3 for $a = -5.5$, $b = 3.5$ and $c = -1$.

conditions within the disk of radius 0.1 and center $[1, 1, 1]$. On Figure 5, we observe that the approximation \mathbf{S}^4 obtained from (28) provides a reasonably good estimate of the chaotic attractor. The results obtained for higher relaxation orders were not satisfying because of the ill-conditioning of the moment matrix.

6 Conclusion and Perspectives

We can summarize the contributions of this article as follows:

- We propose two methods to characterize invariant measures for discrete-time and continuous-time systems with polynomial dynamics and semialgebraic state constraints. Our approach can also be extended to piecewise-polynomial systems;
- The first method allows to approximate as close as desired the moments of an absolutely continuous invariant measure, under the assumption that this density is square integrable. The second method allows to approximate as close as desired the support of a singular invariant measure by using sublevel sets of the Christoffel polynomial constructed from the moment matrix;
- Each method relies on solving a hierarchy of finite-dimensional semidefinite programs. While the convergence of the hierarchy is guaranteed in theory, each program

can be solved in practice, thanks to public-available solvers.

Our two methods for recovering invariant densities or singular invariant measure supports both rely on an extension of Lasserre’s hierarchy of semidefinite relaxations, initially introduced in the context of polynomial optimization. Numerical experiments show that our two methods already yield fairly good approximations of the moments of invariant measures at modest relaxation orders. One further research direction would be to investigate scaling both methods to large-size systems, when either sparsity or symmetry occurs.

Regarding the approximation quality of the invariant densities, our first method could produce more satisfactory results by using Chebyshev polynomials or rational function bases as an alternative to the monomials base. Regarding the approximation quality of the supports of singular invariant measures, our second method would practically yield better estimates by handling the issue related to the ill-conditioning of the moment matrix.

Next, we shall devote research efforts to extend the uniform convergence properties of the Christoffel polynomial studied in [27] to certain classes of singular measures. Another track of investigation would be to develop a similar hierarchy of semidefinite programs to study the support of atomic discrete measures corresponding to finite cycles. A first attempt made in [11] consists in setting the objective function of these programs as particular linear moment combinations. However, the question raised about choosing the adequate objective function to recover a given finite cycle still remains open.

Acknowledgements The authors would like to specially acknowledge the precious help of Milan Korda for providing the example from Section 5.1.1 as well as for his feedback and suggestions.

References

- [1] M. R. Abdalmoaty, D. Henrion, L. Rodrigues. Measures and LMIs for optimal control of piecewise-affine systems. Proceedings of the European Control Conference (ECC), 2013.
- [2] L. Ambrosio, N. Fusco, D. Pallara. Functions of bounded variation and free discontinuity problems. Clarendon Press, 2000.
- [3] E. D. Andersen, K. D. Andersen. The Mosek Interior Point Optimizer for Linear Programming: An Implementation of the Homogeneous Algorithm. In H. Frenk, K. Roos, T. Terlaky, S. Zhang, (Editors) High Performance Optimization, Vol. 33 of Applied Optimization, Springer, 2000.
- [4] A. Arneodo, P. H. Couillet, E. A. Spiegel, C. Tresser. Asymptotic chaos. Physica D Nonlinear Phenomena, 14:327–347, 1985.

- [5] P. Astod, O. Junge. Computing the invariant measure and the Lyapunov exponent for one-dimensional maps using a measure-preserving polynomial basis. *Mathematics of Computation*, 83(288):1869–1902, 2014.
- [6] A. Barvinok. *A Course in Convexity*. American Mathematical Society, 2002.
- [7] M. Dellnitz, G. Froyland, O. Junge. The algorithms behind GAIO—Set oriented numerical methods for dynamical systems. In *Ergodic theory, analysis, and efficient simulation of dynamical systems*, pages 145–174. Springer, 2001.
- [8] M. Dellnitz, A. Hohmann, O. Junge, M. Rumpf. Exploring invariant sets and invariant measures. *CHAOS: An Interdisciplinary Journal of Nonlinear Science*, 7(2):221–228, 1997.
- [9] M. Dellnitz, S. Klus, A. Ziessler. A set-oriented numerical approach for dynamical systems with parameter uncertainty. *SIAM Journal on Applied Dynamical Systems*, 16(1):120–138, 2017.
- [10] M. Hénon. A two-dimensional mapping with a strange attractor. *Communications in Mathematical Physics*, 50(1):69–77, 1976.
- [11] D. Henrion. Semidefinite characterisation of invariant measures for one-dimensional discrete dynamical systems. *Kybernetika*, 48(6):1089–1099, 2012.
- [12] D. Henrion, M. Korda. Convex Computation of the Region of Attraction of Polynomial Control Systems. *IEEE Transactions on Automatic Control*, 59(2):297–312, 2014.
- [13] D. Henrion, J. B. Lasserre, C. Savorgnan. Approximate Volume and Integration for Basic Semialgebraic Sets. *SIAM Review*, 51(4):722–743, 2009.
- [14] D. Henrion, J.-B. Lasserre. Detecting Global Optimality and Extracting Solutions in GloptiPoly. In *Positive Polynomials in Control*, pages 293–310. Springer, 2005.
- [15] D. Henrion, J.-B. Lasserre, J. Löfberg. GloptiPoly 3: moments, optimization and semidefinite programming. *Optimization Methods and Software*, 24(4-5):761–779, 2009.
- [16] D. Henrion, J.-B. Lasserre, M. Mevissen. Mean squared error minimization for inverse moment problems. *Applied Mathematics & Optimization*, 70(1):83–110, 2014.
- [17] T. Kohda, K. Murao. Piecewise polynomial galerkin approximation to invariant densities of one-dimensional difference equations. *Electronics and Communications in Japan (Part I: Communications)*, 65(6):1–11, 1982.
- [18] M. Korda, D. Henrion, C. N. Jones. Convex computation of the maximum controlled invariant set for discrete-time polynomial control systems. In *Proceedings of the IEEE Conference on Decision and Control*, pages 7107–7112, 2013.

- [19] M. Korda, D. Henrion, I. Mezić. Convex computation of extremal invariant measures of nonlinear dynamical systems and Markov processes. Submitted for publication, 2018.
- [20] N. Kryloff, N. Bogoliouboff. La théorie générale de la mesure dans son application à l'étude des systèmes dynamiques de la mécanique non linéaire. *Annals of Mathematics*, 38(1):65–113, 1937.
- [21] A. Lasota, M. C. Mackey. *Chaos, Fractals, and Noise : Stochastic Aspects of Dynamics*. Springer, 1994.
- [22] J.-B. Lasserre. Global Optimization with Polynomials and the Problem of Moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.
- [23] J.-B. Lasserre. *Moments, Positive Polynomials and Their Applications*. Imperial College Press, 2010.
- [24] J.-B. Lasserre. Borel measures with a density on a compact semi-algebraic set. *Archiv der Mathematik*, 101(4):361–371, 2013.
- [25] J.-B. Lasserre. Lebesgue decomposition in action via semidefinite relaxations. *Advances in Computational Mathematics*, 42(5):1129–1148, 2016.
- [26] J.-B. Lasserre, T. Netzer. SOS approximations of nonnegative polynomials via simple high degree perturbations. *Math. Z.*, 256(1):99–112, 2007.
- [27] J. B. Lasserre, E. Pauwels. The empirical Christoffel function in Statistics and Machine Learning. [arXiv:1701.02886](https://arxiv.org/abs/1701.02886). 2017.
- [28] E. H. Lieb, M. Loss. *Analysis*. American Mathematical Society, 2001.
- [29] D. G. Luenberger. *Optimization by Vector Space Methods*. John Wiley & Sons, 1968.
- [30] J. Löfberg. Yalmip : A toolbox for modeling and optimization in MATLAB. In *Proceedings of the IEEE CACSD Symposium*, 2004.
- [31] V. Magron, P.-L. Garoche, D. Henrion, X. Thirioux. Semidefinite Approximations of Reachable Sets for Discrete-time Polynomial Systems. [arXiv:1703.05085](https://arxiv.org/abs/1703.05085). 2017.
- [32] H.L. Royden, P. Fitzpatrick. *Real Analysis*. Prentice Hall, 2010.
- [33] M. Trnovská. Strong duality conditions in semidefinite programming. *Journal of Electrical Engineering*, 56(12/s):1–5, 2005.
- [34] S. M. Ulam, J. von Neumann. On combinations of stochastic and deterministic processes. *Bull. Amer. Math. Soc.*, 53:1120, 1947.
- [35] B. Van der Pol. On relaxation oscillations. *The London, Edinburgh and Dublin Phil. Mag. & J. of Sci.*, 2:978–992, 1926.

- [36] H. Weyl. Über die Gibbs'sche Erscheinung und Verwandte Konvergenzphänomene. Rendiconti del Circolo Matematico di Palermo, 30(1):377–407, 1910.
- [37] K. Yosida. Functional analysis. Springer, 1980.
- [38] K. Yosida, E. Hewitt. Finitely Additive Measures. Transactions of the American Mathematical Society, 72(1):46–66, 1952.