# Computation of attractor dimension and maximal sums of Lyapunov exponents using polynomial optimization

Jeremy P. Parker[1], David Goluskin[2]

[1]Division of Mathematics, University of Dundee, Dundee, DD1 4HN, United Kingdom
[2]Department of Mathematics and Statistics, University of Victoria, Victoria, BC, V8P 5C2, Canada

**Abstract**

Two approaches are presented for computing upper bounds on Lyapunov exponents and their sums, and on Lyapunov dimension, among all trajectories of a dynamical system governed by ordinary differential equations. The first approach expresses a sum of Lyapunov exponents as a time average in an augmented dynamical system and then applies methods for bounding time averages. This generalizes the work of Oeri & Goluskin (Nonlinearity 36:5378–5400, 2023), who bounded the single leading Lyapunov exponent. The second approach considers a different augmented dynamical system, where bounds on sums of Lyapunov exponents are implied by stability of certain sets, and such stability is verified using Lyapunov function methods. Both of our approaches also can be adapted to directly compute bounds on Lyapunov dimension, which in turn implies bounds on the fractal dimension of a global attractor. For systems of ordinary differential equations with polynomial right-hand sides, all of our bounding formulations lead to polynomial optimization problems with sum-of-squares constraints. These sum-of-squares problems can be solved computationally for any particular system to yield numerical bounds, provided the number of variables and degree of polynomials is not prohibitive. Most of our upper bounds are proven to be sharp under relatively weak assumptions. In the case of the polynomial optimization problems, sharpness means that upper bounds converge to the exact values as polynomial degrees are raised. Computational examples demonstrate upper bounds that are sharp to several digits, including for a six-dimensional dynamical system where sums of Lyapunov exponents are maximized on periodic orbits.

## 1 Introduction

The detection and characterization of chaos is a major computational challenge for many dynamical systems. Chaotic properties are often quantified via the spectrum of Lyapunov exponents (LEs), which can be computed over individual trajectories to obtain local information about that trajectory and, perhaps, global implications such as existence of hyperchaos or estimates of attractor dimension. In this paper we consider the problem of finding upper bounds on sums of LEs among all possible trajectories, without knowledge of any particular trajectories. Upper bounds on sums of LEs can be used to construct upper bounds on the fractal dimension of the global attractor. Our upper bounds are complementary to typical trajectory-based computations, which give corresponding lower bounds on sums of LEs and on Lyapunov dimension of global attractors.

The LE spectrum quantifies the asymptotic rates at which linearized perturbations in different directions grow or decay along trajectories of a dynamical system. For dynamics in $n$ dimensions there are $n$ exponents, defined in decreasing order. Partial sums of the first $k$ exponents describe the exponential rates of growth or contraction of $k$-dimensional volumes in state space, so that the first exponent governs the stretching of lines, the sum of the first two governs the growth of areas, and so on. In dissipative systems, these sums become negative beyond a certain value of $k$, reflecting the contraction of higher-dimensional volumes. One way of defining a value at which this contraction begins leads to the Lyapunov dimension, which is non-integer in general, and which provides a natural connection between dynamical stability and the fractal geometry of invariant sets such as strange attractors.

We consider autonomous, continuous-time dynamics governed by an ordinary differential equation (ODE) system,

$$\frac{\mathrm{d}}{\mathrm{d}t}X(t) = F(X(t)), \qquad X(0) = X_0, \tag{1}$$

with $F \in C^1(\Omega, \mathbb{R}^n)$, where $C^k(A, B)$ denotes the space of functions mapping $A$ to $B$ that are $k$ times continuously differentiable. The state space $\Omega \subset \mathbb{R}^n$ may be all of $\mathbb{R}^n$, a subdomain of $\mathbb{R}^n$ or a lower-dimensional manifold embedded in $\mathbb{R}^n$.

In the present work, we apply existing methods for proving global properties of ODE dynamical systems to the problem of bounding LEs and their sums. These existing methods rely on finding auxiliary functions $V \in C^1(\Omega, \mathbb{R})$ subject to certain pointwise inequalities that, in turn, imply global statements about the dynamics. The best known examples of auxiliary functions are Lyapunov functions, whose inequality conditions imply statements about stability. (Our use of Lyapunov functions to study Lyapunov exponents is not typical, despite both objects bearing the same person's name.) Among the lesser known types of auxiliary functions are those whose inequality conditions imply bounds on infinite-time averages. For each quantity studied in the present work, we formulate two different upper bounds: a "shifted spectrum" approach based on Lyapunov function conditions for stability, and a "sphere projection" approach based on auxiliary function conditions for bounding time averages.

The shifted spectrum approach that we present in section 2.3 uses conditions on $V$ that are similar to standard Lyapunov function conditions. In stability theory, the pointwise inequality conditions required of a Lyapunov function may be, for instance,

$$V(X) \geq 0, \qquad F(X) \cdot DV(X) \leq 0 \qquad \forall\, X \in \Omega, \tag{2}$$

where $DV$ denotes the gradient of $V$ with respect to the state space variable $X$. Here $F(X) \cdot DV(X)$ is the Lie derivative of $V$ with respect to the flow since the usual chain rule gives $\frac{\mathrm{d}}{\mathrm{d}t}V(X(t)) = V(X(t)) \cdot DV(X(t))$. The existence of $V$ satisfying (2) would imply that the set of $X \in \Omega$ where $V(X) = 0$ is Lyapunov stable, meaning that trajectories remain arbitrarily close to this set if they start sufficiently close [17]. Until about 25 years ago, there was no broadly applicable method for finding $V$ subject to pointwise inequalities like those in (2). Now, however, it is often possible to construct such $V$ computationally using optimization methods based on sum-of-squares (SOS) polynomials [36, 39]. In particular, if the function $F$ defining the dynamics is polynomial in the components of $X$, and if $V$ is sought from a finite-dimensional space of polynomials, then $F \cdot DV$ is also a polynomial. Deciding whether the polynomials $V$ and $-F \cdot DV$ are nonnegative on $\Omega$ is prohibitively hard in general [32], but this nonnegativity can be ensured by stronger and more tractable SOS conditions. For instance, simply requiring $V$ to be SOS, meaning that it can be represented as a sum of squares of other polynomials, would imply $V \geq 0$ on all of $\mathbb{R}^n$. Other SOS conditions can imply nonnegativity on $\Omega$ but not on all of $\mathbb{R}^n$, as explained in section 2.4.

The sphere projection approach that we present in section 2.2 uses conditions on an auxiliary function $V$ for bounding time averages. For any $\Phi \in C^0(\Omega, \mathbb{R})$ evaluated along a trajectory $X(t)$ with initial condition $X_0$, define the infinite-time average $\overline{\Phi}$ as

$$\overline{\Phi}(X_0) = \limsup_{T \to \infty} \frac{1}{T} \int_0^T \Phi(X(t))\, \mathrm{d}t. \tag{3}$$

On each $X(t)$ that remains bounded forward in time, boundedness of $V(X(t))$ implies $\overline{\frac{\mathrm{d}}{\mathrm{d}t}V} = 0$, and thus $\overline{F \cdot DV} = 0$. Thus, on bounded trajectories,

$$\overline{\Phi} = \overline{\Phi + F \cdot DV} \leq \sup_{X \in \Omega} [\Phi + F \cdot DV]. \tag{4}$$

The right-hand upper bound is useful because it does not involve individual trajectories. Since (4) holds for all bounded trajectories and all $V \in C^1$, one can maximize the left-hand side over $X_0$ and minimize the right-hand side over $V$ to find

$$\sup_{X_0 \in \Omega} \overline{\Phi} \leq \inf_{V \in C^1(\Omega)} \sup_{X \in \Omega} [\Phi + F \cdot DV]. \tag{5}$$

Furthermore, the inequality in (5) is an *equality* when $\Omega$ is a compact set in which trajectories remain [3, 46], and this equality underlies sharpness of our sphere projection approach. When $\Phi$, $V$ and $F$ are polynomial, so is the expression $\Phi + F \cdot DV$ on the right-hand side of (5). In this case the min–max problem can be relaxed into a minimization over polynomial $V$ subject to an SOS constraint, as explained in section 2.4. This constitutes an SOS optimization problem, where the optimization objective is to minimize the resulting upper bound on $\overline{\Phi}$. These ideas were applied by Oeri and Goluskin [34] to bound the single leading LE. Here we extend this approach to sums of LEs and to Lyapunov dimension.

Our two approaches are the first broadly applicable computational methods that can give convergent upper bounds on global maxima of LE sums or Lyapunov dimension. Unlike methods that rely on pointwise optimization over state space [7], our methods remain sharp when the maximizing orbits are periodic, rather than stationary. Furthermore, our methods give a systematic way to search for the maximizing trajectories themselves.

The rest of the paper is organized as follows. Section 2 defines sums of LEs and Lyapunov dimension and then describes our two general approaches to bounding these quantities. After arriving at more general optimization problems, we strengthen their constraints to obtain SOS optimization problems, then we describe how to exploit symmetries. Section 3 presents computational results from applications of our methods to two examples: the chaotic Duffing oscillator, which can be formulated as an autonomous system in 4 variables, and a 3-degree-of-freedom Hamiltonian system adapted from [35], which is an autonomous system in 6 variables. For all LE sums in both examples, we confirm that our computed bounds are sharp to several digits by finding periodic orbits on which the bounds are approximately saturated. Concluding remarks are given in section 4. For completeness, appendix A summarizes necessary material about exterior products, and appendix B proves a non-standard variant of an existence theorem for Lyapunov functions.

# 2 Optimization problems and SOS relaxations

This section defines the objects we seek to bound and formulates our methods for doing so. section 2.1 give expressions for LE sums and Lyapunov dimension that are useful for what follows. In section 2.2 we describe approaches for bounding these quantities based on projecting tangent space vectors onto the unit sphere. In section 2.3 we describe approaches that instead shift the spectrum of the Jacobian so that tangent vectors remain bounded. Section 2.4 describes how each formulation is relaxed to into SOS optimization problems that can be implemented numerically. Finally, section 2.5 describes how symmetries may be exploited in the optimization problems.

## 2.1 Tangent vectors, Lyapunov exponents and Lyapunov dimension

In this section we consider autonomous ODE systems on a domain $\mathcal{B} \subset \mathbb{R}^n$, just as in (1) but denoted here using lowercase $x$ and $f$,

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = f(x(t)), \qquad x(0) = x_0, \tag{6}$$

with $f \in C^1(\mathcal{B}, \mathbb{R}^n)$. Uppercase $X$ and $F$ and domain $\Omega$ will be reserved for associated ODEs on augmented state spaces that combine the dynamics of $x$ with dynamics on its tangent space. Trajectories are assumed to remain in $\mathcal{B}$ for all $t \geq 0$. In cases where $\mathcal{B}$ is unbounded, we further assume that each trajectory remains bounded for $t \geq 0$, meaning there is no forward-time blowup. The differentiability of $f$ and the assumption of no blowup guarantee well-posedness for all positive times, meaning there exists a differentiable flow map $\phi : \mathbb{R}_+ \times \mathcal{B} \to \mathcal{B}$ satisfying $x(t) = \phi^t(x_0)$ for every initial condition $x_0 \in \mathcal{B}$ and all $t \geq 0$.

The LEs associated with a trajectory $x(t)$ quantify the convergence or divergence of infinitesimally nearby trajectories. Linearization of (6) around each point $x(t)$ defines the dynamics of the tangent vector $y(t)$, which evolves in the tangent space $\mathbb{R}^n$ according to

$$\frac{\mathrm{d}}{\mathrm{d}t} y(t) = Df(x(t))\, y(t), \qquad y(0) = y_0, \tag{7}$$

where $Df(x)$ denotes the $n \times n$ Jacobian matrix of $f$ at $x$, and $y_0$ is the initial tangent vector. Equivalently, $y(t)$ is given by the linearized finite-time flow map as

$$y(t) = D\phi^t(x_0)\, y_0. \tag{8}$$

We restrict $y_0$ to the unit sphere,

$$\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n \,:\, |x| = 1\}, \tag{9}$$

because the tangent vector dynamics are linear in $y$ and thus are unaffected by normalization of the initial condition. Considering the separation between trajectories starting at $x_0$ and at $x_0 + \varepsilon y_0$, Taylor expanding the flow map and using (8) gives

$$|\phi^t(x_0) - \phi^t(x_0 + \varepsilon y_0)| = \varepsilon |y(t)| + O(\varepsilon^2 |y_0|^2). \tag{10}$$

The tangent vector typically shrinks or grows exponentially like $|y(t)| \approx e^{\mu t}$ for some average rate $\mu(x_0, y_0)$, in which case (10) suggests

$$|\phi^t(x_0) - \phi^t(x_0 + \varepsilon y_0)| \approx \varepsilon |y_0| e^{\mu(x_0, y_0) t}. \tag{11}$$

If $\mu < 0$ this approximation can be valid for all time. If $\mu > 0$ this approximation can be valid only until the time when separation between trajectories grows too large for (10) to apply, but this time approaches infinity as $\varepsilon \to 0$. Motivated by the expectation that $|y(t)| \approx e^{\mu t}$, the corresponding LE can be defined precisely as

$$\mu(x_0, y_0) = \limsup_{t \to \infty} \frac{1}{t} \log |y(t)|. \tag{12}$$

For each trajectory $x(t)$, which is determined by its initial condition $x_0$, the LE $\mu$ defined by (12) generally depends on the initial tangent vector direction $y_0$. The Oseledets ergodic theorem guarantees that different $y_0$ give at most $n$ different LEs [40], the largest of which is called the leading LE. We define this by definition 2.1 and then give equivalent expressions in sections 2.2 and 2.3 that underlie the methods we introduce there. The literature on LEs uses various definitions that are equivalent in many settings [40].

**Definition 2.1.** For $x(t)$ evolving by the ODE (6), the *leading Lyapunov exponent* is

$$\mu_1(x_0) = \sup_{y_0 \in \mathbb{S}^{n-1}} \limsup_{t \to \infty} \frac{1}{t} \log |y(t)|, \tag{13}$$

where the tangent vector $y(t)$ evolves by (7).

The leading LE is the average growth rate of lengths in the tangent space, maximized over the initial direction. A natural generalization is to consider growth rates of $k$-dimensional volumes in the tangent space for any $k \leq n$. From this one can define leading sums of LEs and, in turn, the full spectrum of LEs [45]. It is also possible to define the spectrum of LEs directly using singular values of $D\phi^t$ [40], but their sums are easier to define, and the sums are what we study in the present work. Let $|y_1 \wedge \cdots \wedge y_k|$ denote the volume of the parallelepiped whose edges coincide with vectors $y_1, \ldots, y_k \in \mathbb{R}^n$. If each $y_i$ is a tangent vector that evolves in time according to (7), the volume of the corresponding parallelepiped generally grows or shrinks exponentially in time, so we can define an average exponential rate analogously to (12). Maximizing this rate over all initial directions of the tangent vectors gives [45]

$$M_k(x_0) = \sup_{\substack{y_1(0), \ldots, y_k(0) \in \mathbb{R}^n \\ |y_1(0) \wedge \cdots \wedge y_k(0)| = 1}} \limsup_{t \to \infty} \frac{1}{t} \log |y_1(t) \wedge \cdots \wedge y_k(t)|, \tag{14}$$

where each $y_i(t)$ evolves as in (7). The value $M_k(x_0)$ is the maximum time-averaged growth rate of $k$-dimensional volumes in the tangent space for the trajectory $x(t)$ starting at $x_0$. This maximum growth rate will be attained by almost every choice of tangent vectors, under the assumptions of certain multiplicative ergodic theorems [40].

The set of all $y_1 \wedge \cdots \wedge y_k$ form a vector space called the exterior product space of $k$-vectors on $\mathbb{R}^n$. The space of $k$-vectors on $\mathbb{R}^n$ has dimension $n_k = \frac{n!}{k!(n-k)!}$. We can therefore identify each $y_1 \wedge \cdots \wedge y_k$ with a vector denoted by $y^k \in \mathbb{R}^{n_k}$, whose Euclidean norm $|y^k|$ gives the volume of the parallelepiped with edges $y_1, \ldots, y_k$. For time-dependent $k$-vectors where each $y_i(t)$ is a tangent vector satisfying the ODE (7), the $k$-vectors satisfy a corresponding ODE (see lemma A.4),

$$\frac{\mathrm{d}}{\mathrm{d}t} y^k(t) = Df^{[k]}(x(t))\, y^k(t). \tag{15}$$

Here $Df^{[k]}$ denotes the $k^{th}$ additive compound for the matrix $Df$, and later $Df^{(k)}$ denotes the $k^{th}$ multiplicative compound. These compounds are linear transformations acting on the space of $k$-vectors, induced by the linear transformation $Df$ on vectors, and they can be computed explicitly as $n_k \times n_k$ matrices. Exterior products and compound matrices are reviewed in appendix A. The expression (14) for $M_k$ can be equivalently stated using $y^k(t)$, resulting in the following definition 2.2 that we choose for $M_k$.

**Definition 2.2.** For $x(t)$ evolving by the ODE (6) from initial condition $x_0$, the Lyapunov exponents are defined such that the *sum of the $k$ leading Lyapunov exponents* is

$$M_k(x_0) = \sup_{y_0^k \in \mathbb{S}^{n_k-1}} \limsup_{t\to\infty} \frac{1}{t} \log |y^k(t)|, \tag{16}$$

where $y^k(t)$ evolves by (15) from initial condition $y_0^k$.

Having defined $M_k$, we can define the LE spectrum $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n$ such the $M_k$ are necessarily LE sums. When $k = 1$ we already have $\mu_1 = M_1$ since definition 2.2 reduces to definition 2.1 when $y^k = y$ and $Df^{[k]} = Df$. For $k \geq 2$, the LE $\mu_k$ can be defined recursively by

$$\mu_k = M_k - M_{k-1}, \tag{17}$$

so indeed $M_k = \sum_{i=1}^k \mu_i$. There are equivalent ways to define $\mu_k$ as the fundamental object and then define the $M_k$ as their partial sums, but for our purposes it is simpler to define $M_k$ directly by (16).

For systems with attractors, a primary use of LE sums is to define and study the dimension of an attractor. Definition 2.3 gives the Kaplan–Yorke formula [16] for the global Lyapunov dimension $d_L$ of an ODE system; see [18] for equivalent definitions. Hausdorff dimension is perhaps a more natural way to define the dimension of a fractal set in general. For the global attractor of a dynamical system, Hausdorff dimension can be very difficult to estimate directly, but it is bounded above by $d_L$ [8]. Thus, the upper bounds we produce for $d_L$ are upper bounds on Hausdorff dimension also.

**Definition 2.3.** For $x(t)$ evolving in $\mathcal{B} \subset \mathbb{R}^n$ by the ODE (6), let $j$ be the smallest integer such that $\sup_{x_0\in\mathcal{B}} M_{j+1}(x_0) < 0$. The *global Lyapunov dimension* over $\mathcal{B}$ is

$$d_L = j + \sup_{x_0\in\mathcal{B}} \left( \frac{M_j(x_0)}{-\mu_{j+1}(x_0)} \right). \tag{18}$$

For particular ODEs, the Kaplan–Yorke formula is usually evaluated numerically along a chaotic trajectory, rather than being maximized over all trajectories. The result indicates a fractal dimension of the chaotic attractor, rather than of the global attractor. Global Lyapunov dimension pertains to the latter, and the so-called Eden conjecture [9, 22] predicts that the maximum in (18) will be attained on predicts that the maximum in (18) is attained on an equilibrium or periodic orbit, not a chaotic orbit. Lemma 2.4 states the fact that $j \leq d_L < j + 1$, which we prove for completeness, along with a restatement of the definition of $d_L$ that we will use in sections 2.2 and 2.3.

**Lemma 2.4.** *The Lyapunov dimension $d_L$ has the following properties.*

1. *For the integer $j$ in definition 2.3, Lyapunov dimension is bounded by $j \leq d_L < j+1$.*
2. *The expression* (18) *for $d_L$ is equivalent to*

$$d_L = \inf_{B \in [0,1)} (j + B) \quad s.t. \quad B\left[M_j(x_0) - M_{j+1}(x_0)\right] - M_j(x_0) \geq 0 \;\; \forall\, x_0 \in \mathcal{B}. \tag{19}$$

*Proof.* The definition of $j$ implies that there exists $\varepsilon > 0$ such that

$$M_{j+1}(x_0) \leq -\varepsilon \quad \text{and} \quad -\mu_{j+1}(x_0) \geq \varepsilon \quad \text{for all } x_0 \in \mathcal{B}, \tag{20}$$

and that $\sup_{x_0 \in \mathcal{B}} M_j(x_0) \geq 0$. From these facts it follows that the supremum in the definition (18) of $d_L$ cannot be negative, which gives the lower bound $j \leq d_L$ in the first part of the lemma. For the upper bound, note that

$$\frac{M_j(x_0)}{-\mu_{j+1}(x_0)} \leq \frac{-\mu_{j+1}(x_0) - \varepsilon}{-\mu_{j+1}(x_0)} < 1, \tag{21}$$

which gives $d_L < j + 1$.

For the second part of the lemma, note that the supremum in the definition (18) of $d_L$ is equivalent to

$$\inf B \quad \text{s.t.} \quad \frac{M_j(x_0)}{-\mu_{j+1}(x_0)} \leq B \;\; \forall\, x_0 \in \mathcal{B}. \tag{22}$$

Since $-\mu_{j+1} = M_j - M_{j+1} > 0$, rearranging the constraint in (22) gives the constraint in (19) as an equivalent form. This infimum is in $[0, 1)$ and so is unchanged by the $B \in [0, 1)$ constraint in (19). Thus expression (19) for $d_L$ is equivalent to expression (18) in the definition. □

**Remark 2.5.** *Combining both parts of lemma 2.4 implies that, if we can find $B \in \mathbb{R}$ and $k \geq j$ where*

$$B\left[M_k(x_0) - M_{k+1}(x_0)\right] - M_k(x_0) \geq 0 \tag{23}$$

*for all $x_0 \in \mathcal{B}$, then the Lyapunov dimension is bounded above by $d_L \leq k + B$. Note that* (23) *cannot hold for all $x_0$ unless $B \geq 0$.*

## 2.2 Sphere projection approach

Our first approach to bounding the LE sums $M_k$ of definition 2.2 uses the formulation (5) for bounding infinite-time averages. This is a generalization of the approach in [34], which was described only for the $k = 1$ case. First note that the time average in the definition of $M_k$ can be expressed as a time integral along trajectories in the augmented state space for $x(t)$ and $y^k(t)$:

$$\limsup_{t\to\infty} \frac{1}{t} \log|y^k(t)| = \limsup_{T\to\infty} \frac{1}{T} \int_0^T \frac{\mathrm{d}}{\mathrm{d}t} \log|y^k(t)| \, \mathrm{d}t \tag{24}$$

$$= \limsup_{T\to\infty} \frac{1}{T} \int_0^T \frac{(y^k)^\mathsf{T} \frac{\mathrm{d}}{\mathrm{d}t} y^k}{|y^k|^2} \, \mathrm{d}t \tag{25}$$

$$= \limsup_{T\to\infty} \frac{1}{T} \int_0^T \frac{(y^k)^\mathsf{T} Df^{[k]}(x) y^k}{|y^k|^2} \, \mathrm{d}t, \tag{26}$$

where $(y^k)^\mathsf{T}$ is the transpose of the vector $y^k$. The formulation (5) for bounding time averages can be applied to the quantity in (26) only if each $(x, y^k)$ remains bounded forward in time, but $y^k(t)$ will be unbounded when $M_k$ is positive. The same tangent space dynamics can be captured with bounded variables by projecting $y^k$ onto a sphere, by introducing $z^k = y^k/|y^k|$. Expressing the right-hand side of (26) in terms of $z^k$ and then maximizing both sides over the initial tangent space directions $z_0^k$ gives

$$M_k(x_0) = \sup_{z_0^k \in \mathbb{S}^{n_k-1}} \overline{\Phi_k}, \tag{27}$$

where the overline denotes a time average of

$$\Phi_k(x, z^k) = (z^k)^\mathsf{T} Df^{[k]}(x) z^k \tag{28}$$

along trajectories in the $(x, z^k)$ state space. The ODE governing $z^k(t)$ follows from the ODE (15) for $y^k(t)$. The $x(t)$ dynamics together with the projected dynamics of $k$-vectors in its tangent space gives the ODE system

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} x \\ z^k \end{bmatrix} = \begin{bmatrix} f(x) \\ \ell_k(x, z^k) \end{bmatrix}, \tag{29}$$

on the augmented state space $\mathcal{B} \times \mathbb{S}^{n_k-1}$, where

$$\ell_k(x, z^k) = Df^{[k]}(x) z^k - \Phi_k(x, z^k) z^k. \tag{30}$$

Representations of $M_k$ as time averages over the projected tangent space dynamics (29) have often appeared in the study of stochastic dynamics, where they are called Furstenberg–Khasminskii formulas [2, chapter 6]. (In the stochastic case, $\Phi_k$ is defined with additional terms and is averaged with respect to stationary measures.)

Auxiliary function methods for bounding time averages can be applied to $M_k$ by using its representation (27) as a time average in the augmented state space for $(x, z^k)$. In particular, we directly apply the formula (5) with state variable $X = (x, z^k)$ on domain $\mathcal{B} \times \mathbb{S}^{n_k-1}$ and with $F(X)$ being the right-hand side of (29). The resulting inequality is the first part of proposition 2.6 below, for which no further proof is needed. If $F \in C^1$ and the domain is compact, then (5) holds with equality, as proved in [46]. This yields the second part of proposition 2.6 since assuming $f \in C^2$ gives that $F \in C^1$. The $k = 1$ special case of proposition 2.6 is proposition 1 in [34].

**Proposition 2.6.** *Let $\frac{\mathrm{d}}{\mathrm{d}t}x(t) = f(x(t))$ with $f \in C^1(\mathcal{B}, \mathbb{R}^n)$ and dynamics forward invariant on $\mathcal{B} \subset \mathbb{R}^n$. Let $\Phi_k(x, z^k)$ and $\ell_k(x, z^k)$ be defined as in* (28) *and* (30)*, respectively.*

1. *Suppose each trajectory $x(t)$ in $\mathcal{B}$ is bounded for $t \geq 0$. For each $k \in \{1, \ldots, n\}$, the maximum leading LE sum $M_k$ among trajectories in $\mathcal{B}$ is bounded above by*

$$\sup_{x_0 \in \mathcal{B}} M_k(x_0) \leq \inf_{V \in C^1(\mathcal{B} \times \mathbb{R}^{n_k})} \sup_{\substack{x \in \mathcal{B} \\ z^k \in \mathbb{S}^{n_k - 1}}} \left[\Phi_k + f \cdot D_x V + \ell_k \cdot D_{z^k} V\right], \tag{31}$$

   *where $D_x V$ and $D_{z^k} V$ denote the gradients of $V(x, z^k)$ with respect to $x$ and $z^k$, respectively.*

2. *Suppose $\mathcal{B}$ is compact and $f \in C^2(\mathcal{B}, \mathbb{R}^n)$. Then* (31) *holds with equality.*

The optimization problem on the right-hand side of (31) may be intractable, but it can be relaxed into computationally tractable optimization problems that also yield upper bounds on $M_k$. Relaxations using SOS polynomials are given in section 2.4 below.

The Lyapunov dimension (18) can be bounded above using similar ideas. One way is to apply upper bounds $M_k \leq B_k$ that have already been found using the right-hand side of (31) or its SOS relaxations. If such bounds are computed with $B_k \geq 0$ but $B_{k+1} < 0$ for some $k \in \{1, \ldots, n\}$, then proposition 2.7 below gives an upper bounds on $d_L$. To apply this result, one does not need to know whether $k$ is the minimal integer $j$ appearing in the definition (18) of $d_L$. It is enough to know that $B_{k+1} < 0$, which implies $k \geq j$.

**Proposition 2.7.** *Suppose that, for some $k \in \{1, \ldots, n\}$ and values $B_k \geq 0$ and $B_{k+1} < 0$,*

$$\sup_{x_0 \in \mathcal{B}} M_k(x_0) \leq B_k \quad \text{and} \quad \sup_{x_0 \in \mathcal{B}} M_{k+1}(x_0) \leq B_{k+1}. \tag{32}$$

*Then, the Lyapunov dimension* (18) *is bounded above by*

$$d_L \leq k + \frac{B_k}{B_k - B_{k+1}}. \tag{33}$$

*Proof.* The meaning of $j$ in definition 2.3 and the fact that $B_{k+1} < 0$ together imply $k \geq j$. According to remark 2.5, it suffices for the inequality (23) to hold for all $x_0$ with $B = \frac{B_k}{B_k - B_{k+1}}$. For any $x_0 \in \mathcal{B}$, the assumptions give $M_k(x_0) \leq B_k$ and $0 < -B_{k+1} \leq -M_{k+1}(x_0)$, so

$$-B_{k+1} M_k(x_0) \leq -B_k M_{k+1}(x_0), \tag{34}$$

and

$$B_k M_k(x_0) - B_{k+1} M_k(x_0) \le B_k M_k(x_0) - B_k M_{k+1}(x_0). \tag{35}$$

Dividing by the positive quantity $B_k - B_{k+1}$ and rearranging gives

$$\frac{B_k}{B_k - B_{k+1}} \left[ M_k(x_0) - M_{k+1}(x_0) \right] - M_k(x_0) \ge 0, \tag{36}$$

which is the required inequality (23). This concludes the proof. $\square$

The upper bound on Lyapunov dimension given by proposition 2.7 cannot be sharp for certain systems, even if $B_k$ and $B_{k+1}$ are each sharp bounds in (32). The reason is that the inequality (34), when maximized over $x_0$, is an equality only if $M_k(x_0)$ and $M_{k+1}(x_0)$ attain their maxima on the same orbits. This is not true in general; see section 3.2 for a counterexample. We thus give a second formulation for upper bounds on $d_L$ that is generally sharper than proposition 2.7. This formulation directly considers the ratio in the Kaplan–Yorke formula (18), rather than separately bounding the $M_k$ and $M_{k+1}$ terms in that ratio. In particular, we consider the $x(t)$ dynamics together with the projected dynamics of $k$-vectors and $(k+1)$-vectors in its tangent space,

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} x \\ z^k \\ z^{k+1} \end{bmatrix} = \begin{bmatrix} f(x) \\ \ell_k(x, z^k) \\ \ell_{k+1}(x, z^{k+1}) \end{bmatrix}, \tag{37}$$

where $\ell_k$ is defined by (30), and likewise for $\ell_{k+1}$. The augmented state space of this system is $(x, z^k, z^{k+1}) \in \mathcal{B} \times S^{n_k - 1} \times \mathbb{S}^{n_{k+1}-1}$. The constraint in (19) from lemma 2.4 can then be expressed in terms of time averages in the ODE system (37). For this we define

$$\Psi^B(x, z^k, z^{k+1}) = (1 - B)\Phi_k(x, z^k) + B\, \Phi_{k+1}(x, z^{k+1}), \tag{38}$$

and we consider the condition

$$\overline{\Psi}^B(x_0, z_0^k, z_0^{k+1}) \le 0 \quad \text{for all } x_0 \in \mathcal{B},\ z_0^k \in \mathbb{S}^{n_k - 1},\ z_0^{k+1} \in \mathbb{S}^{n_{k+1}-1}, \tag{39}$$

where the time average denoted by the overbar is along trajectories of (37). Condition (39) is equivalent to the first condition in (19) of lemma 2.4, as explained below in the proof of proposition 2.8. The condition (39) can be enforced using the general formulation (5) for bounding time averages, applied to the augmented system (37). This yields the first part of proposition 2.8, whose sharpness is addressed by its second part. Proposition 2.8 can be applied only when the LE sum $M_{k+1}$ for certain $k$ is known to be negative on every trajectory. Such negativity can be confirmed by using (31), or its SOS relaxation that we introduce in section 2.4, to obtain a negative upper bound on $M_{k+1}$.

**Proposition 2.8.** *Let $\frac{\mathrm{d}}{\mathrm{d}t}x(t) = f(x(t))$ with $f \in C^1(\mathcal{B}, \mathbb{R}^n)$ and dynamics forward invariant on $\mathcal{B} \subset \mathbb{R}^n$. Let $\Phi_k(x, z^k)$ and $\ell_k(x, z^k)$ be as defined by* (28) *and* (30)*, and likewise for $\Phi_{k+1}(x, z^{k+1})$ and $\ell_{k+1}(x, z^{k+1})$.*

1. *Suppose each trajectory $x(t)$ in $\mathcal{B}$ is bounded for $t \geq 0$. Let the positive integer $k \leq n$ be such that $\sup_{x_0 \in \mathcal{B}} M_{k+1}(x_0) < 0$. Then the global Lyapunov dimension $d_L$ over $\mathcal{B}$ is bounded above by*

$$d_L \leq \inf_{\substack{B \in [0,1] \\ V \in C^1}} (k+B) \quad s.t. \quad \begin{array}{l} \Psi^B + f \cdot D_x V + \ell_k \cdot D_{z^k} V + \ell_{k+1} \cdot D_{z^{k+1}} V \leq 0 \\ \text{for all} \quad x \in \mathcal{B},\ z^k \in \mathbb{S}^{n_k - 1},\ z^{k+1} \in \mathbb{S}^{n_{k+1}-1}, \end{array} \tag{40}$$

*where $V : \mathcal{B} \times \mathbb{S}^{n_k-1} \times \mathbb{S}^{n_{k+1}-1} \to \mathbb{R}$, and $D_x V$, $D_{z^k} V$ and $D_{z^{k+1}} V$ denote the gradients of $V(x, z^k, z^{k+1})$ with respect to each argument.*

2. *Suppose $\mathcal{B}$ is compact and $f \in C^2(\mathcal{B}, \mathbb{R}^n)$. Let $j$ be the minimum admissible $k$, as in definition 2.3 for $d_L$. Then, for all $\varepsilon > 0$,*

$$d_L \ - \ \varepsilon \left| \sup_{x_0 \in \mathcal{B}} M_{j+1}(x_0) \right|^{-1} \leq j + B_\varepsilon \leq d_L, \tag{41}$$

*where*

$$B_\varepsilon = \inf_{\substack{B \in [0,1] \\ V \in C^1}} B \quad s.t. \quad \begin{array}{l} \Psi^B + f \cdot D_x V + \ell_j \cdot D_{z^j} V + \ell_{j+1} \cdot D_{z^{j+1}} V \leq \varepsilon \\ \text{for all} \quad x \in \mathcal{B},\ z^j \in \mathbb{S}^{n_j - 1},\ z^{j+1} \in \mathbb{S}^{n_{j+1}-1}. \end{array} \tag{42}$$

*Proof.* For the first part, the assumption that $\sup_{x_0 \in \mathcal{B}} M_{k+1}(x_0) < 0$ means that $k \geq j$ where $j$ is as defined in definition 2.3. Applying remark 2.5, we therefore have

$$d_L \leq \inf_{B \in [0,1)} (k+B) \quad \text{s.t.} \quad B \left[ M_k(x_0) - M_{k+1}(x_0) \right] - M_k(x_0) \geq 0 \ \text{ for all } x_0 \in \mathcal{B}. \tag{43}$$

The constraint in (43) is equivalent to the $\overline{\Psi}^B \leq 0$ condition (39). To see this, we rearrange the constraint in (43) and use the expression (27) for $M_k$ to obtain an equivalent inequality,

$$(1-B) \sup_{z_0^k \in \mathbb{S}^{n_k}} \overline{\Phi}_k(x_0, z_0^k) + B \sup_{z_0^{k+1} \in \mathbb{S}^{n_{k+1}}} \overline{\Phi}_{k+1}(x_0, z_0^{k+1}) \ \leq \ 0 \quad \forall x_0 \in \mathcal{B}. \tag{44}$$

This inequality holds if and only if it holds without the suprema for all $z_0^k$ and $z_0^{k+1}$, which is exactly the $\overline{\Psi}^B \leq 0$ condition (39). Writing this latter condition in place of the equivalent constraint in (43) gives

$$d_L \leq \inf_{B \in [0,1)} (k+B) \quad \text{s.t.} \quad \overline{\Psi}^B(x_0, z_0^k, z_0^{k+1}) \leq 0 \ \text{ for all } (x_0, z_0^k, z_0^{k+1}) \in \mathcal{B} \times \mathbb{S}^{n_k - 1} \times \mathbb{S}^{n_{k+1}-1}. \tag{45}$$

We want to replace the constraint in (45) with one that does not involve time averages. To do so we upper bound $\overline{\Psi}^B$ using the general formulation (5) for time averages by letting

$X = (x, z^k, z^{k+1})$ on domain $\Omega = \mathcal{B} \times \mathbb{S}^{n_k-1} \times \mathbb{S}^{n_{k+1}-1}$, and letting $F(X)$ be the right-hand side of (37). This gives

$$\sup_{\substack{x_0 \in \mathcal{B} \\ z_0^k \in \mathbb{S}^{n_k-1} \\ z_0^{k+1} \in \mathbb{S}^{n_{k+1}-1}}} \overline{\Psi}_B \leq \inf_{V \in C^1} \sup_{\substack{x \in \mathcal{B} \\ z^k \in \mathbb{S}^{n_k-1} \\ z^{k+1} \in \mathbb{S}^{n_{k+1}-1}}} \left[ \Psi^B + f \cdot D_x V + \ell_k \cdot D_{z^k} V + \ell_{k+1} \cdot D_{z^{k+1}} V \right]. \quad (46)$$

The constraint of (43), which is equivalent to nonpositivity of the left-hand side of (46), can be strengthened by requiring there to exist $V \in C^1$ for which the inner supremum on the right-hand side is nonpositive. This gives the first part of the proposition.

For the second part, note that (43) is an equality when $k = j$, according to lemma 2.4. That equality can be expressed as

$$d_L = \inf_{B \in [0,1)} (j + B) \quad \text{s.t.} \quad \overline{\Psi}^B(x_0, z_0^j, z_0^{j+1}) \leq 0 \;\text{ for all } (x_0, z_0^j, z_0^{j+1}) \in \mathcal{B} \times \mathbb{S}^{n_j-1} \times \mathbb{S}^{n_{j+1}-1}, \quad (47)$$

where equivalence of the constraints in (43) and (47) is as in the first part. To prove the second inequality in the claim (41), let $B$ be any value for which the constraint in (47) holds. It suffices to prove existence of $V$ for which constraint in the definition (42) of $B_\varepsilon$ holds also. This will imply that the infimum over $B$ in (42) is at least as small as in (47).

Under the assumptions that $\mathcal{B}$ is compact and $f \in C^2$, the $X$ domain $\Omega$ is compact, and $F(X)$ defined as the right-hand side of (37) is $C^1$ on an open region containing $\Omega$. For such $X$ and $F$, the general upper bound (5) is an equality [46], so in particular (46) is an equality. Choosing $k = j$, equality in (47) implies that the constraint in (47) is equivalent to

$$0 \geq \inf_{V \in C^1} \sup_{\substack{x \in \mathcal{B} \\ z^k \in \mathbb{S}^{n_k-1} \\ z^{k+1} \in \mathbb{S}^{n_{k+1}-1}}} \left[ \Psi^B + f \cdot D_x V + \ell_k \cdot D_{z^k} V + \ell_{k+1} \cdot D_{z^{k+1}} V \right]. \quad (48)$$

This inequality implies that, for any $\varepsilon > 0$, there exists $V \in C^1$ such that

$$\left[ \Psi^B + f \cdot D_x V + \ell_k \cdot D_{z^k} V + \ell_{k+1} \cdot D_{z^{k+1}} V \right] \leq \varepsilon \quad \forall\, (x, z^k, z^{k+1}) \in \mathcal{B} \times \mathbb{S}^{n_k-1} \times \mathbb{S}^{n_{k+1}-1}. \quad (49)$$

(Such $V$ need not exist with $\varepsilon = 0$ if the infimum over $V$ in the constraint of (46) is not attained.) Condition (49) is the same as the constraint in the minimization (42) defining $B_\varepsilon$, and this implies $B_\varepsilon \leq B$. We have shown that $B_\varepsilon \leq B$ for any $B$ satisfying the constraint in (47), and thus $d_L - j \leq B_\varepsilon$. This establishes the second inequality in (41).

To prove the first inequality in the claim (41), suppose that the inequality constraint in the definition (42) of $B_\varepsilon$ is satisfied for a pair of $B$ and $V$. Averaging this inequality along any trajectory and then maximizing over the initial directions $z_0^k$ and $z_0^{k+1}$ gives

$$M_k(x_0) + B\mu_{k+1}(x_0) \leq \varepsilon, \quad (50)$$

which we divide by $-\mu_{k+1} > 0$ to find

$$\frac{M_k(x_0)}{-\mu_{k+1}(x_0)} - \frac{\varepsilon}{-\mu_{k+1}(x_0)} \leq B. \tag{51}$$

Maximizing the first term over $x_0$, and bounding the second term above using $\mu_{k+1}(x_0) \leq \sup_{x_0 \in \mathcal{B}} M_{k+1} < 0$, we find

$$(d_L - j) - \frac{\varepsilon}{\left|\sup_{x_0 \in \mathcal{B}} M_{j+1}(x_0)\right|} \leq B. \tag{52}$$

Since (52) holds at all $B$ where the constraint set in the definition (42) of $B_\varepsilon$ is not empty, minimizing over such $B$ gives (52) with $B_\varepsilon$ on the right-hand side. This establishes the first inequality in (41), which completes the proof. $\square$

**Remark 2.9.** *The second part of proposition 2.8 implies that*

$$d_L \leq j + B_\varepsilon + \varepsilon \left| \sup_{x_0 \in \mathcal{B}} M_{j+1}(x_0) \right|^{-1} \tag{53}$$

*at each $\varepsilon > 0$, and that this upper bound converges to $d_L$ as $\varepsilon \to 0$.*

### 2.3 Shifted spectrum approach

In the preceding subsection, the possibly unbounded $k$-vectors $y^k$ in the tangent space are captured by $z^k$, which is bounded because it is the projection of $y^k$ onto the unit sphere. In the alternative approach of the present subsection, we capture the dynamics of $y^k$ by letting

$$w^k(t) = w_0^k e^{-Bt} y^k(t), \tag{54}$$

where the initial condition $w_0^k$ can be any $k$-vector. This $w^k(t)$ is bounded forward in time for sufficiently large $B$. The ODE for $w^k$ is the same as the ODE (15) for $y^k$, except the spectrum of the Jacobian is shifted by $B$. Together with the ODE (6) for $x(t)$ this gives the system

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} x \\ w^k \end{bmatrix} = \begin{bmatrix} f(x) \\ m_k^B(x, w^k) \end{bmatrix}, \tag{55}$$

where

$$m_k^B(x, w^k) = Df^{[k]}(x) w^k - B w^k. \tag{56}$$

The following lemma asserts that if $B$ is large enough to prevent unbounded growth of $w^k$, then $B$ is an upper bound on the LE sum $M_k$, and the infimum over such $B$ is equal to $M_k$.

**Lemma 2.10.** *Let $\frac{\mathrm{d}}{\mathrm{d}t} x(t) = f(x(t))$ with $f \in C^1(\mathcal{B}, \mathbb{R}^n)$ and dynamics forward invariant on $\mathcal{B} \subset \mathbb{R}^n$. Let the initial condition $x_0$ give a trajectory that is bounded for $t \geq 0$. Let $w^k(t)$ satisfy the ODE (55) with initial condition $w_0^k$. For each $k \in \{1, \ldots, n\}$:*

1. *The leading LE sum $M_k(x_0)$ is equivalent to*

$$M_k(x_0) = \inf_{B\in\mathbb{R}} B \quad \text{s.t.} \quad \text{for each } w_0^k \in \mathbb{R}^{n_k},\ w^k(t) \text{ is bounded for all } t \geq 0. \tag{57}$$

2. *For any $B > M_k(x_0)$ there exist $K(x_0, B) > 0$ and $\lambda(x_0, B) > 0$ such that*

$$|w^k(t)| \leq K e^{-\lambda t}|w_0^k| \quad \text{for all } t \geq 0 \text{ and } w_0^k \in \mathbb{R}^{n_k}. \tag{58}$$

3. *If $\mathcal{B}$ is compact and $B > \sup_{x_0\in\mathcal{B}} M_k(x_0)$, there exist $K(B)$ and $\lambda(B)$ independent of $x_0$ such that (58) holds for all $x_0 \in \mathcal{B}$.*

*Proof.* If there exists $y_0^k$ such that $|y^k(t)| \to \infty$ in finite time then, according to (54), no $B$ satisfies the condition in (57). Then both sides of (57) are $\infty$, and the second part holds vacuously. Henceforth we assume that $w(t)$ exists for all $t \geq 0$.

We will first show that if $B$ satisfies the condition of (57) then $M_k(x_0) \leq B$. Substituting $y^k = e^{Bt}w^k$ into the time average on the right-hand side of (16) gives

$$\begin{aligned}
\limsup_{t\to\infty} \frac{1}{t} \log \left| y^k(t) \right| &= \limsup_{t\to\infty} \frac{1}{t} \log \left( \left| e^{Bt} w^k(t) \right| / |w_0^k| \right) \\
&= B + \limsup_{t\to\infty} \frac{1}{t} \log \left| w^k(t) \right| \\
&\leq B,
\end{aligned} \tag{59}$$

where the final inequality follows from the boundedness of $w^k(t)$. Taking the supremum of the left-hand side over $y_0^k$ gives $M_k(x_0) \leq B$.

Next we let $B > M_k(x_0)$. Showing that (58) holds will prove the second part of the lemma, and it will complete the proof of the first part by ensuring boundedness of $w^k(t)$. Without loss of generality, we assume $|w_0^k| = 1$. Since

$$\begin{aligned}
M_k(x_0) &\geq \limsup_{t\to\infty} \frac{1}{t} \log |y^k(t)| \\
&= B + \limsup_{t\to\infty} \frac{1}{t} \log |w^k(t)|,
\end{aligned} \tag{60}$$

we can choose some $\lambda < (B - M_k(x_0))$ so that $\limsup_{t\to\infty} \frac{1}{t} \log |w^k(t)| < -\lambda$. Therefore by Grönwall's inequality there exists a minimal $T(w_0^k) \geq 0$ such that $|w^k(t)| \leq e^{-\lambda t}$ for all $t \geq T(w_0^k)$. By compactness, we can let

$$T^* = \max_{w_0^k \in \mathbb{S}^{n_k-1}} T(w_0^k), \tag{61}$$

so that $|w^k(t)| \leq e^{-\lambda t}$ for all $t \geq T^*$ and all $w_0^k \in \mathbb{S}^{n_k-1}$. Finally we let

$$K = \max_{\substack{t\in[0,T^*] \\ w_0^k \in \mathbb{S}^{n_k-1}}} e^{\lambda t}|w^k(t)|. \tag{62}$$

which completes the proof of the first two parts of the lemma.

For the third part, assume that $M_k(x_0)$ is bounded above (otherwise the result is vacuously true), so that we can choose $\lambda < B - \sup_{x_0 \in \mathcal{B}} M_k(x_0)$. Let

$$K = \max_{\substack{x_0 \in \mathcal{B} \\ t \in [0, T^\dagger] \\ w_0^k \in \mathbb{S}^{n_k - 1}}} e^{\lambda t} |w^k(t)|, \tag{63}$$

where $T^\dagger = \max_{x_0 \in \mathcal{B}} T^*(x_0)$ for the $T^*(x_0)$ defined in (61). This maximum is valid because $(t, x_0, w_0^k) \mapsto e^{\lambda t} |w^k(t)|$ is continuous on the compact set $[0, T^\dagger] \times \mathcal{B} \times \mathbb{S}^{n_k - 1}$. With this value of $\lambda$ and $K$, we must then have (58) for every $x_0$, which completes the proof. $\square$

Practical upper bounds on the LE sum $M_k$ are found by combining lemma 2.10 with the specialised Lyapunov theorem given as lemma B.1 applied to the system (55). This yields the following proposition.

**Proposition 2.11.** *Let $\frac{\mathrm{d}}{\mathrm{d}t} x(t) = f(x(t))$ with $f \in C^1(\mathcal{B}, \mathbb{R}^n)$ and dynamics forward invariant on $\mathcal{B} \subset \mathbb{R}^n$. Let $m_k^B(x, w^k)$ be defined as in* (56).

1. *Suppose each trajectory $x(t)$ in $\mathcal{B}$ is bounded for $t \geq 0$. For each $k \in \{1, \dots, n\}$, the maximum leading LE sum $M_k$ among trajectories in $\mathcal{B}$ is bounded above by*

$$\sup_{x_0 \in \mathcal{B}} M_k(x_0) \leq \inf_{V \in C^1(\mathcal{B} \times \mathbb{R}^{n_k})} B \quad \text{s.t.} \quad \begin{aligned} V &\geq \left|w^k\right|^2 \\ f \cdot D_x V + m_k^B \cdot D_{w^k} V &\leq 0, \end{aligned} \tag{64}$$

   *where the right-hand constraints are imposed for all $(x, w^k) \in \mathcal{B} \times \mathbb{S}^{n_k}$.*

2. *Suppose $\mathcal{B}$ is compact and $f \in C^2(\mathcal{B}, \mathbb{R}^n)$. Then* (64) *holds with equality between the supremum and the infimum, and the Lyapunov functions may be restricted to the form $V(x, w^k) = (w^k)^\mathsf{T} U(x) w^k$ with $U \in C^1(\mathcal{B}, \mathbb{R}^{n_k \times n_k})$.*

*Proof.* For the first part, assume $B$ and $V$ satisfy the constraints on the right-hand side of (64). Letting $A(x) = Df^{[k]}(x) - BI_{n_k}$, the first part of lemma B.1 applies to the system (55) and guarantees boundedness of $w^k(t)$ for $t \geq 0$. Boundedness of $w^k(t)$ means that $B$ satisfies the constraint on the right-hand side of lemma 2.10, so $M_k(x_0) \leq B$ for all $x_0 \in \mathcal{B}$. This implies (64), so the first part of the proposition is proven.

For the second part, let $B > \sup_{x_0} M_k(x_0)$. By the third part of Lemma 2.10, there exist $K > 0$ and $\lambda > 0$ such that (135) holds for all $x_0 \in \mathcal{B}$. Observe also that $A \in C^1(\mathcal{B}, \mathbb{R}^{n_k \times n_k})$ since $f \in C^2$, and that $A(x)$ is bounded uniformly in $x$ since $\mathcal{B}$ is compact. Thus the second part of lemma B.1 applies, so there exists $V(x, w^k) = (w^k)^\mathsf{T} U(x) w^k$ satisfying the constraint in (64). This is true for any $B > \sup_{x_0} M_k(x_0)$, so (64) is an equality. $\square$

This method gives bounds for the different sums of LEs, which can then be used with proposition 2.7 to bound the Lyapunov dimension. Analogously to proposition 2.8, a direct bound for the dimension is also possible with a modified shifted spectrum approach, using the augmented system

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x \\ y^{k+1} \\ w \end{bmatrix} = \begin{bmatrix} f(x) \\ Df^{[k+1]}(x)y^{k+1} \\ S_k^B(x, y^{k+1}, w) \end{bmatrix}, \tag{65}$$

where we have defined

$$S_k^B(x, y^{k+1}, w) = Df^{[k]}(x)w + \frac{B}{1-B}\frac{(y^{k+1})^\mathsf{T} Df^{[k+1]}(x)y^{k+1}}{|y^{k+1}|^2}w. \tag{66}$$

No shifting is needed in the $y^{k+1}$ dynamics because this quantity decays exponentially on average when $M_{k+1} < 0$. The $w$ dynamics (66) shift the $Df^{[k]}$ spectrum not by a negative constant, as in (56), but by a factor proportional to the instantaneous growth rate of $y^{k+1}$, which is negative on average.

**Proposition 2.12.** *Let $\frac{\mathrm{d}}{\mathrm{d}t}x(t) = f(x(t))$ with $f \in C^1(\mathcal{B}, \mathbb{R}^n)$ and dynamics forward invariant on $\mathcal{B} \subset \mathbb{R}^n$ with each trajectory $x(t)$ bounded for $t \geq 0$. Let the positive integer $k \leq n$ be such that* $\sup_{x_0 \in \mathcal{B}} M_{k+1}(x_0) < 0$*. Then the global Lyapunov dimension $d_L$ over $\mathcal{B}$ is bounded above by*

$$d_L \leq \inf_{V \in C^1} (k + B) \quad s.t. \quad \begin{aligned} V(x, y^{k+1}, w) &\geq |w|^2 \\ f \cdot D_x V + Df^{[k+1]}y^{k+1} \cdot D_{y^k} V + S_k^B \cdot D_w V &\leq 0, \end{aligned} \tag{67}$$

*where the right-hand constraints are imposed for all $(x, y^{k+1}, w) \in \mathcal{B} \times \mathbb{R}^{n_{k+1}} \times \mathbb{R}^{n_k}$, and $D_x V$, $D_{z^k} V$ and $D_{z^{k+1}} V$ denote the gradients of $V(x, y^{k+1}, w)$ with respect to each argument.*

*Proof.* We will prove that, for any $B$ and $V$ satisfying the constraints in (67), $d_L \leq k + B$. We will use remark 2.5 to do this. First note that $\sup_{x_0 \in \mathcal{B}} M_{k+1}(x_0) < 0$ implies that $k \geq j$ where $j$ is the minimal integer of definition 2.3. Observe that the dynamics of $w$ in the system (65) are of the form $\frac{\mathrm{d}w}{\mathrm{d}t} = A(x, y^{k+1})w$ for some matrix $A$, so we may apply lemma B.1 to this system. This tells us that given any $B$ and $V$ satisfying the constraints, $w(t)$ is bounded for any initial conditions. We now suppose that we have such a $V$ and $B$ and show that (23) holds for all $x_0$.

Consider any initial conditions for (65). Let $y^k = |y^{k+1}|^{\frac{-B}{1-B}} w$, which combined with (66) implies that $y^k$ satisfies the usual tangent dynamics (15). We then have

$$\frac{w^\mathsf{T} Df^{[k]} w}{|w|^2} = \frac{(y^k)^\mathsf{T} Df^{[k]} y^k}{|y^k|^2}. \tag{68}$$

Since $|w(t)|$ is bounded,

$$
\begin{aligned}
0 &\geq \limsup_{T\to\infty} \frac{1}{T} \log \frac{|w(T)|}{|w(0)|} && (69)\\
&= \lim_{T\to\infty} \frac{1}{T} \int_0^T \frac{\mathrm{d}}{\mathrm{d}t} \log \frac{|w(t)|}{|w(0)|} \mathrm{d}t && (70)\\
&= \limsup_{T\to\infty} \frac{1}{T} \int_0^T \frac{1}{2} \left( \frac{\frac{\mathrm{d}|w|^2}{\mathrm{d}t}}{|w|^2} \right) \mathrm{d}t && (71)\\
&= \limsup_{T\to\infty} \frac{1}{T} \int_0^T \left( \frac{w^\mathsf{T}(Df)^{[k]}w}{|w|^2} + \frac{B}{1-B} \frac{(y^{k+1})^\mathsf{T}(Df)^{[k+1]}y^{k+1}}{|y^{k+1}|^2} \right) \mathrm{d}t. && (72)
\end{aligned}
$$

Using (68) and (26), this gives

$$
0 \leq \limsup_{t\to\infty} \frac{1}{t} \log |y^k(t)| + \frac{B}{1-B} \limsup_{t\to\infty} \frac{1}{t} \log |y^{k+1}(t)|. \tag{73}
$$

Since this holds for all $y^{k+1}$ and $y^k$, by (27) we must have

$$
0 \leq M_k(x_0) + \frac{B}{1-B} M_{k+1}(x_0). \tag{74}
$$

which rearranges to

$$
B(M_k(x_0) - M_{k+1}(x_0)) - M_k(x_0) \geq 0. \tag{75}
$$

Remark 2.5 finishes the proof. □

### 2.4 Sum-of-squares relaxations

In sections 2.2 and 2.3 we have formulated upper bounds on LE sums and on Lyapunov dimension that involve pointwise inequality constraints on functions over sets or, equivalently, pointwise suprema of functions over sets. All of these bounds can be expressed as solutions to optimization problems with pointwise nonnegativity constraints. When all expressions are polynomial, nonnegativity can be strengthened into polynomial SOS conditions. This approach is also the basis of the previous paper [34], and we briefly recall the main concepts here.

In order for a function $S : \mathbb{R}^n \to \mathbb{R}$ to satisfy $S(x) \geq 0$ for all $X \in \mathbb{R}^n$, it suffices that $S \in \Sigma$, where $\Sigma$ denotes the set of polynomials that are representable as a sum of squares of other polynomials. The set of nonnegative polynomials strictly contains $\Sigma$, and the relationship between these sets has been extensively studied [43]. Often one wants to enforce $S(x) \geq 0$ on a subset of $\mathbb{R}^n$ without requiring nonnegativity on all of $\mathbb{R}^n$. For instance, one may be interested in a set $\Omega$ with a 'semiagebraic' definition in terms of polynomial equalities and inequalities,

$$
\Omega = \{X \in \mathbb{R}^n \ : \ g_i(X) \geq 0 \text{ for } i = 1, \ldots, I, \ h_j(X) = 0 \text{ for } j = 1, \ldots, J\}, \tag{76}
$$

where the $g_i \in \mathbb{R}[X]$ and $h_i \in \mathbb{R}[X]$ are given. Here $\mathbb{R}[X]$ denotes the ring of polynomials in $X$ with real coefficients. There are various ways to define a set $\Sigma^\Omega$ of polynomials whose nonnegativity on $\Omega$ is implied by SOS conditions. Here we choose a set of polynomials that is called the quadratic module generated by the polynomials that define $\Omega$,

$$\Sigma^\Omega = \left\{ \sigma_0 + \sum_{i=1}^{I} \sigma_i g_i + \sum_{j=1}^{J} \rho_j h_j \ : \ \sigma_i \in \Sigma \text{ for } i = 0, \dots, I, \ \rho_j \in \mathbb{R}[X] \text{ for } j = 1, \dots, J \right\}, \tag{77}$$

where all $\mathbb{R}[X]$ denotes the set of all polynomial functions of $X$ with real coefficients. Any polynomial in $\Sigma^\Omega$ must be nonnegative on $\Omega$ because $\sigma_0(X) \geq 0$ at every $X \in \mathbb{R}^n$, while each $\sigma_i(X)g_i(X) \geq 0$ and each $\rho_j(X)h_j(X) = 0$ at every $X \in \Omega$. When $\Omega = \mathbb{R}^n$, there are no $g_i$ or $\rho_j$, so $\Sigma^\Omega = \Sigma$. In the preceding paper [34], the condition $S \in \Sigma^\Omega$ was expressed in a different but equivalent way: by saying there exist $\sigma_i$ and $\rho_j$ such that $S - \sum \sigma_i g_i - \sum \rho_j h_j$ is SOS.

Computational formulations enforce membership in $\Sigma^\Omega$ via a stronger constraint of membership in a finite-dimensional subspace. The simplest choice is to enforce membership in $\Sigma^\Omega_\nu$, which the truncation of the quadratic module (77) where each $\sigma_i$ and $\rho_j$ has polynomial degree no larger than $\nu$. This gives a hierarchy of computational formulations, where raising $\nu$ cannot worsen results–and typically improves them–but makes computations more expensive. In practice, it is often useful to choose the spaces for $\sigma_i$ and $\rho_i$ more carefully. In our computations of section 3, for instance, we impose symmetries and allow different maximum degrees in different variables. For simplicity in the present subsection, however, we write formulations in terms of the simple truncation $\Sigma^\Omega_\nu$.

The upper bounds on LE sums and Lyapunov dimension formulated in sections 2.2 and 2.3 are stated as, or can be reformulated as, minimizations subject to pointwise inequalities on $V$. When the ODE right-hand side $f(x)$ is polynomial, and when we restrict to auxiliary functions $V$ in some finite-dimensional space of polynomials, then the pointwise inequalities can be enforced using SOS conditions. This gives SOS optimization problems that are computationally tractable when they are not too large, and whose minima give upper bounds on the desired quantities. For the shifted spectrum approach of section 2.3, the SOS relaxations are immediate after choosing a polynomial space for $V$ and enforcing nonnegativity on some set $\Omega$ via membership in $\Sigma^\Omega$. For the sphere projection approach of section 2.2, the relevant minmax problems can be expressed as constrained minimizations, after which the SOS relaxations are simple.

#### 2.4.1 Sphere projection

The sphere projection approach of section 2.2 is based on the general upper bound (5) on time averages, which can be expressed as

$$\max_{X(0)\in\Omega} \overline{\Phi} \leq \inf_{V \in C^1(\Omega)} B \quad \text{s.t.} \quad B - \Phi - F \cdot DV \geq 0 \ \ \forall X \in \Omega. \tag{78}$$

An SOS relaxation of the right-hand side gives an upper bound, stated in the first part of proposition 2.13 below. Furthermore, this upper bound is an equality if $\Omega$ is forward invariant and its semialgebraic specification (76) has the Archimedean property, as stated in the second part of proposition 2.13 below. The Archimedean property implies compactness of $\Omega$ and is only slightly stronger; any specification of compact $\Omega$ that is not Archimedean can be given this property by adding the polynomial $g_i(X) = R^2 - |X|^2$ with a constant $R$ that is large enough to not change $\Omega$ itself. A proof of the second part of proposition 2.13 can be found in [19, Theorem 1]. Briefly, since $\Omega$ is compact and forward invariant, equality in (78) is given by the main theorem of [46], whose applicability on compact manifolds is addressed after (8) in [34]. Then, the Archimedean property guarantees that $\Sigma^\Omega$ contains $B - \Phi - F \cdot DV$ for each suboptimal $B$, by Putinar's Positivstellensatz [41, Lemma 4.1].

**Proposition 2.13.** *Let $\frac{\mathrm{d}}{\mathrm{d}t}X(t) = F(X(t))$ with $F \in \mathbb{R}^n[X]$, and let $\Phi \in \mathbb{R}[X]$. Let the dynamics be forward invariant on a semialgebraic set $\Omega \subset \mathbb{R}^n$, whose specification defines the quadratic module $\Sigma^\Omega$.*

1. *If each trajectory $X(t)$ in $\Omega$ is bounded for $t \geq 0$, then*

$$\max_{X_0 \in \Omega} \overline{\Phi} \leq \inf_{V \in \mathbb{R}[X]} B \quad s.t. \quad B - \Phi - F \cdot DV \in \Sigma^\Omega. \tag{79}$$

2. *If $\Omega$ is compact and its semialgebraic specification is Archimedean, then,*

$$\max_{X_0 \in \Omega} \overline{\Phi} = \inf_{V \in \mathbb{R}[X]} B \quad s.t. \quad B - \Phi - F \cdot DV \in \Sigma^\Omega. \tag{80}$$

All of the upper bounds from section 2.2 can be relaxed into SOS optimization problems by the same reasoning that leads from (5) to (79). Proposition 2.15 below is an SOS relaxation of proposition 2.6 for bounding LE sums. Computations giving such bounds also imply bounds on Lyapunov dimension via proposition 2.7, but potentially sharper bounds on dimension can be computed directly using proposition 2.16, which is an SOS relxation of proposition 2.8. In the proposition statements, $\mathbb{R}^n[x]$ denotes the space of polynomials in $x$ taking vector values in $\mathbb{R}^n$, and the $x(t)$ domain is a semialgebraic set,

$$\mathcal{B} = \{x \in \mathbb{R}^n \ : \ g_i(x) \geq 0 \text{ for } i = 1, \ldots, I, \ h_j(x) = 0 \text{ for } j = 1, \ldots, J\}. \tag{81}$$

This is understood to include the $\mathcal{B} = \mathbb{R}^n$ case when there are no $g_i$ or $h_j$. The $k = 1$ special case of proposition 2.15 is essentially Proposition 2 of [34], although the SOS conditions are more explicit in [34], and the restriction to finite-dimensional polynomial spaces is more explicit here.

**Remark 2.14.** *In the following propositions 2.15 to 2.18, for fixed $k \in \{1, \ldots, n\}$ and polynomial degrees $\nu$ and $\nu'$, the computation of the resulting upper bound $C_i(k, \nu, \nu')$ is an SOS program that can be solved using standard software (see section 3). For certain* *combinations $(k,\nu,\nu')$, there will not exist any $V$ in the chosen space of polynomials that satisfy the SOS constraints, in which case we understand the infimum over bounds to be $+\infty$, meaning that we obtain no upper bound. When bounding Lyapunov dimension, choosing $k > d_L - 1$ is necessary for the constraint set to not be empty. In all cases, raising the polynomial degrees $\nu$ and $\nu'$ cannot shrink the constraint set and will often enlarge it.*

**Proposition 2.15.** *Let $\frac{\mathrm{d}}{\mathrm{d}t}x(t) = f(x(t))$ with $f \in \mathbb{R}^n[x]$, for which a semialgebraic set $\mathcal{B}$ is forward invariant. Let $\Sigma^\Omega$ be the quadratic module* (77) *for the set $\Omega = \mathcal{B} \times \mathbb{S}^{n_k}$ whose semialgebraic specification includes the polynomial constraints on $x$ in* (81) *and the equality $\left|z^k\right| - 1 = 0$. Let $\Phi_k(x,z^k)$ and $\ell_k(x,z^k)$ be defined as in* (28) *and* (30)*, respectively.*

1. *Suppose that each trajectory $x(t)$ in $\mathcal{B}$ is bounded for $t \geq 0$. For each $k \in \{1,\ldots,n\}$, and each pair of polynomial degrees $(\nu,\nu')$, the maximum leading LE sum $M_k$ among trajectories in $\mathcal{B}$ is bounded above by*
$$\sup_{x_0\in\mathcal{B}} M_k(x_0) \leq C_1(k,\nu,\nu'), \tag{82}$$
*where*
$$C_1(k,\nu,\nu') = \inf_{V\in\mathbb{R}[x,z^k]_\nu} B \quad s.t. \quad B - \Phi_k - f\cdot D_x V - \ell_k \cdot D_{z^k} V \in \Sigma^\Omega_{\nu'}. \tag{83}$$

2. *Suppose further that the invariant set $\mathcal{B}$ is compact, and its semialgebraic specification is Archimedean. Then, for each $k$,*
$$\sup_{x_0\in\mathcal{B}} M_k(x_0) = \sup_{\nu,\nu'\in\mathbb{Z}_\geq} C_1(k,\nu,\nu'). \tag{84}$$

*Proof.* The first part is an SOS relaxation of the upper bound (31) on $M_k$ for each $k$. Equivalently, it is a particular case of (79) with $X$, $F$ and $\Phi$ as described above in proposition 2.6, where $V$ is further constrained by having degree no larger than $\nu$, and where $\sigma_i$ and $\rho_j$ in the quadratic module representation (77) have degrees no larger than $\nu'$. The second part is the analogous special case of (80). This equality is applicable because the $\mathcal{B}$ is specified as an Archimedean subset of $\mathbb{R}^n$, and so adding $|z^k| - 1 = 0$ gives an Archimedean specification of $\Omega$ in $\mathbb{R}^{n+n_k}$. All polynomial degrees are admitted in (80), which is equivalent to the supremum over $\nu$ and $\nu'$ in (84). $\square$

**Proposition 2.16.** *Let $\frac{\mathrm{d}}{\mathrm{d}t}x(t) = f(x(t))$ with $f \in \mathbb{R}^n[x]$, for which a semialgebraic set $\mathcal{B}$ is forward invariant, and let the integer $k \leq n$ be such that $\sup_{x_0\in\mathcal{B}} M_{k+1}(x_0) < 0$. Let $\Sigma^\Omega$ be the quadratic module* (77) *for the set $\Omega = \mathcal{B} \times \mathbb{S}^{n_k-1} \times \mathbb{S}^{n_{k+1}-1}$ whose semialgebraic specification includes the polynomial constraints on $x$ in* (81)*, and the equalities $\left|z^k\right| - 1 = 0$ and $\left|z^{k+1}\right| - 1 = 0$. Let $\Phi_k(x,z^k)$ and $\ell_k(x,z^k)$ be defined as in* (28) *and* (30)*, respectively, and likewise for $\Phi_{k+1}(x,z^{k+1})$ and $\ell_{k+1}(x,z^{k+1})$.*

1. *Suppose that each trajectory $x(t)$ in $\mathcal{B}$ is bounded for $t \geq 0$. For each pair of polynomial degrees $(\nu, \nu')$, the the maximum Lyapunov dimension $d_L$ among trajectories in $\mathcal{B}$ is bounded above by*

$$d_L \leq k + C_2(k, \nu, \nu', 0) \tag{85}$$

   *where*

$$C_2(k, \nu, \nu', \varepsilon) = \inf_{\substack{B\in[0,1] \\ V\in\mathbb{R}[x,z^k,z^{k+1}]_\nu}} B \quad s.t. \quad \varepsilon - [\Psi^B + f\cdot D_x V + \ell_k \cdot D_{z^k} V + \ell_{k+1} \cdot D_{z^{k+1}} V] \in \Sigma^\Omega_{\nu'}. \tag{86}$$

2. *Suppose further that the invariant set $\mathcal{B}$ is compact, and its semialgebraic specification is Archimedean. Let $j$ be the minimum admissible $k$, as in definition 2.3 for $d_L$. Then,*

$$d_L = j \; + \inf_{\substack{\nu,\nu'\in\mathbb{Z}_\geq \\ \varepsilon>0}} C_2(j, \nu, \nu', \varepsilon). \tag{87}$$

*Proof.* The first part follows from the upper bound (40) on $d_L$ because the SOS constraint in (86) with $\varepsilon = 0$ implies the nonpositivity constraint on the right-hand side of (40). For the second part, we first establish upper and lower bounds on $C_2$ in terms of $d_L$. For the lower bound, note that the constraint in (86) implies the constraint in the definition (42) of $B_\varepsilon$. This implies $B_\varepsilon \leq C_2(j, \nu, \nu', \varepsilon)$, with which the first inequality in (41) gives

$$d_L \; - \; \varepsilon \left| \sup_{x_0\in\mathcal{B}} M_{j+1}(x_0) \right|^{-1} \leq j + C_2(j, \nu, \nu', \varepsilon) \tag{88}$$

for all $\varepsilon > 0$ and $\nu, \nu' \in \mathbb{Z}_\geq$. Minimizing over these parameters gives

$$d_L \leq j \; + \inf_{\substack{\nu,\nu'\in\mathbb{Z}_\geq \\ \varepsilon>0}} C_2(j, \nu, \nu', \varepsilon) \tag{89}$$

The infimum will be approached or attained as $\nu, \nu' \to \infty$ and $\varepsilon \to 0^+$.

For the upper bound on $C_2$, note that since $\mathcal{B}$ has an Archimedean specification, so does $\Omega$. Thus we can apply the general SOS upper bounds with equality (80), in particular with $F(X)$ being the right-hand side of the ODE system (37), with $\Phi(X)$ being $\Psi^B$, and with $k = j$. This gives

$$\sup_{\substack{x_0\in\mathcal{B} \\ z_0^j\in\mathbb{S}^{n_j-1} \\ z_0^{j+1}\in\mathbb{S}^{n_{j+1}-1}}} \overline{\Psi}^B = \inf_{V\in\mathbb{R}[x,z^j,z^{j+1}]} \varepsilon \quad \text{s.t.} \quad \varepsilon - \left[\Psi^B + f\cdot D_x V + \ell_k \cdot D_{z^j} V + \ell_{j+1} \cdot D_{z^{j+1}} V\right] \in \Sigma^\Omega. \tag{90}$$

Note that the variable called $B$ in (80) is $\varepsilon$ in (90), where $B$ has a different meaning. Let $B$ be such that the constraint holds in expression (47) for $d_L$. This constraint is equivalent to both sides of (90) being nonpositive, and nonpositivity of the right-hand infimum in (90) implies that there exists polynomial $V$ satisfying the SOS constraint for any $\varepsilon > 0$. Since such $V$ exists for any $B > d_L - j$ ,

$$d_L - j \;\geq \inf_{\substack{B\in[0,1] \\ V\in\mathbb{R}[x,z^j,z^{j+1}]}} B \quad \text{s.t.} \quad \varepsilon - \left[\Psi^B + f\cdot D_xV + \ell_k \cdot D_{z^j}V + \ell_{j+1}\cdot D_{z^{j+1}}V\right] \in \Sigma^\Omega \tag{91}$$

for all $\varepsilon > 0$. Restricting the right-hand minimization to $V$ of degree $\nu$ and weights $\sigma_i$ and $\rho_j$ of degree $\nu'$ gives the definition of $C_2(j, \nu, \nu', \varepsilon)$. Thus, for all $\varepsilon > 0$,

$$j \;+ \inf_{\nu,\nu'\in\mathbb{Z}_\geq} C_2(j,\nu,\nu',\varepsilon) \leq d_L. \tag{92}$$

(This inequality need not hold with finite $\nu$ and $\nu'$.) This upper bound of $d_L$, combined with the lower bound of $d_L$ (89), established the claim (87) of the second part and completes the proof. □

The minimization problems defining $C_1$ and $C_2$ for each $(k, \nu, \nu')$ in (82) of proposition 2.15 are SOS programs. That is, polynomials of finite degree are constrained to be SOS, and these polynomials as well as the optimization objective are affine in the optimization parameters. Here the optimization parameters include $B$ as well as the coefficients of $V$ in whatever basis it is represented. Such SOS programs are computationally tractable so long as the dimension of $X$ and the degree of polynomials in the SOS constraints are not too large. Increasing $\nu$ and $\nu'$ gives a hierarchy of SOS programs with increasing computational cost whose minima cannot increase. Typically these minima decrease as $\nu$ and $\nu'$ are raised, and the second part of each theorem guarantees convergence to the sharp bound under mild conditions. More generally, a hierarchy of SOS programs can be defined by choosing different sequences of finite-dimensional polynomial spaces, rather than simply truncating by maximum degrees of $\nu$ and $\nu'$. This often includes imposing symmetries on each polynomial subspace due to symmetries of $f(x)$, as explained in section 2.5 below.

### 2.4.2 Shifted spectrum

Next we give the SOS formulations based on the shifted spectrum approach. Proposition 2.17 below is an SOS relaxation of proposition 2.11 for bounding LE sums. Computations giving such bounds also imply bounds on Lyapunov dimension via proposition 2.7, but potentially sharper bounds on dimension can be computed directly using proposition 2.18, which is an SOS relaxation of proposition 2.12. Unlike our other methods, we do not prove that proposition 2.18 gives sharp results in general.

**Proposition 2.17.** *Let $\frac{\mathrm{d}}{\mathrm{d}t}x(t) = f(x(t))$ with $f \in \mathbb{R}^n[x]$, for which a semialgebraic set $\mathcal{B}$ is forward invariant. Let $m_k^B(x, w^k)$ be defined as in* (56)*. Let $\Sigma^\Omega$ be the quadratic module* (77) *for the set $\Omega = \mathcal{B} \times \mathbb{R}^{n_k}$ whose semialgebraic specification includes the same constraints* (81) *defining $\mathcal{B}$.*

1. *For each $k \in \{1, \dots, n\}$, and each pair of polynomial degrees $(\nu, \nu')$, the maximum leading LE sum $M_k$ among trajectories in $\mathcal{B}$ is bounded above by*

$$\sup_{x_0 \in \mathcal{B}} M_k(x_0) \le C_3(k, \nu, \nu'), \tag{93}$$

*where*

$$C_3(k,\nu,\nu') = \inf_{V \in \mathbb{R}[x,w^k]_\nu} B \quad s.t. \quad \begin{aligned} V - |w^k|^2 &\in \Sigma^\Omega_{\nu'} \\ -[f \cdot D_x V + m_k^B \cdot D_{w^k} V] &\in \Sigma^\Omega_{\nu'}. \end{aligned} \tag{94}$$

2. *Suppose that the invariant set $\mathcal{B}$ is compact, and that its semialgebraic specification is Archimedean. Then, for each $k$,*

$$\sup_{x_0 \in \mathcal{B}} M_k(x_0) = \inf_{\nu,\nu' \in \mathbb{Z}_\ge} C_3(k,\nu,\nu'), \tag{95}$$

*where the infimum in (94) is over $V(x,w^k) = (w^k)^\mathsf{T} U(x) w^k$ with $U \in \mathbb{R}^{n_k \times n_k}[x]$ a symmetric polynomial matrix, and where $\sigma_i(x,w^k)$ and $\rho_j(x,w^k)$ in the SOS representation (77) are likewise quadratic in $w^k$.*

*Proof.* The first part follows from the upper bound (64) on $M_k$ because the SOS constraints in (94) imply the nonnegativity constraints in (64) from proposition 2.11. For the second part, let $B_1$ and $B_2$ be arbitrary with $B_1 > B_2 > \sup_{x_0 \in \mathcal{B}} M_k(x_0)$. It suffices to show that there exists a polynomial matrix $U(x)$ such that $V(x,w^k) = (w^k)^\mathsf{T} U(x) w^k$ satisfies the SOS constraints of (94) for $B = B_1$ with no restrictions on polynomial degree. Since $B_2 > \sup_{x_0 \in \mathcal{B}} M_k(x_0)$, the second part of proposition 2.11 guarantees the existence of a symmetric $U \in C^1(\mathbb{R}^n, \mathbb{R}^{n_k \times n_k})$ for which $V$ satisfies the pointwise inequalities in (64) with $B = B_2$. We will show that this implies existence of the desired polynomial $U$ under the present assumptions.

The pointwise inequalities in (64), with $B = B_2$ in the $V = (w^k)^\mathsf{T} U w^k$ case, can be written in the form $(w^k)^\mathsf{T} \hat{Z}_i(x) w^k \ge 0$, where

$$\hat{Z}_1(x) = U(x) - I_{n_k}, \tag{96}$$
$$\hat{Z}_2(x) = -f(x) \cdot D_x U(x) + 2\, U(x)(B_2\, I_{n_k} - Df^{[k]}(x)), \tag{97}$$

with $I_{n_k}$ denoting the $n_k \times n_k$ identity matrix. The pointwise inequalities on $\mathcal{B} \times \mathbb{R}^{n_k}$ are equivalent to the condition that both $\hat{Z}_i(x) \succeq 0$ for all $x \in \mathcal{B}$. In fact, by scaling $U$ we can satisfy both of these equations with the additional restriction that $\hat{Z}_1 \succ 0$ strictly on $\mathcal{B}$. Let

$$Z_1(x) = U(x) - I_{n_k}, \tag{98}$$
$$Z_2(x) = -f(x) \cdot D_x U(x) + 2\, U(x)(B_1\, I_{n_k} - Df^{[k]}(x)), \tag{99}$$

which are the same definitions with $B_2$ replaced by $B_1$. Since $B_1 > B_2$, we now have both $Z_i(x) \succ 0$ strictly for all $x \in \mathcal{B}$.

Because there exists $U \in C^1$ such that both $Z_i(x) \succ 0$ on $\mathcal{B}$, and $\mathcal{B}$ is compact, there also exists a polynomial matrix $U \in \mathbb{R}^{n_k \times n_k}[x]$ for which both $Z_i(x) \succ 0$ on $\mathcal{B}$. This follows from a strengthened version of the Stone-Weierstrass approximation theorem [25]. This polynomial $U$ satisfies the SOS constraints of (94) due to $\mathcal{B}$ being Archimedean and $Z_i$ being strictly positive definite on $\mathcal{B}$. Recalling that $\mathcal{B}$ is the set where all $g_i(x) \geq 0$ and $h_j(x) = 0$, applying a matrix Positivstellensatz [44, theorem 2] guarantees existence of polynomial matrices $S_i(x)$ and $T_i(x)$, with the $S_i(x)$ being squares of other polynomial matrices, such that

$$Z_1(x) = S_0(x) + \sum_{i=1}^{I} g_i(x) S_i(x) + \sum_{j=1}^{J} h_j(x) T_j(x), \tag{100}$$

and likewise for $Z_2$. The existence of such representations implies that both $(w^k)^\mathsf{T} Z_i w^k$ belong to the quadratic module $\Sigma^\Omega$. Therefore $V = (w^k)^\mathsf{T} U w^k$ satisfies the constraints of (94) with $B = B_1$, and the second part of the proposition is proven. □

**Proposition 2.18.** *Let $\frac{\mathrm{d}}{\mathrm{d}t} x(t) = f(x(t))$ with $f \in \mathbb{R}^n[x]$, for which a semialgebraic set $\mathcal{B}$ is forward invariant on $\mathcal{B} \subset \mathbb{R}^n$ with each trajectory $x(t)$ bounded for $t \geq 0$. Let the positive integer $k \leq n$ be such that $\sup_{x_0 \in \mathcal{B}} M_{k+1}(x_0) < 0$. Let $\Sigma^\Omega$ be the quadratic module (77) for the set $\Omega = \mathcal{B} \times \mathbb{R}^{n_{k+1}} \times \mathbb{R}^{n_k}$ whose semialgebraic specification includes the polynomial constraints on $x$ in (81). Let each trajectory $x(t)$ in $\mathcal{B}$ be bounded for $t \geq 0$. For each $k \in \{1, \ldots, n\}$, and each pair of polynomial degrees $(\nu, \nu')$, the the maximum Lyapunov dimension $d_L$ among trajectories in $\mathcal{B}$ is bounded above by*

$$d_L \leq k + C_4(k, \nu, \nu'), \tag{101}$$

*where*

$$C_4(k, \nu, \nu') = \inf_{\substack{B \in [0,1] \\ V \in \mathbb{R}[x, y^{k+1}, w]_\nu}} B \quad s.t. \quad \begin{aligned} V - |w|^2 &\in \Sigma^\Omega_{\nu'} \\ (1 - B) \left|y^{k+1}\right|^2 S_k^B &\in \Sigma^\Omega_{\nu'}. \end{aligned} \tag{102}$$

## 2.5 Symmetry exploitation

If an ODE being studied using auxiliary function methods has symmetry, then in many cases the same symmetry can be imposed on the auxiliary functions, leading to symmetries in SOS programs that can be exploited computationally by block diagonalizing the corresponding semidefinite programs (SDPs) [11]. More precisely, an ODE system (1) with right-hand side $F(X)$ is invariant under a map $\Lambda : \mathbb{R}^n \to \mathbb{R}^n$ if and only if $F$ is $\Lambda$-equivariant, meaning that $F(\Lambda X) = \Lambda F(X)$ for all $X$. In such cases, $X(t)$ solves the ODE if and only if $\Lambda X(t)$ does. The trajectory itself may be $\Lambda$-invariant, meaning that $X(t) = \Lambda X(t)$, or it may not be. For any ODE, the set of such transformations forms a group. We focus on the common situation where each $\Lambda$ is an orthogonal linear transformation on $\mathbb{R}^n$, meaning the group of such $\Lambda$ is a subgroup of the orthogonal group $O(n)$. (In some contexts the set of $\Lambda$ would be called the linear representation of a group, rather than a group itself, but we do not draw this distinction.) In many formulations of auxiliary function methods for ODEs, if the ODE is invariant under $\Lambda$ then the results are provably unchanged if the optimization is over the restricted set of $V$ which are $\Lambda$-invariant, meaning $V(X) = V(\Lambda X)$ for all $X$. In the context of bounding time averages, see the appendix of [34] for a proof that $V(X)$ can be taken as $\Lambda$-invariant under certain conditions. For our shifted spectrum methods, the Lyapunov function in lemma B.1 can be taken as $\Lambda$-invariant also. We omit the proof, which closely follows the argument in [34].

In order to apply such results in the context of bounding LE sums and Lyapunov dimension, we must determine the symmetries of ODEs on the augmented state spaces for $(x, z^k)$ and $(x, w^k)$, governed by (29) and (55). First, regardless of whether the $x$ ODE has any symmetry, the augmented ODEs have a sign symmetry due to linearity of the $y^k$ dynamics. These sign symmetries are summarized by proposition 2.19. Second, if the right-hand side $f(x)$ of the $x$ ODE is equivariant under an orthogonal transformation, this induces equivariance on the right-hand sides of the augmented ODEs. This is summarized by proposition 2.20, which invokes the $k^{th}$ multiplicative compound $\Lambda^{(k)}$ of $\Lambda$ (cf. appendix A). Note that $I_n$ denotes the $n \times n$ identity matrix, and $\mathrm{diag}(A, \dots, B)$ is a block diagonal matrix with square blocks $A, \dots, B$.

**Proposition 2.19.** *Let $\frac{\mathrm{d}}{\mathrm{d}t}x(t) = f(x(t))$ with $f : \mathcal{B} \to \mathbb{R}^n$. The augmented ODEs are invariant under the following orthogonal transformations.*

1. *The ODE system* (29) *is invariant under* $\mathrm{diag}(I_n, -I_{n_k})$.
2. *The ODE system* (37) *is invariant under* $\mathrm{diag}(I_n, \pm I_{n_k}, \mp I_{n_{k+1}})$.
3. *The ODE system* (55) *is invariant under* $\mathrm{diag}(I_n, -I_{n_k})$.
4. *The ODE system* (65) *is invariant under* $\mathrm{diag}(I_n, \pm I_{n_{k+1}}, \mp I_{n_k})$.

*Proof.* All four parts have very similar proofs, so we prove only the last. The claim amounts to the right-hand side of (65) being equivariant under negation of $y^{k+1}$. For $f(x)$ this statement is trivial. The requirement that $Df^{[k+1]}(x)(-y^{k+1}) = -Df^{[k+1]}(x)y^{k+1}$ is also immediate. The final requirement that $S_k^B(x, -y^{k+1}, w) = S_k^B(x, y^{k+1}, w)$ holds because $y^{k+1}$ appears quadratically in both the numerator and denominator in (66). □

**Proposition 2.20.** *Let $\mathcal{B} \subset \mathbb{R}^n$ be invariant under $\Lambda \in O(n)$. If $\frac{\mathrm{d}}{\mathrm{d}t}x(t) = f(x(t))$ with $f : \mathcal{B} \to \mathbb{R}^n$ is invariant under $\Lambda$, then the augmented ODEs are invariant under the following orthogonal transformations.*

1. *The ODE system* (29) *is invariant under* $\mathrm{diag}(\Lambda, \Lambda^{(k)})$.

2. *The ODE system* (37) *is invariant under* $\mathrm{diag}(\Lambda, \Lambda^{(k)}, \Lambda^{(k+1)})$.

3. *The ODE system* (55) *is invariant under* $\mathrm{diag}(\Lambda, \Lambda^{(k)})$.

4. *The ODE system* (65) *is invariant under* $\mathrm{diag}(\Lambda, \Lambda^{(k+1)}, \Lambda^{(k)})$.

*Proof.* By assumption $f(x)$ is equivariant under the orthogonal transformation $\Lambda$, meaning $f(\Lambda x) = \Lambda f(x)$. The multiplicative compound $\Lambda^{(k)}$ is orthogonal by proposition A.2, thus all transformations in the proposition are orthogonal. The claim is proved by showing that the right-hand sides of the augmented ODEs are equivariant under the claimed transformations. All four parts are very similar, so we prove only the first. For this we must show that $\ell_k(\Lambda x, \Lambda^{(k)} z^k) = \Lambda^{(k)} \ell_k(x, z^k)$. It is sufficient to show that

$$Df^{[k]}(\Lambda x)\Lambda^{(k)} z^k = \Lambda^{(k)} Df^{[k]}(x) z^k, \tag{103}$$

so that

$$\Phi_k(\Lambda x, \Lambda^{(k)} z^k) = \Phi_k(x, z^k) \tag{104}$$

by orthogonality of $\Lambda^{(k)}$. Oeri and Goluskin [34] showed that $Df(\Lambda x) = \Lambda Df(x)\Lambda^{\mathsf{T}}$. Taking the $k$th additive compound of this equation gives, via proposition A.3,

$$Df^{[k]}(\Lambda x) = \Lambda^{(k)} Df^{[k]}(x) \left(\Lambda^{(k)}\right)^{\mathsf{T}}. \tag{105}$$

Orthogonality then gives (103). The proofs of the other parts follow similarly, noting that (103) remains valid with $y^k$ or $w^k$ in place of $z^k$. □

Combining both propositions gives any finite symmetry group for the augmented systems. For example, if $\mathcal{B}$ is invariant and $f$ is equivariant under a group generated by $\{\Lambda_1, \dots, \Lambda_j\}$ then the right-hand sides of the augmented systems (29) and (55) are equivariant under the group generated by

$$\left\{\mathrm{diag}(\Lambda_1, \Lambda_1^{(k)}), \dots, \mathrm{diag}(\Lambda_j, \Lambda_j^{(k)}), \mathrm{diag}(I_n, -I_{n_k})\right\}, \tag{106}$$

and we can restrict to $V$ that are invariant under the same transformations. This can be enforced by representing $V$ in a symmetry-invariant basis. For sign symmetries this can be done with monomial bases, but general rotations require more sophisticated choices of symmetry-invariant polynomial basis. Note that for the shifted spectrum method proposition 2.17 it is sufficient to consider $V$ that are is quadratic in $w^k$, which automatically gives invariance of $V$ with respect to the symmetry of proposition 2.19.

## 3 Computational examples

We now present computational examples in which our methods based on SOS optimization are applied to particular ODEs. It has been conjectured that, for all systems with equilibria embedded in chaotic attractors, the supremum in the formula for the Lyapunov dimension (18) is attained on an equilibrium [9, 22]. This indeed is the case for many simple systems, including Lorenz's 1963 model [23]. In order to demonstrate the power of our methods, we focus on two example systems for which we know there are no equilibria.

For all of our examples, the code to replicate the results, implemented in Julia, is available at `https://github.com/jeremypparker/Parker_Goluskin_LEs`. We used SumOfSquares.jl [20] to reformulate SOS optimization problems as SDPs, along with SymbolicWedderburn.jl [15] for symmetry exploitation. To solve the resulting SDPs, we primarily used Mosek [1]. Certain SDPs were also solved using Clarabel.jl [14] and Hypatia.jl [5], which gave very similar results with longer runtimes. In applications of the sphere projection approach, the bounds $B$ can be minimized as the objective of the SDPs, and the bounds we report are the numerical optima returned by Mosek. In applications of the shifted spectrum approach, the value of $B$ must be fixed during each SDP solution because it multiplies coefficients of $V$ in the SOS constraints, and the coefficients to be optimized over can appear only linearly. In such computations, if there exists a polynomial $V$ satisfying the SOS constraints, then the SDP will have a nonempty constraint set, and the SDP solver should report *feasibility*. We thus perform a bisection search in $B$ to find the boundary between large $B$ values where the SDPs are feasible and small $B$ values where they are infeasible. Close to this boundary the numerical conditioning of the SDPs becomes increasingly poor, making it difficult to precisely identify the infimum over admissible $B$. Confirming that near-optimal bounds are valid calls for rigorous and/or multiple-precision computations [12, 37], but for simplicity we have used only double precision arithmetic in our examples. We call a bound 'feasible' if Mosek reports a dual objective value of less than $10^{-6}$, 'infeasible' if the final dual objective value is greater than 1 and 'indeterminate' otherwise, but we stress that these computations are not rigorously validated.

To confirm that accuracy of our computed bounds, we computed various period orbits and the LE sums along each of these orbits. We used the Tsit5 solver from the OrdinaryDiffEq.jl package [42] to solve integrate the ODE systems, find periodic orbits and compute LEs. For a given periodic orbit of period $T$, the LEs $\mu_k$ were found using the eigenvalues of the Jacobian $D\phi^T$. To compute the sums $M_k$, we simply sum the $\mu_k$ in descending order.

### 3.1 The Duffing Oscillator

The Duffing oscillator is a second-order nonautonomous ODE governing amplitude of oscillations that are damped and sinusoidally forced. Different versions exist, but we choose

$$\ddot{A} + \delta\dot{A} + \beta A + \alpha A^3 = \gamma\cos\omega t \tag{107}$$

with parameters $(\alpha, \beta, \gamma, \delta, \omega) = (1, -1, 0.3, 0.2, 1)$. We numerically estimate the Lyapunov exponents on the chaotic attractor to be approximately 0.158 and $-0.358$, which gives a local Lyapunov dimension of $d_L \approx 1.441$. Kuznetsov and Reitmann [18] derive a rigorous bound on the Lyapunov dimension of the global attractor in the Duffing oscillator which, for these parameter values, gives a bound $d_L \leq 1.9910$. Note that we have not defined LEs or Lyapunov dimension for nonautonomous systems. This is straightforward to do, but instead we convert (107) to an autonomous system.

The ODE (107) is not a polynomial system, as required for the formulations presented in section 2, but it can be transformed into a polynomial system by a standard trick (cf. Parker et al. [38]). We let $x_1 = A$ and $x_2 = \dot{A}$, and we introduce $x_3 = \sin \omega t$ and $x_4 = \cos \omega t$. Then the $x(t)$ vector evolves according to

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \begin{bmatrix} x_2 \\ -\delta x_2 - \beta x_1 - \alpha x_1^3 + \gamma x_4 \\ \omega x_4 \\ -\omega x_3 \end{bmatrix}. \tag{108}$$

Recalling that $x_3$ and $x_4$ are sine and cosine of the same argument, we are interested only in dynamics on the manifold

$$\mathcal{B} = \{x \in \mathbb{R}^4 \,|\, x_3^2 + x_4^2 = 1\}. \tag{109}$$

Since only odd powers of $x$ appear in (108), the right-hand side is equivariant under the simple sign symmetry $\Lambda = -I$, and $\mathcal{B}$ is invariant, so we can also impose that our polynomials $V$ are invariant under the same symmetry. This is exploited in our code using the methods described in section 2.5.

For the system (108), the Jacobian is

$$Df(x) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\beta - 3\alpha x_1^2 & -\delta & 0 & \gamma \\ 0 & 0 & 0 & \omega \\ 0 & 0 & -\omega & 0 \end{bmatrix}. \tag{110}$$

From this, we compute the additive compound matrices

$$Df^{[2]}(x) = \begin{bmatrix} -\delta & 0 & \gamma & 0 & 0 & 0 \\ 0 & 0 & \omega & 1 & 0 & 0 \\ 0 & -\omega & 0 & 0 & 1 & 0 \\ 0 & -\beta - 3\alpha x_1^2 & 0 & -\delta & \omega & \gamma \\ 0 & 0 & -\beta - 3\alpha x_1^2 & \omega & -\delta & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{111}$$

$$Df^{[3]}(x) = \begin{bmatrix} -\delta & \omega & -\gamma & 0 \\ -\omega & -\delta & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -\beta - 3\alpha x_1^2 & -\delta \end{bmatrix}, \tag{112}$$

$$\text{and} \quad Df^{[4]}(x) = -\delta. \tag{113}$$

In this case, the fourth additive compound matrix—the trace of the Jacobian—is independent of $x$, and therefore either of our methods gives the sharp result that $M_4 = -\delta$ on any orbit. This implies that the dimension $d_L$ is optimized on the same orbit as the optimizer for $M_3$ and, given a sharp bound for $M_3$, proposition 2.7 will give a sharp bound for $d_L$. Furthermore, since there are no equilibria, every trajectory must have a zero LE associated with the flow, and an additional vanishing LE corresponding to amplitude change in the $x_3$ and $x_4$ variables. Therefore $M_1 = M_2 = M_3$ for all orbits of interest, and it suffices to compute bounds on $M_1$ only in order to deduce bounds for $M_2$, $M_3$ and $d_L$. Nevertheless, to test our methods we computed bounds using our formulations for $M_1$, $M_2$, $M_3$ and $d_L$.

To formulate SOS optimization problems, we must choose finite dimensional spaces for the polynomials to be optimized over. In all cases we choose a maximum degree of $V$ and then construct a basis from symmetry-invariant monomials. For each SOS constraint, the restriction that $h_j(x) = 0$, where $h_j = x_3^2 + x_4^2 - 1$, is enforced via additional unknown polynomials $\rho_j$, whose maximum degrees are chosen so that $h_j \rho_j$ matches the maximum degree of other terms in the expression. See the accompanying code for full details.

For the sphere projection methods proposition 2.15 and proposition 2.16, computational results are reported in tables 1 and 2 respectively. In both cases, clearly converged bounds are achieved with $V$ of degree 4 in $x$ and 2 in $z^k$. For the shifted spectrum methods proposition 2.17 and proposition 2.18, numerical errors are large when the degree of $V$ in $x$ is 4 or smaller, so we let $V$ have degree 6 in $x$. In this case, we find feasible bounds $M_1 \leq 0.887898$ and $d_L \leq 3.8162$, and we find infeasible bounds when each final digit is decreased by 1. These bounds perfectly match those found via the sphere projection methods.

To find an orbit that maximises the leading LE $M_1$, and consequently $d_L$ also, we follow a strategy similar to that of Tobasco et al. [46]. We restrict to a Poincaré section of $\mathcal{B}$ with $x_3 = 0$ and $x_4 = 1$, then we attempt to find the global minimum of the expression

$$B - \Phi_k - f \cdot D_x V - \ell_k \cdot D_{z^k} V \tag{114}$$

| $d_x$ | $d_z$ | $k=1$ | $k=2$ | $k=3$ | $k=4$ |
|---|---|---|---|---|---|
| 2 | 0 | 27.3625 | 31.8231 | 29.7823 | $\underline{-0.2000}$ |
| 2 | 2 | 14.8295 | 15.5894 | 13.9366 | $\underline{-0.2000}$ |
| 4 | 0 | $\underline{1.}3666$ | $\underline{1.}3668$ | $\underline{1.}3666$ | $\underline{-0.2000}$ |
| 4 | 2 | $\underline{0.8879}$ | $\underline{0.888}0$ | $\underline{0.8879}$ | $\underline{-0.2000}$ |
| 4 | 4 | $\underline{0.8879}$ | $\underline{0.8879}$ | $\underline{0.8879}$ | $\underline{-0.2000}$ |

Table 1: Upper bounds on the sum $M_k$ of the $k$ leading LEs, computed by solving the SOS optimization problem in proposition 2.15 for the Duffing oscillator in the variables of (108). Polynomial degrees are at most $d_x$ in $x$ and $d_z$ in $z^k$. Tabulated values are rounded up to the precision shown, and the underlines indicate the digits that agree (after rounding) with the values computed by numerical integration for the periodic orbit of figure 1. On this orbit $M_4 = 0.2$ and $M_1 = M_2 = M_3 \approx 0.88785$.

| $d_x$ | $d_z$ | Bound |
|---|---|---|
| 0 | 0 | $\underline{4.}0000$ |
| 0 | 2 | $\underline{4.}0000$ |
| 0 | 4 | $\underline{4.}0000$ |
| 2 | 0 | $\underline{3.}9912$ |
| 2 | 2 | $\underline{3.}9774$ |
| 2 | 4 | $\underline{3.}9614$ |
| 4 | 0 | $\underline{3.8}724$ |
| 4 | 2 | $\underline{3.8162}$ |
| 4 | 4 | $\underline{3.8162}$ |
| 6 | 0 | $\underline{3.8}555$ |
| 6 | 2 | $\underline{3.8162}$ |

Table 2: Upper bounds on the Lyapunov dimension $d_L$, computed by solving the SOS optimization problem in proposition 2.16 with $k = 3$ for the Duffing oscillator in the variables of (108). Polynomial degrees are at most $d_x$ in $x$ and $d_z$ in $z^k$ and $z^{k+1}$. Tabulated values are rounded up to the precision shown, and underlines indicate digits that agree (after rounding) with the value $d_L \approx 3.81615$ computed by numerical integration for the periodic orbit of figure 1.

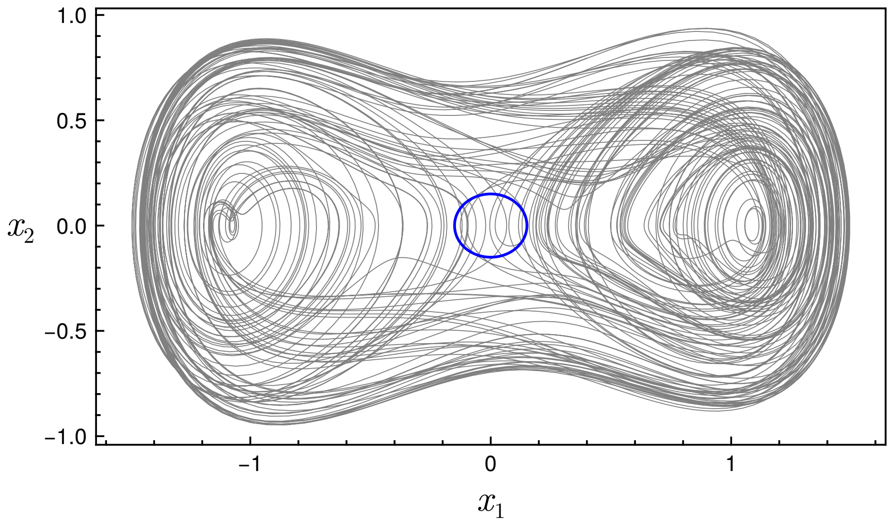

Figure 1: A chaotic trajectory (grey) in the system (108), and the unstable periodic orbit of period $2\pi$ discussed in the text (blue).

in the constraint of (83), where $B$ and $V$ are the approximately optimal values returned by the SDP solver for the sphere projection method. This expression must vanish on average over the optimal orbit, and it is everywhere non-negative, so close to the orbit it must be very small in value. The minimization strategy is to plot, at each $x_1$ and $x_2$ over the approximate range of the chaotic attractor, the minimum of this expression over all $z^k \in \mathbb{S}^{n_k-1}$, which is approximated by random sampling on the unit sphere. This minimum value is then passed to a Newton-shooting algorithm to converge the true periodic orbit, which approximately intersects the point

$$(x_1, x_2, x_3, x_4) = (-0.14984, 0.01520, 0, 1) \tag{115}$$

and has a period of $2\pi$. This is the small, approximately circular periodic orbit depicted in figure 1. We computed LEs on this orbit to be approximately

$$\mu_1 = 0.88785, \quad \mu_2 = 0, \quad \mu_3 = 0, \quad \mu_4 = -1.08785, \tag{116}$$

giving

$$M_1 = M_2 = M_3 = 0.88785, \quad M_4 = -0.2. \tag{117}$$

The Kaplan-Yorke formula (18) then gives $d_L = 3.81615$, given that the supremum must be attained on this orbit. These values perfectly match the bounds given by $C_1$, $C_2$, $C_3$ and $C_4$. In the limit $\gamma \to 0$ of (107), the orbit becomes an equilibrium point at the origin. This matches previous observations that global Lyapunov dimension is attained on equilibria in systems that have such points. Finally, we note that our bounds on $d_L$ given by $C_2$ and $C_4$ of (108) are equivalent to $d_L \leq 1.8162$ for (107), a significant improvement over the bound of Kuznetsov and Reitmann [18], and that this result must be sharp as an equivalent periodic orbit exists within the nonautonomous system (107).

### 3.2 A three-degree-of-freedom Hamiltonian system

For an example where LE sums for different $k$ are maximized on different orbits, we consider a higher-dimensional system whose large number of symmetries keeps the computational cost tractable. In particular, we consider a Hamiltonian system with Hamiltonian function

$$H(x) = \frac{1}{2}|x|^2 + x_1^2x_2^2 + x_2^2x_3^2 + x_3^2x_1^2, \tag{118}$$

the first three components of $x$ being the position variables, and the final three the corresponding momenta. This is a specific form of a triaxially symmetric potential studied by Palacián et al. [35], which has applications in galactic dynamics [4, 6]. The Hamiltonian (118) gives rise to a system of ODEs in 6 variables,

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \begin{bmatrix} x_4 \\ x_5 \\ x_6 \\ -x_1 - 2x_1x_2^2 - 2x_1x_3^2 \\ -x_2 - 2x_2x_1^2 - 2x_2x_3^2 \\ -x_3 - 2x_3x_1^2 - 2x_3x_2^2 \end{bmatrix}. \tag{119}$$

The right-hand side of (119) is equivariant under $x \mapsto \Lambda x$ where $\Lambda$ is any matrix in the subgroup of $O(6)$ generated by

$$\Lambda_1 = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \Lambda_2 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \Lambda_3 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

We consider the fixed energy level $H = 1$. At this value of $H$, there are no equilibria, so $\mu_3 = \mu_4 = 0$ on all orbits. We consider only orbits whose position variables remain in the region where $\frac{1}{2}\left(x_1^2 + x_2^2 + x_3^2\right) + \left(x_1^2x_2^2 + x_2^2x_3^2 + x_3^2x_1^2\right) \leq 1$, depicted in figure 2, and whose full state vector $x$ remains in the compact sphere $\frac{1}{2}|x|^2 \leq 1$. Two families of particularly short periodic orbits, which are also depicted in figure 2, were found by a brute-force search over random initial conditions. The first family, which we call $S_1$, has a period of approximately 5.278 and passes near

$$x = (0.83871, 0, 0, 0, 0, 1.13867),$$

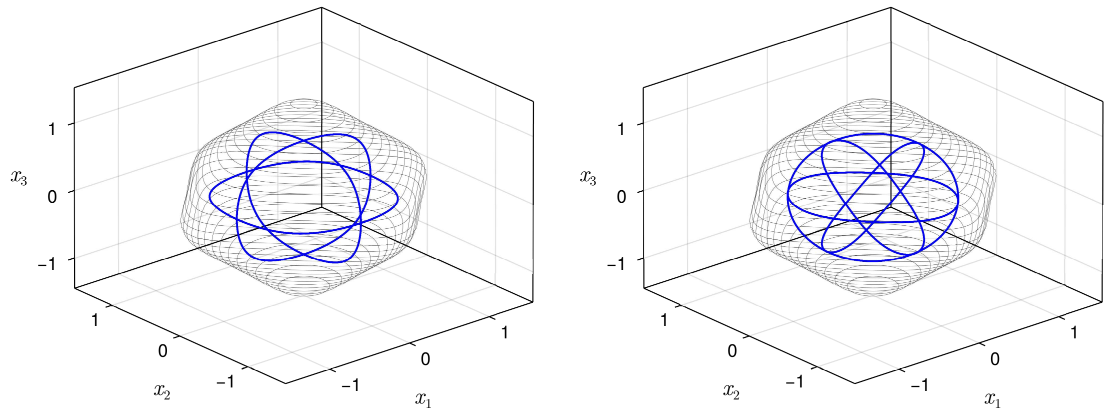


Figure 2: Plots of the position variables $(x_1, x_2, x_3)$ for the Hamiltonian system (118) at the energy level $H = 1$. Left: The family of periodic orbits $S_1$. Right: The family of periodic orbits $S_2$. Trajectories are confined to stay within the region $\frac{1}{2}\left(x_1^2 + x_2^2 + x_3^2\right) + \left(x_1^2x_2^2 + x_2^2x_3^2 + x_3^2x_1^2\right) \le 1$, which is bounded by the shaded surface. On $S_1$, the LE sums are $M_1 = M_2 = M_3 = M_4 = M_5 = 0.42913$ and on $S_2$, $M_1 = M_5 = 0.25919$ and $M_2 = M_3 = M_4 = 0.51837$.

or near points that map to this $x$ under a symmetry. The $S_1$ orbits have LEs of approximately

$$(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6) = (0.42913, 0, 0, 0, 0, -0.42913),$$

corresponding to LE sums of

$$(M_1, M_2, M_3, M_4, M_5, M_6) = (0.42913, 0.42913, 0.42913, 0.42913, 0.42913, 0).$$

The second family of orbits, which we call $S_2$, has a period of approximately 4.857 and passes near

$$x = (-0.57566, 0.57566, 0, 0.39004, 0.39004, 0.90186),$$

or near a symmetry-related point. These $S_2$ orbits have LEs of approximately

$$(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6) = (0.25919, 0.25919, 0, 0, -0.25919, -0.25919),$$

corresponding to sums of

$$(M_1, M_2, M_3, M_4, M_5, M_6) = (0.25919, 0.51837, 0.51837, 0.51837, 0.25919, 0).$$

Because this system is Hamiltonian, volumes in state space are conserved, and our methods for bounding Lyapunov dimension are not relevant. We carried out SOS computations to bound LE sums using the sphere projection and shifted spectrum approaches,

| $d_x$ | $d_z$ | $k=1$ | $k=2$ | $k=3$ | $k=4$ | $k=5$ | $k=6$ |
|---|---|---|---|---|---|---|---|
| 2 | 0 | 1.2361 | 2.0232 | 2.1082 | 2.0232 | 1.2361 | $\underline{0.0000}$ |
| 2 | 2 | $\underline{0.4}843$ | $\underline{0.}9216$ | | $\underline{0.}9216$ | 0.4843 | $\underline{0.0000}$ |
| 4 | 0 | 1.1855 | 2.0046 | 2.0611 | 2.0046 | 1.1855 | $\underline{0.0000}$ |
| 4 | 2 | $\underline{0.4292}$ | | | | $\underline{0.4292}$ | $\underline{0.0000}$ |

Table 3: Upper bounds on the sum $M_k$ of the $k$ leading LEs, computed by solving the SOS optimization problem in proposition 2.15 for the Hamiltonian system (119). Polynomial degrees are at most $d_x$ in $x$ and $d_z$ in $z^k$. Tabulated values are rounded up to the precision shown, and the underlines indicate the digits that agree (after rounding) with the values computed by numerical integration for the periodic orbits $S_1$ ($k = 1, 5, 6$) or $S_2$ ($k = 2, 3, 4$) of figure 2. Values are missing in cases where computations were prohibitively memory-intensive.

| $d_x$ | $k=1$ | $k=2$ | $k=3$ | $k=4$ | $k=5$ | $k=6$ |
|---|---|---|---|---|---|---|
| 2 | $\underline{0.43}11$ | $\underline{0.}6141$ | $\underline{0.}7579$ | $\underline{0.}6141$ | $\underline{0.43}11$ | $\underline{0.0000}$ |
| 4 | $\underline{0.4292}$ | $\underline{0.5184}$ | $\underline{0.5}711$ | $\underline{0.5184}$ | $\underline{0.4292}$ | $\underline{0.0000}$ |
| 6 | $\underline{0.4292}$ | $\underline{0.5184}$ | $\underline{0.5}300$ | $\underline{0.5184}$ | $\underline{0.4292}$ | $\underline{0.0000}$ |

Table 4: Upper bounds on the sum $M_k$ of the $k$ leading LEs, computed by solving the SOS optimization problem in proposition 2.17 for the Hamiltonian system (119). Polynomial degrees are at most $d_x$ in $x$ and 2 in $w^k$. The values in this table are not rounded. The underlines indicate the digits that agree (after rounding) with the values computed by numerical integration for the periodic orbits $S_1$ ($k = 1, 5, 6$) or $S_2$ ($k = 2, 3, 4$) of figure 2.

and tables 3 and 4 give the results. In both cases we add the explicit constraint $|x|^2 \leq 2$ along with $H(x) = 1$ in the definition of the set $\mathcal{B}$ because this improves the numerical results. Using the shifted spectrum approach, we find sharp bounds for $M_1$, $M_2$, $M_4$, $M_5$ and $M_6$. Using the sphere projection approach, we found sharp bounds on $M_1$, $M_5$ and $M_6$. Sharp bounds for the other $k$ values seemed to require larger polynomial degrees that were beyond our computational resources.

Given the Hamiltonian structure of the system, bounding $M_5$ is directly equivalent to bounding $M_1$, bounding $M_4$ is equivalent to bounding $M_2$, and bounding $M_6$ is trivial, since this is always zero. These facts are borne out by the results. Bounding the sum $M_3$ is not *a priori* identical to any other case, but the lack of equilibria means that we expect the true value to be equal to $M_2$ and $M_4$. Furthermore, bounds for $M_3$ require a particularly high-dimensional augmented system of dimension $n + n_k = 26$, which is computationally challenging at even modest polynomial degree despite exploiting all symmetries of the system. This is especially true for the sphere projection method, which has higher polynomial degree in the tangent space dynamics. In table 3, the missing values were not possible to compute when solving SDPs with Mosek due to insufficient memory, despite having 3TB of memory. This could be surmounted by using a first-order SDP solver at the expense of very long runtimes and/or lower precision. In table 4, the $M_3$ result is not sharp at the maximum degree $d_x = 6$ that was computationally feasible.

# 4 Conclusion

We have described two distinct approaches to bounding sums of Lyapunov exponents as well as Lyapunov dimension in ODE systems. The general versions of our formulations require optimizing over auxiliary functions in the class $C^1$. We prove that most of these formulations give sharp bounds under weak assumptions, but the optimization problems are not tractable in general. In the case of ODEs with polynomial right-hand sides, our bounding formulations can be relaxed using SOS constraints, which we again prove are convergent, and which are often computationally tractable using SOS optimization. Carrying out such SOS computations for particular ODEs, we obtained nearly perfect bounds in challenging examples. For one of the examples, the sphere projection method gave much better conditioned SDPs in practice. For the other example, the shifted spectrum method was more appropriate due to onerous memory requirements of the sphere projection method. In section 3.1, we demonstrated how the results of our bounding methods could be used to find the optimal, potentially previously undetected, periodic orbits which maximize the bounds.

All of our methods could be used to construct fully rigorous bounds via a computer assisted proof, making use of interval arithmetic or rational projection [12, 37]. In both of our examples, the results of the SOS programs allowed us to find periodic orbits that attained our bounds to within several digits, thus confirming their sharpness. In principle, computer-assisted proofs based on our framework could also give rigorous upper bounds on LEs in partial differential equations, by combining our methods with some in Goluskin and Fantuzzi [13], but such computations would be challenging.

In the thesis of Oeri [33], an analogous approach to the sphere projection method was formulated for bounding the leading LE of discrete-time dynamical systems, with an SOS implementation for polynomial maps. This can be directly adapted to sums of LEs by using additive (rather than multiplicative) compound matrices. An approach analogous to our shifted spectrum method is also possible for bounding sums of LEs for discrete-time systems, and methods can additionally be derived for the Lyapunov dimension. In general, these are computationally difficult because the equivalent of the Lie derivative is a composition of the dynamics with an auxiliary function, leading to very high polynomial degrees. This may be the focus of a future publication.

## Data availability statement

Code to reproduce all the examples in this paper is available at `https://github.com/jeremypparker/Parker_`

## Acknowledgements

The authors thank Samuel Punshon-Smith and Anthony Quas for useful discussions. During this work DG was supported by the NSERC Discovery Grants Program (awards RGPIN-2018-04263 and RGPIN-2025-06823). Computational resources were provided by the Digital Research Alliance of Canada.

## A Exterior products and compound matrices

The exterior (or *wedge* or *Grassmann*) product is a multilinear map of vectors, defined by the property that $x_2 \wedge x_1 = -x_1 \wedge x_2$, and $x_1 \wedge x_2 = 0$ only if $x_1 = \alpha x_2$ for some $\alpha$, or $x_2 = 0$. An exterior product of $k$ vectors in a vector space of dimension $n \geq k$ is called a $k$-vector. The properties of the exterior product imply that permuting the $x_i$ in $x_1 \wedge \cdots \wedge x_k$ either leaves the $k$-vector unchanged or negates it, and that the $k$-vector is nonzero if and only if the $x_i$ are linearly independent. The set of $k$-vectors in $\mathbb{R}^n$ is therefore a linear space of dimension $n_k = \frac{n!}{(n-k)!}$. We identify this space with $\mathbb{R}^{n_k}$ and choose as a basis the set of $k$-vectors of the form

$$e_{i_1} \wedge \cdots \wedge e_{i_k}, \tag{120}$$

where $i_1 < \ldots < i_k$ and $e_i$ is the $i^{th}$ standard basis vector in $\mathbb{R}^{n_k}$. The standard Euclidean norm on this space satisfies the property that

$$|x_1 \wedge \cdots \wedge x_k| \tag{121}$$

is the volume of the parallepiped spanned by $x_1, \dots, x_k \in \mathbb{R}^n$ [47].

Linear maps on vectors in $\mathbb{R}^n$ induce linear maps on corresponding $k$-vectors in $\mathbb{R}^{n_k}$. Likewise, linear ODEs on $\mathbb{R}^n$ induce linear ODEs on $k$-vectors. As defined below, the map on $k$-vectors induced by a map $A$ is captured by the multiplicative compound $A^{(k)}$, and the ODE for $k$-vectors induced by an ODE with right-hand side $A$ is captured by the additive compound $A^{[k]}$. The multiplicative compound is common in many different areas of mathematics, and is known variously as the *compound representation*, *exterior power representation*, *wedge power representation* or simply the *induced action* of a linear map. Compound matrices have been applied previously in the calculation and bounding of LEs [18, 26, 29, 30] and also stability of dynamical systems [31]. Here we state key definitions and theorems; a more detailed introduction can be found in Marshall et al. [28, Chapter 19F], in more generality for complex non-square matrices.

**Definition A.1.** For any real matrix $A \in \mathbb{R}^{n\times n}$ and integer $k \in \{1, \dots, n\}$:

1. The $k^{th}$ *multiplicative compound* of $A$ is the real matrix $A^{(k)} \in \mathbb{R}^{n_k\times n_k}$ whose entries are the $k \times k$ subdeterminants of $A$, taken in lexicographic order.

2. The $k^{th}$ *additive compound* of $A$ is the real matrix $A^{[k]} \in \mathbb{R}^{n_k\times n_k}$ such that

$$A^{[k]} = \frac{\mathrm{d}}{\mathrm{d}\delta} \left(I + \delta A\right)^{(k)}\Big|_{\delta=0} . \tag{122}$$

The lexicographic order in definition A.1 refers to the ordering of all $n_k$ permutations of $k$ increasing integers chosen from $\{1, \dots, n\}$. As an example, the lexicographic order of $k = 3$ integers in the case $n = 4$ is

$$(1, 2, 3), \quad (1, 2, 4), \quad (1, 3, 4), \quad (2, 3, 4). \tag{123}$$

Therefore, the $(2, 3)$ entry of $A^{(3)}$ is

$$A^{(3)}_{2,3} = \det \begin{bmatrix} a_{1,1} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,3} & a_{2,4} \\ a_{4,1} & a_{4,3} & a_{4,4} \end{bmatrix}, \tag{124}$$

where $a_{p,q}$ denotes the $(p, q)$ element of $A$. Note that the row and column indices come from the second and third permutations, respectively. It follows from the definitions that $A^{(1)} = A^{[1]} = A$, and $A^{(n)} = \det A$ whereas $A^{[n]} = \mathrm{tr}A$. In fact, the eigenvalues of multiplicative and additive compounds are respectively the products and sums of the eigenvalues of the original matrix, which gives rise to their names. An explicit example of additive compound matrices is given in (111).

The next two propositions list additional properties that will be useful to us for multiplicative and additive compounds, respectively. In proposition A.2, part 1 is a special case of the Cauchy–Binet theorem. Parts 2–4 follow quickly from the definition, and the last part is theorem 6.2.10 of Fiedler [10].

**Proposition A.2** (Properties of multiplicative compound matrices)**.** *For any real square matrices $A$ and $B$:*

1. $(AB)^{(k)} = A^{(k)} B^{(k)}$.
2. $\left(A^{(k)}\right)^{\mathsf{T}} = \left(A^{\mathsf{T}}\right)^{(k)}$.
3. *If $A$ is invertible, then* $\left(A^{-1}\right)^{(k)} = \left(A^{(k)}\right)^{-1}$.
4. *$A$ is orthogonal if and only if $A^{(k)}$ is orthogonal.*
5. *For $k$-vectors identified with $\mathbb{R}^{n_k}$ and for $A \in \mathbb{R}^{n \times n}$,*

$$A^{(k)} \left(x_1 \wedge \cdots \wedge x_k\right) = (Ax_1) \wedge \cdots \wedge (Ax_k). \tag{125}$$

**Proposition A.3** (Properties of additive compound matrices)**.** *For any real square matrices:*

1. $(A + B)^{[k]} = A^{[k]} + B^{[k]}$.
2. $A^{[k]} \left(x_1 \wedge \cdots \wedge x_k\right) = (Ax_1 \wedge x_2 \wedge \cdots \wedge x_k) + (x_1 \wedge Ax_2 \wedge \cdots \wedge x_k) + \cdots + (x_1 \wedge x_2 \wedge \cdots \wedge Ax_k)$.
3. *If $Q$ is invertible,* $\left(QAQ^{-1}\right)^{[k]} = Q^{(k)} A^{[k]} \left(Q^{(k)}\right)^{-1}$.

*Proof.* Part 1 is immediate from the definition. For part 2, see theorem 2f of [24]. Part 3 follows by calculation,

$$\begin{aligned}
\left(QAQ^{-1}\right)^{[k]} &= \frac{\mathrm{d}}{\mathrm{d}\delta} \left(I + \delta QAQ^{-1}\right)^{(k)}\Big|_{\delta=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}\delta} \left(Q\left(I + \delta A\right) Q^{-1}\right)^{(k)}\Big|_{\delta=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}\delta} Q^{(k)} \left(I + \delta A\right)^{(k)} \left(Q^{(k)}\right)^{-1}\Bigg|_{\delta=0}
\end{aligned}$$

where the last equality uses properties 1 and 2 from proposition A.2. ☐

Multiplicative compound matrices are straightforward to compute directly, at least when $n$ is small. Additive compound matrices can be computed with symbolic manipulation. In our computational examples we compute additive compounds using Julia with the DynamicPolynomials.jl package [21]. Explicit formulas for the elements of additive compounds also exist [27].

We now have the ingredients to prove (15):

**Lemma A.4.** *If*

$$\frac{\mathrm{d}}{\mathrm{d}t} y_i(t) = Df(x(t))\, y_i(t) \tag{126}$$

*for $i = 1, \dots, k$, and we define*

$$y^k(t) = y_1(t) \wedge \cdots \wedge y_k(t), \tag{127}$$

*then*

$$\frac{\mathrm{d}}{\mathrm{d}t} y^k(t) = Df^{[k]}(x(t))\, y^k(t). \tag{128}$$

*Proof.* We have

$$y_i(t) = D\phi^t(x_0)\, y_i(0), \tag{129}$$

so, using proposition A.2,

$$y^k(t) = \left(D\phi^t(x_0)\, y_1(0)\right) \wedge \cdots \wedge \left(D\phi^t(x_0)\, y_k(0)\right) = \left(D\phi^t x_0\right)^{(k)} y^k(0). \tag{130}$$

Observe that $\phi^0(x_0) = x_0$, and $\phi^{t+\delta}(x_0) = \phi^\delta(\phi^t(x_0))$, so that

$$D\phi^{t+\delta}(x_0) = D\phi^\delta(x(t))\, D\phi^t(x_0). \tag{131}$$

Then by proposition A.2,

$$\left[D\phi^{t+\delta}(x_0)\right]^{(k)} = \left[D\phi^\delta(x(t))\right]^{(k)} \, \left[D\phi^t(x_0)\right]^{(k)}, \tag{132}$$

so by the usual definition of the derivative,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} y^k &= \lim_{\delta \to 0} \left( \frac{\left(D\phi^{t+\delta}(x_0)\right)^{(k)} y^k(0) - \left(D\phi^t(x_0)\right)^{(k)} y^k(0)}{\delta} \right) \\
&= \lim_{\delta \to 0} \left( \frac{\left[D\phi^\delta(x(t))\right]^{(k)} - I}{\delta} \right) y^k(t).
\end{aligned}$$

Furthermore, we have

$$\left[D\phi^\delta(x(t))\right]^{(k)} = [I + \delta Df(x(t)) + o(\delta)]^{(k)}$$

and so

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} y^k &= \lim_{\delta \to 0} \left( \frac{[I + \delta Df(x(t)) + o(\delta)]^{(k)} - I}{\delta} \right) y^k(t) \\
&= \frac{\mathrm{d}}{\mathrm{d}\delta} \, [I + \delta Df(x(t))]^{(k)} \Big|_{\delta=0} y^k(t) \\
&= (Df(x(t)))^{[k]}\, y^k(t),
\end{aligned}$$

by definition A.1. $\square$

# B A specialised Lyapunov theorem

Here we present a Lyapunov theorem for the stability of one variable $w \in \mathbb{R}^m$ whose linear dynamics is controlled by a second variable $x$. This is applicable to both proposition 2.11 and proposition 2.12. The second part, whose proof follows that of similar results in Khalil [17], is the so-called converse Lyapunov theorem, which states that under stronger conditions, it is always possible to find a Lyapunov function and that this $V$ is quadratic in $w$.

**Lemma B.1.** *For a domain $\mathcal{B} \subset \mathbb{R}^n$, let $x(t)$ and $w(t)$ solve the ODE system*

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x \\ w \end{bmatrix} = \begin{bmatrix} f(x) \\ A(x)w \end{bmatrix} \tag{133}$$

*on $\mathcal{B} \times \mathbb{R}^m$, with $f \in C^0(\mathcal{B}, \mathbb{R}^n)$ and $A \in C^0(\mathcal{B}, \mathbb{R}^{m\times m})$.*

1. *If there exists $V \in C^1(\mathcal{B} \times \mathbb{R}^m)$ such that*

$$V(x,w) \geq |w|^2 \quad \text{and} \quad f(x)\cdot D_x V(x,w) + [A(x)w]\cdot D_w V(x,w) \leq 0 \tag{134}$$

*for all $(x,w) \in \mathcal{B} \times \mathbb{R}^m$, then $w(t)$ is bounded forward in time for each $w(0) \in \mathbb{R}^m$.*

2. *If $A \in C^1(\mathcal{B}, \mathbb{R}^{m\times m})$ is bounded and $f \in C^1(\mathcal{B}, \mathbb{R}^n)$, and if there exist $K > 0$ and $\lambda > 0$ such that*

$$|w(t)| \leq K e^{-\lambda t} |w(0)| \tag{135}$$

*for all solutions to (133), then there exists $U \in C^1(\mathcal{B}, \mathbb{R}^{m\times m})$ such that (134) is satisfied with $V(x,w) = w^\mathsf{T} U(x) w$.*

*Proof.* For the first part, let $V \in C^1$ satisfy the two inequality conditions. For contradiction, suppose that there is a solution with $|w(t)|$ unbounded. The first $V$ condition then implies that $V(x(t), w(t))$ is unbounded forward in time. The second $V$ condition implies

$$\frac{\mathrm{d}}{\mathrm{d}t}V(x(t),w(t)) = f(x(t))\cdot D_x V(x(t),w(t)) + [A(x(t))w(t)]\cdot D_w V(x(t),w(t)) \leq 0. \tag{136}$$

This expression is continuous in $t$, so we can integrate to show that $V$ remains bounded for all time, which is a contradiction.

For the second part, fix a time $T > 0$, to be chosen later. Define

$$V(x_0, w_0) = \frac{2L}{1 - e^{-2LT}} \int_0^T |w(t)|^2 \, dt, \tag{137}$$

where $(x(t), w(t))$ evolves under (133) from initial condition $(x(0), w(0)) = (x_0, w_0)$, and $L$ is an upper bound for the operator norm $\|A(x)\|$ over $x \in \mathcal{B}$. Observe that, by linearity of $w(t)$ with respect to $w_0$, this is quadratic in $w_0$ and thus can be written

$$V(x_0, w_0) = w_0^\mathsf{T} U(x_0) w_0 \tag{138}$$

for some symmetric matrix $U(x_0)$. Since $A \in C^1(\mathcal{B}, \mathbb{R}^{m\times m})$, it follows that $U \in C^1(\mathcal{B}, \mathbb{R}^{m\times m})$.

Note that

$$\begin{aligned}\left|\frac{\mathrm{d}}{\mathrm{d}t}|w|^2\right| &\le 2|w|\,|A(x)w| \\ &\le 2\|A(x)\||w|^2 \\ &\le 2L|w|^2\end{aligned} \tag{139}$$

i.e. $\frac{\mathrm{d}}{\mathrm{d}t}|w|^2 \ge -2L|w|^2$. Thus Grönwall's inequality gives us

$$|w(t)|^2 \ge e^{-2Lt}|w_0|^2 \tag{140}$$

and so

$$V(x_0, w_0) \ge \frac{2L}{1-e^{-2LT}}\int_0^T e^{-2Lt}|w_0|^2 dt = |w_0|^2, \tag{141}$$

satisfying the first condition of (134).

For the second condition of (134), we define

$$v(t) = V(x(t), w(t)) = \frac{2L}{1-e^{-2LT}}\int_t^{t+T}|w(s)|^2\,\mathrm{d}s. \tag{142}$$

Then,

$$\begin{aligned}&f(x(t))\cdot D_x V(x(t), w(t)) + [A(x(t))w(t)]\cdot D_w V(x(t), w(t)) \\ &\quad= \frac{\mathrm{d}v}{\mathrm{d}t} \\ &\quad= \frac{2L}{1-e^{-2LT}}\left(|w(t+T)|^2 - |w(t)|^2\right) \\ &\quad\le \frac{2L}{1-e^{-2LT}}\left(K^2 e^{-2\lambda T} - 1\right)|w(t)|^2,\end{aligned} \tag{143}$$

so if $T \ge \frac{1}{\lambda}\log K$ the condition is satisfied. □

# References


[1] M. ApS. *MOSEK Optimizer API for Julia*, 2025.

[2] L. Arnold. *Random dynamical systems.* Springer Monographs in Mathematics. Springer-Verlag, 1998.

[3] J. Bochi. Ergodic optimization of Birkhoff averages and Lyapunov exponents. In *Proc. Int. Congr. Math.*, pages 1825–1846, 2018.

[4] N. Caranicolas. 1:1:1 resonant periodic orbits in 3-dimensional galactic-type Hamiltonians. *Astronomy and Astrophysics*, 282:34–36, 1994.

[5] C. Coey, L. Kapelevich, and J. P. Vielma. Solving natural conic formulations with Hypatia.jl. *INFORMS Journal on Computing*, 34(5):2686–2699, 2022.

[6] T. de Zeeuw. Motion in the core of a triaxial potential. *Monthly Notices of the Royal Astronomical Society*, 215(4):731–760, 1985.

[7] C. R. Doering and J. Gibbon. On the shape and dimension of the Lorenz attractor. *Dynamics and Stability of Systems*, 10(3):255–268, 1995.

[8] A. Eden, C. Foias, and R. Temam. Local and global Lyapunov exponents. *Journal of Dynamics and Differential Equations*, 3:133–177, 1991.

[9] A. O. Eden. *An abstract theory of L-exponents with applications to dimension analysis*. Indiana University, 1989.

[10] M. Fiedler. *Special matrices and their applications in numerical mathematics*. Courier Corporation, 2008.

[11] K. Gatermann and P. A. Parrilo. Symmetry groups, semidefinite programs, and sums of squares. *Journal of Pure and Applied Algebra*, 192(1-3):95–128, 2004.

[12] D. Goluskin. Bounding averages rigorously using semidefinite programming: mean moments of the Lorenz system. *Journal of Nonlinear Science*, 28:621–651, 2018.

[13] D. Goluskin and G. Fantuzzi. Bounds on mean energy in the Kuramoto–Sivashinsky equation computed using semidefinite programming. *Nonlinearity*, 32(5):1705, 2019.

[14] P. J. Goulart and Y. Chen. Clarabel: An interior-point solver for conic programs with quadratic objectives, 2024.

[15] M. Kaluba, D. Kielak, and P. W. Nowak. On property $(T)$ for $Aut(F_n)$ and $SL_n(Z)$. *Annals of Mathematics*, 193(2):539–562, 2021.

[16] J. Kaplan and J. Yorke. Chaotic behavior of multidimensional difference equations. *Functional Differential Equations and Approximation of Fixed Points*, pages 204–227, 1979.

[17] H. Khalil. *Nonlinear systems*. Prentice Hall, 2002.

[18] N. Kuznetsov and V. Reitmann. *Attractor dimension estimates for dynamical systems: theory and computation*. Springer, 2020.

[19] M. Lakshmi, G. Fantuzzi, J. Fernández-Caballero, Y. Hwang, and S. Chernyshenko. Finding extremal periodic orbits with polynomial optimisation, with application to a nine-mode model of shear flow. *SIAM J. Appl. Dyn. Syst.*, 19:763–787, 2020.

[20] B. Legat, C. Coey, R. Deits, J. Huchette, and A. Perry. Sum-of-squares optimization in Julia. In *The First Annual JuMP-dev Workshop*, 2017.

[21] B. Legat, S. Timme, T. Weisser, L. Kapelevich, N. Sajko, Manuel, A. Demin, C. Hansknecht, C. Rackauckas, J. TagBot, M. Forets, and T. Holy. JuliaAlgebra/DynamicPolynomials.jl: v0.5.7, May 2024.

[22] G. A. Leonov and S. Lyashko. Eden's hypothesis for a Lorenz system. *Vestnik St. Petersburg University: Mathematics*, 26(3):15–18, 1993.

[23] G. A. Leonov, N. V. Kuznetsov, N. Korzhemanova, and D. Kusakin. Lyapunov dimension formula for the global attractor of the Lorenz system. *Communications in Nonlinear Science and Numerical Simulation*, 41:84–103, 2016.

[24] D. London. On derivations arising in differential equations. *Linear and Multilinear Algebra*, 4(3):179–189, 1976.

[25] G. G. Lorentz. *Approximation of Functions*. Chelsea Publishing Company, 2nd edition, 1986.

[26] D. Manika. Application of the compound matrix theory for the computation of Lyapunov exponents of autonomous Hamiltonian systems. Master's thesis, Aristotle University of Thessaloniki, 2013.

[27] M. Margaliot and E. D. Sontag. Revisiting totally positive differential systems: A tutorial and new results. *Automatica*, 101:1–14, 2019.

[28] A. Marshall, I. Olkin, and B. Arnold. *Inequalities: Theory of Majorization and Its Applications*. Springer Series in Statistics. Springer New York, 2010.

[29] D. Martini, D. Angeli, G. Innocenti, and A. Tesi. Ruling out positive Lyapunov exponents by using the Jacobian's second additive compound matrix. *IEEE Control Systems Letters*, 6:2924–2928, 2022.

[30] D. Martini, D. Angeli, G. Innocenti, and A. Tesi. Bounding Lyapunov exponents through second additive compound matrices: Case studies and application to systems with first integral. *International Journal of Bifurcation and Chaos*, 33(10):2350114, 2023.

[31] J. S. Muldowney. Compound matrices and ordinary differential equations. *The Rocky Mountain Journal of Mathematics*, pages 857–872, 1990.

[32] K. G. Murty and S. N. Kabadi. Some NP-complete problems in quadratic and nonlinear programming. *Mathematical Programming: Series A and B*, 39(2):117–129, 1987.

[33] H. Oeri. *Convex optimization methods for bounding Lyapunov exponents*. PhD thesis, University of Victoria, 2023.

[34] H. Oeri and D. Goluskin. Convex computation of maximal Lyapunov exponents. *Nonlinearity*, 36(10):5378, 2023.

[35] J. F. Palacián, C. Vidal, J. Vidarte, and P. Yanguas. Periodic solutions and KAM tori in a triaxial potential. *SIAM Journal on Applied Dynamical Systems*, 16(1):159–187, 2017.

[36] A. Papachristodoulou and S. Prajna. On the construction of Lyapunov functions using the sum of squares decomposition. In *Proceedings of the 41st IEEE Conference on Decision and Control, 2002.*, volume 3, pages 3482–3487. IEEE, 2002.

[37] J. P. Parker. The Lorenz system as a gradient-like system. *Nonlinearity*, 37(9):095022, 2024.

[38] J. P. Parker, D. Goluskin, and G. Vasil. A study of the double pendulum using polynomial optimization. *Chaos*, 31(10):103102, 2021.

[39] P. Parrilo. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization.* PhD thesis, California Institute of Technology, 2000.

[40] A. Pikovsky and A. Politi. *Lyapunov exponents: a tool to explore complex dynamics.* Cambridge University Press, 2016.

[41] M. Putinar. Positive polynomials on semi-algebraic sets. *Indiana University Mathematics Journal*, 452:969–984, 1993.

[42] C. Rackauckas and Q. Nie. Differentialequations.jl–a performant and feature-rich ecosystem for solving differential equations in Julia. *Journal of Open Research Software*, 5(1):15, 2017.

[43] B. Reznick. Some concrete aspects of hilbert's 17th problem. *Contemporary Mathematics*, 253(251-272), 2000.

[44] C. W. Scherer and C. W. J. Hol. Matrix sum of squares relaxations for robust semidefinite programs. *Mathematical Programming*, 107:189–211, 2006.

[45] R. Temam. *Infinite-dimensional dynamical systems in mechanics and physics*, volume 68. Springer Science & Business Media, 2012.

[46] I. Tobasco, D. Goluskin, and C. R. Doering. Optimal bounds and extremal trajectories for time averages in nonlinear dynamical systems. *Physics Letters A*, 382(6):382–386, 2018.

[47] S. Winitzki. *Linear algebra via exterior products.* Sergei Winitzki, 2009.