EXPLOITING TERM SPARSITY IN MOMENT-SOS HIERARCHY FOR DYNAMICAL SYSTEMS

JIE WANG, CORBINIAN SCHLOSSER, MILAN KORDA, AND VICTOR MAGRON

Abstract. In this paper we present term sparsity sum-of-squares (TSSOS) methods applied to several problems from dynamical systems, such as region of attraction, maximum positively invariant sets and global attractors. We combine the TSSOS algorithm of Wang, Magron and Lasserre [SIAM J. Optim., 31(1):30-58, 2021] with existing infinite dimensional linear program representations of those sets. This leads to iterative schemes in the moment-sum-of-squares hierarchy which allows less expensive computations while keeping convergence guarantees. Finally this procedure is closely related to sign symmetries of the dynamical system as was already revealed for polynomial optimization. Numerical examples demonstrate the efficiency of the approach in the presence of appropriate sparsity.

1. Introduction

The idea of translating problems from dynamical systems to infinite dimensional linear programming problems dates back to at least the work of Rubio [10] and Lewis and Vinter [7, 14] concerned with optimal control problems. More recently this idea was extended to other problems: maximum positively invariant (MPI) set [4], region of attraction (ROA) [3], reachable set [8], global attractors (GA) [12] and invariant measures [5] among others. These problems then can be solved in the spirit of [6] using a convergent sequence of finite dimensional convex optimization problems was proposed. This procedure results in a hierarchy of moment-sum-of-squares (moment-SOS) relaxations leading to a sequence of semidefinite programs (SDPs).

However, the size of these SDPs scales rapidly with the degree of the polynomials involved and the state-space dimension. As a consequence, despite being convex, these SDP relaxations may be challenging to solve even for problems of modest size. To this end, different speed-up techniques have been proposed to reduce the size of SDPs via exploiting structure of the dynamical system. Among these are symmetries (see [9] for symmetry exploitation of polynomial optimization, as well as [2] exploiting symmetries in the context of dynamical systems) or correlative sparsity as in [15] for polynomial optimization, or in [11] where a specific sparsity structure was used to decompose the SDP while preserving convergence guarantees. In this paper we present the use of a recent term sparsity approach [19, 20] which has been already proven useful for a wide range of polynomial optimization problems,

Date: November 18, 2021.

2020 *Mathematics Subject Classification.* 93C10, 93C15, 93C20, 90C22, 90C23, 34C11, 37C10.

Key words and phrases. moment-SOS hierarchy, term sparsity, dynamical system, region of attraction, maximum positively invariant set, global attractor, convex relaxation, semidefinite programming.

involving complex [17] or noncommutative [18] variables, and fast approximation of joint spectral radius of sparse matrices [16]. For all above problems, one is able to formulate computationally cheaper hierarchies with still good convergence properties.

Whereas the approach of [11] is concerned with the sparsity among the state variables themselves, the approach proposed here exploits sparsity in the *algebraic description* of the dynamics, in particular among the monomial terms appearing in the components of the polynomial vector field. The method proceeds by searching non-negativity certificates comprised of polynomials with specific sets of terms only which in turn are enlarged in an iterative scheme. From an operator-theoretic perspective, the proposed term sparsity approach exploits term sparsity of the data and dynamics by algebraic (or graph theoretic) properties of the Liouville-operator associated to the dynamics. Interestingly, the proposed approach intrinsically comprises the sign symmetry reduction.

We demonstrate the approach on a number of examples (including randomly generated instances and a 16-state fluid mechanics example) and observe a very promising trade-off between speed-up and the solution accuracy. The results also confirm our theoretical analysis that this iterative procedure automatically retrieves the sign symmetry reduction. A Julia implementation of our approach is freely available online[1].

The rest of this paper is organized as follows. In Section 2, we introduce the notation and give some preliminaries. In Section 3, we show how to exploit term sparsity in the moment-SOS hierarchy by taking the computation of MPI sets as an example and reveal its relation with the sign symmetry reduction. Section 4 illustrates the approach by numerical examples. Conclusions are given in Section 5.

2. Notation and preliminaries

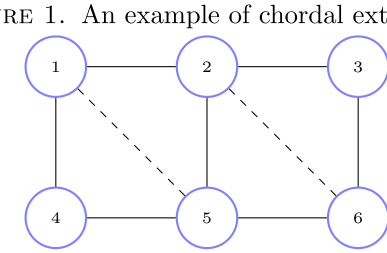
Let $\mathbf{x} = (x_1, \dots, x_n)$ be a tuple of variables and $\mathbb{R}[\mathbf{x}] = \mathbb{R}[x_1, \dots, x_n]$ be the ring of real n-variate polynomials. For $d \in \mathbb{N}$, the subset of polynomials in $\mathbb{R}[\mathbf{x}]$ of degree no more than $2d$ is denoted by $\mathbb{R}_{2d}[\mathbf{x}]$. A polynomial $f \in \mathbb{R}[\mathbf{x}]$ can be written as $f(\mathbf{x}) = \sum_{\boldsymbol{\alpha}\in\mathscr{A}} f_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}}$ with $\mathscr{A} \subseteq \mathbb{N}^n$, $f_{\boldsymbol{\alpha}} \in \mathbb{R}$ and $\mathbf{x}^{\boldsymbol{\alpha}} = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. The support of f is then defined by $\operatorname{supp}(f) := \{\boldsymbol{\alpha} \in \mathscr{A} \mid f_{\boldsymbol{\alpha}} \neq 0\}$. For $\mathscr{A} \subseteq \mathbb{N}^n$, let $\mathbb{R}[\mathscr{A}]$ be the set of polynomials whose supports are contained in $\mathscr{A}$. The notation $Q \succeq 0$ for a matrix Q indicates that Q is positive semidefinite (PSD). For a positive integer r, the set of $r \times r$ symmetric matrices is denoted by $\mathbf{S}^r$ and the set of $r \times r$ PSD matrices is denoted by $\mathbf{S}^r_+$. For $d \in \mathbb{N}$, let $\mathbb{N}^n_d := \{\boldsymbol{\alpha} = (\alpha_i)_{i=1}^n \in \mathbb{N}^n \mid \sum_{i=1}^n \alpha_i \leq d\}$. For $\boldsymbol{\alpha} \in \mathbb{N}^n, \mathscr{A}, \mathscr{B} \subseteq \mathbb{N}^n$, let $\boldsymbol{\alpha} + \mathscr{B} := \{\boldsymbol{\alpha} + \boldsymbol{\beta} \mid \boldsymbol{\beta} \in \mathscr{B}\}$ and $\mathscr{A} + \mathscr{B} := \{\boldsymbol{\alpha} + \boldsymbol{\beta} \mid \boldsymbol{\alpha} \in \mathscr{A}, \boldsymbol{\beta} \in \mathscr{B}\}$. We use $|\cdot|$ to denote the cardinality of a set. For two vectors $\mathbf{a} = (a_i)_{i=1}^n$ and $\mathbf{b} = (b_i)_{i=1}^n$, let $\mathbf{a} \cdot \mathbf{b} := \sum_{i=1}^n a_i b_i$ and $\mathbf{a} \circ \mathbf{b} := (a_1 b_1, \dots, a_n b_n)$.

Given a polynomial $f(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$, if there exist polynomials $f_1(\mathbf{x}), \dots, f_t(\mathbf{x})$ such that $f(\mathbf{x}) = \sum_{i=1}^t f_i(\mathbf{x})^2$, then we call $f(\mathbf{x})$ a *sum of squares (SOS)* polynomial. The set of SOS polynomials is denoted by $\Sigma[\mathbf{x}]$. Assume that $f \in \Sigma_{2d}[\mathbf{x}] := \Sigma[\mathbf{x}] \cap \mathbb{R}_{2d}[\mathbf{x}]$ and $\mathbf{x}^{\mathbb{N}^n_d}$ is the $\binom{n+d}{d}$-dimensional column vector consisting of elements $\mathbf{x}^{\boldsymbol{\alpha}}, \boldsymbol{\alpha} \in \mathbb{N}^n_d$ (fix any ordering on $\mathbb{N}^n$). Then f is an SOS polynomial if and only if there exists a PSD matrix Q (called a Gram matrix) such that $f = (\mathbf{x}^{\mathbb{N}^n_d})^T Q \mathbf{x}^{\mathbb{N}^n_d}$. For convenience, we abuse notation in the sequel and denote by $\mathbb{N}^n_d$ instead of $\mathbf{x}^{\mathbb{N}^n_d}$ the standard monomial basis and use the exponent $\boldsymbol{\alpha}$ to represent a monomial $\mathbf{x}^{\boldsymbol{\alpha}}$.

[1] `https://github.com/wangjie212/SparseDynamicSystem`.

An *(undirected) graph* $G(V,E)$ or simply G consists of a set of nodes V and a set of edges $E \subseteq \{\{u,v\} \mid u \neq v, (u,v) \in V \times V\}$. For a graph G, we use $V(G)$ and $E(G)$ to indicate the node set of G and the edge set of G, respectively. For two graphs G, H, we say that G is a *subgraph* of H, denoted by $G \subseteq H$, if both $V(G) \subseteq V(H)$ and $E(G) \subseteq E(H)$ hold. A graph is called a *chordal graph* if all its cycles of length at least four have a chord[2]. The notion of chordal graphs plays an important role in sparse matrix theory. Any non-chordal graph $G(V,E)$ can be always extended to a chordal graph $G'(V,E')$ by adding appropriate edges to E, which is called a *chordal extension* of $G(V,E)$. As an example, in Figure 1 the two dashed edges are added to obtain a chordal extension. The chordal extension of G is usually not unique and the symbol G' is used to represent any specific chordal extension of G throughout the paper. For a graph G, there is a particular chordal extension which makes every connected component of G to be a complete subgraph, which is called the *maximal* chordal extension. Typically, we consider only chordal extensions that are subgraphs of the maximal chordal extension. For graphs $G \subseteq H$, we assume that $G' \subseteq H'$ holds throughout the paper.

FIGURE 1. An example of chordal extensions



Given a graph $G(V,E)$, a symmetric matrix Q with rows and columns indexed by V is said to have sparsity pattern G if $Q_{uv} = Q_{vu} = 0$ whenever $u \neq v$ and $\{u,v\} \notin E$. Let $\mathbf{S}_G$ be the set of symmetric matrices with sparsity pattern G. The PSD matrices with sparsity pattern G form a convex cone $\mathbf{S}_+^{|V|} \cap \mathbf{S}_G = \{Q \in \mathbf{S}_G \mid Q \succeq 0\}$. When the sparsity pattern graph $G(V,E)$ is a chordal graph, the cone $\mathbf{S}_+^{|V|} \cap \mathbf{S}_G$ can be decomposed as a sum of simple convex cones by virtue of the following theorem and hence the related optimization problem can be solved more efficiently. Recall that a *clique* of a graph is a subset of nodes that induces a complete subgraph. A *maximal clique* is a clique that is not contained in any other clique.

Theorem 2.1 ([1], Theorem 2.3)**.** *Let $G(V,E)$ be a chordal graph and assume that $C_1, \ldots, C_l$ are the list of maximal cliques of $G(V,E)$. Then a matrix $Q \in \mathbf{S}_+^{|V|} \cap \mathbf{S}_G$ if and only if Q can be written as $Q = \sum_{i=1}^{l} Q_i$, where $Q_i \in \mathbf{S}_+^{|V|}$ has nonzero entries only with row and column indices coming from C_i for $i = 1, \ldots, l$.*

Given a graph G with $V = \mathbb{N}_d^n$, let $\Sigma[G]$ be the set of SOS polynomials that admit a PSD Gram matrix with sparsity pattern G, i.e., $\Sigma[G] := \{(\mathbf{x}^{\mathbb{N}_d^n})^T Q \mathbf{x}^{\mathbb{N}_d^n} \mid Q \in \mathbf{S}_+^{|V|} \cap \mathbf{S}_G\}$.

[2]A chord is an edge that joins two nonconsecutive nodes in a cycle.

3. Exploiting term sparsity

In this section, we propose an iterative procedure to exploit term sparsity for the moment-SOS hierarchy of certain computational problems related to dynamical systems. The intuition behind this procedure is the following: starting with a minimal initial support set, we expand the support set that is taken into account in the SOS relaxation by iteratively performing chordal extension to the related sparsity pattern graphs inspired by Theorem 2.1. For the ease of understanding, we illustrate the approach by considering the computation of MPI sets. But there is no difficulty to extend the approach to other situations, e.g., the computations of ROA [3] and GA [12], bounding extreme events [2].

Suppose that the dynamical system we are considering is given by

$$\begin{cases} \dot{x}_1 &= f_1(\mathbf{x}), \\ \dot{x}_2 &= f_2(\mathbf{x}), \\ &\vdots \\ \dot{x}_n &= f_n(\mathbf{x}), \end{cases} \tag{3.1}$$

where $\mathbf{f} \coloneqq \{f_1, \ldots, f_n\} \subseteq \mathbb{R}[\mathbf{x}]$. Moreover, let the constraint set

$$X \coloneqq \{\mathbf{x} \in \mathbb{R}^n \mid p_j(\mathbf{x}) \geq 0 \text{ for } j = 1, \ldots, m\} \tag{3.2}$$

with $p_1, \ldots, p_m \in \mathbb{R}[\mathbf{x}]$. For the sake of convenience, we set $p_0 \coloneqq 1$. Let $d_f \coloneqq \max\{\deg(f_i) : i = 1, \ldots, n\}$ and $d_j \coloneqq \deg(p_j)$ for $j = 0, 1, \ldots, m$, $d_p \coloneqq \max\{d_j : j = 1, \ldots, m\}$.

Definition 3.1. *For a dynamical system the* maximum positively invariant (MPI) set *is the set of initial conditions* $\mathbf{x}_0$ *such that the solutions* $\varphi_t(\mathbf{x}_0)$ *stay in* X *for all* $t \in \mathbb{R}_+$.

As proposed in [4], given a positive integer d, the d-th order SOS relaxation for approximating the MPI set is defined by:

$$\theta_d \coloneqq \begin{cases} \inf & \int_X w(\mathbf{x})\, \mathrm{d}\mathbf{x} \\ \text{s.t.} & v \in \mathbb{R}[\mathbf{x}]_{2d+1-d_f}, w \in \mathbb{R}[\mathbf{x}]_{2d}, \\ & \beta v - \nabla v \cdot \mathbf{f} = a_0 + \sum_{j=1}^m a_j p_j, \\ & w = b_0 + \sum_{j=1}^m b_j p_j, \\ & w - v - 1 = c_0 + \sum_{j=1}^m c_j p_j, \end{cases} \tag{3.3}$$

where $\beta > 0$ is a preassigned discount factor (say, $\beta = 1$), ∇ is the gradient with respect to $\mathbf{x}$, and $a_j, b_j, c_j \in \Sigma_{2d-d_j}[\mathbf{x}]$ for $j = 0, 1, \ldots, m$. The dynamics enter through the discounted Liouville operator $v \mapsto \beta v - \nabla v \cdot \mathbf{f}$. By [4], the sequence of optima of (3.3) converges monotonically from above to the volume of the MPI set provided the polynomials $(p_j)_{j=1}^m$ satisfy the Archimedianity condition (e.g., one of the p_j's is a, possibly redundant, ball constraint). Furthermore, the set

$$S_d \coloneqq w^{-1}([1, +\infty]) = \{\mathbf{x} \in X : w(\mathbf{x}) \geq 1\}$$

provides an outer approximation for the MPI set.

For a graph $G(V, E)$ with $V \subseteq \mathbb{N}^n$, define $\mathrm{supp}(G) \coloneqq \{\boldsymbol{\beta} + \boldsymbol{\gamma} \mid \{\boldsymbol{\beta}, \boldsymbol{\gamma}\} \in E\}$. Now we give the iterative procedure to exploit term sparsity. Fix a relaxation order d.

Let $\mathscr{A} = \bigcup_{j=1}^{m} \operatorname{supp}(p_j)$ and v be a polynomial with generic coefficients supported on $\mathscr{A}$. Let

$$\mathscr{A}_d^1 := \mathscr{A} \cup \operatorname{supp}(\nabla v \cdot \mathbf{f}) \cup 2\mathbb{N}_d^n \tag{3.4}$$

with $2\mathbb{N}_d^n := \{2\boldsymbol{\alpha} \mid \boldsymbol{\alpha} \in \mathbb{N}_d^n\}$. Intuitively, the set $\mathscr{A}_d^1$ is the minimal support that has to be involved in the SOS relaxation (3.3)[3]. Next, we expand the support set by iteratively performing chordal extension on some related graphs. More formally, for every integer $s \geq 1$, we iteratively define the graph $G_{d,j}^s$ which will be imposed as the sparsity pattern graph for a Gram matrix of a_j with $V(G_{d,j}^s) := \mathbb{N}_{d-d_j}^n$ and

$$E(G_{d,j}^s) := \{\{\boldsymbol{\beta}, \boldsymbol{\gamma}\} \mid \boldsymbol{\beta} + \boldsymbol{\gamma} + \operatorname{supp}(p_j) \cap (\mathscr{A}_d^s \cup \operatorname{supp}(\nabla v_d^s \cdot \mathbf{f})) \neq \emptyset\},$$

where v_d^s is a polynomial with generic coefficients supported on $\mathscr{A}_d^s \cap \mathbb{N}_{2d+1-d_f}^n$ for $j = 0, 1, \ldots, m$, and further let $\mathscr{A}_d^{s+1} = \operatorname{supp}((G_{d,0}^s)')$. In doing so, we get a finite ascending chain of support sets

$$\mathscr{A}_d^1 \subseteq \cdots \subseteq \mathscr{A}_d^{\tilde{s}} = \mathscr{A}_d^{\tilde{s}+1} = \cdots$$

and a finite ascending chain of graphs

$$(G_{d,j}^1)' \subseteq \cdots \subseteq (G_{d,j}^{\tilde{s}})' = (G_{d,j}^{\tilde{s}+1})' = \cdots$$

for each $j = 0, 1, \ldots, m$.

For a given $s \geq 1$, with $\mathscr{B}_d^{s,1} := \mathscr{A}_d^s$ for every $l \geq 1$ we also iteratively define the graph $H_{d,j}^{s,l}$ which will be imposed as the sparsity pattern graph for a Gram matrix of b_j or c_j with $V(H_{d,j}^{s,l}) := \mathbb{N}_{d-d_j}^n$ and

$$E(H_{d,j}^{s,l}) := \{\{\boldsymbol{\beta}, \boldsymbol{\gamma}\} \mid \boldsymbol{\beta} + \boldsymbol{\gamma} + \operatorname{supp}(p_j) \cap \mathscr{B}_d^{s,l} \neq \emptyset\}$$

for $j = 0, 1, \ldots, m$, and further let

$$\mathscr{B}_d^{s,l+1} = \bigcup_{j=0}^{m} (\operatorname{supp}(p_j) + \operatorname{supp}((H_{d,j}^{s,l})')). \tag{3.5}$$

In doing so, we get a finite ascending chain of support sets

$$\mathscr{B}_d^{s,1} \subseteq \cdots \subseteq \mathscr{B}_d^{s,\tilde{l}} = \mathscr{B}_d^{s,\tilde{l}+1} = \cdots$$

and a finite ascending chain of graphs

$$(H_{d,j}^{s,1})' \subseteq \cdots \subseteq (H_{d,j}^{s,\tilde{l}})' = (H_{d,j}^{s,\tilde{l}+1})' = \cdots$$

for each $j = 0, 1, \ldots, m$.

The two indices s and l are used to control the size of support sets of v respectively w. For a pair $s, l \geq 1$, we may thereby consider the following sparse SOS relaxation for approximating the MPI set:

$$\theta_d^{s,l} := \begin{cases} \inf & \int_X w(\mathbf{x})\,\mathrm{d}\mathbf{x} \\ \text{s.t.} & v \in \mathbb{R}[\mathscr{A}_d^s \cap \mathbb{N}_{2d+1-d_f}^n], w \in \mathbb{R}[\mathscr{B}_d^{s,l}], \\ & \beta v - \nabla v \cdot \mathbf{f} = a_0 + \sum_{j=1}^m a_j p_j, \\ & w = b_0 + \sum_{j=1}^m b_j p_j, \\ & w - v - 1 = c_0 + \sum_{j=1}^m c_j p_j, \end{cases} \tag{3.6}$$

[3] Here the subset $2\mathbb{N}_d^n$ is included in the definition of $\mathscr{A}_d^1$ to guarantee convergence; see [20].

where $a_j \in \Sigma[(G^s_{d,j})']$, $b_j, c_j \in \Sigma[(H^{s,l}_{d,j})']$ for $j = 0, 1, \ldots, m$. Notice that the Gram matrix of any SOS involved in (3.6) admits a block decomposition because of Theorem 2.1. Hence the corresponding SDP could be easier to solve.

Proposition 3.2. *With the above notations, $\theta^{s,l}_{d+1} \le \theta^{s,l}_d$, $\theta^{s,l}_d \le \theta^{s+1,l}_d$, $\theta^{s,l}_d \le \theta^{s,l+1}_d$, and $\theta^{s,l}_d \ge \theta_d$ hold true for $d \ge \max\{\lceil d_f/2\rceil, \lceil d_p/2\rceil\}, s \ge 1, l \ge 1$.*

Proof. The first three inequalities follow the fact that we always take more supports into account in the corresponding sparse SOS relaxation when increasing d, s or l. The last inequality $\theta^{s,l}_d \ge \theta_d$ holds because the feasible set of the sparse SOS relaxation is a subset of the corresponding dense SOS relaxation. □

Sign symmetry. The *sign symmetries* of the system (3.1) with the constraint set X (3.2) consist of all vectors $\mathbf{r} = (r_i) \in \mathbb{Z}_2^n := \{0, 1\}^n$ that satisfy

$$f_i((-1)^{\mathbf{r}} \circ \mathbf{x}) = (-1)^{r_i} f_i(\mathbf{x}) \text{ for } i = 1, \ldots, n,$$

and

$$p_j((-1)^{\mathbf{r}} \circ \mathbf{x}) = p_j(x_1, \ldots, x_n) \text{ for } j = 1, \ldots, m,$$

where $(-1)^{\mathbf{r}} := ((-1)^{r_1}, \ldots, (-1)^{r_n})$. Given a set of sign symmetries $R \subseteq \mathbb{Z}_2^n$, we define $R^\perp := \{\boldsymbol{\alpha} \in \mathbb{N}^n \mid \mathbf{r} \cdot \boldsymbol{\alpha} \equiv 0 \,(\text{mod } 2), \forall \mathbf{r} \in R\}$. Then a polynomial g is invariant under the sign symmetries R[4] if and only if $\text{supp}(g) \subseteq R^\perp$. The following theorem[5] tells us that the SOS relaxation (3.3) inherits the sign symmetries of the dynamical system (3.1).

Theorem 3.3. *Let R be the set of sign symmetries of the system* (3.1) *with the constraint set X* (3.2)*. If we additionally impose the constraints that* $\text{supp}(v), \text{supp}(w) \subseteq R^\perp$ *and* $\text{supp}(a_j), \text{supp}(b_j), \text{supp}(c_j) \subseteq R^\perp$ *for $j = 0, 1, \ldots, m$ in* (3.3)*, the resulting program has the same optimum with* (3.3)*.*

Proof. Suppose that $v, w, \{a_j\}_{j=0}^m, \{b_j\}_{j=0}^m, \{c_j\}_{j=0}^m$ are an optimal solution to (3.3). We remove the terms of v, w with exponents not belonging to $R^\perp$ from the expression of v, w and denote the resulting polynomials by $\tilde{v}, \tilde{w}$ respectively. Let Q_j be a PSD Gram matrix of a_j for any j such that $a_j = (\mathbf{x}^{\mathbb{N}^n_{d-d_j}})^T Q_j \mathbf{x}^{\mathbb{N}^n_{d-d_j}}$. We then define $\widetilde{Q}_j \in \mathbf{S}^{|\mathbb{N}^n_{d-d_j}|}$ by

$$[\widetilde{Q}_j]_{\boldsymbol{\beta\gamma}} := \begin{cases} [Q_j]_{\boldsymbol{\beta\gamma}}, & \text{if } \boldsymbol{\beta} + \boldsymbol{\gamma} \in R^\perp, \\ 0, & \text{otherwise,} \end{cases}$$

and let $\tilde{a}_j := (\mathbf{x}^{\mathbb{N}^n_{d-d_j}})^T \widetilde{Q}_j \mathbf{x}^{\mathbb{N}^n_{d-d_j}}$. One can easily check that $\widetilde{Q}_j$ is block diagonal (after an appropriate permutation on rows and columns). So $Q_j \succeq 0$ implies $\widetilde{Q}_j \succeq 0$ and it follows that $\tilde{a}_j$ is an SOS polynomial. In a similar way, we define $\tilde{b}_j$,$\tilde{c}_j$ for $j = 0, 1, \ldots, m$ which are all SOS polynomials by a similar argument as for $\tilde{a}_j$. As we remove exactly the terms with exponents not belonging to $R^\perp$ from both sides of the equations in (3.3), $\tilde{v}, \tilde{w}, \{\tilde{a}_j\}_{j=0}^m, \{\tilde{b}_j\}_{j=0}^m, \{\tilde{c}_j\}_{j=0}^m$ are again a feasible solution

[4]That is, $g((-1)^{\mathbf{r}} \circ \mathbf{x}) = g(\mathbf{x})$ for any $\mathbf{r} \in R$.

[5]A similar symmetry reduction already appeared in [2] in the study of bounding extreme events in dynamical systems.

to (3.3). It remains to show $\int_X \tilde{w}(\mathbf{x})\,\mathrm{d}\lambda = \int_X w(\mathbf{x})\,\mathrm{d}\lambda$. Take any $\boldsymbol{\alpha} \in \mathbb{N}^n_{2d}\backslash R^{\perp}$. Then there exists $\mathbf{r} = (r_i)_i \in R$ such that $\boldsymbol{\alpha} \cdot \mathbf{r} \not\equiv 0 \,(\mathrm{mod}\ 2)$. We have

$$\int_X \mathbf{x}^{\boldsymbol{\alpha}}\,\mathrm{d}\lambda = \int_X ((-1)^{\mathbf{r}} \circ \mathbf{x})^{\boldsymbol{\alpha}}\,\mathrm{d}\lambda = \int_X (-1)^{\boldsymbol{\alpha}\cdot\mathbf{r}}\mathbf{x}^{\boldsymbol{\alpha}}\,\mathrm{d}\lambda = -\int_X \mathbf{x}^{\boldsymbol{\alpha}}\,\mathrm{d}\lambda,$$

where the first equality follows from the fact that X is invariant under the sign symmetry $\mathbf{r}$. This immediately gives $\int_X \mathbf{x}^{\boldsymbol{\alpha}}\,\mathrm{d}\lambda = 0$ from which we deduce $\int_X \tilde{w}(\mathbf{x})\,\mathrm{d}\lambda = \int_X w(\mathbf{x})\,\mathrm{d}\lambda$ as desired. □

From the proof of Theorem 3.3, we see that the sign symmetries of the dynamical system (3.1) endow any SOS polynomial involved in (3.3) with a block structure. Our iterative procedure to exploit term sparsity actually produces block structures that are compatible with the sign symmetries of the dynamical system. Furthermore, when maximal chordal extensions are used in the construction, the block structures converge to the one given by the sign symmetries of the system. The key observation is the following lemma.

Lemma 3.4. *For any relaxation order $d \geq \max\{\lceil d_f/2\rceil, \lceil d_p/2\rceil\}$, the sign symmetries of the system (3.1) with the constraint set X (3.2) coincide with the sign symmetries of $\mathscr{A}^1_d$ (3.4), i.e., the set $\{\mathbf{r} \in \mathbb{Z}^n_2 \mid \mathbf{r}\cdot\boldsymbol{\alpha} \equiv 0 \ mod\ 2, \forall\, \boldsymbol{\alpha} \in \mathscr{A}^1_d\}$.*

Proof. Let us denote the set of sign symmetries of the system (3.1) by R and the set of sign symmetries of $\mathscr{A}^1_d$ by R'. For any $\mathbf{r} \in R$, we wish to show $\mathbf{r} \in R'$. It suffices to show $\mathbf{r}\cdot\boldsymbol{\alpha} \equiv 0 \bmod 2$ for any $\boldsymbol{\alpha} \in \mathscr{A}^1_d = \mathscr{A} \cup \mathrm{supp}(\nabla v \cdot \mathbf{f}) \cup 2\mathbb{N}^n_d$ where $\mathscr{A} = \bigcup_{j=1}^m \mathrm{supp}(p_j)$ and v is a polynomial with generic coefficients supported on $\mathscr{A}$. If $\boldsymbol{\alpha} \in \mathscr{A} \cup 2\mathbb{N}^n_d$, we clearly have $\mathbf{r}\cdot\boldsymbol{\alpha} \equiv 0 \bmod 2$ by definition. Noting $\nabla v(\mathbf{x}) = \nabla v((-1)^{\mathbf{r}} \circ \mathbf{x}) = (-1)^{\mathbf{r}} \circ (\nabla v)((-1)^{\mathbf{r}} \circ \mathbf{x})$, we have $(\nabla v \cdot \mathbf{f})((-1)^{\mathbf{r}} \circ \mathbf{x}) = (\nabla v)((-1)^{\mathbf{r}} \circ \mathbf{x}) \cdot \mathbf{f}((-1)^{\mathbf{r}} \circ \mathbf{x}) = ((-1)^{\mathbf{r}} \circ \nabla v(\mathbf{x})) \cdot ((-1)^{\mathbf{r}} \circ \mathbf{f}(\mathbf{x})) = (\nabla v \cdot \mathbf{f})(\mathbf{x})$. As a result, $\mathbf{r}\cdot\boldsymbol{\alpha} \equiv 0 \bmod 2$ for any $\boldsymbol{\alpha} \in \mathrm{supp}(\nabla v \cdot f)$.

Conversely, for any $\mathbf{r} \in R'$, we wish to show $\mathbf{r} \in R$. Since $\mathscr{A} \subseteq \mathscr{A}^1_d$, we have $p_j((-1)^{\mathbf{r}} \circ \mathbf{x}) = p_j(\mathbf{x})$ for $j = 1, \ldots, m$. Let $\boldsymbol{\alpha} = (\alpha_i) \in \mathscr{A}$ with $\alpha_1 > 0$. Then $\mathrm{supp}(\mathbf{x}^{\boldsymbol{\alpha}} f_1/x_1) \subseteq \mathscr{A}^1_d$. The condition $\mathbf{r} \in R'$ implies $(-1)^{\mathbf{r}\cdot\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} f_1((-1)^{\mathbf{r}} \circ \mathbf{x})/((-1)^{r_1}x_1) = \mathbf{x}^{\boldsymbol{\alpha}} f_1/x_1$, which gives $f_1((-1)^{\mathbf{r}} \circ \mathbf{x}) = (-1)^{r_1} f_1(\mathbf{x})$ as $(-1)^{\mathbf{r}\cdot\boldsymbol{\alpha}} = 1$. Similarly, we can prove $f_i((-1)^{\mathbf{r}} \circ \mathbf{x}) = (-1)^{r_i} f_i(\mathbf{x})$ for $i = 2, \ldots, n$. Thus $\mathbf{r} \in R$. □

Based on Lemma 3.4, we can prove the following theorem by a similar argument as for [20, Theorem 6.5] and so we omit the proof.

Theorem 3.5. *For any relaxation order $d \geq \max\{\lceil d_f/2\rceil, \lceil d_p/2\rceil\}$, if maximal chordal extensions are used in the construction, then the block structures produced by the above iterative procedure converge to the one given by the sign symmetries of the system (3.1) (with the constraint set X (3.2)) as s, l increase. As a corollary, in this case $\theta^{s,l}_d = \theta_d$ when s, l are sufficiently large*[6].

Remark 3.6. *The flexibility of choosing chordal extensions in the construction allows one to balance between the computation cost and the quality of approximation.*

---

[6]The values of s, l for $\theta^{s,l}_d = \theta_d$ to be valid are a priori unknown, but they are typically no greater than 3 in practice.

4. Illustrative examples

In this section, we give two numerical examples to illustrate the iterative procedure to exploit term sparsity. The procedure has been implemented as a Julia package `SparseDynamicSystem` which is available at:

`https://github.com/wangjie212/SparseDynamicSystem`.

All numerical examples were computed on an Intel Core i5-8265U@1.60GHz CPU with 8GB RAM memory and `Mosek` 9.0 is used as an SDP solver.

4.1. **Randomly generated models ([13]).** We consider the following sparse dynamical system for varying n:

$$\dot{x}_i = (\mathbf{x}^T B\mathbf{x} - 1)x_i, \quad i = 1, \ldots, n, \tag{4.1}$$

where $B \in \mathbf{S}_G$ is a random positive definite matrix satisfying:

(1) G is a random graph with n nodes and $n-4$ edges;
(2) For $1 \le i \le n$, $B_{ii} \in [1, 2]$ and for $1 \le i < j \le n$, $B_{ij} \in [-0.5, 0.5]$.

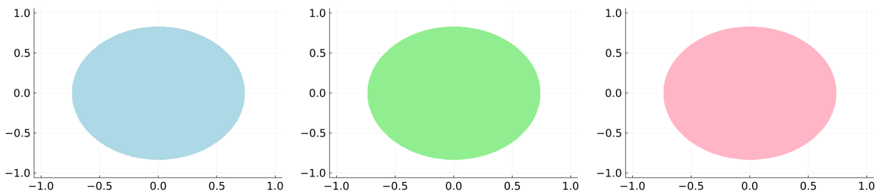
The constraint set is $X = \{\mathbf{x} \in \mathbb{R}^n \mid 1 - x_i^2 \ge 0 \text{ for } i = 1, \ldots, n\}$. For each system, we approximate the MPI set via solving the dense relaxation (3.3) and the sparse relaxations (3.6) with $s = 1, 2$[7] respectively. Figures 2a–2d show outer approximations of the MPI set for $n = 6, 8, 10, 12$ respectively, where TS (term sparsity) corresponds to the sparse relaxation with $s = 1$, SS (sign symmetry) corresponds to the sparse relaxation with $s = 2$[8], and FD (fully dense) corresponds to the dense relaxation. In Table 1, we list the optima and the running time for solving the corresponding SDPs. From Figures 2a-2d and Table 1, we can conclude that SS is significantly faster than FD by one or two orders of magnitude without sacrificing any accuracy; TS is several times faster than SS at the cost of possibly providing slightly weaker approximations.

Table 1. The results for randomly generated models, TS: term sparsity, SS: sign symmetry, FD: fully dense, opt: optimum, time: running time in seconds.

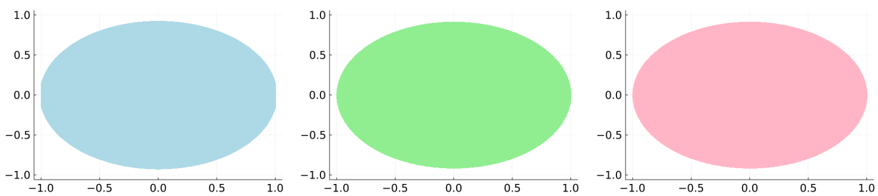
n	$2d$	TS		SS		FD	
		opt	time	opt	time	opt	time
6	10	6.60	6.41	6.23	26.7	6.23	2633
	12	5.69	159	5.05	289	-	-
8	8	15.0	34.5	14.5	61.8	14.5	3288
	10	12.4	486	11.8	1861	-	-
10	6	93.3	7.21	93.3	15.6	93.3	518
	8	30.7	899	29.8	1368	-	-
12	4	510	0.39	438	1.46	438	8.84
	6	312	125	311	288	-	-

[7]The value of l can be either 1 or 2, depending on problems.

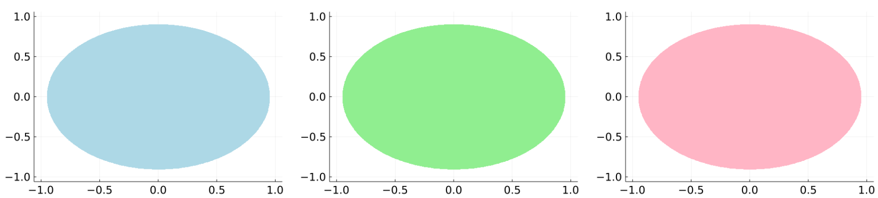
[8]Namely, the block structures of the sparse hierarchy converge to the one determined by the sign symmetries of the system when $s = 2$.



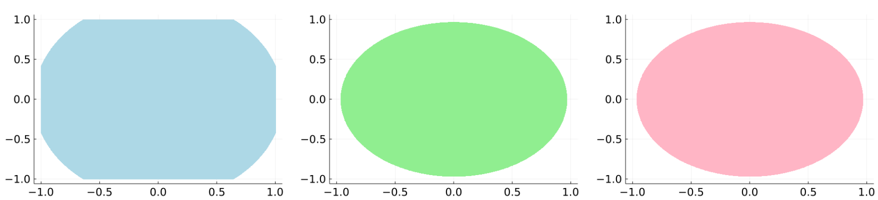
(A) $n = 6, 2d = 10$. The projection shown is for $x_i = 0, i = 3, \ldots, 6$.



(B) $n = 8, 2d = 8$. The projection shown is for $x_i = 0, i = 3, \ldots, 8$.



(C) $n = 10, 2d = 6$. The projection shown is for $x_i = 0, i = 3, \ldots, 10$.



(D) $n = 12, 2d = 4$. The projection shown is for $x_i = 0, i = 3, \ldots, 12$.

FIGURE 2. Outer approximations for the MPI set of randomly generated models. Term sparsity ● Sign symmetries ● Fully dense ●

4.2. **The 16 mode fluid model ([2, Example 4.2]).** This model is given by the dynamical system:

$$\dot{x}_n = -(2\pi n)^2 x_n + \sqrt{2}\pi n \left(\sum_{i=1}^{16-n} x_i x_{i+n} - \frac{1}{2} \sum_{i=1}^{n-1} x_i x_{n-i} \right), n = 1, \ldots, 16. \tag{4.2}$$

Let $\Phi(\mathbf{x}) = 2\pi^2 \sum_{i=1}^{16} i^2 x_i^2$, $X_0 = \{\mathbf{x} \in \mathbb{R}^{16} \mid \Phi(\mathbf{x}) = \Phi_0\}$ and $X = \{\mathbf{x} \in \mathbb{R}^{16} \mid ||\mathbf{x}||_2^2 \leq \Phi_0/(2\pi^2)\}$, where $\Phi_0 \in \mathbb{R}$.

Let Φ_∞^* denote the largest value attained by $\Phi(\mathbf{x}(t; t_0, \mathbf{x}_0))$ among all trajectories that start from $X_0 \subseteq X$ and evolve forward over the time interval $[0, \infty)$. The SOS relaxations were proposed in [2] to bound Φ_∞^* from above. We can adapt the iterative procedure presented in this paper to derive sparse SOS relaxations for bounding Φ_∞^*. We solve the dense relaxation and the sparse relaxations with $2d = 4$ and with $s = 1, 2$ respectively. The results are shown in Figure 3 and Figure 4, where TS corresponds to the sparse relaxation with $s = 1$, SS corresponds to the sparse relaxation with $s = 2$, and FD corresponds to the dense relaxation. As we can see, all three approaches give almost the same upper bounds for different Φ_0; on average, TS is six times faster than SS and SS is five times faster than FD.

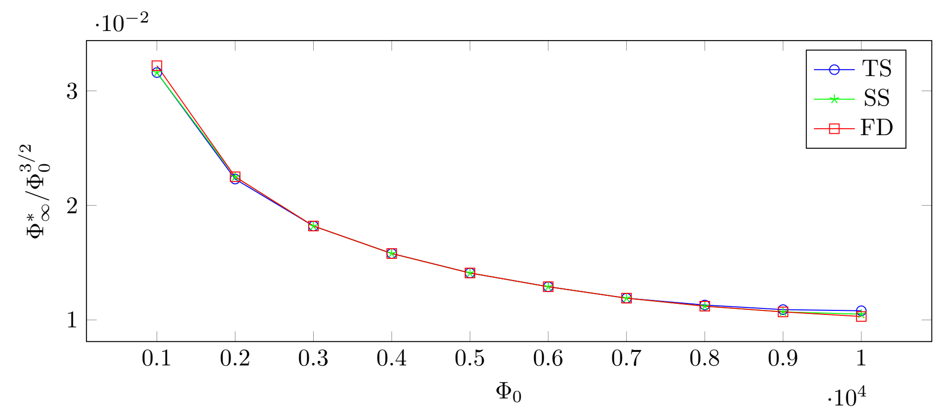


Figure 3. Upper bounds on Φ_∞^* for the 16 mode fluid model with $2d = 4$, TS: term sparsity, SS: sign symmetry, FD: fully dense.

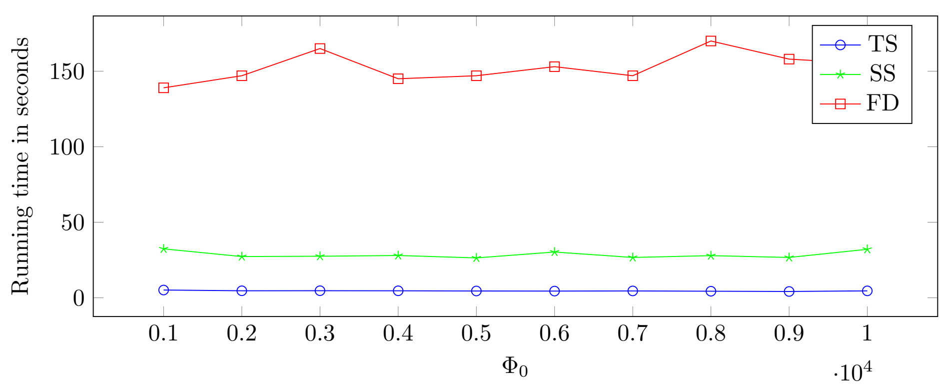


Figure 4. Running time for the 16 mode fluid model with $2d = 4$, TS: term sparsity, SS: sign symmetry, FD: fully dense.

5. Conclusions

This paper presents a reduction approach by exploiting term sparsity for the moment-SOS hierarchy of problems arising from the study of dynamical systems. As demonstrated by the numerical examples, this approach provides a trade-off between computational costs and the solution accuracy. Moreover, it is able to guarantee convergence under certain conditions and recover the sign symmetry reduction. Our next plan is to apply this approach for analyzing properties of dynamical systems, e.g., coming from power systems.

References

[1] Jim Agler, William Helton, Scott McCullough, and Leiba Rodman. Positive semidefinite matrices with a given sparsity pattern. *Linear Algebra and Its Applications*, 107:101–149, 1988.

[2] Giovanni Fantuzzi and David Goluskin. Bounding extreme events in nonlinear dynamics using convex optimization. *SIAM Journal on Applied Dynamical Systems*, 19(3):1823–1864, 2020.

[3] Didier Henrion and Milan Korda. Convex computation of the region of attraction of polynomial control systems. *IEEE Transactions on Automatic Control*, 59(2):297–312, 2014.

[4] Milan Korda, Didier Henrion, and Colin N Jones. Convex computation of the maximum controlled invariant set for polynomial control systems. *SIAM Journal on Control and Optimization*, 52(5):2944–2969, 2014.

[5] Milan Korda, Didier Henrion, and Igor Mezić. Convex computation of extremal invariant measures of nonlinear dynamical systems and markov processes. *Journal of Nonlinear Science*, 31(1):1–26, 2021.

[6] J.-B. Lasserre, D. Henrion, C. Prieur, and E. Trélat. Nonlinear optimal control via occupation measures and LMI relaxations. *SIAM Journal on Control and Optimization*, 47:1643–1666, 2008.

[7] RM Lewis and RB Vinter. Relaxation of optimal control problems to equivalent convex programs. *Journal of Mathematical Analysis and Applications*, 74(2):475–493, 1980.

[8] Victor Magron, Pierre-Loic Garoche, Didier Henrion, and Xavier Thirioux. Semidefinite approximations of reachable sets for discrete-time polynomial systems. *SIAM Journal on Control and Optimization*, 57(4):2799–2820, 2019.

[9] Cordian Riener, Thorsten Theobald, Lina Jansson Andrén, and Jean B Lasserre. Exploiting symmetries in sdp-relaxations for polynomial optimization. *Mathematics of Operations Research*, 38(1):122–141, 2013.

[10] J. E. Rubio. Generalized curves and extremal points. *SIAM Journal on Control*, 13(1):28–47, 1975.

[11] Corbinian Schlosser and Milan Korda. Sparse moment-sum-of-squares relaxations for nonlinear dynamical systems with guaranteed convergence. *arXiv preprint arXiv:2012.05572*, 2020.

[12] Corbinian Schlosser and Milan Korda. Converging outer approximations to global attractors using semidefinite programming. *Automatica*, 134, 2021.

[13] Weehong Tan and Andrew Packard. Stability region analysis using polynomial and composite polynomial lyapunov functions and sum-of-squares programming. *IEEE Transactions on Automatic Control*, 53(2):565–571, 2008.

[14] Richard B Vinter and Richard M Lewis. The equivalence of strong and weak formulations for certain problems in optimal control. *SIAM Journal on Control and Optimization*, 16(4):546–570, 1978.

[15] H. Waki, S. Kim, M. Kojima, and M. Muramatsu. Sums of Squares and Semidefinite Programming Relaxations for Polynomial Optimization Problems with Structured Sparsity. *SIAM Journal on Optimization*, 17(1):218–242, 2006.

[16] Jie Wang, Martina Maggio, and Victor Magron. SparseJSR: A fast algorithm to compute joint spectral radius via sparse SOS decompositions. In *2021 American Control Conference (ACC)*, pages 2254–2259. IEEE, 2021.

[17] Jie Wang and Victor Magron. Exploiting sparsity in complex polynomial optimization. *arXiv preprint arXiv:2103.12444*, 2021.

[18] Jie Wang and Victor Magron. Exploiting term sparsity in noncommutative polynomial optimization. *Computational Optimization and Applications*, 80(2):483–521, 2021.

[19] Jie Wang, Victor Magron, and Jean B. Lasserre. Chordal-TSSOS: a moment-SOS hierarchy that exploits term sparsity with chordal extension. *SIAM Journal on Optimization*, 31(1):114–141, 2021.

[20] Jie Wang, Victor Magron, and Jean B. Lasserre. TSSOS: A moment-SOS hierarchy that exploits term sparsity. *SIAM Journal on Optimization*, 31(1):30–58, 2021.

Jie Wang, Academy of Mathematics and Systems Science, CAS
Email address: wangjie212@amss.ac.cn
URL: https://wangjie212.github.io/jiewang

Corbinian Schlosser, Laboratory for Analysis and Architecture of Systems, CNRS
Email address: corbinian.schlosser@gmail.com

Milan Korda, Laboratory for Analysis and Architecture of Systems, CNRS
Email address: korda@laas.fr
URL: https://homepages.laas.fr/mkorda

Victor Magron, Laboratory for Analysis and Architecture of Systems, CNRS
Email address: vmagron@laas.fr
URL: https://homepages.laas.fr/vmagron