Exploiting Term Sparsity in Moment-SOS Hierarchy for Dynamical Systems

Jie Wang, Corbinian Schlosser, Milan Korda, and Victor Magron

Abstract—In this article, we develop a dynamical system counterpart to the term sparsity sum-of-squares algorithm proposed for static polynomial optimization. This allows for computational savings and improved scalability while preserving convergence guarantees when sum-of-squares methods are applied to problems from dynamical systems, including the problems of approximating region of attraction, the maximum positively invariant set, and the global attractor. At its core, the method exploits the algebraic structure of the data, thereby complementing existing methods that exploit causality relations among the states of the dynamical system. The procedure encompasses 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 this type of sparsity.

Index Terms—Convex relaxation, dynamical system, global attractor (GA), maximum positively invariant (MPI) set, moment-sum-of-squares (moment-SOS) hierarchy, region of attraction (ROA), semidefinite programming (SDP), term sparsity.

I. INTRODUCTION

The idea of translating problems from dynamical systems to infinite dimensional linear programming problems dates back to, at least, the work of Rubio [1] and Lewis and Vinter [2] concerned with optimal control problems. More recently, this idea was extended to other problems: approximations of maximum positively invariant (MPI) set [3], region of attraction (ROA) [4], reachable set [5], global attractors (GA) [6], and invariant measures [7]. These problems, then, can be solved in the spirit of [8] using a convergent sequence of finite dimensional convex optimization problems. This procedure results in a hierarchy of moment-sum-of-squares (moment-SOS) relaxations leading to a sequence of semidefinite programs (SDPs), as was done for optimal control problems in [8].

However, the size of these SDPs scales rapidly with the relaxation order and the state-space dimension. As a consequence, despite being convex, these SDP relaxations may be challenging to solve even for problems of modest state-space dimension. In the context of polynomial optimization, the similar problem of scalability has been extensively studied in recent years via exploiting structure of the system, e.g., [9] by exploiting symmetries, [10] by exploiting correlative sparsity, [11] and [12] by exploiting term sparsity. In this article, we present the use of the recent term-sparsity approach, which has already proven useful for a wide range of polynomial optimization problems, involving noncommuting variables [13], and fast approximation of joint spectral radius of sparse matrices [14]. For these problems, one is able to formulate computationally cheaper hierarchies by exploiting term sparsity with strong convergence properties.

Whereas the approaches of [15] and [16] are concerned with the sparsity in the couplings between 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 for nonnegativity certificates comprised of polynomials with only specific sets of terms, 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 with the dynamics. Interestingly, this approach intrinsically comprises the sign-symmetry reduction. In general, the term sparsity approach allows a tradeoff between the computational cost and the accuracy of approximation, whereas exploiting sign symmetries does not sacrifice any accuracy. So the term sparsity approach would provide an alternative reduction when no nontrivial sign symmetry is available or exploiting sign symmetries is still too expensive.

To summarize, the main contributions of this article are as follows.

1) We develop a term sparsity approach for SOS methods in dynamical systems, allowing for computational reduction beyond existing sparsity exploitation methods.

2) The method provides a sequence of SDPs of increasing complexity, and we show that this sequence converges in finitely many steps to the sign-symmetry reduction of the original problem.

3) We demonstrate the approach on a number of examples (including the extended Lorenz system, randomly generated instances, and a 16-state fluid mechanics example) and observe a promising tradeoff between speed-up and the solution accuracy.

The rest of this article is organized as follows. In Section II, we introduce the notation and give some preliminaries. In Section III, we show how to exploit term sparsity in the moment-SOS hierarchy by taking the computation of MPI sets as an example and revealing its relation with the sign-symmetry reduction. Section IV illustrates the approach by numerical examples. Finally, Section V concludes this article.

II. NOTATION AND PRELIMINARIES

Let $x = (x_1, \ldots, x_n)$ be a tuple of variables and $\mathbb{R}[x] = \mathbb{R}[x_1, \ldots, x_n]$ be the ring of real n-variate polynomials. For $d \in \mathbb{N}$, the subset of polynomials in $\mathbb{R}[x]$ of degree no more than $2d$ denoted by $\mathbb{R}_{2d}[x]$. A polynomial, $f \in \mathbb{R}[x]$ can be written as $f(x) = \sum_{\alpha \in \mathcal{M}} f_{\alpha} x^{\alpha}$ with $\mathcal{M} \subseteq \mathbb{N}^n$, $f_{\alpha} \in \mathbb{R}$, and $x^{\alpha} = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. The support of $f$ is then defined by $\text{supp}(f) = \{ \alpha \in \mathcal{M} \mid f_{\alpha} \neq 0 \}$. For $\mathcal{M} \subseteq \mathbb{N}^n$, let $\mathbb{R}[\mathcal{M}]$ be the set of polynomials whose supports are
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 $S^r_+$, and the set of $r \times r$ PSD matrices is denoted by $S^r_+$. For $d \in \mathbb{N}$, let $N^d_\beta := \{ \alpha = (\alpha_{ij})_{i,j=1}^d \in \mathbb{R}^d \mid \sum_{i=1}^d \alpha_i \leq d \}$. For $\alpha \in \mathbb{R}^d$, $\beta \in \mathbb{R}^d$, let $\alpha + \beta := (\alpha_i + \beta_i)_{i=1}^d$ and $\alpha \cdot \beta := (\alpha_i \beta_i)_{i=1}^d$. We use $| \cdot |$ to denote the cardinality of a set. For two vectors $a = (a_i)_{i=1}^d$ and $b = (b_i)_{i=1}^d$, let $a \cdot b := \sum_{i=1}^d a_i b_i$ and $a := (a_1, b_1, \ldots, a_n, b_n)$.

Given a polynomial $f(x) \in \mathbb{R}[x]$, if there exist polynomials $f_1(x), \ldots, f_s(x)$ such that $f(x) = \sum_{i=1}^s f_i(x)^2$, then we call $f(x)$ a sum of squares (SOS) polynomial. The set of SOS polynomials is denoted by $\Sigma[x]$. Assume that $f \in \Sigma_{2\mathbb{R}} := \Sigma[\mathbb{R}] \cap \mathbb{R}_{\geq 0}$, and $\mathbb{R}^d \times \mathbb{R}^d$ is the $(n+d)$-dimensional column vector consisting of elements $x^\top \alpha$, $\alpha \in \mathbb{R}^d$ (fix any ordering on $\mathbb{R}^d$). Then, $f$ is an SOS polynomial if and only if there exists a PSD matrix $Q$ called a Gram matrix, such that $f = (x^\top N_{\beta})^\top Q x^\top N_{\beta}$. For convenience, we abuse notation in the following and denote by $N_{\beta}^d$ instead of $x^\top N_{\beta}$ the standard monomial basis and use the exponent $\alpha$ to represent a monomial $x^\alpha$.

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 $G$ and $H$, 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. The notion of chordal graphs plays an important role in sparse matrix theory. Any nonchordal 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 $G(V,E)$. 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 article. 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 article. Given a graph $G(V,E)$, a symmetric matrix $Q$ with rows and columns indexed by $V$ is said to have sparsity graph $G$ if $Q_{uv} = Q_{vu} = 0$ whenever $u \neq v$ and $(u,v) \notin E$. Let $G_Q$ be the set of symmetric matrices with sparsity graph $G$. The PSD matrices with sparsity graph $G$ form a convex cone $S^{|V|}_+ \cap S_G = \{ Q \in S^{|V|}_+ \mid Q \succeq 0 \}$. When the sparsity graph $G(V,E)$ is a chordal graph, the cone $S^{|V|}_+ \cap S_G$ can be decomposed as a sum of simple convex cones by virtue of the following theorem. 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** ([17, Th. 2.3]): Let $G(V,E)$ be a chordal graph and assume that $C_1, \ldots, C_r$ are all maximal cliques of $G(V,E)$. Then, a matrix $Q \in S^{|V|}_+ \cap S_G$ if and only if $Q$ can be written as $Q = \sum_{i=1}^r Q_i$, where $Q_i \in S^{C_i}_+$ has nonzero entries only with row and column indices coming from $C_i$.

The chordal decomposition, stated in Theorem 2.1, has enabled significant progress in large-scale semidefinite programming; see for instance [18] and [19]. Given a graph $G$ with $V = N^d_\beta$, let $\Sigma[G]$ be the set of SOS polynomials that admit a PSD Gram matrix with sparsity graph $G$, i.e., $\Sigma[G] := \{(x^\top N_{\beta})^\top Q x^\top N_{\beta} \mid Q \in S^{|V|}_+ \cap S_G \}$. Note that, in general, $\Sigma[G]$ is a strict subset of $\Sigma_{2d}[x]$, and therefore, the sparse SOS strengthening $f \in \Sigma[G]$ of the inequality $f \geq 0$ is generally more conservative than the corresponding dense strengthening $f \in \Sigma_{2d}[x]$. \footnote{A chord is an edge that joins two nonconsecutive nodes in a cycle.}

### III. 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 support extension and chordal extension to the related sparsity graphs, inspired by Theorem 2.1. In doing so, we obtain ascending chains of support sets, which will lead to a hierarchy of sparse SDP relaxations. 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 [4] and GA [6], and bounding extreme events [20]. Suppose that the dynamical system we are considering is given by

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

(1)

with $f := (f_1, \ldots, f_n) \in \mathbb{R}[x]$, and the constraint set

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

(2)

with $p_1, \ldots, p_m \in \mathbb{R}[x]$. For the sake of convenience, we set $p_0 := 1$. Let $d_j := \max\{\deg(f_i) : i = 1, \ldots, n\}$, and $d_j := \deg(p_j)$ for $j = 0, 1, \ldots, m$. $d_j := \max\{d_j : j = 1, \ldots, m\}$. For $x_0 \in \mathbb{R}^n$ and $t \geq 0$, let $\varphi_t(x_0)$ denote the solution of (1) with initial condition $\varphi_0(x_0) = x_0$.

**Definition 3.1:** For a dynamical system (1) with the constraint set $X$, the MPI set is the set of initial conditions $x_0 \in X$, such that the solutions $\varphi_t(x_0)$ stay in $X$ for all $t \geq 0$.

As proposed in [3], given a positive integer $d$, the $d$th order SOS relaxation for approximating the MPI set is defined by

$$\hat{\theta}_d := \inf_{a_j, b_j, c_j, v, w} \int_{x_0}^x u(x) \, dx$$

s.t.

{\begin{aligned}
\beta &- \nabla_v \cdot f & = a_0 + \sum_{j=1}^m a_j p_j \\
\beta &- \nabla_w \cdot f & = b_0 + \sum_{j=1}^m b_j p_j \\
a_j, b_j, c_j & \in S^{|V|}_+ \cap S_G, j = 0, 1, \ldots, m
\end{aligned}}$$

(3)

where $\beta > 0$ is a preassigned discount factor (say, $\beta = 1$), and $\nabla$ is the gradient with respect to $x$. The dynamics enter through the discounted Liouville operator $v \mapsto \beta \cdot \nabla_v - \nabla_w \cdot f$. By [3], the sequence of optima of (3) converges monotonically from above to the volume of the MPI set, provided the polynomials $p_j$ satisfy a technical compactness condition, which is satisfied, for example, if $p_m = R - ||x||^2$ for some $R > 0$. Furthermore, the set $S_d := \{ w^{-1} \{ [1, +\infty) \} \mid x \in X : w(x) \geq 1 \}$ provides an outer approximation for the MPI set.

For a graph $G(V,E)$ with $V \subseteq N^d_\beta$, define $\text{supp}(G) := \{ \beta + \gamma \mid \beta, \gamma \in V \} \subseteq E$. Now, we give the iterative procedure to exploit term sparsity. Fix a relaxation order $d$. Let $\mathcal{A} = \bigcup_{j=1}^m \text{supp}(p_j)$ and $v$ be a polynomial with generic coefficients supported on $\mathcal{A}$. Let $\mathcal{A}^d := \mathcal{A} \cup \supp(\nabla_v \cdot f) \cup 2N^d_\beta$ with $2N^d_\beta := \{ 2 \alpha \mid \alpha \in N^d_\beta \}$. Intuitively, the set $\mathcal{A}^d$ is the minimal support that has to be involved in the SOS relaxation (3).\footnote{Here, the subset $2N^d_\beta$ is included in the definition of $\mathcal{A}^d$ to guarantee convergence; see [12].} Now, we expand the support set by iteratively performing support extension and chordal extension on some related graphs. In the following, we perform a construction, which is motivated by the desire to impose sparsity patterns to Gram matrices of $\beta v - \nabla \cdot f, w$, and $w - v - 1$, which are required to be SOS polynomials over $X$ with a prescribed support. We begin with the...
construction for \( \beta - \nabla v \cdot f \). For every integer \( s \geq 1 \), we iteratively define the graph \( G_{a,j}^s \), which will be imposed as the sparsity graph for a Gram matrix of \( a_j \) with \( V(G_{a,j}^s) := \mathbb{N}_{d-j}^s \) and

\[
E(G_{a,j}^s) := \{ (\beta, \gamma) \mid \beta + \gamma + \text{supp}(p_j) \cap (s a^s_j \cup \text{supp}(\nabla v^s \cdot f)) \neq \emptyset \}
\]

(4)

where \( v^s \) is a polynomial with generic coefficients supported on \( s a^s_j \cap \mathbb{N}_{d-j}^{s+1} \) for \( j = 0, 1, \ldots, m \), and further let \( s a^{s+1} = \text{supp}(G_{a,0}^s) \). Doing so, we get a finite ascending chain of support sets \( s a^s_j \subseteq \cdots \subseteq s a^{s+1} = \cdots \), and a finite ascending chain of graphs \( (G_{a,j}^s) \subseteq \cdots \subseteq (G_{a,j}^{s+1}) = \cdots \) for each \( j = 0, 1, \ldots, m \).

Remark 3.2: There is term sparsity to exploit for the SOS relaxation (3) if the graph \( G_{a,0}^s \) is not complete.

Example 3.3: Let us consider the classical Lorenz system:

\[
x_1 = 10(x_2 - x_1), \quad x_2 = x_1(28 - x_3) - x_2, \quad x_3 = x_1x_2 - \frac{8}{3}x_3
\]

with the constraint set \( X = \{ x \in \mathbb{R}^3 \mid 1 - x_1^2 \geq 0, 1 - x_2^2 \geq 0, 1 - x_3^2 \geq 0 \} \). Take the relaxation order \( d = 2 \) and then \( s a^2_j = \{ 1, x^2_1, x^2_2, x^2_3, x_1x_2, x_1x_3, x_2x_3, x^2_1x^2_2, x^2_1x^2_3, x^2_2x^2_3, x^2_1x^2_2x^2_3, x^2_1x^2_3, x^2_2x^2_3, x_1x^2_2x^2_3 \} \).

The sparsity graph \( G_{a,0}^2 \) is displayed in Fig. 1. If we use maximal chordal extensions, then \( (G_{a,j}^2) \) consists of two cliques, respectively \( 6 \) and \( 4 \), and \( (G_{a,j}^2) = (G_{a,j}^2) \) for each \( j = 0, 1, 2, 3 \).

Now, we perform similar constructions for \( w \) and \( v - v \). For a given \( s \geq 1 \), with \( s a^s_j := s a^s_j \) for every \( l \geq 1 \), we iteratively define the graph \( H_{a,j}^l \), which will be imposed as the sparsity graph for a Gram matrix of \( b_j \) or \( c_j \) with \( V(H_{a,j}^l) := \mathbb{N}_{d-j}^l \) and

\[
E(H_{a,j}^l) := \{ (\beta, \gamma) \mid \beta + \gamma + \text{supp}(p_j) \cap (s a^s_j \cup \text{supp}(H_{a,j}^l)) \neq \emptyset \}
\]

(5)

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

\[
H_{a,j}^{l+1} = \bigcup_{j=0}^m (\text{supp}(p_j) + \text{supp}(H_{a,j}^l))
\]

(6)

Doing so, we also get a finite ascending chain of support sets \( s a^s_j \subseteq \cdots \subseteq s a^{s+1} = \cdots \), and a finite ascending chain of graphs \( (H_{a,j}^s) \subseteq \cdots \subseteq (H_{a,j}^{s+1}) = \cdots \) for each \( j = 0, 1, \ldots, m \).

Remark 3.4: To handle polynomial inequalities imposed over another constraint set (say, defined by \( q_j \geq 0, j = 1, \ldots, t \)), we simply replace \( p_j \) by \( q_j \) in (5) and (6) to obtain the related sparsity graphs.

Example 3.5: Continue considering Example 3.3. Take \( s = 1 \). The sparsity graph \( H_{a,0}^1 \) is displayed in Fig. 2. If we use maximal chordal extensions, then \( (H_{a,0}^1) \) consists of two cliques of size \( 6 \) and \( 4 \), respectively, and \( H_{a,1}^2 = H_{a,0}^2 \). In fact, we have \( H_{a,0}^2 = H_{a,1}^3 = \cdots \) and \( (H_{a,0}^1) = (H_{a,1}^1) = \cdots \) for \( j = 0, 1, 2, 3 \).

Remark 3.6: For the reason for using two different support sets for \( v \) and \( w \) and that \( v \) and \( w \) are of different degrees when one fixes a relaxation order \( d \) for (3).

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

\[
\begin{align*}
\theta_d^{s,l} := \inf_{a_j, b_j, c_j, v, w} & \int_X w(x) \, dx \\
\text{s.t.} & \quad v, w = w + \sum_{j=1}^m b_j p_j \\
& \quad w = \sum_{j=1}^m c_j p_j \\
& \quad a_j \in \Sigma([H_{a,j}^s]) \cup \Sigma([H_{a,j}^l]) \\
& \quad j = 0, 1, \ldots, m.
\end{align*}
\]

(7)

Notice that the Gram matrix of any SOS involved in (7) admits a block decomposition because of Theorem 2.1. Hence, the corresponding SDP could be easier to solve.

Proposition 3.7: With the above notations, we have \( \theta_d^{s,l} \leq \theta_d^s \leq \theta_d^{s,l+1} \leq \theta_d^l \). \( \theta_d^{s+1} \leq \theta_d^s \leq \theta_d^l \) for \( d \geq \max \{ [d/2], [d/2] \} \), \( s \geq 1 \), and \( l \geq 1 \).

Proof: First, note that \( \text{supp}(G_1) \subseteq \text{supp}(G_2), \Sigma[G_1] \subseteq \Sigma[G_2] \) if \( G_1 \) and \( G_2 \) are two graphs with \( G_1 \subseteq G_2 \). As a result, the feasible set of the corresponding sparse SOS relaxation becomes larger when we increase \( d, s \), or \( l \), from which the first three inequalities then follow. The last inequality \( \theta_d^{s,l} \geq \theta_d^l \) 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 (1) with the constraint set \( X \) constant set \( X \) (2) consists of all vectors \( r \in \mathbb{R}^m \) satisfying \( f_i((-1)^r \cdot x) = (-1)^{r_i} f_i(x) \) for each \( i = 1, \ldots, m \), and \( p_j((-1)^r \cdot x) = p_j(x) \) for each \( j = 1, \ldots, m \), where \((-1)^{r_i} > 0 \) (1). Given a set of sign symmetries \( R \subseteq \mathbb{R}^m \), we define \( R^+ := \{ \alpha \in \mathbb{R}^m \mid r \cdot \alpha \equiv 0 \text{ (mod 2\})} \). Then, a polynomial \( g \) is invariant under the sign symmetries \( R^+ \) if and only if \( \text{supp}(g) \subseteq R^2 \). The following theorem tells us that the SOS relaxation (3) inherits the sign symmetries of the dynamical system (1).

Theorem 3.8: Let \( R \) be the set of sign symmetries of system (1) with the constraint set \( X \) (2). If we additionally impose the constraints that \( \text{supp}(v) \subseteq R^+ \) and \( \text{supp}(a_j) \cup \text{supp}(b_j) \cup \text{supp}(c_j) \subseteq R^l \) for \( j = 0, 1, \ldots, m \), in (3), the resulting program has the same optimum with (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). We remove the terms of \( v \) and \( w \) with exponents

\[ g((-1)^r \cdot x) = g(x) \]

for all \( r \in R \), and a similar symmetry reduction already appeared in (20) in the study of bounding extreme events in dynamical systems.
not belonging to \( R^c \) from the expression of \( v \) and \( w \) and denote the resulting polynomials by \( \tilde{v} \) and \( \tilde{w} \), respectively. Let \( \tilde{Q}_j \) be a PSD Gram matrix of \( a_j \) for any \( j \) such that \( a_j = (x_{N_j \cap d}^T \cdots a_{N_j \cap d-1}) \). We then define \( \tilde{Q} \) \( \in S_{N_2} \) by letting \( \tilde{Q} \tilde{a}_j \tilde{a}_j^T = \tilde{Q}_j \tilde{a}_j \tilde{a}_j^T \), and putting zeros elsewhere, and let \( \tilde{Q} \) be block diagonal (after an appropriate permutation on rows and columns). So \( \tilde{Q} \) \( \geq 0 \) and it follows that \( \tilde{Q}_j \) 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 \( R^c \) from both sides of the equations in (3), \( \tilde{v}, \tilde{w}, \{ \tilde{a}_j \}, \{ \tilde{b}_j \}, \{ \tilde{c}_j \} \) are again a feasible solution to (3). It remains to show that \( f_1 \) \( \geq 0 \) if \( f_2 \) \( \geq 0 \). Then, there exists \( \alpha = (\alpha_i) \geq 0 \) such that \( \alpha \cdot r \neq 0 \) (mod 2). We have \( f_1 \alpha^\top \alpha \leq f_2 (1 - r)^\top \alpha \alpha \leq f_2 \alpha^\top \alpha \), where the first equality follows from the fact that \( X \) is invariant under the sign symmetry \( r \). This immediately gives \( f_1 \alpha^\top \alpha \leq 0 \) from which it follows \( f_1 \alpha^\top \alpha \). ■

From the proof of Theorem 3.8, we see that the sign symmetries of the dynamical system (1) endow any SOS polynomial involved in (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 3.9:** Given \( d \geq \max \{ [d_1/2], [d_p/2] \} \), the sign symmetries of the system (1) with the constraint set \( X \) (2) coincide with the sign symmetries of \( \alpha \), i.e., the set \( \{ \alpha \in \mathbb{Z}^2_\alpha | r \cdot \alpha \equiv 0 \) (mod 2) \( \forall \alpha \in \alpha \} \).

**Proof:** Let us denote the set of sign symmetries of system (1) by \( R \) and the set of sign symmetries of \( \alpha \) by \( R' \). For any \( r \in R \), we wish to show that \( r \in R' \). It suffices to show that \( r \cdot \alpha \equiv 0 \) (mod 2) for any \( \alpha \in \alpha \). Suppose \( r \cdot \alpha \equiv 0 \) (mod 2) for any \( \alpha \in \alpha \). Then, \( \tilde{v}, \tilde{w}, \{ \tilde{a}_j \}, \{ \tilde{b}_j \}, \{ \tilde{c}_j \} \) are again a feasible solution to (3). It remains to show that \( f_1 \alpha^\top \alpha \leq 0 \) if \( f_2 \alpha^\top \alpha \). Then, there exists \( \alpha = (\alpha_i) \geq 0 \) such that \( \alpha \cdot r \neq 0 \) (mod 2). We have \( f_1 \alpha^\top \alpha \leq f_2 (1 - r)^\top \alpha \alpha \leq f_2 \alpha^\top \alpha \), where the first equality follows from the fact that \( X \) is invariant under the sign symmetry \( r \). This immediately gives \( f_1 \alpha^\top \alpha \leq 0 \) from which it follows \( f_1 \alpha^\top \alpha \). ■

Based on Lemma 3.9, we can prove the following theorem by a similar argument as for [12, Th. 6.5] and so we omit the proof.

**Theorem 3.10:** Given \( d \geq \max \{ [d_1/2], [d_p/2] \} \), if maximal chordal extensions are used in the construction, then the block structures produced by the iterative procedure stated above Proposition 3.7 converge to the one given by the sign symmetries of system (1) (with the constraint set \( X \) (2)) as \( s \), \( l \) increase. As a corollary of this fact and Theorem 3.8, in this case, we have \( \theta_d \geq \theta_d \) for \( s, l \) sufficiently large.

**Remark 3.11:** Except maximal chordal extensions, other choices of chordal extensions may result in moment-SOS hierarchies with a lower complexity, but that generally produce more conservative estimates for the problem at hand.

**Remark 3.12:** From a more general point of view, the core elements of the strategy for exploiting term sparsity consist of (i) an algorithm to construct an increasing sequence \( G_1 \subseteq G_2 \subseteq \cdots \) of sparsity graphs for the Gram matrices of SOS polynomials, and (ii) a way to impose the positive semidefiniteness of a large Gram matrix via decomposition into smaller matrix constraints. The strategy we suggest in this article is just one particular choice (perhaps the most natural one as it converges to sign symmetries automatically).

### IV. ILLUSTRATIVE EXAMPLES

In this section, we give several 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 examples were computed on an Intel i9-10900@2.80 GHz CPU with 64 GB RAM memory and Mosek is used as an SDP solver. The timing is recorded in seconds. “opt” denotes the optimum. The index \( l \) is fixed to 1.

**A. Comparison With the Method of [16]**

Consider \( s_1 = (x_1^2 + x_2^2 - 1/2)x_1, s_2 = (x_2^2 + x_2^2 - 1/2)x_2, x_4 = (x_2^2 + x_2^2 - 1/2)x_3 \) with the constraint set \( X = [-1, 1]^3 \). According to [16], the system is decoupled w.r.t. \( \{x_1, x_2\} \) and \( \{x_2, x_3\} \). This system cannot be decoupled by [15]. We compare the results for approximating the MPI set obtained by exploiting term sparsity \( (s = 1) \), sign symmetries \( (s = 2) \), the dense method and the method of [16] in Table I. We see that TS is the fastest and provides medium bounds; SS and FD provide the best bounds whereas the method of [16] provides the worst bounds.

**B. Comparison With the Method of [15]**

Consider the system \( s_1 = (x_1^2 + x_2^2 - 1/2)x_1, s_2 = (x_2^2 + x_2^2 - 1/2)x_2, s_3 = (x_2^2 + x_2^2 - 1/2)x_3 \) with the constraint set \( X = [-1, 1]^3 \). According to [15], the system is decoupled w.r.t. \( \{x_1, x_2\} \) and \( \{x_2, x_3\} \). This system cannot be decoupled by [16]. We compare the results for approximating the MPI set obtained by exploiting term sparsity \( (s = 1) \), sign symmetries \( (s = 2) \), the dense method, and the method of [15] in Table II. We can conclude that SS and FD provide the same bounds while the former is more efficient; TS and the method of [16] are even more efficient but provide weaker bounds.

**C. Extended Lorenz System**

Consider the extended Lorenz system \( x_1 = 10r_1 - 12x_2, x_2 = -\frac{20}{3} r_1 x_1 + x_2 + \frac{4}{3} x_3, x_3 = -\frac{20}{3} x_3 - 15x_1 x_2, x_4 = 10(x_1 - x_1), x_5 = x_1(28 - x_3 - x_3) \) with the constraint set \( X = [1, 1]^5 \).

| Table I RESULTS FOR SECTION IV-A: TS: TERM SPARSITY, SS: SIGN SYMMETRY, FD: FULLY DENSE |
|---|---|---|---|---|---|
| 2d | TS | SS | PD | [16] |
| opt | time | opt | time | opt | time |
| 18 | 2.86 | 0.15 | 1.66 | 2.55 | 1.66 | 33.0 | 6.21 | 0.57 |
| 20 | 2.86 | 0.21 | 1.55 | 3.72 | 1.55 | 56.2 | 5.89 | 1.31 |
| 22 | 2.86 | 0.33 | 1.49 | 6.00 | 1.49 | 132 | 5.64 | 2.38 |

5Values of \( s, l \) for \( \theta_d \) are valid are a priori unknown, but they are typically no greater than 3 in practice.

6Namely, the block structures of the sparse hierarchy converge to the one determined by the sign symmetries of the system when \( s = 2 \).
According to [15], the system is decoupled w.r.t. \( \{x_1, x_2, x_3, x_4\} \) and \( \{x_1, x_2, x_3, x_5\} \). This system cannot be decoupled by [16]. In Table III, we report the results for approximating the MPI set by exploiting term sparsity (\( s = 2 \)), sign symmetries (\( s = 3 \)), the dense method, and the method of [15]. Here, we see that SS is several times faster than FD without sacrificing any accuracy; TS is slightly faster than SS while providing somewhat weaker bounds; and the method of [15] is the fastest but provides weaker bounds.

### D. Randomly Generated Models [21]

We consider the following sparse dynamical system for varying \( n \):
\[
    \dot{x}_i = (x^T B x - 1)x_i, \quad i = 1, \ldots, n,
\]
where \( B \in \mathbb{S}_G \) is a random positive definite matrix satisfying the following:

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

The constraint set is \( X = [-1, 1]^n \). For each system, we approximate the MPI set by exploiting term sparsity (\( s = 1 \)), sign symmetries (\( s = 2 \)), and the dense method. Fig. 3 shows outer approximations of the MPI set for \( n = 6, 8, 10, 12 \), respectively. In Table IV, we list optima and running time for solving corresponding SDPs. From Fig. 3 and Table IV, we can conclude the following: SS is faster than FD by one or two orders of magnitude without sacrificing any accuracy; TS is several times faster than SS while providing slightly weaker bounds.

### E. Bounding Extreme Events for a 16-Mode Fluid Model

This model ([20, Example 4.2]) is given by the following system:
\[
    \dot{x}_n = -\left(2\pi n\right)^2 x_n + \sqrt{2\pi n} \sum_{i=1}^{16-n} x_i x_{i+n} - \frac{1}{2} \sum_{i=1}^{n-1} x_i x_{n-i}
\]
for \( n = 1, \ldots, 16 \). Let \( \Phi(x) = 2\pi^2 \sum_{i=1}^{16} x_i^2 \), \( X_0 = \{x \in \mathbb{R}^{16} \mid \Phi(x) \leq \Phi_0 \} \), and \( X = \{x \in \mathbb{R}^{16} \mid ||x||^2 \leq \Phi_0/(2\pi^2) \} \), where \( \Phi_0 \in \mathbb{R} \). Let \( \Phi_\infty \) denote the largest value attained by \( \Phi(x(t; t_0, x_0)) \) among all trajectories that start from \( X_0 \subseteq X \) and evolve forward over the time interval \([0, \infty)\). SOS relaxations were proposed in [20] to bound \( \Phi_\infty \) from above. This system cannot be decoupled by either [15] or [16].
We can adapt the iterative procedure presented in this article to derive sparse SOS relaxations for bounding $\Phi$. Taking the relaxation order $d = 2$, we solve the dense relaxation (FD) and the sparse relaxations with $s = 1$ and $s = 2$ (SS). In particular, for $s = 1$, we include the results obtained either with approximately smallest chordal extensions [TS(S)] or with maximal chordal extensions [TS(M)]. The results are shown in Fig. 4. It turns out that the upper bounds given by the four methods agree to within 4% for different $\Phi$. On average, TS(S) is six times faster than SS, TS(M) is slightly faster than SS, and SS is five times faster than FD, which indicates that there is significant speed-up by exploiting term sparsity.

V. CONCLUSION

This article 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 tradeoff between computational costs and the solution accuracy. Moreover, it is able to guarantee convergence under certain conditions and recover the sign-symmetry reduction.

REFERENCES

[1] J. E. Rubio, “Generalized curves and extremal points,” SIAM J. Control, vol. 13, no. 1, pp. 28–47, 1975.
[2] R. M. Lewis and R. B. Vinter, “Relaxation of optimal control problems to equivalent convex programs,” J. Math. Anal. Appl., vol. 74, no. 2, pp. 475–493, 1980.
[3] M. Korda, D. Henrion, and C. N. Jones, “Convex computation of the maximum controlled invariant set for polynomial control systems,” SIAM J. Control Optim., vol. 52, no. 5, pp. 2944–2969, 2014.
[4] D. Henrion and M. Korda, “Convex computation of the region of attraction of polynomial control systems,” IEEE Trans. Autom. Control, vol. 59, no. 2, pp. 297–312, Feb. 2014.
[5] V. Magron, P.-L. Garoche, D. Henrion, and X. Thirioux, “Semidefinite approximations of reachable sets for discrete-time polynomial systems,” SIAM J. Control Optim., vol. 57, no. 4, pp. 2799–2820, 2019.
[6] C. Schlosser and M. Korda, “Converging outer approximations to global attractors using semidefinite programming,” Automatica, vol. 134, 2021, Art. no. 109900.
[7] M. Korda, D. Henrion, and I. Mezić, “Convex computation of extremal invariant measures of nonlinear dynamical systems and Markov processes,” J. Nonlinear Sci., vol. 31, no. 1, pp. 1–26, 2021.
[8] J.-B. Lasserre, D. Henrion, C. Prieur, and E. Trélat, “Nonlinear optimal control via occupation measures and LMI relaxations,” SIAM J. Control Optim., vol. 47, pp. 1643–1666, 2008.
[9] C. Riener, T. Theobald, L. J. Andrén, and J. B. Lasserre, “Exploiting symmetries in SDP-relaxations for polynomial optimization,” Math. Operations Res., vol. 38, no. 1, pp. 122–141, 2013.
[10] H. Waki, S. Kim, M. Kojima, and M. Muramatsu, “Sums of squares and semidefinite programming relaxations for polynomial optimization problems with structured sparsity,” SIAM J. Optim., vol. 17, no. 1, pp. 218–242, 2006.
[11] J. Wang, V. Magron, and J. B. Lasserre, “Chordal-TSSOS: A moment-SOS hierarchy that exploits term sparsity with chordal extension,” SIAM J. Optim., vol. 31, no. 1, pp. 114–141, 2021.
[12] J. Wang, V. Magron, and J. B. Lasserre, “TSSOS: A moment-SOS hierarchy that exploits term sparsity,” SIAM J. Optim., vol. 31, no. 1, pp. 30–58, 2021.
[13] J. Wang and V. Magron, “Exploiting term sparsity in noncommutative polynomial optimization,” Comput. Optim. Appl., vol. 80, no. 2, pp. 483–521, 2021.
[14] J. Wang, M. Maggio, and V. Magron, “SparseJSR: A fast algorithm to compute joint spectral radius via sparse SOS decompositions,” in Proc. IEEE Amer. Control Conf., 2021, pp. 2254–2259.
[15] C. Schlosser and M. Korda, “Sparse moment-sum-of-squares relaxations for nonlinear dynamical systems with guaranteed convergence,” 2020, arXiv:2012.05572.
[16] M. Tacchi, C. Cardozo, D. Henrion, and J. B. Lasserre, “Approximating regions of attraction of a sparse polynomial differential system,” IFAC-PapersOnLine, vol. 53, no. 2, pp. 3266–3271, 2020.
[17] J. Agler, W. Helton, S. McCullough, and L. Rodman, “Positive semidefinite matrices with a given sparsity pattern,” Linear Algebra Its Appl., vol. 107, pp. 101–149, 1988.
[18] M. Fukuda, M. Kojima, K. Murota, and K. Nakata, “Exploiting sparsity in semidefinite programming via matrix completion I: General framework,” SIAM J. Optim., vol. 11, no. 3, pp. 647–674, 2001.
[19] L. Vandenberghe and M. S. Andersen, “Chordal graphs and semidefinite optimization,” Found. Trends Optim., vol. 1, no. 4, pp. 241–433, 2015.
[20] G. Fantuzzi and D. Golinski, “Bounding extreme events in nonlinear dynamics using convex optimization,” SIAM J. Appl. Dynamical Syst., vol. 19, no. 3, pp. 1823–1864, 2020.
[21] W. Tan and A. Packard, “Stability region analysis using polynomial and composite polynomial Lyapunov functions and sum-of-squares programming,” IEEE Trans. Autom. Control, vol. 53, no. 2, pp. 565–571, Mar. 2008.