Sparse moment-sum-of-squares relaxations for nonlinear dynamical systems with guaranteed convergence

Corbinian Schlosser, Milan Korda

Sparse moment-sum-of-squares relaxations for nonlinear dynamical systems with guaranteed convergence

December 12, 2020

Corbinian Schlosser[1], Milan Korda[1,2]

Abstract

This paper develops sparse moment-sum-of-squares approximations for three problems from nonlinear dynamical systems: region of attraction, maximum positively invariant set and global attractor. We prove general results allowing for a decomposition of these sets provided that the vector field and constraint set posses certain structure. We combine these decompositions with the methods from [10], [12] and [20] based on infinite-dimensional linear programming. For polynomial dynamics, we show that these problems admit a sparse sum-of-squares (SOS) approximation with guaranteed convergence such that the number of variables in the largest SOS multiplier is given by the dimension of the largest subsystem appearing in the decomposition. The dimension of such subsystems depends on the sparse structure of the vector field and the constraint set and can allow for a significant reduction of the size of the semidefinite program (SDP) relaxations, thereby allowing to address far larger problems without compromising convergence guarantees. The method is simple to use and based on convex optimization. Numerical examples demonstrate the approach.

Introduction

Many tasks concerning dynamical systems are of computationally complex nature and often not tractable in high dimension. Among these are the computations of the region of attraction (ROA), maximum positively invariant (MPI) set and global attractors (GA), which are the focus of this work. These sets are ubiquitous in the study of dynamical systems and have numerous applications. For example the ROA is the natural object to certify which

[1]CNRS; LAAS; 7 avenue du colonel Roche, F-31400 Toulouse; France. cschlosser@laas.fr, korda@laas.fr

[2]Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, CZ-16626 Prague, Czech Republic.

initial values will be steered to a desired configuration after a finite time T while the solution trajectory satisfies the state constraint at all times. The question of which initial values will stay in the constraint set for all positive times is answered by the MPI set. The GA describes which configurations will be reached by the solutions of the dynamical system asymptotically. This is of importance for controlled systems with a given feedback control where one might be interested if the given feedback control forces the solution to converge to a specific point or whether a more complex limiting behavior may occur. Since these objects are complex in nature computations of these are challenging tasks. Computational methods for the ROA have been pioneered by Zubov [25] in the 1960s and have a long history, summarized in [3]. A survey on the (controlled) MPI set and computational aspects can be found in [1]. Computations of the GA are typically approached via Lyapunov functions [7], via finite-time truncation or set oriented methods [5].

Given the curse of dimensionality problem present in computation of these sets, it is important to exploit structure in order to reduce the complexity. There are several concepts used for reducing the complexity, as for example symmetries (see, e.g.,[6]) or knowledge of Lyapunov or Hamilton functions (see, e.g., [22]). Here we investigate specific type of sparsity found in dynamical systems.

The central concept in this text is decoupling of the dynamical system into smaller subsystems, thereby allowing for computational time reduction, building on the work [2]. Even though our main goal is to exploit this decoupling computationally, we study the sparse structure at a rather general level, allowing for our results to be used within other computational frameworks and for other problems than those encountered in this work. The main novelty is the following: (i) We generalize the method of [2] to far more general graph structures. (ii) We use the proposed decoupling scheme within the moment sum-of-squares hierarchy framework, obtaining a sparse computational scheme with a guaranteed convergence from the outside to the sets of interest; to the best of our knowledge this is the first time sparsity is exploited in the moment-sos hierarchy for dynamical systems without compromising convergence. (iii) We treat different problems than [2], namely the computation of the ROA, MPI and GA rather than the reachable set computation problem.

We follow the approach from [10], [12] and [20] where outer approximations of the ROA, MPI set and GA are based on infinite dimensional linear programs on continuous functions approximated via the moment-sum-of-squares hierarchy (see [16] for a general introduction and [9] for recent applications).

Sparsity exploitation in static polynomial optimization goes back to the seminal work of [23], providing convergence results based on the so-called running intersection property. The situation in dynamical systems is more subtle and so far sparsity exploitation came at the cost of convergence such as in [21] where a different sparsity structure, not amenable to our techniques, was considered.

The computational benefit of decoupling into lower dimensional subsystems is quantified by the complexity of the resulting semidefininte programms (SDPs) in the moment-sum-of-squares hierarchy. As a measure of complexity of such an SDP, we use the number of variables involved in the largest sum-of-squares (SOS) multiplier. This number is determined by the dimension of the largest subsystem in the decoupling. To determine the subsystems we represent the interconnection between the dynamics of the states by the directed sparsity

graph of the dynamics f where the nodes are weighted by the dimension of the corresponding state space. We call a node x_j an ancestor of another node x_i if there exists a directed path from x_j to x_i in the (dimension weighted) sparsity graph of f. With this notation we can informally state our main result:

Theorem 1 (informal) *There exists a convergent hierarchy of sum-of-squares problems with the largest sum-of-squares multiplier containing ω variables, where ω is the largest weighted number of ancestors of one node in the dimension weighted sparsity graph of the dynamics.*

This allows for a potentially dramatic reduction in computation time when the dynamics is very spare in the sense considered in this work.

Not only the moment-sum-of-squares approach can benefit from decoupling into smaller subsystems but also other methods such as the set oriented methods [5] enjoy less computational complexity in lower dimensions. The decoupling procedure proposed here (Algorithm 1) also applies to any other method for approximating the ROA, MPI set or GA that satisfies certain convergence properties, as is the case for the set-oriented methods [5] for the GA.

1 Notations

The natural numbers are with zero included and denoted by $\mathbb{N}$. For a subset $J \subset \mathbb{N}$ we denote by $|J|$ its cardinality. The non-negative real numbers $[0, \infty)$ are denoted by $\mathbb{R}_+$. For two sets K_1, K_2 we denote their symmetric difference given by $K_1 \setminus K_2 \cup K_2 \setminus K_1$ by $K_1 \Delta K_2$. The function $\operatorname{dist}(\cdot, K)$ denotes the distance function to K and $\operatorname{dist}(K_1, K_2)$ denotes the Hausdorff distance of two subsets of $\mathbb{R}^n$ (with respect to a given metric or norm). The space of continuous functions on X is denoted by $\mathcal{C}(X)$ and the space of continuously differentiable functions on $\mathbb{R}^n$ by $\mathcal{C}^1(\mathbb{R}^n)$. The Lebesgue measure will always be denoted by λ. The ring of multivariate polynomials in variables $x = (x_1, \ldots, x_n)$ is denoted by $\mathbb{R}[x] = \mathbb{R}[x_1, \ldots, x_n]$ and for $k \in \mathbb{N}$ the ring of multivariate polynomials of total degree at most k is denoted by $\mathbb{R}[x]_k$. We will denote the open ball of radius r with respect to the euclidean metric by $B_r(0)$.

2 Setting and preliminary definitions

We consider a nonlinear dynamical system

$$\dot{x} = f(x) \tag{1}$$

with the state $x \in \mathbb{R}^n$ and a locally Lipschitz vector field $f : \mathbb{R}^n \to \mathbb{R}^n$. The following graph is a key tool in exploiting sparsity of f.

Definition 1 (Dimension weighted sparsity graph) *Let the variable $x \in \mathbb{R}^n$ and the function $f : \mathbb{R}^n \to \mathbb{R}^n$ be partitioned (after a possible permutation of indices) as $x = (x_1, \ldots, x_N)$ and $f = (f_1, \ldots, f_N)$ with $x_i \in \mathbb{R}^{n_i}$, $f_i : \mathbb{R}^{n_i} \to \mathbb{R}^{n_i}$ and $\sum_{i=1}^N n_i = n$. The dimension weighted sparsity graph associated to f induced by this partition is defined by:*

1. *The set of nodes is* $(x_1, \ldots, x_N)$.

2. (x_i, x_j) *is an edge if the function* f_j *depends on* x_i .

3. *The weight of a node* x_i *is equal to* n_i.

.

Remark 1 *Without putting weights on nodes we call the graph just sparsity graph of* f *(induced by the partitioning). The (dimension weighted) sparsity graph is not unique as it depends on the partition of* x *and* f*. Choosing a good partition is key to maximizing the computational savings obtained from the sparse SDP relaxations developed in this work in section 7.*

Remark 2 *For a dynamical system a sparsity graph describes the dependence of the dynamics of a state on other states. More precisely, there exists a directed path from* i *to* j *in the sparsity graph of* f *if and only if the dynamics of* x_j *depend (indirectly via other states) on the state* x_i.

As an example consider the following function $f : \mathbb{R}^{10} \to \mathbb{R}^{10}$

$$f(y_1, \ldots, y_{10}) = (y_1^2 y_2, y_1 y_2, y_3 y_2 + y_3^2, y_7 - y_4^4, y_7 y_5^2, y_2 y_6, y_2^3 y_6 y_7, y_3^2 y_6 y_8^2, y_6 y_9^5, y_7^2).$$

The grouping $x_1 = (y_1, y_2)$, $x_2 = y_3$, $x_3 = (y_4, y_5)$, $x_4 = (y_6, y_7)$ and $x_5 = (y_8, y_9, y_{10})$ induces the functions $f_1(y_1, y_2) = (y_1^2 y_2, y_1 y_2)$, $f_2(y_3) = (y_3 y_2 + y_3^2)$, $f_3(y_4, y_5) = (y_7 - y_4^4, y_7 y_5^2)$, $f_4(y_6, y_7) = (y_2 y_6, y_2^3 y_6 y_7)$ and $f_5(y_8, y_9, y_{10}) = (y_3^2 y_6 y_8^2, y_6 y_9^5, y_7^2)$ according to Definition 1. Figure 1 shows its dimension weighted sparsity graph.

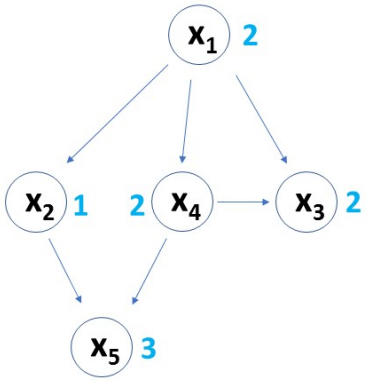

Figure 1: The dimension weighted sparsity graph of the function (2) with respect to the partitioning $x_1 = (y_1, y_2)$, $x_2 = y_3$, $x_3 = (y_4, y_5)$, $x_4 = (y_6, y_7, y_8, y_9)$ and $x_5 = (y_{10}, y_{11}, y_{12})$

Definition 2 (Predecessor, leaf, Past) 1. *For a sparsity graph we call a node* x_i *a predecessor of node* x_j *if either* $x_i = x_j$ *or if there is a directed path from* x_i *to* x_j.

2. *A node* x_i *is called a leaf if it does not have a successor (i.e., all nodes connected to* x_i *are its predecessors).*

3. *The set of all predecessors of x_i is called the past of x_i and denoted by $\mathbf{P}(x_i)$.*

4. *The largest dimension weighted past ω in a directed graph with weights n_i and nodes x_i is given by*

$$\omega := \max_i \sum_{x_j \in \mathbf{P}(x_i)} n_j. \tag{2}$$

For the graph from figure 1, the node x_5 has the largest weighted path. Its past is colored in blue in Figure 2.

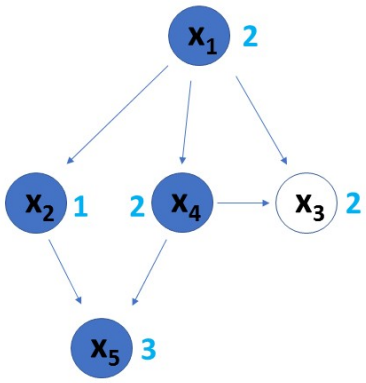

Figure 2: Past of x_5 in the sparsity graph from Figure 1

In Remark 2 we have seen that the past of a node x_i determines all the nodes the dynamics of x_i (indirectly) depend on. Therefore the following definition is closely related to the notion of the past of a node.

Definition 3 *For a dynamical system $\dot{x} = f(x)$ on $\mathbb{R}^n$ we call a set of states $(x_i)_{i\in I}$ for some index set $I \subset \{1, \dots, n\}$ a subsystem of $\dot{x} = f(x)$ if we have*

$$f_I \circ P_I = P_I \circ f \tag{3}$$

where $f_I := (f_i)_{i\in I}$ denotes the components of f according to the index set I and P_I denotes the canonical projection onto the states x_I, i.e. $P_I(x) := x_I$.

Since f_I formally depends on $x \in \mathbb{R}^n$ we mean by the term $f_I \circ P_I$ that f_I only depends on the variables x_I.

If φ_t denotes the flow of the dynamical system and φ_t^I the flow of the subsystem, condition (3) translates to

$$\varphi_t^I \circ P_I = P_I \circ \varphi_t^I. \tag{4}$$

For a given node x_i the past $\mathbf{P}(x_i)$ of this node determines the states of the smallest subsystem of the dynamical system containing x_i, and we refer to this subsystem by the subsystem induced by $\mathbf{P}(x_i)$. In acyclic sparsity graphs the nodes with maximal past are leafs, i.e. nodes with no successor, because a successor has a larger past than its predecessor.

2.1 Main result summary

With these notations, the main result, informally written, reads

Theorem 2 (informal) *There exists a hierarchy of SDPs with all sum-of-squares multipliers depending on at most ω variables such that their solutions provide outer approximations of the ROA, MPI set and GA with guaranteed convergence as the degree of the sum-of-squares multipliers tends to infinity.*

In order to prove this theorem we first consider a prototype sparse dynamics which allows a decoupling of the dynamical system. For this prototype system we get decompositions of the ROA, MPI set and GA, inspired by [2]. Those decompositions allow us to decouple the computations of those sets as well. We apply this to the linear programs for those sets from [10], [12] and [20], which describe the desired sets up to a discrepancy of Lebesgue measure zero.

Dynamics that are not of that prototype type are approached by a generalization of the decoupling derived from the prototype setting.

3 Sparse dynamics: the prototype setting

The dynamical system we are considering first is as in [2] and has the following form

$$\begin{aligned} \dot{x}_1 &= f_1(x_1) \\ \dot{x}_2 &= f_2(x_1, x_2) \\ \dot{x}_3 &= f_3(x_1, x_3) \end{aligned} \tag{5}$$

on the state space $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{n_3}$ and we consider a constraint set $X \subset \mathbb{R}^{n_1+n_2+n_3}$ and locally Lipschitz continuous functions $f_1 : X \to \mathbb{R}^{n_1}$, $f_2 : X \to \mathbb{R}^{n_2}$ and $f_3 : X \to \mathbb{R}^{n_3}$ where f_1 only depends on x_1, i.e. is constant in (x_2, x_3), f_2 only depends on (x_1, x_2), i.e. is constant in x_3 and f_3 only depends on (x_1, x_3), i.e. is constant in x_2. The sparsity graph of the system (5) has the "cherry" structure depicted in Figure 3. This indicates that the

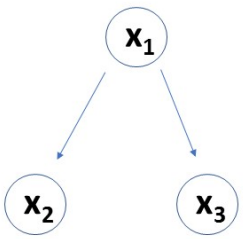

Figure 3: 2-cherry structure.

system splits into the decoupled dynamics

$$\frac{\mathrm{d}}{\mathrm{dt}}(x_1, x_2) = (f_1, f_2)(x_1, x_2) \tag{6}$$

with corresponding flow $\varphi_t^{(2)}$ and

$$\frac{\mathrm{d}}{\mathrm{dt}}(x_1, x_3) = (f_1, f_3)(x_1, x_3) \tag{7}$$

with corresponding flow $\varphi_t^{(3)}$ and

$$\dot{x}_1 = f_1(x_1) \tag{8}$$

with corresponding flow $\varphi_t^{(1)}$. Let $P_{1,i}$ denote the canonical projection onto (x_1, x_i) and P_1 the canonical projection onto the x_1 component. Then the subsystem relations (6), (7) and (8) read

$$(f_1, f_2) \circ P_{1,2} = P_{1,2} \circ f \quad , \quad (f_1, f_3) \circ P_{1,3} = P_{1,3} \circ f \quad , \quad f_1 \circ P_1 = P_1 \circ f$$

and we have for the corresponding flows

$$\begin{aligned} \varphi_t^{(2)} \circ P_{1,2} &= P_{1,2} \circ \varphi_t \\ \varphi_t^{(3)} \circ P_{1,3} &= P_{1,3} \circ \varphi_t \\ \varphi_t^{(1)} \circ P_1 &= P_1 \circ \varphi_t \end{aligned} \tag{9}$$

for all $t \in \mathbb{R}_+$. Note that the x_1-component of the flows $\varphi^{(2)}$ and $\varphi^{(3)}$ are given by $\varphi^{(1)}$ due to the decoupled dynamics of x_1.

The equations in (9) state that the subsystems behave like factor systems, i.e. the projections map solutions of the whole system to solutions of the subsystems.

The state constraints need to be taken into account more carefully. For instance the constraint set for (6) for a fixed x_3 is given by

$$X_{1,2}(x_3) := \{(x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} : (x_1, x_2, x_3) \in X\}. \tag{10}$$

In a similar way we define

$$X_{1,3}(x_2) := \{(x_1, x_3) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_3} : (x_1, x_2, x_3) \in X\}. \tag{11}$$

and

$$X_1(x_2, x_3) := \{x_1 \in \mathbb{R}^{n_1} : (x_1, x_2, x_3) \in X\}. \tag{12}$$

In order to get that the subsystems (6), (7) and (8) are completely decoupled, we need a splitting also in the constraint sets, i.e. the sets $X_{1,2}(x_3), X_{1,3}(x_2)$ and $X_1(x_2, x_3)$ do not depend on x_2 and x_3.

Proposition 1 *For variable $(x_1, x_2, x_3) \in X$ the sets $X_{1,2}(x_3), X_{1,3}(x_2)$ and $X_1(x_2, x_3)$ are independent of (x_2, x_3) if and only if X is of the form*

$$X = X_1 \times X_2 \times X_3 \tag{13}$$

for some $X_1 \subset \mathbb{R}^{n_1}$, $X_2 \subset \mathbb{R}^{n_2}$ and $X_3 \subset \mathbb{R}^{n_3}$. In particular if (13) holds then (9) holds with corresponding constraint sets $X_1 \times X_2$, $X_1 \times X_3$ and X_1.

Proof: If X is of the form (13) then we have for arbitrary $x_3 \in X_3$

$$\begin{aligned} X_{1,2}(x_3) &= \{(x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} : (x_1, x_2, x_3) \in X\} \\ &= \{(x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} : x_1 \in X_1, x_2 \in X_2\} \end{aligned}$$

and we see that this is independent of x_3. The same argument works also for the sets $X_{1,3}(x_2)$ and $X_1(x_2, x_3)$. On the other hand let all the sets $X_{1,2}(x_3), X_{1,3}(x_2)$ and $X_1(x_2, x_3)$ be independent of (x_2, x_3). Let us denote those sets by $X_{1,2}, X_{1,3}$ and X_1 and let P_2, P_3 and $P_{2,3}$ be the canonical projections onto the x_2, x_3 and (x_2, x_3) component respectively. We have by

$$\begin{aligned} X &= \bigcup_{x_2 \in P_2(X)} \{(x_1, x_2, x_3) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{n_3} : (x_1, x_3) \in X_{1,3}\} \\ &= \bigcup_{x_3 \in P_3(X)} X_{1,2}(x_3) \times \{x_3\} = X_{1,2} \times P_3(X). \end{aligned} \tag{14}$$

In a similar way we see that

$$X_1 \times P_{2,3}(X) = X = \{(x_1, x_2, x_3) : x_2 \in P_2(X), (x_1, x_3) \in X_{1,3}\} \tag{15}$$

We claim

$$X = P_1(X) \times P_2(X) \times P_3(X).$$

To check this it suffices to check $X \supset P_1(X) \times P_2(X) \times P_3(X)$. Therefore let $x_1 \in P_1(X), x_2 \in P_2(X)$ and $x_3 \in P_3(X)$. By (15) there exists a pair (x_2', x_3') such that $(x_1, x_2', x_3') \in X$. From $(x_1, x_2', x_3') \in X$ it follows $(x_1, x_2') \in X_{1,2}$. Hence by (14) $(x_1, x_2', x_3) \in X$. It follows $(x_1, x_3) \in X_{1,3}$ and so $(x_1, x_2, x_3) \in X$ by (15). □

The last proposition states that we can only completely decouple systems if the constraint set X decomposes as a product. The reason is that otherwise the constraint sets of the subsystems varies with changing states x_2, x_3 and x_1. We give an example that illustrates this issue on the maximum positively invariant set defined in Definition 5. Consider the following system

$$\dot{x}_1 = 0 \ , \ \dot{x}_2 = -x_1 x_2 \ , \ \dot{x}_3 = x_1 x_3 (1 - x_3) \tag{16}$$

on $\mathbb{R}^3$ with constraint set $X = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1, x_3 \in [\frac{1}{2}, 1], x_2 \in [0, 1], x_2 \geq x_3 - \frac{1}{2}\}$. Here X does not factor into a product because the x_2 component in X depends on the state x_3. Because $x_2(t)$ converges to 0 as $t \to \infty$ and $x_3(t)$ converges to 1 as $t \to \infty$ for any initial value coming from X it follows that eventually any trajectory starting in X leaves the constraint set X. But for fixed $x_3 = \frac{1}{2}$ we have $X_{1,2}(x_3) = [\frac{1}{2}, 1] \times [0, 1]$ and any solution for the subsystem induced by (x_1, x_2) starting in $X_{1,2}(\frac{1}{2})$ stays in this set for all times $t \in \mathbb{R}_+$. This different behaviour is due the varying of $x_3(t)$ and hence the constraint set for (x_1, x_2), namely $X_{1,2}(x_3(t))$, is changing in time, which in this case causes that any trajectory with initial value in X to leave X eventually. This is why we will have the following assumption for the rest of this text.

Assumption 1 *The constraint set $X \subset \mathbb{R}^n$ is given by $X = X_1 \times \cdots \times X_k$ where each $X_i \subset \mathbb{R}^{n_i}$ is compact.*

If needed it is possible to perform a permutation of coordinates of $\mathbb{R}^n$ beforehand in order to allow a (better) factorization.

In the next chapter we will see that the sparse prototype dynamics lead to a sparse description of the reachable set, the ROA, MPI set and GA.

4 Sparse representations for the reachable set, the RoA, MPI set and GA

As long as we do not explicitly refer to the sparse dynamics from (5) by a dynamical system we will refer to solutions of the ODE (1) with a compact constraint set $X \subset \mathbb{R}^n$.

Definition 4 (Region of attraction) *For a dynamical system, a finite time $T \in \mathbb{R}_+$ and a target set $X_\mathrm{T} \subset X$ the region of attraction (ROA) of X_T is defined as*

$$R_T := \{x_0 \in X : \exists x(\cdot) \text{ s.t. } \dot{x}(t) = f(x(t)),\ x(0) = x_0,\ x(t) \in X \text{ on } [0,T],\ x(T) \in X_\mathrm{T}\}\,. \tag{17}$$

Remark 3 *The reachable set from an initial set $X_\mathrm{I} \subset X$ in time T*

$$S_T := \{x \in X : \dot{x}(t) = f(x(t)), x(t) \in X \text{ on } [0,T],\ x(T) = x, x(0) \in X_\mathrm{I}\} \tag{18}$$

can be obtained by time reversal, i.e. by $S_T = R_T$ for $X_\mathrm{T} := X_\mathrm{I}$ and the dynamics given by $\dot{x} = -f(x)$.

Definition 5 (Maximum positively invariant set) *For a dynamical system the maximum positively invariant (MPI) set is the set of initial conditions x_0 such that the solutions $\varphi_t(x_0)$ stay in X for all $t \in \mathbb{R}_+$.*

The MPI set will be denoted by M_+ in the following.

Definition 6 (Global attractor) *A compact set $\mathcal{A} \subset X$ is called the global attractor (GA) if it is minimal uniformly attracting, i.e., it is the smallest compact set $\mathcal{A}$ such that*

$$\lim_{t\to\infty} \operatorname{dist}(\varphi_t(M_+), \mathcal{A}) = 0.$$

Remark 4 *An important property of the global attractor is that it is characterized by being invariant, i.e. $\varphi_t(\mathcal{A}) = \mathcal{A}$ for all $t \in \mathbb{R}_+$, and attractive see [19].*

In the following we will see that the sparse structure of (5) is inherited to the ROA, MPI set and GA. We will start with the ROA and see that it decomposes into the ROA sets for the subsystems glued together along the (decoupled) x_1 component.

Proposition 2 *Let the dynamical system be of the form (5). Suppose that Assumption 1 holds, i.e.* $X = X_1 \times X_2 \times X_3$*, and assume* $X_{\mathrm{T}} = X_{1,\mathrm{T}} \times X_{2,\mathrm{T}} \times X_{3,\mathrm{T}}$*. Let* $R_T^{(2)}$*,* $R_T^{(3)}$ *and* $R_T^{(1)}$ *be the regions of attraction for the subsystems (6) with target set* $X_{1,\mathrm{T}} \times X_{2,\mathrm{T}}$*, (7) with target set* $X_{1,\mathrm{T}} \times X_{3,\mathrm{T}}$ *and (8) with target set* $X_{1,\mathrm{T}}$*. Then* R_T *is given by*

$$\{(x_1, x_2, x_3) \in X : (x_1, x_2) \in R_T^{(2)}, (x_1, x_3) \in R_T^{(3)}\}. \tag{19}$$

Further $P_1 R_T \subset R_T^{(1)}$*, in general they are not equal.*

Proof: Let R denote the set from (19). Let $(x_1, x_2, x_3) \in R$ with corresponding solutions $(\hat{x}_1(\cdot), x_2(\cdot))$ and $(\tilde{x}_1(\cdot), x_3(\cdot)), (x_1(\cdot))$ as in the definition of $R_T^{(2)}$, $R_T^{(3)}$ and $R_T^{(1)}$. We have by uniqueness of solutions of (8) that $x_1(\cdot) = \hat{x}_1(\cdot) = \tilde{x}_1(\cdot)$. Hence we see that $(x_1(\cdot), x_2(\cdot), x_3(\cdot))$ is a solution of the whole system with initial value $(x_1(0), x_2(0), x_3(0)) = (x_1, x_2, x_3)$, $(x_1(t), x_2(t), x_3(t)) \in X$ for all $t \in [0, T]$ and $(x_1(T), x_2(T), x_3(T)) \in X_{1,\mathrm{T}} \times X_{2,\mathrm{T}} \times X_{3,\mathrm{T}} = X_{\mathrm{T}}$. Hence $(x_1, x_2, x_3) \in R_T$. On the other hand let $(x_1, x_2, x_3) \in R_T$ be initial values and the corresponding solution of the differential equation $(x_1(\cdot), x_2(\cdot), x_3(\cdot))$. Since X decomposes into a product by Assumption (1) we have $(x_1(0), x_2(0)) = (x_1, x_2)$, $(x_1(t), x_2(t)) \in X_1 \times X_2$, $(x_1, x_3), (x_1(t), x_3(t)) \in X_1 \times X_3$ for all $t \in [0, T]$ and $(x_1(T), x_2(T)) \in X_{1,\mathrm{T}} \times X_{2,\mathrm{T}}$, $(x_1(T), x_3(T)) \in X_{1,\mathrm{T}} \times X_{3,\mathrm{T}}$. All in all we have $(x_1, x_2) \in R_T^{(2)}$ and $(x_1, x_3) \in R_T^{(3)}$, i.e. $(x_1, x_2, x_3) \in R$. The inclusion $P_1 R_T \subset R_T^{(1)}$ follows again directly from (19) and (9). For an example where those sets do not coincide take $\dot{x}_2 = 1, \dot{x}_3 = 1, \dot{x}_1 = 0$ (without control) and $X = X_{\mathrm{T}} = [0, 1]^3$ and $T > 0$. Then $R_T = \emptyset$ and $R_T^{(1)} = [0, 1] = X_1$. □

We will refer to connecting two sets along a common component as in (20) by saying we glue the sets together along the x_1 component.

And by very similar arguments we get a similar result for the MPI set and the GA.

Proposition 3 *Let* $X = X_1 \times X_2 \times X_3$ *be as in (13) for the sparse dynamical system (5) and let* $M_+^{(2)}$ *and* $M_+^{(3)}$ *denote the MPI sets for the subsystems (6) and (7) then the MPI set* M_+ *of (5) is given by*

$$\{(x_1, x_2, x_3) \in X : (x_1, x_2) \in M_+^{(2)}, (x_1, x_3) \in M_+^{(3)}\}. \tag{20}$$

Further $P_1 M_+ \subset M_+^{(1)}$ *for the MPI set* $M_+^{(1)}$ *for (8), in general they are not equal.*

Proof: Let M denote the set from (20). Let $(x_1, x_2, x_3) \in M$ and $t \in \mathbb{R}_+$. We have $y_2 := \varphi_t^{(2)}(x_1, x_2) \in X_1 \times X_2$ and $y_3 := \varphi_t^{(3)}(x_1, x_3) \in X_1 \times X_3$. Further by (9) the x_1 component of y_2 and y_3 coincide. Hence it follows from the second statement of Proposition 1 that $\varphi_t(x_1, x_2, x_3) \in X_1 \times X_2 \times X_3 = X$. That means M is invariant and hence M is contained in the MPI set. On the other hand let (x_1, x_2, x_3) be in the MPI set. Again by Proposition 1 we have for all $t \in \mathbb{R}_+$ that $\varphi_t^{(2)}(x_1, x_2) \in P_{1,2}(X) = X_1 \times X_2$ and $\varphi_t^{(3)}(x_1, x_3) \in P_{1,3}(X) = X_1 \times X_3$. Hence $(x_1, x_2) \in M_+^{(2)}$ and $(x_1, x_3) \in M_+^{(3)}$, i.e. $(x_1, x_2, x_3) \in M$. Since $\varphi_t(x_1, x_2, x_3) \in X$ implies $\varphi_t^{(1)}(x_1) \in X_1$ the inclusion $P_1 M_+ \subset P_1 M_+^{(1)}$ follows immediately. An example for which $M_+ = \emptyset$ while $M_+^{(1)} \neq \emptyset$ is again given by $\dot{x}_2 = 1, \dot{x}_2 = 1, \dot{x}_1 = 0$ on $[0, 1]^3$. Here $M_+, M_+^{(2)}$ and $M_+^{(3)}$ are all empty while $M_+^{(1)} = [0, 1] = X_1$. □

And a similar result for GAs is stated next.

Proposition 4 *Let $X = X_1 \times X_2 \times X_3$ be as in (13) and $\mathcal{A}^{(2)}$ be the GA for (6) and $\mathcal{A}^{(3)}$ the GA for (7). Then gluing them together along x_1 gives the GA for the whole system, i.e.*

$$\mathcal{A} = \{(x_1, x_2, x_3) \in X : (x_1, x_2) \in \mathcal{A}^{(2)}, (x_1, x_3) \in \mathcal{A}^{(3)}\}. \tag{21}$$

Further $P_1\mathcal{A} \subset \mathcal{A}^{(1)}$ for the GA $\mathcal{A}^{(1)}$ for (8), in general they are not equal..

Proof: Let the set from (21) be denoted by A. Since A is given by

$$X \cap P_{1,2}^{-1}(\mathcal{A}^{(2)}) \cap P_{1,3}^{-1}(\mathcal{A}^{(3)}) \tag{22}$$

we see that it is a closed subset of X, and hence compact. By [19] Definition 10.4. and Theorem 10.6. it suffices to show that $A = M_+ \cap M_-$ where M_- denotes the maximal negatively invariant set, or in other words, the maximum positively invariant set for the dynamics given by $\dot{x} = -f(x)$. Hence we can apply Proposition 3 to get

$$\begin{aligned}
\mathcal{A} &= M_+ \cap M_- \\
&= \{(x_1, x_2, x_3) \in X : (x_1, x_2) \in M_+^{(2)}, (x_1, x_3) \in M_+^{(3)}\} \cap \\
&\quad \{(x_1, x_2, x_3) \in X : (x_1, x_2) \in M_-^{(2)}, (x_1, x_3) \in M_-^{(3)}\} \\
&= \{(x_1, x_2, x_3) \in X : (x_1, x_2) \in M_+^{(2)}, (x_1, x_2) \in M_-^{(2)}, (x_1, x_3) \in M_+^{(3)}, (x_1, x_3) \in M_-^{(3)}\} \\
&= \{(x_1, x_2, x_3) \in X : (x_1, x_2) \in \mathcal{A}^{(2)}, (x_1, x_3) \in \mathcal{A}^{(3)}\}
\end{aligned}$$

where also the last equation follows from $\mathcal{A}^{(i)} = M_+^{(i)} \cap M_-^{(i)}$ ([19] Definition 10.4. and Theorem 10.6.). The inclusion $P_1\mathcal{A} \subset \mathcal{A}^{(1)}$ follows from the corresponding inclusion for the MPI set in Proposition 3. An example for systems for which $P_1\mathcal{A} \subsetneq \mathcal{A}^{(1)}$ are systems with empty MPI sets but non-empty $\mathcal{A}^{(1)}$. □

5 A decoupling procedure for approximating the ROA, MPI set and GA

The approach in this section is based on the structure of the ROA, MPI set and GA for sparse settings from Propositions 2, 3 and 4 and a (non-specified) method for computing or approximating those sets. The idea is simple, we use a method for computing or approximating those sets for the subsystems and Propositions 2, 3 and 4 allow to glue the resulting sets together to obtain (an approximation of) the desired sets.

Remark 5 (Decoupling procedure) *Given a sparse dynamical system and a method for approximating/computing the ROA, MPI set or GA for an arbitrary dynamical system, we can compute/approximate the desired set for the subsystems (6) and (7) with corresponding state constraints $X_1 \times X_2$ and $X_1 \times X_3$ (and repsectively $X_{1,T} \times X_{2,T}$, $X_{1,T} \times X_{3,T}$ for the ROA) due to the given method and glue the resulting sets together along their common component x_1 as in Propositions 2, 3 and 4.*

It is clear that, if a method is able to exactly compute the ROA, MPI set or GA for an arbitrary dynamical system, the gluing procedure from Remark 5 using this method as well produces the corresponding set exactly due to the decomposition of those sets by Propositions 2, 3 and 4.
Typically exact representations of those sets are not computable and hence we have to work with approximations. Depending on how the approximation is measured or what the chosen methods produces the decoupling procedure from Remark 5 might fail. For example it might not always be guaranteed that the set

$$\{(x_1, x_2, x_3) \in X : (x_1, x_2) \in K_1,\ (x_1, x_3) \in K_2\} \tag{23}$$

is non-empty where K_1 and K_2 denote approximations of the desired sets for the subsystems constructed by the chosen method . But if the methods produce outer approximations, i.e. $K_i \supset R_{\mathrm{T}}$ or M_+ or $\mathcal{A}$ respectively, then of course (23) gives a superset of the desired set and hence can only be empty if the desired set is empty. We give two typical examples of a metric for the quality of the (outer) approximation, one is the Hausdorff distance, i.e. for $S_2 \subset S_1 \subset \mathbb{R}^n$

$$d_{H,\|\cdot\|_\infty}(S_1, S_2) := \sup\{\mathrm{dist}(s_1, S_2) : s_1 \in S_1\} \tag{24}$$

where the distance is induced by a norm on $\mathbb{R}^n$; the specific choice of norm will not be important in the following due to equivalence of norms on $\mathbb{R}^n$. The other metric we consider is the Lebesgue measure of the symmetric difference of two sets $S_1, S_2 \subset \mathbb{R}^n$

$$d_\lambda(S_1, S_2) := \lambda(S_1 \Delta S_2) \tag{25}$$

where Δ denotes the symmetric difference.

In the following proposition, the term "desired set" refers to the ROA or the MPI set or the GA.

Proposition 5 *Let $S_k^1 \subset X_1 \times X_2 \subset \mathbb{R}^{n_1+n_2}$ and $S_k^2 \subset X_1 \times X_3 \subset \mathbb{R}^{n_1+n_3}$ and let*

$$S_k^{1,2} := \{(x_1, x_2, x_3) \mid (x_1, x_2) \in S_k^1, (x_1, x_3) \in S_k^2\}.$$

For $i = 1, 2$ let S^i denote the desired set for the subsystems (6) and (7) and S denotes the desired set for (5) . Then the following holds:

1. *Hausdorff distance: If $S_k^i \to S^i$ with respect to the Hausdorff distance, i.e. for $i = 1, 2$*

$$d_H(S_k^i, S^i) \to 0\ , \quad \text{as } k \to \infty. \tag{26}$$

Then

$$d_H(S_k^{1,2}, S) \to 0\ , \quad \text{as } k \to \infty. \tag{27}$$

2. *Lebesgue measure: We have*

$$d_\lambda(S, S_k^{1,2}) \leq \lambda(S^1 \Delta S_k^1)\lambda(X_3) + \lambda(S^2 \Delta S_k^2)\lambda(X_2). \tag{28}$$

In particular if S_k^i converges to S^i with respect to d_λ for $i = 1, 2$ then $S_k^{1,2}$ converges to S with respect to d_λ.

Proof: Assume (27) does not hold. Then there exists a $\varepsilon > 0$ and an unbounded subsequence $(k_m)_{m\in\mathbb{N}}$ such that

$$d_H(S^{1,2}_{K_m}, S) > \varepsilon. \tag{29}$$

Since by assumption S^1_k and S^2_k are outer approximations we get by Propositions 2, 3 and 4 that $S^{1,2}_k \supset S$ for all $k \in \mathbb{N}$ and (29) gives a sequence of points $x_{k_m} = (x^1_{k_m}, x^2_{k_m}, x^3_{k_m}) \in S^{1,2}_{k_m}$ such that $\operatorname{dist}(x_{k_m}, S) > \varepsilon$. Since $S^{1,2}_k \subset X$ and by compactness of X there exists $x = (x^1, x^2, x^3) \in X$ and a subsequence of $(k_m)_{m\in\mathbb{N}}$ which we will still denoted by $(k_m)_{m\in\mathbb{N}}$ such that $x_{k_m} \to x$ as $m \to \infty$. By assumption (26) there exist $y_{k_m} = (y^1_{k_m}, y^2_{k_m}) \in S^1$ and $z_{k_m} = (z^1_{k_m}, z^3_{k_m}) \in S^2$ with $\|y_{k_m} - (x^1_{k_m}, x^2_{k_m})\|, \|z_{k_m} - (x^1_{k_m}, x^3_{k_m})\| \to 0$ as $m \to \infty$. Hence also $y^1_{k_m} \to (x^1, x^2)$, $z_{k_m} \to (x^1, x^3)$ as $m \to \infty$. Because S^1, S^2 are closed it follows from Propositions 2, 3 or 4 respectively that $(x^1, x^2, x^3) \in S$. In particular we get

$$\varepsilon < \operatorname{dist}(x_{k_m}, S) \leq \|x_{k_m} - (x^1, x^2, x^3)\| \to 0$$

as $m \to \infty$, contradiction. For the second statement from $S^i_k \subset X_1 \times X_{i+1}$ for $i = 1, 2$ it follows $S^{1,2}_k \subset X$. From the decoupling of S by Propositions 2, 3 and 4 and the definition of $S^{1,2}_k$ we get

$$\begin{aligned} S\Delta S^{1,2}_k \quad \subset \quad & \{(x_1, x_2, x_3) \in X : (x_1, x_2) \in S^1 \Delta S^1_k\} \cup \\ & \{(x_1, x_2, x_3) \in X : (x_1, x_3) \in S^2 \Delta S^2_k\}. \end{aligned}$$

Applying the Lebesgue measure to this inclusion gives

$$\begin{aligned} \lambda(S\Delta S^{1,2}_k) \quad \leq \quad & \lambda\left(\{(x_1, x_2, x_3) \in X : (x_1, x_2) \in S^1 \Delta S^1_k\}\right) + \\ & \lambda\left(\{(x_1, x_2, x_3) \in X : (x_1, x_3) \in S^2 \Delta S^2_k\}\right) \\ = \quad & \lambda(S^1 \Delta S^1_k)\lambda(X_3) + \lambda(S^2 \Delta S^2_k)\lambda(X_2). \end{aligned}$$

□

Proposition 5 says that one can compute a converging sequence of approximations to the desired set by computing converging approximations for each subsystem separately, thereby reducing the computational effort. Specific computational methods to do so based on semidefinite programming are discussed in Section 7.

In the next chapter we will generalize the decoupling approach based on the sparsity graph to general dynamics induced by a function f.

6 More general graph structures

By the same arguments we used for the simple cherry structure of the prototype setting (5), gluing together along x_1 works also for more subsystems, i.e. for dynamics of the form $\dot{x}_i = f_i(x_1, x_i)$ for $i = 2, \ldots, r$ with $\dot{x}_1 = f_1(x_1)$. Induction on the branching allows also more tree-like structures.

But instead of following this specific approach, in this section we are lead by the observation that Propositions 2, 4 and 3 can be rephrased as

$$S = \{x \in X : P_i(x) \in S_i \text{ for } i \in I\} \tag{30}$$

where S denotes the desired set – i.e. the ROA, MPI set or GA – and some index set I where for $i \in I$ the set S_i denotes the desired set for a (maximal) subsystem and $P_i(x)$ denotes the projection on $\mathbb{R}^{n_i}$ for the corresponding subsystem.

We will see that such a result is true for general dynamical systems. To be able to state the result from Theorem 3 in a more convenient way we assume that the sparsity graph is acyclic. It follows that the subsystems we need to consider are induced by leafs (Lemma 1), i.e. the subsystem's nodes are given by the pasts of the corresponding leafs. We can always achieve acyclic sparsity graph by choosing a suitable partition. For example, it suffices to choose the partition in such a way that for each circle all its nodes are assigned to one element of the partition. This is illustrated in Figure 4.

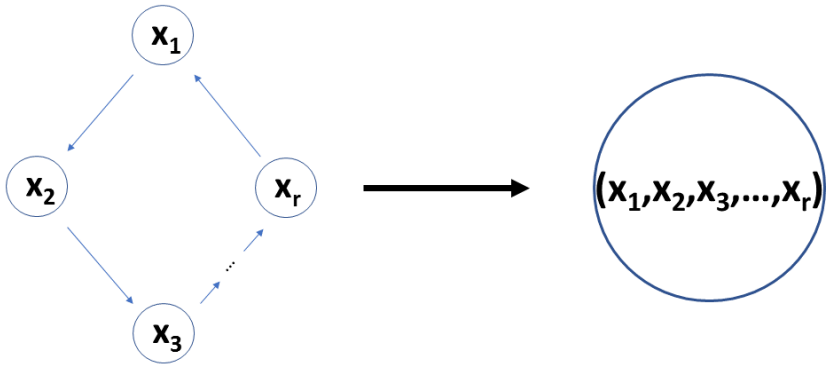

Figure 4: A circle reduces to one node.

To be more precise we define the reduction of a circle to one node formally in the following remark.

Remark 6 (Circle reduction) *Let $J_1, \dots, J_N$ be a partition of $\{1, \dots, n\}$ with corresponding states $\mathbf{x}_1, \dots, \mathbf{x}_N$. Let $\mathbf{x}_{i_1}, \dots, \mathbf{x}_{i_l}$ form a cycle in the sparsity graph of f with respect to the partition $J_1, \dots, J_N$. Then grouping $\mathbf{x}_{i_1}, \dots, \mathbf{x}_{i_l}$ together means considering the new partition consisting of $\tilde{J} := \bigcup_{r=1}^{l} J_{i_r}$ and J_i for $i \in \{1, \dots, N\} \setminus \{i_1, \dots, i_l\}$.*

A circle can be detected for example by the depth first traversal algorithm with complexity $\mathcal{O}(|V| + |E|)$ (see for example [4] p. 543) where V denotes the set of nodes and E the set of edges. Each detected circle reduces the number of nodes and edges so that the complexity of finding all circles is less than $\mathcal{O}(|V|(|V| + |E|))$.

Reducing a circle to one node does not affect our approach. This is because all nodes in the circle necessarily occur always together in a subsystem containing any of the nodes from the circle. Hence the subsystems obtained from a sparsity graph and the same sparsity graph where circles have been reduced to single nodes coincide.

Similar arguments reveal that a system cannot be reduced if for example its corresponding graph is a path in which each branching is contained in a circle; see Figure 5. For straight

lines this is even more obvious: The nodes are connected only by one incoming and one outgoing edge. Exploiting such sparse structures for the ROA is investigated by [21].

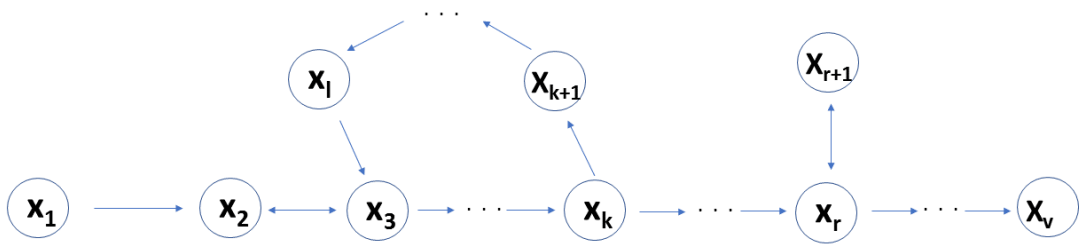

Figure 5: Straight paths with circles.

The following lemma contains the basic properties of leafs and their pasts that we need for the proof of the main theorem.

Lemma 1 *Any directed graph without circles has at least one leaf. Furthermore, for directed graphs without circles we have for the set V of nodes that $V = \bigcup_{x \text{ leaf}} \mathbf{P}(x)$.*

Proof: Let W be a maximal path in the graph, i.e. a path that can't be extended in G. Let x be the last node in W. We claim that x is a leaf. If x is not a leaf then there exists an edge (x, y) in G for some node y. By maximality of W we can't add y to W, that means the edge (x, y) has been used before in W. This means that W has visited x before, i.e. there is a part of W that connects x to itself, i.e. a circle – contradiction. For the remaining statement let y be an arbitrary node. We can choose a longest path containing this node which has to end in a leaf x, hence y is contained in the past of x. □

Before proving our main result we proceed as we did before and first establish a description of the ROA, MPI set and GA by decomposing into subsystems according to the sparse structure of the dynamics, but now for general dynamics and not just trees as described in the previous sections.

Theorem 3 (Decomposition of the ROA, MPI set and GA) *Assume $X = X_1 \times \cdots \times X_r$ and (for the ROA) $X_{\mathrm{T}} = X_{1,\mathrm{T}} \times \cdots \times X_{r,\mathrm{T}}$ for compact sets $X_j, X_{j,T} \subset \mathbb{R}^{n_j}$ for $j = 1, \ldots, r$. Assume the sparsity graph has no circles. Let $x_1, \ldots, x_l$ be the leafs of the sparsity graphs of f with corresponding pasts $\mathbf{P}(x_1), \ldots, \mathbf{P}(x_l)$. For the ROA let $T \in \mathbb{R}_+$. Then the ROA R_T, MPI set M_+, GA $\mathcal{A}$ are given by*

$$R_T = \{x \in X : x_{\mathbf{P}(x_i)} \in R_T^i \text{ for } i = 1, \ldots, l\} \tag{31}$$

and

$$M_+ = \{x \in X : x_{\mathbf{P}(x_i)} \in M_+^i \text{ for } i = 1, \ldots, l\} \tag{32}$$

and

$$\mathcal{A} = \{x \in X : x_{\mathbf{P}(x_i)} \in \mathcal{A}^i \text{ for } i = 1, \dots, l\} \tag{33}$$

where $R_T^i, M_+^i, \mathcal{A}^i$ denote the ROA, MPI set and the GA for the subsystem induced by the past of the leaf x_i and $x_{\mathbf{P}(x_i)}$ denotes the vector of states of x that corresponds to the past of x_i.

Proof: We proceed in the same way as for the basic example (5) in Proposition 2, Proposition 3 and Proposition 4. Hence we only argue for the ROA because the arguments for the other sets are just adaptions according to the proofs of Propositions 3 and 4. Let R denote the right hand side of (31). Let $x \in R$. We have to show that for the solution $x(\cdot)$ of the dynamical system with initial value x we have $x(t) \in X$ for $t \in [0, T]$ and $x(T) \in X_T$. If we write $x(t) = (x_1(t), \dots, x_r(t))$ this means we have to show $x_k(t) \in X_k$ for $t \in [0, T]$ and $x_k(T) \in X_{k,\mathrm{T}}$ for all $k = 1, \dots, r$. Fix $k \in \{1, \dots, r\}$, by Lemma 1 and the assumption that the sparsity graph has no circles it follows that $x_k \in \mathbf{P}(x_i)$ for some leaf x_i. By definition of R it follows $x_k(t) \in X_k$ for all $t \in [0, T]$ and $x_k(T) \in X_{k,\mathrm{T}}$ from $x \in R$. Hence $x \in R_{\mathrm{T}}$. For an element $x \in R_{\mathrm{T}}$ we have $x(t) = (x_1(t), \dots, x_r(t)) \in X_1 \times \dots X_r = X$ for all $t \in [0, T]$ and $x(T) = (x_1(T), \dots, x_r(T)) \in X_{1,\mathrm{T}} \times \dots X_{r,\mathrm{T}}$. Let x_i be a leaf then clearly $x_{\mathbf{P}(x_i)}(t) \in \prod_{x_j \in \mathbf{P}(x_i)} X_j$ for $t \in [0, T]$ and $x_{\mathbf{P}(x_i)}(T) \in \prod_{x_j \in \mathbf{P}(x_i)} X_{j,\mathrm{T}}$, which exactly means $x_{\mathbf{P}(x_i)} \in R_{\mathrm{T}}^i$. □

This allows us to compute the desired sets based on computing them for the subsystems induced by the leafs.

Algorithm 1 (Decoupling procedure) *Given a dynamical system induced by f and a method for approximating/computing the ROA, MPI set or GA for an arbitrary dynamical system. Let $J_1, \dots, J_N$ be any partition of $\{1, \dots, n\}$.*

i. Reduce the circles in the corresponding sparsity graph of f as in Remark 6.

ii. Compute approximations for subsystems: Let $x_{i_1}, \dots, x_{i_l}$ be the leafs of the corresponding sparsity graph after reducing the circles. Use the given method to compute approximations $S_{i_1}, \dots, S_{i_l}$ of the ROAs, MPI sets or GAs respectively for the subsystems induced by the pasts of the leafs $x_{i_1}, \dots, x_{i_l}$.

iii. Glue $S_{i_1}, \dots, S_{i_l}$ together as in Theorem 3 by

$$S := \{x \in X : x_{\mathbf{P}(x_{i_r})} \in S_{i_r} \text{ for } r = 1, \dots, l\}.$$

The last ingredient before stating the main theorem is a generalization of the convergence property of the decoupling procedure from Proposition 5.

Theorem 4 *Let a dynamical system on $\mathbb{R}^n$ be induced by f with state constraint $X = \prod_{j=1}^N X_j$ for compact sets $X_j \subset \mathbb{R}^{n_j}$ and for the ROA $X_{\mathrm{T}} = \prod_{j=1}^N X_{j,\mathrm{T}}$ for a partition $J_1, \dots, J_N$ of $\{1, \dots, n\}$ with $|J_j| = n_j$. Given a method for approximating the ROA, MPI set or GA for an arbitrary dynamical system such that*

1. *in case of Hausdorff distance (induced by any norm on $\mathbb{R}^n$): If the method gives a convergent sequence of outer approximations S_k of the desired set S, i.e. $S_k \supset S$ and*

$$d_H(S_k, S) \to 0 \ , \quad as\ k \to \infty. \tag{34}$$

Then the decoupling procedure, Algorithm 1, produces a sequence of sets S'_k with

$$d_H(S'_k, S) \to 0 \ , \quad as\ k \to \infty. \tag{35}$$

for S denoting the desired set for the (sparse) dynamical system.

2. *In case of Lebesgue measure: Let the sparsity graph of f be acyclic and let $x_{i_1}, \ldots, x_{i_l}$ be the leafs. Let S'^r denote an approximation of the desired set S^r for $r = 1, \ldots, l$ for the subsystems induced by the leaf x_{i_r}. Then we have*

$$d_\lambda(S, S') := \lambda(S \Delta S') \leq \sum_{r=1}^{l} \lambda(S^r \Delta S'^r) \lambda(\prod_{x_k \notin \mathbf{P}(x_{i_r})} X_k) \tag{36}$$

where S is the desired set for the sparse dynamical system and S' the set obtained from Algorithm 1. In particular if a method produces approximations of S^i that converge to S^i with respect to d_λ then the decoupling method produces a set that converges to S with respect to d_λ.

Proof: We proceed in the same way as for the prototype setting. Let $x_{i_1}, \ldots, x_{i_l}$ be the leaf in the sparsity graph obtained from the decoupling procedure and $S^1_k, \ldots, S^l_k$ be the corresponding (converging outer) approximations of the desired sets for the subsystems induced by the leafs. For the first statement assume (35) does not hold. Then there exists a $\varepsilon > 0$ and an unbounded subsequence $(k_m)_{m \in \mathbb{N}}$ such that

$$d_H(S'_{K_m}, S) > \varepsilon \tag{37}$$

and we find points $x_{k_m} \in S'_{k_m}$ with $\operatorname{dist}(x_{k_m}, S) > \varepsilon$. By construction of S'_k, boundedness of $S^1, \ldots, S^l$ and the assumption (34) it follows that there exists $x \in \mathbb{R}^n$ and a subsequence of $(k_m)_{m \in \mathbb{N}}$ which we will still denote by $(k_m)_{m \in \mathbb{N}}$ such that $x_{k_m} \to x$ as $m \to \infty$. By assumption (34) there exist $y^i_{k_m} \in S^i$ for $i = 1, \ldots, l$ with $\|y^i_{k_m} - (x_{k_m})_{\mathbf{P}(x_i)}\| \to 0$ as $m \to \infty$. Hence also $y^i_{k_m} \to x_{\mathbf{P}(x_i)}$ as $m \to \infty$ for $i = 1, \ldots, l$. Because $S^1, \ldots, S^l$ are closed it follows $x_{\mathbf{P}(x_i)} \in S^i$ for $i = 1, \ldots, l$ and by Theorem 3 we get $x \in S$. In particular we get

$$\varepsilon < \operatorname{dist}(x_{k_m}, S) \leq \|x_{k_m} - x\| \to 0$$

as $m \to \infty$, which is a contradiction. For the second statement we get by the decoupling procedure Algorithm 1 that $S' \subset X$ and

$$S \Delta S' \ \subset \ \bigcup_{r=1}^{l} \{x \in X : x_{\mathbf{P}(x_{i_r})} \in S^r \Delta S'^r\}.$$

Applying the Lebesgue measure to this inclusion gives

$$\begin{aligned}\lambda(S\Delta S') &\leq \sum_{r=1}^{l} \lambda\left(\{x \in X : x_{\mathbf{P}(x_{i_r})} \in S^r \Delta S'^r\}\right) \\ &= \sum_{r=1}^{l} \lambda(S^r \Delta S'^r)\lambda(\prod_{x_k \notin \mathbf{P}(x_r)} X_k).\end{aligned}$$

□

In the next section we will state methods from [10], [12] and [20] that give converging (with respect to d_λ) approximations of the ROA, MPI set and GA. Then we have everything we need to state and prove our main theorem.

Before doing so we first describe how to choose a good partition of nodes for the sparsity graph of a function f.

6.1 Selecting a partition

The choice of a partition of the states can influence the performance of the method strongly.

First we start with factorizing the state space as fine as possible in order to decouple the dynamical system as much as possible.

Definition 7 *We say $X \subset \mathbb{R}^n$ factors with respect to a partition $J_1, \ldots, J_N$ of $\{1, \ldots, n\}$ if there exist sets $X_i \subset \mathbb{R}^{n_i}$ where $n_i = |J_i|$ for $i = 1, \ldots, n$ such that*

$$X = \{x \in \mathbb{R}^n : P_{J_i} x \in X_i \text{ for } i = 1, \ldots, N\}.$$

We say $J_1, \ldots, J_N$ induces a factorization; the sets X_i are given by $P_{j_i} X$.

This is the natural generalization of the factorization we needed for the prototype setting. The following Lemma allows us to find a finest factorization of X which will be useful in order to group only as much nodes in the sparsity graph together as needed.

Lemma 2 *There exists a minimal factorization for X; that is a factorization induced by $J_1, \ldots, J_N$ of X, such that for any other factorization induced by $I_1, \ldots, I_M$ we have for all $i = 1, \ldots, M$ that $I_i = \bigcup_{k: J_k \subset I_i} J_k$.*

Proof: We give a proof in the Appendix. □

In order to satisfy Assumption 1 for this factorization, namely that $X = \prod_{i=1}^{N} X_i$, we only need to perform a permutation of coordinates of $\mathbb{R}^n$.

It is now clear that the partition obtained from Lemma 2 allows the finest decoupling of the dynamical system into subsystems, i.e. a decoupling into subsystems of smallest dimension.

In the following we focus on finding outer approximations of the ROA, MPI set and GA based on convex optimization proposed in [10], [12] and [20].

7 Structured semidefinite programming outer approximations

For the region of attraction, maximum positively invariant set and global attractors there exist representations in terms of solutions of infinite dimensional linear programs (see for example [21], [10], [13], [12] and [20]). Those provide converging outer approximations. We will follow those methods here and combine them with the decoupling procedure which then allows faster computation. Further we propose similar LPs that exploit the sparse structure even further but they have the disadvantage that they do not provide guaranteed convergence which is why we suggest to pair them with the convergent approach obtained from a hierarchy of SDPs from [10], [12] and [20] with the decoupling procedure, thereby guaranteeing convergence by design.

At the beginning of this section we consider again general dynamical system on $\mathbb{R}^n$ with compact state constraint set $X \subset \mathbb{R}^n$ and no sparse structure. Sparse structures will be considered in the subsections 7.3 and 7.4.

7.1 Linear program representations for the ROA, MPI set and GA

To state the LP from [10] for the ROA we need the Liouville operator $\mathcal{L} : \mathcal{C}^1([0,T] \times X) \to \mathcal{C}([0,T] \times X)$ that captures the dynamics, which is given by

$$\mathcal{L}v := \frac{\partial}{\partial t} v + \nabla v \cdot f. \tag{38}$$

The dual LP from [10] is given by

$$\begin{array}{lll} d^* := & \inf \int\limits_X w(x)\, d\lambda(x) & \\ \text{s.t.} & v \in \mathcal{C}^1([0,T] \times \mathbb{R}^n), w \in \mathcal{C}(X) & \\ & \mathcal{L}v(t,x) \geq 0 & \text{on } [0,T] \times X \\ & v(T,x) \geq 0 & \text{on } X_{\mathrm{T}} \\ & w(x) \geq 0 & \text{on } X \\ & w(x) \geq v(0,x) + 1 & \text{on } X \end{array} \tag{39}$$

In [12] an LP that relates to the MPI set was presented. This LP with discounting factor $\beta > 0$ is given by

$$\begin{array}{lll} d^* := & \inf \int\limits_X w(x)\, d\lambda(x) & \\ \text{s.t.} & v \in \mathcal{C}^1(\mathbb{R}^n), w \in \mathcal{C}(X) & \\ & \nabla v \cdot f \leq \beta v & \text{on } X \\ & w \geq 0 & \text{on } X \\ & w \geq v + 1 & \text{on } X \end{array} \tag{40}$$

Based on the (dual) LP for the MPI set the following LP for the GA was proposed in [20] with discounting factors $\beta_1, \beta_2 > 0$

$$\begin{array}{lll} d^* := & \inf \int_X w \, d\lambda & \\ \text{s.t.} & p, q \in \mathcal{C}^1(\mathbb{R}^n), w \in \mathcal{C}(X) & \\ & -p - q + w \geq 1 & \text{on X} \\ & w \geq 0 & \text{on X} \\ & \beta_1 p - \nabla p \cdot f \geq 0 & \text{on X} \\ & \beta_2 q + \nabla q \cdot f \geq 0 & \text{on X} \end{array} \tag{41}$$

Remark 7 *The dual problem (39), (40) and (41) have the advantage that they give rise to outer approximations by the sets $w^{-1}([1, \infty)$, which get tight as feasible points (v, w) or (p, q, w) respectively get optimal. But this is typically not the case for primal feasible elements, this is why we don't state the primal LPs here. Inner approximations can be approached in a similar way by using the LPs for inner approximations from [14] and [17]*

7.2 Semidefinite programs for the ROA, MPI set and GA

In the previous subsection we have presented infinite dimensional LPs on the space of continuous functions – whose minimizers, or more precisely minimizing sequences, allow representations of the ROA, MPI set and GA. In this section we state a well known approach to such LPs that reduces the LP to a hierarchy of semidefinite programs (SDPs). Those SDP tightenings for the dual problems can be found in the corresponding papers (for example [10], [12], [20]). Combining the SDP approach with the decoupling procedure from section 5 we get a sparse approach towards approximating the ROA, MPI set and GA. We state the SDP procedure here to have a selfcontained sparse approach to convergent approximations for those sets.

For this approach it is necessary to assume additional algebraic structure of the problem because the dual LP tightens to a sum-of-squares problem, which leads to hierarchy of SDPs. This is a standard procedure and we refer to [16] or [15] for details.

Assumption 2 *The vector field f is polynomial and $X \subset \mathbb{R}^n$ is a compact basic semi-algebraic set, that is, there exist polynomials $p_1, \ldots, p_i \in \mathbb{R}[x]$ such that $X = \{x \in \mathbb{R}^n : p_j(x) \geq 0$ for $j = 1, \ldots, i\}$. Further we assume that one of the p_j is given by $p_j(x) = R^2 - \|x\|_2^2$ for some large enough $R \in \mathbb{R}$. And similar for $X_T \subset \mathbb{R}^n$ for polynomials q_j for $j = 1, \ldots, l$.*

If there are no such polynomials of the form $R^2 - \|x\|_2^2$ then by compactness of X, X_T we can add the redundant inequality $R - \|x\|_2^2 \geq 0$ for the smallest radius R such that $\overline{B_R(0)}$ contains X, X_T. This will be needed in order to apply Putinar's Positivstellensatz (see [18]).

The idea for the SDP tightenings is first to reduce the space of continuous functions to the space of polynomials. The fact that the optimal value for the LP is not affected is justified by the Stone-Weierstraß theorem (and the existence of strictly feasible points). For the space of polynomials there is a natural way of reducing to a finite dimensional space, namely by

bounding the total degree. That gives a sequence of finite dimensional optimization problems (in the coefficients of the polynomials). But those optimization problems are not tractable because testing non-negativity is a difficult task. The replacement of non-negativity as a sum-of-squares conditions allows a representation as an SDP. Finally convergence is guaranteed by Putinar's positivstellensatz.

In the following k will always denote the maximal total degree of the occurring polynomials and d_f the total degree of the polynomial f.

We start with the SDP tightening for the ROA for a (non-sparse) dynamical system with constraint set X with finite time horizon $[0, T]$

$$\begin{aligned} d_k^* := \quad & \inf \langle \mathbf{w}, \boldsymbol{\lambda} \rangle \\ \text{s.t.} \quad & v \in \mathbb{R}[t,x]_{k+1-d_f}, w \in \mathbb{R}[x]_k \\ & -\mathcal{L}v = s_1 + \sum_{j=1}^{i} a_j p_j + b\mathbf{t}(T - \mathbf{t}) \\ & v(T,\cdot,\cdot) = s_2 + \sum_{j=1}^{l} c_j q_j \\ & w = s_3 + \sum_{j=1}^{i} d_l p_j \\ & w - v(0,\cdot,\cdot) - 1 = s_4 + \sum_{j=1}^{i} e_j p_j \end{aligned} \tag{42}$$

for sum-of-squares polynomials $s_1, a_j, b_l^1, b \in \mathbb{R}[t,x]$, $s_2, s_3, s_4, c_j, e_j \in \mathbb{R}[x]$ for $j = 1, \ldots, i$; such that all occurring polynomials in the SDP (42) have degree at most k. The vector $\boldsymbol{\lambda}$ denotes the vector of moments of the Lebesgue measure on X and $\mathbf{w}$ denotes the coefficients of the polynomial w, such that

$$\langle \mathbf{w}, \boldsymbol{\lambda} \rangle = \int_X w(x)\, d\lambda.$$

And similar for the MPI set.

$$\begin{aligned} d_k^* := \quad & \inf \langle \mathbf{w}, \boldsymbol{\lambda} \rangle \\ \text{s.t.} \quad & v \in \mathbb{R}[x]_{k+1-d_f}, w \in \mathbb{R}[x]_k \\ & \beta v - \nabla v \cdot f = s_1 + \sum_{j=1}^{i} a_j p_j \\ & w = s_2 + \sum_{j=1}^{i} b_j p_l \\ & w - v - 1 = s_3 + \sum_{j=1}^{i} c_j p_j \end{aligned} \tag{43}$$

with sum-of-squares polynomials $s_1, s_2, s_3, a_j, b_j, c_j \in \mathbb{R}[x]$, for $j = 1, \ldots, i$, such that all occurring polynomials in (43) are of degree at most k.

And analogue for the GA.

$$\begin{aligned} d_k^* := \quad & \inf\langle \mathbf{w}, \boldsymbol{\lambda} \rangle \\ \text{s.t.} \quad & p, q \in \mathbb{R}[x]_{k+1-d_f}, w \in \mathbb{R}[x]_k \\ & \beta_1 p - \nabla p \cdot f = s_1 + \sum_{j=1}^{i} a_j p_j \\ & \beta_2 q - \nabla q \cdot f = s_2 + \sum_{j=1}^{i} b_j p_j \\ & w = s_3 + \sum_{j=1}^{i} c_j p_l \\ & w - p - q - 1 = s_4 + \sum_{j=1}^{i} d_j p_j \end{aligned} \tag{44}$$

with sum-of-squares polynomials $s_1, s_2, s_3, s_4, a_j, b_j, c_j, d_j \in \mathbb{R}[x]$, for $j = 1, \ldots, i$ such that all occurring polynomials in (43) are of degree at most k.

By [10], [12] and [20] the sequences d_k^* from (42) for the ROA, (43) for the MPI set and (44) for the GA converge monotonically from above to the Lebesgue measure of the corresponding sets. Further the sets

$$S_k := w^{-1}([1, \infty)] = \{x \in X : w(x) \geq 1\} \tag{45}$$

are outer approximations that get tight (with respect to Lebesgue measure discrepancy) when (v, w), respectively (p, q, w) gets optimal.

7.3 The main algorithm and main theorem

Now we have everything we need to state our main algorithm and prove our main theorem. The main ingredients are Theorem 4 and convergence properties for the hierarchy of SDPs.

Algorithm 2 *Let $J_1, \ldots, J_N$ be a partition of $\{1, \ldots, n\}$ with $|J_j| = n_j$ and a dynamical system on $\mathbb{R}^n$ be induced by a polynomial f with state constraint $X = \prod_{j=1}^{N} X_j$, for compact basic semialgebraic sets $X_j \subset \mathbb{R}^{n_j}$ satisfying Assumption 2 (and for the ROA $X_T = \prod_{j=1}^{N} X_{j,T}$ for compact basic semialgebraic $X_{T,j} \subset \mathbb{R}^{n_j}$) for $j = 1, \ldots, N$. Fix the maximum degree $k \in \mathbb{N}$ of polynomials occurring in the SDPs.*

i. Reduce the circles in the corresponding dimension weighted sparsity graph of f as in Remark 6.

ii. Compute outer approximations of the ROA, MPI set or GA for subsystems by the SDPs (42), respectively (43) or respectively (44): Let $x_{i_1}, \ldots, x_{i_l}$ be the leafs of the corresponding sparsity graph after reducing the circles. Use the SDPs (42), (43) or (44) respectively for polynomials up to degree k to compute approximations $S_1^{(k)}, \ldots, S_l^{(k)}$ of the ROAs, MPI sets or GA respectively for the subsystems induced by the pasts of the leafs $x_{i_1}, \ldots, x_{i_l}$.

iii. Glue $S_1^{(k)}, \ldots, S_l^{(k)}$ together as in Theorem 3 by

$$\begin{aligned} S^{(k)} &:= \{x \in X : x_{\mathbf{P}(x_{i_r})} \in S_r^{(k)} \text{ for } r = 1, \ldots, l\} \\ &= \{x \in X : w_k^i(x_{\mathbf{P}(x_i)}) \geq 1 \text{ for } i = 1, \ldots, l\}. \end{aligned} \tag{46}$$

The second equality in (46) follows immediately from (45).

Before stating the main theorem we remind of the definition of the largest dimension weighted past ω defined in (2).

Theorem 5 *Algorithm 2 produces converging outer approximations of the ROA, MPI set or GA respectively, i.e.*

$$S^{(k)} \supset S \text{ for all } k \in \mathbb{N} \text{ and } d_\lambda(S^{(k)}, S) = \lambda(S^{(k)} \Delta S) \to 0 \text{ as } k \to \infty$$

where S denotes the ROA, MPI set or GA respectively for the dynamical system. The complexity of the corresponding SDPs that need to be solved in Algorithm 2 is determined by ω.

Proof: This follows immediately from the convergence results of [10], [12], [20] and Theorem 4 because the largest SDP, i.e. the SDP involving the most variables, that occurs is induced by the subsystem whose leaf has the largest weighted past and this SDP acts on sum-of-squares multipliers on ω variables. □

That the complexity of the SDPs is determined by ω is the reason why this approach is useful to reduce complexity. The SDPs obtained by SOS hierarchies grow combinatorically in the number of variables and the degree bound k. The number of variables used in each branch of the tree reduces the number of variables for the remaining problems. To make this more precise let us have a look at the basic branching as in Figure (3). Let n_1, n_2, n_3 be the number of variables in x_1, x_2, x_3. Let k be the degree used for the SDPs. Then the size of the largest sum-of-squares multiplier for the full system is

$$\binom{n_1 + n_2 + n_3 + \frac{k}{2}}{\frac{k}{2}}$$

while for the subsystems it is

$$\binom{n_1 + n_2 + \frac{d}{2}}{\frac{k}{2}} \text{ and } \binom{n_1 + n_3 + \frac{k}{2}}{\frac{k}{2}}.$$

For general graphs it follows similarly that the more the graph separates into subsystems the more effective this approach gets.

Let's precisely count the number of variables in the SDP for degree bound k. Here in case of the MPI set. Let $x_1, \ldots, x_l$ be the leafs of the dimension weighted sparsity graph of f and the total set of nodes $\{x_1, \ldots, x_N\}$. Let $n(\mathbf{P}(x_i)) = \sum_{j \in \mathbf{P}(x_i)} n_j$ be the dimension of the state space for the past of x_i. Let ν_j be the number of constraints defining the set X_j and

$\phi_i := \sum_{x_j \in \mathbf{P}(x_i)} \nu_j$ the number of constraints defining the state constraint for the subsystem induced by by the past of x_i. Let $\Phi := \sum_{j=1}^{N} \nu_j$ be the total number of constraints defining X. The number of variables in the non-sparse SDP for the full system is given by

$$3\,(1+\Phi) \cdot \binom{n+\frac{k}{2}}{\frac{k}{2}} \tag{47}$$

while for the sparse SDP we get

$$3\sum_{i=1}^{l}(1+\phi_i) \cdot \binom{n(\mathbf{P}(x_i))+\frac{k}{2}}{\frac{k}{2}} \tag{48}$$

Hence we see that the reduction in the number of variables is significant if the dynamics is strongly separated, i.e. pasts of the leafs overlap less, i.e. n_i and ϕ_i are small compared to n and Φ. Which is what we would expect because strong separation tells us that fewer interactions are needed in order to describe the system.

Remark 8 *Treating the subsystems separately by the decoupling procedure has another advantage. Namely it allows to take properties of the subsystems into account. Particularly for the SDP approach this allows for example the use of different degrees for the hierarchies of different subsystems. This can be useful if the hierarchy for some subsystems allow the use of low degrees to already capture the dynamics well while for other subsystems high degrees are required to obtain accurate approximations. For the whole system this typically means that also a high degree for the SDP hierarchy is needed (in order to capture the dynamics of the more complex subsystem).*

7.4 Sparse improvement

We propose a slightly adapted LP that allows a further (sparse) improvement on the outer approximation while maintaining the reduced computational complexity.

For the rest of this section assume that the sparsity graph of f with respect to a given partition is acyclic and has leafs $x_{i_1}, \ldots, x_{i_l}$. Let $I_1, \ldots, I_l$ be the set of indices corresponding to the nodes in the past of $x_{i_1}, \ldots, x_{i_l}$. And let $X_{I_r} := \prod_{j \in I_r} X_j$ denote the constraint space for the subsystem induced by the past of x_{i_r} for $r = 1, \ldots, l$. The set x_{I_r} denotes the projection of $x \in X$ onto X_{I_r}, i.e. the components of x corresponding to I_r, similar for the function f let f_{I_r} denote the components of f corresponding to the index set I_r. Let n_r be the dimension of the state space for the subsystem induced by the past of x_{i_r}, i.e. $X_{I_r} \subset \mathbb{R}^{n_r}$.

It is possible to combine the LPs for the subsystems but such that the constraints only act on functions on X_{I_r} for $r = 1, \ldots, l$.

We propose the following dual sparse LP for the ROA

$$
\begin{aligned}
d_s^* := \quad & \inf \sum_{r=1}^{l} \int_{X_{I_r}} w^r(y)\, d\lambda(y) \\
\text{s.t.} \quad & v^r \in \mathcal{C}^1([0,T] \times \mathbb{R}^{n_r}), w^r \in \mathcal{C}(X_{I_r}) && 1 \le r \le l \\
& \sum_{r=1}^{l} \mathcal{L}_r v^r(t, x_{I_r}) \ge 0 && \text{on } [0,T] \times X \\
& \sum_{r=1}^{l} v^r(T, x_{I_r}) \ge 0 && \text{on } X_{\mathrm{T}} \\
& \sum_{r=1}^{l} w^r(x_{I_r}) \ge 0 && \text{on } X \\
& \sum_{r=1}^{l} w^r(x_{I_r}) - v^r(0, x_{I_r}) \ge l && \text{on } X
\end{aligned}
\tag{49}
$$

Where $\mathcal{L}_r$ denotes the Liouville operator (38) on the subsystem induced by the past of x_{i_r}.

The LP is sparse because the functions w^r, v^r only depend on x_{I_r} instead of x. For the corresponding SDP we choose the SOS multiplier to only depend on the variables x_{I_r}.

Remark 9 *We have summed the corresponding inequalities of the LP (39) for the subsystems. This has the advantage that the set of feasible points for the LP (and the corresponding SDP) is larger. On the other hand it enforces less structure on the feasible points. This can potentially hamper convergence of the approximations. This undesirable property can be avoided by intersecting with the approximations coming from the fully decoupled approach; this is formally stated in Theorem 6.*

Similar to the set constructed by the decoupling based on the SDP hierarchy in (45) we can construct a superset of the ROA based on feasible sets for the sparse LP (49).

Proposition 6 *Let $(w^1, v^1, w^2, v^2, \ldots, w^l, v^l)$ be feasible. Then*

$$
\{x \in X : \sum_{r=1}^{l} w^r(x_{I_r}) \ge l\} \supset R_T. \tag{50}
$$

Proof: We can apply Lemma 2 from [10] to the functions $v(x) := \sum_{r=1}^{l} v^r(x_{I_r})$ and $w(x) := \sum_{r=1}^{l} w^r(x_{I_r})$ and it follows the conclusion. □

Similar arguments for the LPs (40) and (41) for the MPI set and the GA lead to the following

sparse LPs for discounting factors $\beta_1, \ldots, \beta_l > 0$

$$\begin{array}{lll} d_s^* := & \inf \sum_{r=1}^{l} \int_{X_{I_r}} w^r(y)\, d\lambda(y) & \\ \text{s.t.} & v^r \in \mathcal{C}^1(\mathbb{R}^{n_r}), w^r \in \mathcal{C}(X_{I_r}) & 1 \le r \le l \\ & \sum_{r=1}^{l} \left(\nabla v^r \cdot f_{I_r} - \beta_r v^r\right)(x_{I_r}) \le 0 & \text{on } X \\ & \sum_{r=1}^{l} w^r(x_{I_r}) \ge 0 & \text{on } X \\ & \sum_{r=1}^{l} (w^r - v^r)(x_{I_r}) \ge l & \text{on } X \end{array} \tag{51}$$

for the MPI set and

$$\begin{array}{lll} d_s^* := & \inf \sum_{r=1}^{l} \int_{X_{I_r}} w^r(y)\, d\lambda(y) & \\ \text{s.t.} & p^r, q^r \in \mathcal{C}^1(\mathbb{R}^{n_r}), w^r \in \mathcal{C}(X_{I_r}) & \\ & \sum_{r=1}^{l} \left(\nabla p^r \cdot f_{I_r} - \beta_r p^r\right)(x_{I_r}) \le 0 & \text{on } X \\ & \sum_{r=1}^{l} \left(\nabla q^r \cdot f_{I_r} + \beta_r q^r\right)(x_{I_r}) \le 0 \text{ on } X & \text{on } X \\ & \sum_{r=1}^{l} w^r(x_{I_r}) \ge 0 & \text{on } X \\ & \sum_{r=1}^{l} (w^r - p^r - q^r)(x_{I_r}) \ge l & \text{on } X \end{array} \tag{52}$$

for the GA.

And as for the region of attraction we get the following proposition.

Proposition 7 *Let $(w^1, v^1, w^2, v^2, \ldots, w^l, v^l)$ be feasible for (51) or $(w^1, p^1, q^1, \ldots, w^l, p^l, q^l)$ be feasible for (52) respectively. Then*

$$\{x \in X : \sum_{r=1}^{l} w^r(x_{I_r}) \ge l\} \supset M_+ \tag{53}$$

or respectively

$$\{x \in X : \sum_{r=1}^{l} w^r(x_{I_r}) \ge l\} \supset \mathcal{A}. \tag{54}$$

Proof: For the MPI set we can apply Lemma 3 from [12] to the functions $v(x) := \sum_{r=1}^{l} v^r(x_{I_r})$ and $w(x) := \sum_{r=1}^{l} w^r(x_{I_r})$ and for the GA we apply Lemma 1 from [20] to the functions $v^1(x) := \sum_{r=1}^{l} p^r(x_{I_r})$, $v^2(x) := \sum_{r=1}^{l} q^r(x_{I_r})$ and $w(x) := \sum_{r=1}^{l} w^r(x_{I_r})$. □

We can enforce the sparse structure of the LPs (49), (51) and (52) to the corresponding hierarchy of SDPs; by that we mean that instead of replacing the non-negativity constraint by an SOS constraint with polynomials on X we only use SOS polynomials on the spaces X_{I_r}. This reduces the complexity due to the possibility to work with the smaller spaces $\mathbb{R}[x_{I_1}], \dots, \mathbb{R}[x_{I_l}] \subset \mathbb{R}[x]$ similar to treating the subsystems separately as in the previous subsection.

Even though this approach has similar computational complexity – because the largest SOS multiplier acts on ω variables – we can't guarantee convergence. This is why we need to pair this method with the convergent method based on the decoupling the dynamical systems to obtain a convergent sequence of outer approximation.

Theorem 6 *Under the assumption of Theorem 5 let $S^{(k)}$ for $k \in \mathbb{N}$ be the outer approximation of the ROA, respectively the MPI set, respectively the GA from (46) and $Y^{(k)}$ be the sets obtained from (50), respectively (53) or respectively (54) by optimal points $(w^r, v^r)_{r=1,\dots,l}$, respectively $(w^r, p^r, q^r)_{r=1,\dots,l}$, of the corresponding sparse SDPs for (49), respectively (51), respectively (52). Then $S^{(k)} \cap Y^{(k)}$ is a converging (with respect to d_λ) outer approximation of the ROA, respectively the MPI set, respectively the GA. The largest occurring SOS multiplier acts on ω variables.*

Proof: By Propositions 6 and 7 we have $S^{(k)} \supset S^{(k)} \cap Y^{(k)} \supset S$ where S denotes the desired set. Hence convergence follows from convergence of $S^{(k)}$ stated in Theorem 5. By the enforced sparse structure of the SDPs for the sparse LPs (49), (51) and (52) the largest SOS multiplier occurs corresponding to the subsystem induced by a leaf with the state space of largest dimension; hence it acts on ω variables. □

8 Discrete systems

All concepts from the previous chapters can also be applied to discrete systems

$$x_{k+1} = f(x_k)\ ,\ x_0 \in X \tag{55}$$

for a function $f : \mathbb{R}^n \to \mathbb{R}^n$ and a compact constraint set $X \subset \mathbb{R}^n$.

The definition of a subsystem is the same and the (dimension weighted) sparsity also is defined the same way.

By the same arguments as for continuous time systems the decomposition of the region of attraction, maximum positive invariant set and GA is proven. And the convergence property for the decoupling procedure in Theorem 4 follows immediately because the only ingredients there are the decomposition of these sets and a well working method for approximating them.

Also the LPs and corresponding hierarchies of SDPs for those sets have counterparts for discrete systems [10], [12] and [20].

9 Numerical examples

9.1 Cherry structure

As an example we consider the interconnection of Van der Pol oscillators as in Figure 6. For

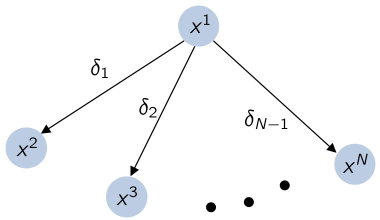

Figure 6: Interconnection of Van-Der Pol oscillators in a cherry structure.

the leaf nodes $x^2, \ldots, x^N$, the dynamics is

$$\begin{aligned}\dot{x}_1^i &= 2x_2^i \\ \dot{x}_2^i &= -0.8x_1^i - 10[(x_1^i)^2 - 0.21]x_2^i + \delta_i x_1^1.\end{aligned}$$

For the root note x^1, the dynamics is

$$\begin{aligned}\dot{x}_1^1 &= 2x_2^1 \\ \dot{x}_2^1 &= -0.8x_1^1 - 10[(x_1^1)^2 - 0.21]x_2^1.\end{aligned}$$

We illustrate the decoupling procedure by computing outer approximations of the MPI set of this system with respect to the constraint set $[-1.2, 1.2]^{2N}$. We carry out the computation for degree $k = 8$ and $N = 10$, resulting in a total dimension of the state-space equal to 20. The optimal decoupling in this case is into subsystems (x^1, x^i), $i = 2, \ldots, N$, each of dimension four. Figure 7 shows the sections of the MPI set outer approximations when the value at the root node is fixed at [0.5,-0.1]. The computation time was 12 seconds.[1] Next we carried out the the computation with $k = 8$ and $N = 26$, resulting in state-space dimension of 52. Figure 8 shows the sections of the MPI set outer approximations when the value at the root node is fixed at [0.5,-0.1]. The total computation time was 40.3 seconds. It should be mentioned that these problems in dimension 20 or 52 are currently intractable without structure exploitation. Here the sparse structure allowed for decoupling in 9 respectively 25 problems in 4 variables, which were solved in less than a minute in total.

9.2 Tree structure

As our second example we consider a network of Van der Pol oscillators as in Figure 9. The coupling is as in the previous example from the first component of the predecessor state to the second component of the successor state. The coupling intensity δ is set to 0.1 for each

[1] All computations were carried out using Yalmip and MOSEK running on Matlab and 4.2 GHz Intel Core i7, 32 GB 2400MHz DDR4.

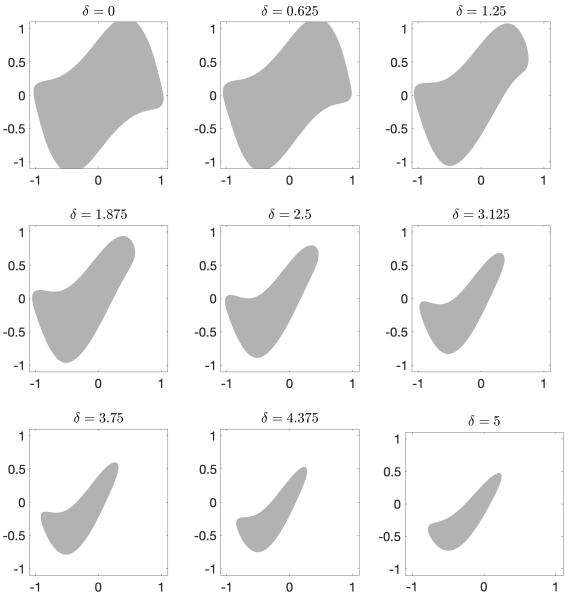

Figure 7: Van der Pol oscillators in a cherry structure: Sections of the outer approximations of the MPI set for $k = 8$ and $N = 10$.

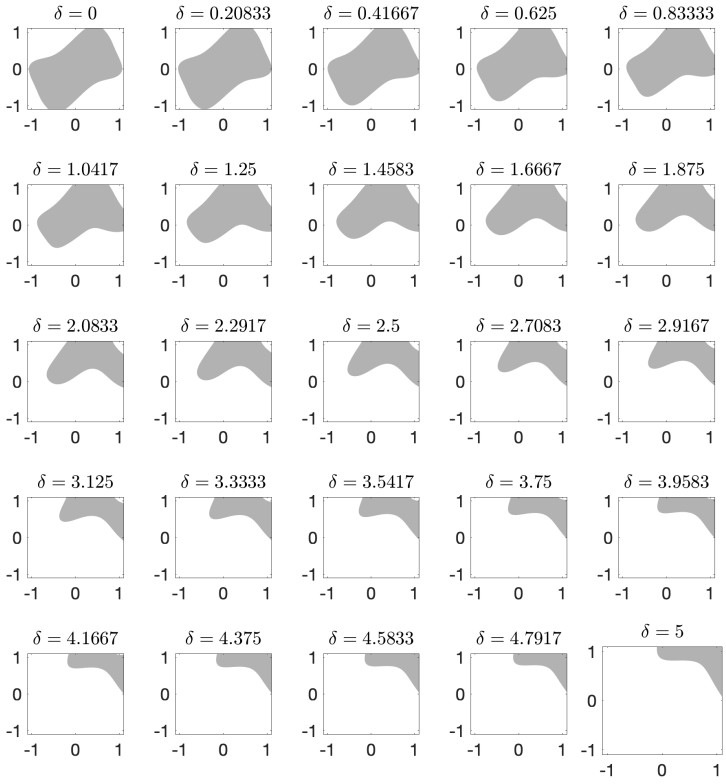

Figure 8: Van der Pol oscillators in a cherry structure: Sections of the outer approximations of the MPI set for degree $k = 8$ and $N = 26$.

edge. The goal is to compute the MPI set with respect to the constraint set $[-1.2, 1.2]^{10}$. The optimal decoupling is now into 3 subsystems given by (x^1, x^2, x^4), (x^1, x^2, x^5), (x^1, x^3); the respective dimensions are 6, 6 and 4. Figure 10 shows six random sections of the ten dimensional MPI set outer approximation computed by our approach with degree $k = 8$. Even though the the overall state-space dimension 10 is less than it was in our previous example, the computation time of 285 seconds is higher since the maximum dimension of the subsystems is higher.

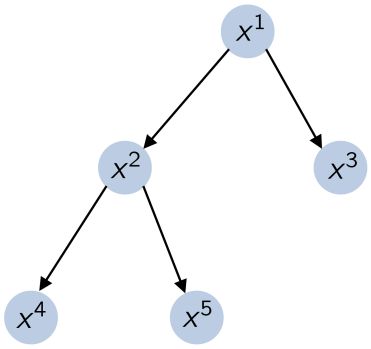

Figure 9: Interconnection of Van-Der Pol oscillators in a tree structure.

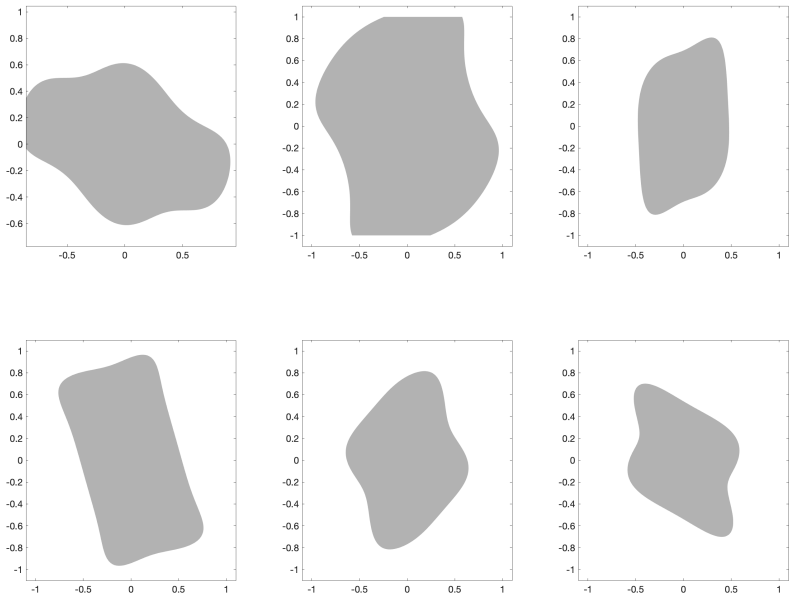

Figure 10: Van der Pol oscillators in a tree structure: Random projections of the outer approximation to the ten dimensional MPI set.

10 Conclusion

We presented a decomposition of several important sets related to nonlinear dynamical systems based on their correspondences for subsystems of the dynamical system. This was motivated by [2] and extended from the region of attraction to also the maximum positively invariant set as well as GA. We focused on the uncontrolled but state constraint case and showed how this concept can be generalized for general dynamical systems on $\mathbb{R}^n$. We showed that this decomposition gives rise to methods for computing these sets from their correspondences for the subsystems. Using the works [10], [12] and [20] we presented a method that provides a converging sequence of outer approximations based on convex optimization problems, while exploiting the underlying structure.

We believe that decomposing the dynamical system into subsystems as presented here can be beneficial for other objectives such as constructions of Lyapunov functions or invariant measures to name just two. It may also be of interest to exploit sparsity for extreme value computation, building on [8]. Another direction of future work is the inclusion of control, i.e., the computation of the region of attraction with control, the maximum controlled invariant set and optimal control. Utilizing this approach in a data-driven setting, building on [11], is another possible generalization.

Sparsity in the dependence of the dynamics of the states is not the only structure of f that can be exploited. If for example f is a polynomial, then the algebraic structure of f can be investigated, leading possibly to further reduction of complexity, in analogy to the static polynomial optimization (e.g., [24]). In addition, more general sparse structures should be investigated as we have seen that our approach treats straight paths or circles as subsystems – in the same way as if all the corresponding nodes were fully connected. Work in this direction was done in [21].

Finally this approach depends on the explicit description of f – it is not coordinate free. This can be seen for example by a linear dynamical system $\dot{x} = Ax$ for diagonalizable matrix $A \in \mathbb{R}^{n\times n}$ with non-zero entries. Since every entry of A is non-zero the sparsity graph is the complete graph while after a change of coordinates that diagonalizes A the corresponding sparsity graph for this dynamical system consists of isolated nodes, i.e. there are no edges at all. Therefore it would be interesting to understand the intrinsic, coordinate free situation, taking into account the presence of constraints that need to decouple simultaneously with the dynamics.

11 Appendix: proof of Lemma 2

Proof: *of Lemma 2.*
We look at the set $T := \{J \subset \{1,\dots,n\} : J \text{ and } \{1,\dots,n\}\setminus J \text{ induces a factorization of } X\}$. The set T is the collection of all partitions consisting of only two sets, such that they induce a factorization of X. We will see that T contains minimal elements (with respect to inclusion); these will give rise to the desired factorization of X. We start with the following properties of T.

1. T is non-empty.
   $J = \{1, \dots, n\}$ is contained in X because in induces the trivial factorization X of factoring into itself.

2. T is closed with respect to taking the complement in $\{1, \dots, n\}$.
   Let $J \in T$ then $J^c := \{1, \dots, n\} \setminus J \in T$ because J^c, J is a partition that induces the same factorization as J, J^c.

3. T is closed with respect to intersections.
   Let $J_1, J_2 \in T$ with corresponding sets $X_1 := P_{J_1}(X)$, $X_2 := P_{J_1^c}(X)$ and $Y_1 := P_{J_2}(X), Y_2 := P_{J_2^c}(X)$. Let $J := J_1 \cap J_2$ and $I := \{1, \dots, n\} \setminus J$. We claim J, I induces a factorization. Therefore let $Z_1 := P_J(X)$ and $Z_2 := P_I(X)$. We need to show that we have
   $$X = X' := \{x \in \mathbb{R}^n : P_J(x) \in Z_1, P_I(x) \in Z_2\}. \tag{56}$$
   For any $x \in X$ we have $x \in X'$ by definition of Z_1 and Z_2. Let $x' \in X'$. By definition of Z_1 there exists $x_1 \in X$ with $P_J(x_1) = P_J(x')$. From $J_1 \in T$ it follows $P_{J_1}(x_1) \in X_1$. Since $I \supset J_1^c$ it follows $P_{J_1^c} x_1 \in P_{J_1^c}(P_I(X)) = P_{J_1^c}(X) = X_2$. Since J_1, J_1^c induces a factorization we get that the element $x_2 \in \mathbb{R}^n$ with $P_{J_1}(x_2) = P_{J_1}(x_1)$ and $P_{J_1^c}(x_2) = P_{J_1^c}(P_I(x'))$ belongs to X. If we repeat this process with J_1 replaced by J_2 we find an element $x_3 \in X$ such that $P_J(x_3) = P_{J_1 \cap J_2}(x_3) = P_{J_1 \cap J_2}(x')$ and $P_I(x_3) = P_{J_1^c \cup J_2^c}(x_3) = P_{J_1^c \cup J_2^c}(x')$, i.e. $x' = x_3 \in X$.

4. T is closed with respect to taking union.
   Let $J_1, J_2 \in T$. Then $J_1 \cup J_2 = (J_1^c \cap J_2^c)^c \in T$.

It follows that T is a (finite) topology and hence there exists a minimal basis of T (consisting of the smallest neighbourhoods of each point), i.e. for each $i \in \{1, \dots, n\}$ define $U_i := \bigcap_{J \in T : i \in J} J \in T$. Those U_i are minimal elements in T containing i, and hence their unions covers $\{1, \dots, n\}$. Further for $i \neq k$ the sets U_i and U_k are either identical or disjoint, otherwise intersecting them would create smaller non-empty elements in T. Let $J_1, \dots, J_N$ be the partition induced by the sets U_i, i.e. for all $k = 1, \dots, N$ the set J_k is given by some U_i and $J_1, \dots, J_N$ is a partition. We claim that this defines the finest partition that factorizes X. First let $I_1, \dots, I_M$ induce a factorization of X. Let $1 \leq k \leq M$. Then I_k, I_k^c induces a partition because $I_1, \dots, I_M$ already induces a partition. That means $I_k \in T$ and since the U_i build a basis we have $I_k = \bigcup_{i \in J_k} U_i$. It remains to show that $J_1, \dots, J_N$ defines a partition. For each $1 \leq k \leq N$ there exist sets X_k (and X_k') such that

$$X = \{x \in \mathbb{R}^n : P_{J_k}(x) \in X_k, P_{J_k^c}(x) \in X_k'\}. \tag{57}$$

We claim $X = \{x \in \mathbb{R}^n : P_{J_i}(x) \in X_i \text{ for } i = 1, \dots, N\}$. It suffices to show that $\{x \in \mathbb{R}^n : P_{J_i}(x) \in X_i \text{ for } i = 1, \dots, N\} \subset X$. Therefore let $x \in \mathbb{R}^n$ such that $P_{J_i}(x) \in X_i$. From $J_2 \in T$ it follows from $P_{J_2}(x) \in X_2$ that there exists a $x^2 \in X$ with $P_{J_2}(x^2) = P_{J_2}(x)$ because $J_2 \in T$. Hence it follows $P_{J_1^c}(x^2) \in X_1'$. In particular the element

$$\tilde{x}^2 = (\tilde{x}_i^2)_{i=1,\dots,n} \text{ with } \tilde{x}_i^2 = \begin{cases} x_i, & i \in J_1 \\ x_i^2, & i \in J_1^c \end{cases} \tag{58}$$

belongs to X and satisfies $\tilde{x}_i^2 = x_i$ for $i \in J_1 \cup J_2$. Now we can continue this process for the new partition $(J_1 \cup J_2), J_3, \ldots, J_N$ and find an element $\tilde{x}^3 \in X$ with $\tilde{x}_i^3 = x_i$ for $i \in J_1 \cup J_2 \cup J_3$. Continuing until we have reached J_N we find that finally $x = \tilde{x}^N \in X$. □

12 Acknowledgement

This work has been supported by European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Actions, grant agreement 813211 (POEMA), by the Czech Science Foundation (GACR) under contract No. 20-11626Y and by the AI Interdisciplinary Institute ANITI funding, through the French "Investing for the Future PIA3" program under the Grant agreement n° ANR-19-PI3A-0004.

References

[1] Blanchini, F. (1999). Set invariance in control. Automatica, 35(11):1747-1767.

[2] Chen, M., Herbert, S. L., Vashishtha, M. S., Bansal, S., & Tomlin C. J. (2001). "Decomposition of Reachable Sets and Tubes for a Class of Nonlinear Systems," in IEEE Transactions on Automatic Control, vol. 63, no. 11, pp. 3675-3688.

[3] Chesi, G. (2011). Domain of attraction; analysis and control via SOS programming. Lecture Notes in Control and Information Sciences, Vol. 415, Springer-Verlag, Berlin.

[4] Cormen T. H., Leiserson C. E., Rivest R. L. & Stein C. (2001). Introduction to Algorithms, Second Edition. MIT Press and McGraw-Hill.

[5] Dellnitz, M., & Junge, O. (2002). Set oriented numerical methods for dynamical systems, in Handbook of Dynamical Systems, Vol. 2 , North-Holland, Amsterdam, 221-264.

[6] Fantuzzi, G., & Goluskin, D. (2020). Bounding Extreme Events in Nonlinear Dynamics Using Convex Optimization. SIAM J. Appl. Dyn. Syst., 19(3), 1823–1864.

[7] Giesl, P., & Hafstein, S. (2015). Review on computational methods for Lyapunov functions. Discrete & Continuous Dynamical Systems - B, 2015, 20 (8), 2291-2331.

[8] Goluskin, D. (2018). Bounding extreme values on attractors using sum-of-squares optimization, with application to the Lorenz attractor. arXiv preprint arXiv:1807.09814.

[9] Henrion, D., Korda, M., & Lasserre J. B. "Moment-sos Hierarchy, The: Lectures In Probability, Statistics, Computational Geometry, Control And Nonlinear PDEs". Vol. 4, World Scientific, 2020.

[10] Henrion, D., & Korda, M. (2014). "Convex Computation of the Region of Attraction of Polynomial Control Systems" in IEEE Transactions on Automatic Control, vol. 59, no. 2, pp. 297-312, Feb. doi: 10.1109/TAC.2013.2283095.

[11] Korda, M. (2020). Computing controlled invariant sets from data using convex optimization. SIAM Journal on Control and Optimization, vol. 58, no. 5, pp. 2871-2899.

[12] Korda, M., Henrion, D., & Jones, J. N. (2014). Convex computation of the maximum controlled invariant set for polynomial control systems. SIAM J. CONTROL OPTIM., 52(5), pp. 2944-2969.

[13] Korda, M., Henrion, D., & Jones, C. N. (2014). Controller design and region of attraction estimation for nonlinear dynamical systems. IFAC Proceedings Volumes Volume 47, Issue 3, Pages 2310-2316.

[14] Korda, M., Henrion, D., & Jones, C. N. (2013). Inner approximations of the region of attraction for polynomial dynamical systems. IFAC Proceedings Volumes Volume 46, Issue 23, pages 534-539.

[15] Lasserre, J. B. (2001). Global optimization with polynomials and the problem of moments. SIAM Journal on optimization, 11(3), 796-817.

[16] Lasserre, J. B. (2009). Moments, positive polynomials and their applications. Imperial College Press, London, UK.

[17] Oustry, A., M. Tacchi, M., & Henrion, D. (2019). Inner approximations of the maximal positively invariant set for polynomial dynamical systems. arXiv preprint arXiv:1903.04798.

[18] Putinar, M. (1993). Positive polynomials on compact semi-algebraic sets. Indiana Univ. Mathematics Journal, 42, pp. 969-984.

[19] Robinson, J. C. (2001). Infinite-Dimensional Dynamical Systems. An Introduction to Dissipative Parabolic PDEs and the Theory of Global Attractors. Cambridge University Press.

[20] Schlosser, C., & Korda, M., (2020). Converging outer approximations to global attractors using semidefinite programming. arXiv preprint arXiv:2005.03346

[21] Tacchi, M., Cardozo, C., Henrion, D., & Lasserre, J. B. (2019). Approximating regions of attraction of a sparse polynomial differential system. arXiv preprint arXiv:1911.09500.

[22] Valmorbida, G., & Anderson, J. (2017). Region of attraction estimation using invariant sets and rational Lyapunov functions. Automatica, Elsevier, 2017, 75, pp.37-45.

[23] Waki, H., Kim, S., Kojima, M., & Muramatsu, M. (2006). "Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity". SIAM Journal on Optimization, 17(1), 218-242.

[24] Wang, J., Magron, V., Lasserre, J.B., & Mai, N.H.A. (2020). CS-TSSOS: Correlative and term sparsity for large-scale polynomial optimization. arXiv preprint arXiv:2005.02828.

[25] Zubov, V. I. (1964). Methods of A. M. Lyapunov and their application. Noordhoff, Groningen.