Bounding extreme events in nonlinear dynamics using convex optimization*

Giovanni Fantuzzi[†] and David Goluskin[‡]

Abstract. We study a convex optimization framework for bounding extreme events in nonlinear dynamical systems governed by ordinary or partial differential equations (ODEs or PDEs). This framework bounds from above the largest value of an observable along trajectories that start from a chosen set and evolve over a finite or infinite time interval. The approach needs no explicit trajectories. Instead, it requires constructing suitably constrained auxiliary functions that depend on the state variables and possibly on time. Minimizing bounds over auxiliary functions is a convex problem dual to the non-convex maximization of the observable along trajectories. This duality is strong, meaning that auxiliary functions give arbitrarily sharp bounds, for sufficiently regular ODEs evolving over a finite time on a compact domain. When these conditions fail, strong duality may or may not hold; both situations are illustrated by examples. We also show that near-optimal auxiliary functions can be used to construct spacetime sets that localize trajectories leading to extreme events. Finally, in the case of polynomial ODEs and observables, we describe how polynomial auxiliary functions of fixed degree can be optimized numerically using polynomial optimization. The corresponding bounds become sharp as the polynomial degree is raised if strong duality and mild compactness assumptions hold. Analytical and computational ODE examples illustrate the construction of bounds and the identification of extreme trajectories, along with some limitations. As an analytical PDE example, we bound the maximum fractional enstrophy of solutions to the Burgers equation with fractional diffusion.

Key words. Extreme events, nonlinear dynamics, auxiliary functions, bounds, differential equations, polynomial optimization

AMS subject classifications. 93C10, 93C15, 93C20, 90C22, 34C11, 37C10, 49M29

1. Introduction.

Predicting the magnitudes of extreme events in deterministic dynamical systems is a fundamental problem with a wide range of applications. Examples of practical relevance include estimating the amplitudes of rogue waves in fluid or optical systems [62], the fastest possible mixing by incompressible fluid flows [23, 56], and the largest load on a structure due to dynamical forcing. In addition, extreme events relating to finite-time singularity formation are central to mathematical questions about the well-posedness and regularity of partial differential equations (PDEs). One such question is the Millennium Prize Problem concerning regularity of the three-dimensional Navier–Stokes equations [8], for which finite bounds on various quantities that grow transiently would imply the global existence of smooth solutions [22, 17, 18, 15].

*Submitted to the editors 29 July 2019.

Funding: The first author (GF) was supported in part by an EPSRC Doctoral Prize Fellowship and in part by an Imperial College Research Fellowship. The second author (DG) was supported by the NSERC Discovery Grants Program through awards RGPIN-2018-04263, RGPAS-2018-522657, and DGECR-2018-00371. Both authors gratefully acknowledge the hospitality of the 2018 Geophysical Fluid Dynamics program at the Woods Hole Oceanographic Institution, where part of this work was completed.

†Department of Aeronautics, Imperial College London, South Kensington Campus, London, SW7 2AZ, United Kingdom. (giovanni.fantuzzi10@imperial.ac.uk, https://www.imperial.ac.uk/people/giovanni.fantuzzi10).

‡Department of Mathematics and Statistics, University of Victoria, Victoria, BC, V8P 5C2, Canada (goluskin@uvic.ca, https://web.uvic.ca/~goluskin).

This work studies extreme events in dynamical systems governed by ordinary differential equations (ODEs) or PDEs. Specifically, given a scalar quantity of interest Φ, we seek to bound its largest possible value along trajectories that evolve forward in time from a prescribed set of initial conditions. This maximum, denoted by Φ^* and defined precisely in the next section, may be considered over all forward times or up to a finite time. Our definition of extreme events as maxima applies equally well to minima since a minimum of Φ is a maximum of $-\Phi$.

Bounding Φ^* from above and from below are fundamentally different tasks. A lower bound is implied by any value of Φ on any relevant trajectory, whereas upper bounds are statements about whole classes of trajectories and require a different approach. Analytical bounds of both types appear in the literature for many systems with complicated nonlinear dynamics, but often they are far from sharp. More precise lower bounds on Φ^* have sometimes been obtained using numerical integration, for instance to study extreme transient growth, optimal mixing, and transition to turbulence in fluid mechanics [5, 6, 21, 23, 56, 37]. In such computations, adjoint optimization [29] is used to search for an initial condition that locally maximizes Φ at a fixed terminal time, and a second level of optimization can vary the terminal time. Since both optimizations are non-convex, they give a local maximum of Φ but do not give a way to know whether it coincides with the global maximum Φ^* or is strictly smaller. Thus, adjoint optimization cannot give upper bounds on Φ^*, even when made rigorous by interval arithmetic. To find such an upper bound using numerical integration, one could use verified computations to find an outer approximation to the reachable set of trajectories starting from a bounded set [12], and then bound Φ^* from above by the global maximum of Φ on this approximating set. However, the latter is hard to compute if either Φ or the set on which it must be maximized are not convex.

The present study describes a general framework for bounding Φ^* from above that does not rely on numerical integration. This framework can be implemented analytically, computationally, or both, depending on what is tractable for the equations being studied. It falls within a broad family of methods, dating back to Lyapunov's work on nonlinear stability [53], whereby properties of dynamical systems are inferred by constructing *auxiliary functions*, which depend on the system's state and possibly on time, and which satisfy suitable inequalities. Lyapunov functions [53, 14], which often are used to verify nonlinear stability, are one type of auxiliary functions. Other types can be used to approximate basins of attraction [69, 40, 31, 75] and reachable sets [54, 36], estimate the effects of disturbances [83, 13, 3], guarantee the avoidance of certain sets [66, 4], design nonlinear optimal controls [47, 32, 55, 41, 85, 42], bound infinite-time averages or stationary stochastic expectations [10, 20, 44, 25, 71, 43, 27], and bound extreme values over global attractors [26]. Some of these works refer to auxiliary functions as Lyapunov, Lyapunov-like, storage, or barrier functions, or as subsolutions to the Hamilton–Jacobi equation. Others do not use auxiliary functions explicitly but characterize nonlinear dynamics using invariant or occupation measures; the two approaches are related by Lagrangian duality and are equivalent in many cases. Furthermore, many proofs about differential equations that rely on monotone quantities can be viewed as special cases of various auxiliary function methods. For instance, as we explain in Example 2.2, the bounds on transient growth in fluid systems proved in [5, 6] fit within the general framework described here. Similarly, the "background method" introduced in [16] to bound infinite-time averages in fluid dynamics is equivalent to using quadratic auxiliary functions in a different framework [9, 27].

In this paper, we describe how to use auxiliary functions to bound extreme values among nonlinear ODE or PDE trajectories starting from a specified set of initial conditions. Precisely, any differentiable auxiliary function satisfying two inequalities given in section 2 provides an *a priori* upper bound on Φ^*, without any trajectories being known. In the field of PDE analysis, these inequality conditions have been used implicitly to bound extreme events (e.g., [5, 6]), but the unifying framework we describe often has gone unrecognized. In the field of control theory, generalizations of our framework appear as convex relaxations of deterministic optimal control problems (e.g., [81, 80, 48, 79]) and of stochastic optimal stopping problems [11]. In these works, constraints on auxiliary functions are deduced using convex duality after replacing the maximization of Φ over trajectories with a convex maximization over occupation measures. Here we derive the same constraints using elementary calculus, and we illustrate their application using numerous ODE examples and one PDE example.

Unlike the maximization over trajectories that defines Φ^*, seeking the smallest upper bound among all admissible auxiliary functions defines a convex minimization problem. In general these two optimization problems are weakly dual: the minimum is an upper bound on the maximum but may not be equal to it. In some cases they are strongly dual, meaning that the maximum over trajectories coincides with the minimum over auxiliary functions, and these functions act as Lagrange multipliers that enforce the dynamics when maximizing Φ over trajectories. In such cases there exist auxiliary functions giving arbitrarily sharp upper bounds on Φ^*. Strong duality holds for a large class of sufficiently regular ODEs where the maximum of Φ is taken over a finite time horizon. This strong duality has been proved for a more general class of optimal control problems using measure theory and convex duality [48], and Appendix D gives a simpler proof for our present context that shows existence of near-optimal auxiliary functions using a mollification argument similar to [33].

In many practical applications, constructing auxiliary functions that yield explicit upper bounds on Φ^* is difficult regardless of whether strong duality holds. We illustrate various constructions here but do not have an approach that works universally. However, in the important case of dynamical systems governed by polynomial ODEs, polynomial auxiliary functions can be constructed using computational methods for polynomial optimization. With an infinite time horizon, this approach is applicable if the only invariant trajectories are algebraic sets, which is always true of steady states and is occasionally true of periodic orbits. With a finite time horizon, there is no such restriction. Polynomial ODEs are computationally tractable because the inequality constraints on auxiliary functions amount to nonnegativity conditions on certain polynomials. Polynomial nonnegativity is NP-hard to decide [59] but can be replaced by the stronger constraint that the polynomial is representable as a sum of squares (SOS). Optimization problems subject to SOS constraints can be reformulated as semidefinite programs (SDPs) [60, 45, 64] and solved using algorithms with polynomial-time complexity [78]. Thus, one can minimize upper bounds on Φ^* for polynomial ODEs by numerically solving SOS optimization problems. Moreover, we prove that bounds computed with SOS methods becomes sharp as the degree of the polynomial auxiliary function is raised, provided that the time horizon is finite, certain compactness properties hold, and the minimization over general auxiliary functions is strongly dual to the maximization of Φ over trajectories. We illustrate the computation of very sharp bounds using SOS methods for several ODE examples, including a 16-dimensional system.

In addition to methods for bounding Φ^* above, we describe a way to locate trajectories on which the observable Φ attains its maximum value of Φ^*. Specifically, auxiliary functions that prove sharp or nearly sharp upper bounds on Φ^* can be used to define regions in state space where each such trajectory must lie prior to its extreme event. We illustrate this using an ODE for which nearly optimal polynomial auxiliary functions can be computed by SOS methods.

The rest of this paper is organized as follows. Section 2 explains how auxiliary functions can be used to bound the magnitudes of extreme events in nonlinear dynamical systems. We construct bounds in several ODE examples and one PDE example; some but not all of these bounds are sharp. Section 3 explains how auxiliary functions can be used to locate trajectories leading to extreme events. Section 4 describes how polynomial optimization can be used to construct auxiliary functions computationally for polynomial ODEs. Bounds computed in this way for various ODE examples appear in that section and others. Section 5 extends the framework to give bounds on extreme values at particular times or integrated over time, rather than maximized over time, giving a more direct derivation of bounding conditions that have appeared in [81, 80, 48, 79]. Conclusions and open questions are offered in section 6. Appendices contain details of calculations and an alternative proof of the strong duality result that follows from [48].

2. Bounds using auxiliary functions.

Consider a dynamical system on a Banach space $\mathcal{X}$ that is governed by the differential equation

$$\dot{x} = F(t, x), \quad x(t_0) = x_0. \tag{2.1}$$

Here, $F : \mathbb{R} \times \mathcal{X} \to \mathcal{X}$ is continuous and possibly nonlinear, the initial time t_0 and initial condition x_0 are given, and $\dot{x}$ denotes $\partial_t x$. When $\mathcal{X} = \mathbb{R}^n$, (2.1) defines an n-dimensional system of ODEs. When $\mathcal{X}$ is a function space and F a differential operator, (2.1) defines a parabolic PDE, which may be considered in either strong or weak form [70, 68]. The trajectory of (2.1) that passes through the point $y \in \mathcal{X}$ at time s is denoted by $x(t; s, y)$. We assume that, for every choice of $(s, y) \in \mathbb{R} \times \mathcal{X}$, this trajectory exists uniquely on an open time interval, which can depend on both s and y and might be unbounded.

Suppose that $\Phi : \mathbb{R} \times \mathcal{X} \to \mathbb{R}$ is a continuous function that describes a quantity of interest for system (2.1). Let Φ^* denote the largest value attained by $\Phi[t, x(t; t_0, x_0)]$ among all trajectories that start from a prescribed set $X_0 \subset \mathcal{X}$ and evolve forward over a closed time interval $\mathcal{T}$ that is either finite, $\mathcal{T} = [t_0, T]$, or infinite, $\mathcal{T} = [t_0, \infty)$:

$$\Phi^* := \sup_{\substack{x_0 \in X_0 \\ t \in \mathcal{T}}} \Phi[t, x(t; t_0, x_0)]. \tag{2.2}$$

We write Φ^*_T and Φ^*_∞ instead of Φ^* when necessary to distinguish between finite and infinite time horizons. Our objective is to bound Φ^* from above without knowing trajectories of (2.1).

Let $\Omega \subset \mathcal{T} \times \mathcal{X}$ be a region of spacetime in which the graphs $(t, x(t; t_0, x_0))$ of all trajectories starting from X_0 remain up to the time horizon of interest. In applications one may be able to identify a set Ω that is strictly smaller than $\mathcal{T} \times \mathcal{X}$, otherwise it suffices to choose $\Omega = \mathcal{T} \times \mathcal{X}$. The maximum (2.2) that we aim to bound depends only on trajectories within Ω.

To derive upper bounds on Φ^* we employ auxiliary functions $V : \Omega \to \mathbb{R}$. In most cases we require V to be differentiable along trajectories of (2.1), so that its Lie derivative

$$\mathcal{L}V(s,y) := \lim_{\varepsilon \to 0} \frac{V\left[s+\varepsilon, x(s+\varepsilon; s, y)\right] - V(s,y)}{\varepsilon} \tag{2.3}$$

is well defined. By design the function $\mathcal{L}V : \Omega \to \mathbb{R}$ coincides with the rate of change of V along trajectories, meaning $\frac{\mathrm{d}}{\mathrm{d}t}V(t,x(t)) = \mathcal{L}V(t,x(t))$ if $x(t)$ solves (2.1) and all derivatives exist. Crucially, an expression for $\mathcal{L}V$ can be derived without knowing the trajectories. In practice one differentiates $V[t,x(t;s,y)]$ with respect to t and uses the differential equation (2.1). For example, when $\mathcal{X} = \mathbb{R}^n$ and (2.1) is a system of ODEs, the chain rule gives

$$\mathcal{L}V(t,x) = \partial_t V(t,x) + F(t,x) \cdot \nabla_x V(t,x). \tag{2.4}$$

Subsection 2.1 presents inequality constraints on V and $\mathcal{L}V$ that imply upper bounds on Φ^*, as well as a convex framework for optimizing these bounds. Both can be obtained as particular cases of a general relaxation framework for optimal control problems [81, 80, 48], but we give an elementary derivation. Subsection 2.2 compares bounds obtained when $\Omega = \mathcal{T} \times \mathcal{X}$, meaning that the constraints on V are imposed globally in spacetime, to bounds obtained when a strictly smaller Ω containing all relevant trajectories can be found. Finally, subsection 2.3 discusses conditions under which arbitrarily sharp upper bounds on Φ^* can be proved.

2.1. Bounding framework. Assume that for each initial condition $x_0 \in X_0$ a trajectory $x(t;t_0,x_0)$ exists on some open time interval where it is unique and absolutely continuous. This does not preclude trajectories that are unbounded in infinite or finite time. To bound Φ^* we define a class $\mathcal{V}(\Omega)$ of admissible auxiliary functions as the subset of all differentiable functions, $C^1(\Omega)$, that do not increase along trajectories and bound Φ from above pointwise. Precisely, $V \in \mathcal{V}(\Omega)$ if and only if

$$\mathcal{L}V(t,x) \le 0 \quad \forall (t,x) \in \Omega, \tag{2.5a}$$

$$\Phi(t,x) - V(t,x) \le 0 \quad \forall (t,x) \in \Omega. \tag{2.5b}$$

The system dynamics enter only in the derivation of $\mathcal{L}V$; conditions (2.5a,b) are imposed pointwise in the spacetime domain Ω and can be verified without knowing any trajectories. If $\Omega = \mathcal{T} \times \mathcal{X}$ we call V a *global* auxiliary function, otherwise it is *local* on a smaller chosen Ω.

We claim that

$$\Phi^* \le \inf_{V \in \mathcal{V}(\Omega)} \sup_{x_0 \in X_0} V(t_0,x_0), \tag{2.6}$$

with the convention that the righthand side is $+\infty$ if $\mathcal{V}(\Omega)$ is empty. To see that (2.6) holds when $\mathcal{V}$ is not empty, consider fixed $V \in \mathcal{V}(\Omega)$ and $x_0 \in X_0$. For any $t \ge t_0$ up to which the trajectory $x(t;t_0,x_0)$ exists and is absolutely continuous, the fundamental theorem of calculus can be combined with (2.5a,b) to find

$$\Phi[t,x(t;t_0,x_0)] \le V[t,x(t;t_0,x_0)] = V(t_0,x_0) + \int_{t_0}^{t} \mathcal{L}V[\xi,x(\xi;t_0,x_0)]\,\mathrm{d}\xi \le V(t_0,x_0). \tag{2.7}$$

Thus, the existence of any $V \in \mathcal{V}(\Omega)$ implies that $\Phi[t, x(t; t_0, x_0)]$ is bounded uniformly on $\mathcal{T}$ for each x_0. Conversely, if Φ blows up before the chosen time horizon for any $x_0 \in X_0$, then no auxiliary functions exist. Maximizing both sides of (2.7) over $t \in \mathcal{T}$ and $x_0 \in X_0$ gives

$$\Phi^* \leq \sup_{x_0 \in X_0} V(t_0, x_0), \tag{2.8}$$

and then minimizing over $\mathcal{V}(\Omega)$ gives (2.6) as claimed.

The minimization problem on the righthand side of (2.6) is convex and gives a bound on the (generally non-convex) maximization problem defining Φ^* in (2.2). Despite convexity of the minimization, it usually is difficult to construct an optimal or near-optimal auxiliary function, even with computer assistance. Nevertheless, any auxiliary function satisfying (2.5a,b) gives a rigorous upper bound on Φ^* according to (2.8). This framework therefore can be useful for analysis, and sometimes for computation, even when the dynamics are very complicated. Analytically, one often can find a suboptimal auxiliary function that yields fairly good bounds. Computationally, for certain systems including polynomial ODEs, one can optimize V over a finite-dimensional subset of $\mathcal{V}(\Omega)$ to obtain bounds that are very good and sometimes perfect. However, the inequality in (2.6) is strict in general, meaning that there are cases where the optimal bounds provable using conditions (2.5a,b) are not sharp. Local auxiliary functions can sometimes produce sharp bounds when global ones fail, although this depends on the spacetime set Ω inside which the graphs of trajectories are known to remain. This is illustrated by examples in subsection 2.2, while subsection 2.3 discusses sufficient conditions for bounds from auxiliary functions to be arbitrarily sharp. First, however, we present two examples where global auxiliary functions work well.

Example 2.1 concerns a simple ODE where the optimal upper bound (2.6) produced by global V appears to be sharp. We conclude this by constructing V increasingly near to optimal, obtaining bounds that are extremely close to Φ^*. These V are constructed computationally using polynomial optimization methods, the explanation of which is postponed until section 4. Example 2.2 proves bounds for the Burgers equation with ordinary and fractional diffusion. We analytically construct V giving bounds that are finite, but unlikely to be sharp. The bounds for fractional diffusion are novel, while those for ordinary diffusion show that the proof of the same result in [5] can be seen as an instance of the auxiliary function framework.

Example 2.1. Consider the nonautonomous ODE system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 t - 0.1 x_1 - x_1 x_2 \\ -x_1 t - x_2 + x_1^2 \end{bmatrix}. \tag{2.9}$$

All trajectories eventually approach the origin, but various quantities can grow transiently. For example, consider the maximum of $\Phi = x_1$ over an infinite time horizon. Let the initial time be $t_0 = 0$ and the set of initial conditions X_0 contain only the point $x_0 = (0, 1)$. Then, Φ^*_∞ is the largest value of x_1 along the trajectory with $x(0) = (0, 1)$, and it is easy to find by numerical integration. Doing so gives $\Phi^* \approx 0.30056373$, and this value can be used to judge the sharpness of upper bounds on Φ^*_∞ that we produce using global auxiliary functions.

The quadratic polynomial

$$V(t, x) = \tfrac{1}{2}\left(1 + x_1^2 + x_2^2\right) \tag{2.10}$$

Table 1

Upper bounds on Φ^*_∞ for Example 2.1, computed using polynomial optimization with V of various polynomial degrees. For the single initial condition $x_0 = (0,1)$, numerical integration gives $\Phi^* \approx 0.30056373$ for all time horizons larger than $T = 1.6635$, which agrees with the degree-8 bound to the tabulated precision. For the set X_0 of initial conditions on the shifted unit circle with center $(-\frac{3}{4}, 0)$, nonlinear optimization of the initial angular coordinate yields $\Phi^*_\infty \approx 0.49313719$, which agrees with the degree-10 bound to the tabulated precision.

	Upper bounds	
deg(V)	$X_0 = \{(0,1)\}$	X_0 circle
2	1	1.75
4	0.41381042	0.80537235
6	0.30056854	0.49808038
8	0.30056373	0.49313760
10	”	0.49313719

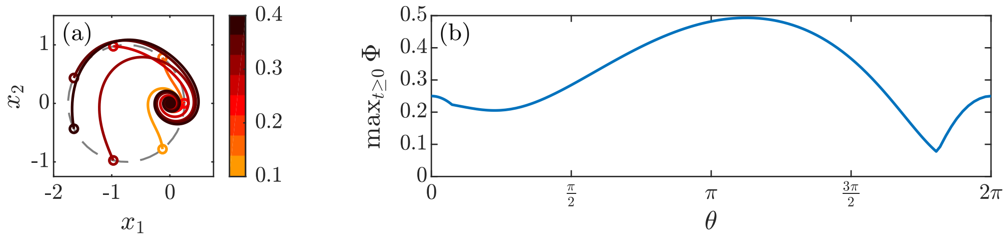

Figure 1. *(a) Sample trajectories starting from the circle with center $(-\frac{3}{4}, 0)$ and unit radius (- - -). The initial conditions are marked with a circle, while the color scale reflects the maximum value of Φ along each trajectory. (b) Numerical approximation to the maximum of Φ along single trajectories with initial condition on the shifted unit circle $(\cos\theta - \frac{3}{4}, \sin\theta)$ as a function of the angular coordinate θ.*

is an admissible global auxiliary function, meaning that it satisfies the inequalities (2.5a,b) on $\Omega = [0,\infty) \times \mathbb{R}^2$. For this V and the chosen X_0 and t_0, the bound (2.8) yields

$$\Phi^*_\infty \leq V(0, x_0) = 1. \tag{2.11}$$

This is the best bound that can be proved using global quadratic V, as shown in Appendix A, but optimizing polynomial V of higher degree produces better results. For instance, the best global quartic V that can be constructed using polynomial optimization is

$$\begin{aligned} V(t,x) = 0.2353 + 0.7731\,x_1^2 + 0.1666\,x_1x_2 + 0.4589\,x_2^2 + 0.5416\,x_1^3 + 0.05008\,tx_1^2 \\ + 0.1616\,tx_1x_2 + 0.2505\,tx_2^2 - 0.1058\,x_1^2x_2 + 0.1730\,x_1x_2^2 - 0.5766\,x_2^3 \\ + 0.2962\,x_1^4 + 0.1888\,t^2x_1^2 + 0.1888\,t^2x_2^2 + 0.5923\,x_1^2x_2^2 + 0.2962\,x_2^4, \end{aligned} \tag{2.12}$$

where numerical coefficients have been rounded. The bound on Φ^*_∞ that follows from the above V is reported in Table 1, along with bounds that follow from computationally optimized V of polynomial degrees 6, 8, and 10 (omitted for brevity). The bounds improve as the degree of V is raised, and the optimal degree-8 bound is sharp up to nine significant figures. The numerical approach used for such computations is described in section 4.

Unlike searching among particular trajectories, bounding Φ^* from above is not more difficult when the set X_0 of initial conditions is larger than a single point. For example, consider initial conditions on the shifted unit circle centered at $(-\frac{3}{4}, 0)$,

$$X_0 = \left\{(x_1, x_2) : \left(x_1 + \tfrac{3}{4}\right)^2 + x_2^2 = 1\right\} = \left\{ \left(\cos\theta - \tfrac{3}{4}, \sin\theta\right) : \theta \in [0, 2\pi)\right\}. \tag{2.13}$$

Sample trajectories and the variation of $\max_{t\geq 0} \Phi$ with the angular position θ in X_0 are shown in Figure 1. Finding the trajectory that attains Φ^* requires numerical integration, combined with nonlinear optimization over initial conditions in X_0. Starting MATLAB's optimizer `fmincon` from initial guesses with angular coordinate $\theta = \frac{3\pi}{4}$ and $\theta = \frac{\pi}{10}$ yields locally optimal initial conditions of $\theta \approx 1.125\pi$ and $\theta = 2\pi$, which lead to Φ values of 0.49313719 and 0.25, respectively. Figure 1(b) confirms that the former initial condition is globally optimal, meaning $\Phi^* \approx 0.49313719$. On the other hand, polynomial auxiliary functions can be optimized by the methods of section 4 using exactly the same algorithms as when X_0 contains a single point. For initial conditions on the shifted unit circle X_0, Table 1 lists upper bounds on Φ^* implied by numerically optimized polynomial V of degrees up to 10. We omit the computed V for brevity. The optimal degree-10 V gives a bound that is sharp to eight significant figures. ■

Example 2.2. To illustrate the analytical use of global auxiliary functions for PDEs, we consider mean-zero period-1 solutions $u(t, x)$ of the Burgers equation with fractional diffusion,

$$\begin{gathered} \dot{u} = -uu_x - (-\Delta)^\alpha u, \\ u(0, x) = u_0(x), \quad u(t, x+1) = u(t, x), \quad \int_0^1 u(t, x)\,\mathrm{d}x = 0. \end{gathered} \tag{2.14}$$

Following standard PDE notation, in this example the state variable in $\mathcal{X}$ is denoted by $u(t, \cdot)$, whereas $x \in [0, 1]$ is the spatial variable. Discussion of this equation and a definition of the fractional Laplacian $(-\Delta)^\alpha$ can be found in [84]. Ordinary diffusion is recovered when $\alpha = 1$. For each $\alpha \in (\frac{1}{2}, 1]$, solutions exist and remain bounded when the Banach space $\mathcal{X}$ in which solutions evolve is the Sobolev space H^s with $s > \frac{3}{2} - 2\alpha$ [38]. Let us consider a quantity that is called fractional enstrophy in [84],

$$\Phi(u) := \frac{1}{2}\int_0^1 \left[(-\Delta)^{\frac{\alpha}{2}} u\right]^2 \mathrm{d}x. \tag{2.15}$$

We aim to bound Φ^*_∞ among trajectories whose initial conditions u_0 have a specified value Φ_0 of fractional enstrophy, so the set of initial conditions is

$$X_0 = \{u \in \mathcal{X} : \Phi(u) = \Phi_0\}. \tag{2.16}$$

Here we prove Φ_0-dependent upper bounds on Φ^*_∞ for $\alpha \in (\frac{3}{4}, 1]$. Such bounds have been reported for ordinary diffusion ($\alpha = 1$) [5] but not for $\alpha < 1$. We employ global auxiliary functions of the form

$$V(u) = \left[\Phi(u)^\beta + C\|u\|_2^2\right]^{1/\beta}, \tag{2.17}$$

where $\|u\|_2^2 = \int_0^1 u^2\,\mathrm{d}x$ and the constants $\beta, C > 0$ are to be chosen. This ansatz is guided by the realization that the analysis of the $\alpha = 1$ case [5] is equivalent to the auxiliary function framework with $\beta = 1/3$ in (2.17).

To be an admissible auxiliary function, V must satisfy (2.5a,b). The inequality $V(u) \geq \Phi(u)$ holds for every positive C, while the inequality $\mathcal{L}V(u) \leq 0$ constrains β and C. To derive an expression for $\mathcal{L}V(u)$ we first note that differentiating along trajectories of (2.14) and integrating by parts gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\|u(t,\cdot)\|_2^2 = -4\Phi[u(t,\cdot)], \tag{2.18a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi[u(t,\cdot)] = R[u(t,\cdot)] := -\int_0^1 [(-\Delta)^\alpha u]^2\,\mathrm{d}x - \int_0^1 u u_x (-\Delta)^\alpha u\,\mathrm{d}x. \tag{2.18b}$$

Differentiating $V[u(t,\cdot)]$ in time thus gives

$$\mathcal{L}V(u) = \frac{1}{\beta}\left[\Phi(u)^\beta + C\|u\|_2^2\right]^{\frac{1}{\beta}-1}\left[\beta\Phi(u)^{\beta-1}R(u) - 4C\Phi(u)\right]. \tag{2.19}$$

The sign of $\mathcal{L}V$ is that of the expression in the rightmost brackets, so an estimate for $R(u)$ is needed. Theorem 2.2 in [84] provides $R(u) \leq \sigma_\alpha \Phi(u)^{\gamma_\alpha}$, with $\gamma_\alpha = \frac{8\alpha-3}{6\alpha-3}$ and explicit prefactors σ_α that blow up as $\alpha \to \frac{3}{4}^+$. By fixing $\beta = 2-\gamma_\alpha$ and $C = (2-\gamma_\alpha)\sigma_\alpha/4$, we guarantee that (2.19) is nonpositive. Thus, V is a global auxiliary function yielding the bound

$$\Phi_\infty^* \leq \sup_{u_0\in X_0}\left[\Phi_0^{2-\gamma_\alpha} + \frac{(2-\gamma_\alpha)\sigma_\alpha}{4}\|u_0\|_2^2\right]^{\frac{1}{2-\gamma_\alpha}} \tag{2.20}$$

according to (2.8). Finally, the righthand maximization over u_0 can be carried out analytically by calculus of variations to bound Φ_∞^* in terms of only the initial fractional enstrophy Φ_0,

$$\Phi_\infty^* \leq \left[\Phi_0^{2-\gamma_\alpha} + \frac{(2-\gamma_\alpha)\sigma_\alpha}{2(2\pi)^{2\alpha}}\Phi_0\right]^{\frac{1}{2-\gamma_\alpha}}. \tag{2.21}$$

The bound (2.21) is finite for every $\alpha \in (\frac{3}{4}, 1]$. The coefficient on Φ_0 is bounded uniformly for α in this range, but the exponent $\frac{1}{2-\gamma_\alpha}$ blows up as $\alpha \to \frac{3}{4}^+$. When $\alpha = 1$ we can replace σ_α with a smaller prefactor from [52] to find

$$\Phi_\infty^* \leq \left(\Phi_0^{1/3} + 2^{-10/3}\pi^{-8/3}\Phi_0\right)^3. \tag{2.22}$$

The above estimate is identical to the result of [5],[1] and their argument is equivalent to ours in that it implicitly relies on our V being nonincreasing along trajectories. Similarly, in [6] the same authors bound a quantity called palinstrophy in the two-dimensional Navier–Stokes equations, and that proof can be seen as using (in their notation) the global auxiliary function $V(u) = \left[\mathcal{P}(u)^{1/2} + (4\pi\nu^2)^{-2}\mathcal{K}(u)^{1/2}\mathcal{E}(u)\right]^2$.

The bound (2.21) is unlikely to be sharp. For $\alpha = 1$ it scales like $\Phi_\infty^* \leq \mathcal{O}(\Phi_0^3)$ when $\Phi_0 \gg 1$, whereas numerical and asymptotic evidence suggests that $\Phi_\infty^* = \mathcal{O}(\Phi_0^{3/2})$ [5, 65]. It is an open question whether going beyond the V ansatz (2.17) can produce sharper analytical bounds, and whether the optimal bound (2.6) that can be proved using global auxiliary functions would be sharp in this case. ■

[1]Expression (5) in [5] is claimed to hold with $\mathcal{E}$ being identical to our $\Phi(u)$, but in fact it holds with $\mathcal{E} = 2\Phi(u)$ because their derivation uses estimate (3.7) from [52]. With this correction, and with $L = 1$ and $\nu = 1$, the expression in [5] agrees with our bound (2.22).

2.2. Global versus local auxiliary functions. In various cases, such as Example 2.1 above, global auxiliary functions can produce arbitrarily sharp upper bounds on Φ^*. Other times they cannot. In Example 2.3 below, global auxiliary functions give bounds that are finite but not sharp. In Example 2.4, no global auxiliary functions exist. Sharp bounds can be recovered in both examples by using local auxiliary functions, meaning that we enforce constraints (2.5a,b) only on a subset $\Omega \subsetneq \mathcal{T} \times \mathcal{X}$ of spacetime that contains all trajectories of interest.

There are various ways to determine that trajectories starting from the initial set X_0 remain in a spacetime set Ω during the time interval $\mathcal{T}$. One option is to choose a function $\Psi(t, x)$ and use global auxiliary functions to show that $\Psi^* \le B$ for initial conditions in X_0. This implies that trajectories starting from X_0 remain in the set

$$\Omega := \{(t, x) \in \mathcal{T} \times \mathcal{X} : \, \Psi(t, x) \le B\}. \tag{2.23}$$

Any Ψ that can be bounded using global auxiliary functions can be used, including $\Psi = \Phi$, and Ω can be refined by considering more than one Ψ. Another way to show that trajectories never exit a prescribed set Ω is to construct a barrier function that is nonpositive on $\{t_0\} \times X_0$, positive outside Ω, and whose zero level set cannot be crossed by trajectories. Barrier functions can be constructed analytically in some cases, and computationally for ODEs with polynomial righthand sides; see [66, 4] and references therein. Finally, in the polynomial ODE case the computational methods of [31] can produce a spacetime set $\Omega = \mathcal{T} \times X$, where $X \subsetneq \mathcal{X}$ is an outer approximation for the evolution of the initial set X_0 over the time interval $\mathcal{T}$. The next two examples demonstrate the differences between global and local auxiliary functions for a simple ODE where a suitable choice of Ω is apparent.

Example 2.3. Consider the autonomous one-dimensional ODE

$$\dot{x} = x^2, \qquad x(0) = x_0. \tag{2.24}$$

Trajectories $x(t) = x_0/(1 - x_0 t)$ with nonzero initial conditions grow monotonically. If $x_0 < 0$, then $x(t) \to 0$ as $t \to \infty$; if $x_0 > 0$, then $x(t)$ blows up at the critical time $t = 1/x_0$. Suppose the set of initial conditions X_0 includes only a single point x_0, the time interval is $\mathcal{T} = [0, \infty)$, and the quantity to be bounded is

$$\Phi(x) = \frac{4x}{1 + 4x^2}. \tag{2.25}$$

Since $|\Phi(x)| \le 1$ uniformly, Φ^*_∞ is finite for each x_0 despite the blowup of trajectories starting from positive initial conditions. Explicit solutions give

$$\Phi^*_\infty = \begin{cases} 0, & x_0 \le 0, \\ 1, & 0 < x_0 \le \frac{1}{2}, \\ \dfrac{4x_0}{1 + 4x_0^2}, & x_0 > \frac{1}{2}. \end{cases} \tag{2.26}$$

Here X_0 contains only one initial condition, so the optimal bound (2.6) simplifies to

$$\Phi^*_\infty \le \inf_{V \in \mathcal{V}(\Omega)} V(0, x_0). \tag{2.27}$$

The constant function $V \equiv 1$ belongs to $\mathcal{V}$ for each x_0 and implies the trivial bound $\Phi^*_\infty \le 1$, which is sharp for $x_0 \in (0, 1/2]$. For all other $x_0 \ne 0$ there exist different V providing sharp bounds on Φ^*_∞, regardless of whether the domain Ω of auxiliary functions is global or local. This is shown in Appendix B. At the semistable point $x_0 = 0$, however, sharp bounds are possible only with local auxiliary functions on certain Ω.

In the $x_0 = 0$ case, the resulting trajectory is simply $x(t) \equiv 0$. Thus it suffices to enforce the auxiliary function constraints (2.5a,b) locally on $\Omega = [0,\infty) \times \{0\}$. On this Ω, the constant function $V \equiv 0$ is a local auxiliary function giving the sharp bound $\Phi^* \le 0$. In fact, the same is true with $\Omega = [0,\infty) \times X$ for any X with $0 \in X \subseteq (-\infty, 0]$. On the other hand, if the chosen set X contains any open neighborhood of 0, then sharp bounds are not possible. This is true in particular for global auxiliary functions, which must satisfy constraints (2.5a,b) on $\Omega = [0,\infty) \times \mathbb{R}$. The righthand minimum in (2.27) over global auxiliary functions is attained by the constant function $V = 1$. No better bound is possible with global V because they must satisfy $V(0, 0) \ge 1$. To prove this, recall that every $V(t, x)$ is continuous by definition. Thus for any $\delta > 0$ there exists $y > 0$ such that $V(0, 0) \ge V(0, y) - \delta$. The trajectory of (2.24) with initial condition $x(0) = y$ blows up in finite time and must therefore pass through $x = \frac{1}{2}$ at some time t^*. Condition (2.5b) requires that $V(t^*, \frac{1}{2}) \ge \Phi(\frac{1}{2}) = 1$, while (2.5a) implies that V decays along trajectories, so

$$V(0,0) \ge V(0,y) - \delta \ge V(t^*, \tfrac{1}{2}) - \delta \ge 1 - \delta \tag{2.28}$$

for every $\delta > 0$. Thus $V(0, 0) \ge 1$, so when $x_0 = 0$ the righthand minimum over global V in (2.27) is indeed attained by $V \equiv 1$. Local auxiliary functions can prove better bounds, but a similar argument shows that the sharp bound $\Phi^* \le 0$ for $X_0 = \{0\}$ is possible only if $0 \in X \subseteq (-\infty, 0]$. That is, the upper limit of X must coincide with the boundary of the basin of attraction of the semistable point at 0. In more complicated systems it may not be possible to locate X so precisely. In such cases, if global auxiliary functions do not give sharp bounds, local ones might not either, at least for spacetime sets Ω that one can identify in practice. ■

Example 2.4. In some cases, global auxiliary functions can fail to exist even if Φ^* is finite. Again consider the ODE (2.24) from Example 2.3 with $\mathcal{T} = [0, \infty)$ and a single initial condition $X_0 = \{x_0\}$, but now consider the quantity

$$\Phi(t, x) = x^2 \mathrm{e}^x. \tag{2.29}$$

Recalling that $x(t)$ approaches zero if $x_0 \le 0$ and blows up otherwise, we find

$$\Phi^*_\infty = \begin{cases} 4\,\mathrm{e}^{-2}, & x_0 \le -2, \\ x_0^2\,\mathrm{e}^{x_0}, & -2 < x_0 \le 0, \\ \infty, & x_0 > 0. \end{cases} \tag{2.30}$$

For auxiliary functions satisfying (2.5a,b) globally on $\Omega = [0, \infty) \times \mathbb{R}$, $\mathcal{V}(\Omega)$ must be empty when $x_0 > 0$ since $\Phi^*_\infty = \infty$. However, $\mathcal{V}(\Omega)$ is empty also when $x_0 \le 0$, despite Φ^*_∞ being finite. This is because any global V satisfying (2.5a,b) must be nonincreasing for trajectories starting at all $y \in \mathbb{R}$, not only for initial conditions in the set of interest X_0. In particular,

$$V(0, y) \ge V[t, x(t; 0, y)] \ge \Phi[t, x(t; 0, y)] = x(t; 0, y)^2\,\mathrm{e}^{x(t;0,y)} \tag{2.31}$$

for all $y \in \mathbb{R}$ and all $t \geq 0$, where the second inequality follows from (2.5b). No V that is continuous on $[0, \infty) \times \mathbb{R}$ can satisfy (2.31) because, for each $y > 0$, the rightmost expression becomes infinite as t approaches the blowup time $1/x_0$. Thus, $\mathcal{V}(\Omega)$ is empty.

Sharp bounds on finite Φ^* become possible with local rather than global auxiliary functions, much as in Example 2.3. Since Φ^* is finite only when $X_0 \subseteq (-\infty, 0]$, and trajectories starting from any such X_0 stay within $X = (-\infty, 0]$, conditions (2.5a,b) can be enforced locally on $\Omega = [0, \infty) \times X$. As in Example 2.3, it is crucial that X contains no points outside the basin of the semistable equilibrium at the origin. A local V giving sharp bounds is

$$V(t,x) = \begin{cases} 4\,\mathrm{e}^{-2}, & x \leq -2, \\ x^2\,\mathrm{e}^{x}, & x > -2. \end{cases} \tag{2.32}$$

At each $x_0 \leq 0$ this V is equal to the value (2.30) of Φ^*_∞ for the single trajectory starting at x_0. Thus, this V gives a sharp bound on Φ^*_∞ for every possible initial set $X_0 \subseteq (-\infty, 0]$. ■

2.3. Sharpness of optimal bounds.

The best bounds on Φ^* provable using auxiliary functions are often but not always sharp. Examples 2.3 and 2.4 above show that the upper bound (2.6) can be strict, at least for infinite time horizons and global auxiliary functions. For finite time horizons and local auxiliary functions, on the other hand, arguments in [48] prove that (2.6) is an equality provided trajectories remain in a compact set over the finite time interval of interest. Section 2.3.1 states this result and gives an explicit counterexample for infinite time horizons. Section 2.3.2 explains why sharp bounds are always possible if one allows V to be discontinuous, a fact which is useful for theory but not for explicitly bounding quantities in particular systems.

2.3.1. Sharp bounds for ODEs with finite time horizon.

Local auxiliary functions can produce arbitrarily sharp bounds on Φ^*_T with finite time horizon T for well posed ODEs, provided the initial set X_0 is compact and trajectories that start from it remain inside a compact set X up to time T. Precisely, Theorem 2.1 and equation (5.3) in [48] imply the following result.

Theorem 2.5 ([48]). *Let $\dot{x} = F(t,x)$ be an ODE with F locally Lipschitz in both arguments. Given $\Phi : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ continuous, an initial time t_0, a finite time interval $\mathcal{T} = [t_0, T]$, and a compact set of initial conditions X_0, define Φ^*_T as in (2.2). Assume that:*

(A.1) All trajectories starting from X_0 at time t_0 remain in a compact set X for $t \in \mathcal{T}$;

(A.2) There exist a time $t_1 > T$ and a bounded open neighborhood Y of X such that, for all initial points $(s,y) \in [t_0, t_1] \times Y$, a unique trajectory $x(t; s, y)$ exists for all $t \in [s, t_1]$.

Then, letting $\mathcal{V}(\Omega)$ denote the set of differentiable auxiliary functions that satisfy (2.5a,b) on the compact set $\Omega := \mathcal{T} \times X$,

$$\Phi^*_T = \inf_{V \in \mathcal{V}(\Omega)} \sup_{x_0 \in X_0} V(t_0, x_0). \tag{2.33}$$

In Appendix D we give an alternative proof of this theorem that uses mollification to construct near-optimal V. This construction does not yield explicit bounds on Φ^*_T for particular ODEs because it invokes trajectories, which generally are not known. Both the original proof in [48] and our proof rely on assumptions (A.1) and (A.2) to ensure that trajectories starting in a neighborhood of X remain bounded past the time horizon T and are regular in the sense that the map $(s, y) \mapsto x(t; s, y)$ is locally Lipschitz on $[t_0, t_1] \times Y$. Regularity over a spacetime set slightly larger than Ω is used to construct smooth uniform approximations to certain functions on Ω via mollification. However, the assumptions are not necessary for the equality (2.33) to hold. For instance, the example in Appendix B violates assumption (A.1) when $x_0 > 0$ and $T = 1/x_0$, yet the V in (B.1) implies sharp bounds on Φ^*_T.

It is an open challenge to weaken the assumptions of Theorem 2.5. With infinite time horizons, for instance, auxiliary functions give sharp bounds in some examples but not others. Sharp bounds for an infinite time horizon are illustrated in Appendix B. In the next example, on the other hand, there exists a set X such that infinite-time analogues of assumptions (A.1) and (A.2) hold, yet differentiable local auxiliary functions cannot give sharp bounds on Φ^*_∞.

Example 2.6. Consider the one-dimensional ODE

$$\dot{x} = x^2 - x^3, \tag{2.34}$$

which has two equilibria: the semistable point $x_s = 0$ and the attractor $x_a = 1$. Although no explicit analytical solution is available, trajectories exist for all times. As $t \to \infty$, they approach x_s if $x_0 \leq 0$ and approach x_a if $x_0 > 0$. We let

$$\Phi(x) = 4x(1 - x) \tag{2.35}$$

and seek upper bounds on Φ^*_∞ for initial conditions in the set $X_0 = [-1, 0]$. All trajectories starting in X_0 approach x_s from below, so

$$\Phi^*_\infty = \sup_{\substack{x_0 \in X_0 \\ t \in [t_0, \infty)}} \Phi[x(t; x_0)] = 0. \tag{2.36}$$

Trajectories with initial conditions in $X_0 = [-1, 0]$ remain there, so the smallest X we could choose is $X = X_0$. With this choice, $V \equiv 0$ gives a sharp upper bound. However, suppose we choose $X = [-1, 1]$, which is the smallest connected set that is globally attracting and contains X_0. For this X, assumptions analogous to (A.1) and (A.2) in Theorem 2.5 hold on the infinite time interval $[0, \infty)$, yet any upper bound on $\Phi^*_\infty = 0$ provable with differentiable local V cannot be smaller than 1. Indeed, any such V must be continuous at $(t, x) = (0, 0)$ and arguing as in Example 2.3 shows that $V(0, 0) \geq 1$, so any V subject to (2.5a,b) satisfies

$$\max_{x \in [-1, 0]} V(0, x) \geq 1. \tag{2.37}$$

Thus, with $X = [-1, 1]$, any bound implied by (2.6) is no smaller than 1 as claimed above. ■

The inability of differentiable auxiliary functions to produce sharp bounds in Examples 2.3 and 2.6 is due to the map $x_0 \mapsto x(t; 0, x_0)$ from initial conditions to trajectories not being locally Lipschitz near the saddle point $x_s = 0$. Because the time horizon is infinite, a fixed distance from x_s is eventually reached by trajectories starting arbitrarily close to x_s. This does not happen when the time horizon is finite. We cannot say whether the strong duality result of Theorem 2.5 applies with an infinite time horizon when the map $x_0 \mapsto x(t; 0, x_0)$ is Lipschitz; both the original proof in [48] and our alternative in Appendix D rely on the time interval $\mathcal{T}$ being compact.

2.3.2. Nondifferentiable auxiliary functions. One way to guarantee that optimization over V gives sharp bounds on Φ^*, regardless of whether the time horizon is finite or infinite, is to weaken the local sufficient condition (2.5a,b) by removing the requirement that V is differentiable. Since the Lie derivative $\mathcal{L}V$ may not be defined in this case, condition (2.5a) must be replaced with the direct constraint that V does not increase along trajectories,

$$V[s+\tau, x(s+\tau; s, y)] \le V(s,y) \quad \forall \tau \ge 0 \text{ and } (s,y) \in \Omega. \tag{2.38}$$

Slight modification of the argument leading to (2.8) then proves

$$\Phi^*_\infty \le \min_{\substack{V:\,(2.5b),\\ (2.38)}} \sup_{x_0 \in X_0} V(t_0, x_0). \tag{2.39}$$

Condition (2.38) cannot be checked when trajectories are not known exactly.[2] Differentiability of V therefore is crucial to find explicit bounds for particular systems because the Lie derivative $\mathcal{L}V$ gives a way to check that V is nonincreasing without knowing trajectories.

For theoretical purposes, on the other hand, nondifferentiable V are useful because

$$V^*(s,y) := \sup_{t \ge s} \Phi[t, x(t; s, y)] \tag{2.40}$$

is optimal and attains equality in (2.39), meaning

$$\Phi^*_\infty = \min_{\substack{V:\,(2.5b),\\ (2.38)}} \sup_{x_0 \in X_0} V(t_0, x_0) = \sup_{x_0 \in X_0} V^*(t_0, x_0). \tag{2.41}$$

This V^* is discontinuous in general because of the maximization over time. It follows directly from the definition of Φ^*_∞ that V^* satisfies (2.5b) globally and gives a sharp bound when substituted into (2.41). To see that (2.38) holds, observe that the trajectory starting from y at time s is the same as that starting from $x(s+\tau; s, y)$ at time $s+\tau$. Then, since $\tau \ge 0$,

$$\begin{aligned} V^*[s+\tau, x(s+\tau; s, y)] &= \sup_{t \ge s+\tau} \Phi\{t, x[t; s+\tau, x(s+\tau; s, y)]\} \\ &= \sup_{t \ge s+\tau} \Phi[t, x(t; s, y)] \\ &\le \sup_{t \ge s} \Phi[t, x(t; s, y)] \\ &= V^*(s,y). \end{aligned} \tag{2.42}$$

Example 2.7 below gives V^* in a case where trajectories are known.

[2]For systems with discrete-time dynamics, on the other hand, discontinuous V may be practically useful. This work focuses on continuous-time dynamics, but the convex bounding framework of subsection 2.1 readily extends to maps $x_{n+1} = F(n, x_n)$ when the continuous-time decay condition (2.5a) is replaced by the discrete version of (2.38), namely that $V[n+1, F(n,x_n)] \le V(n, x_n)$ for all $n \in \mathbb{N}$ and $x_n \in \mathcal{X}$. This can be checked directly without knowing trajectories. In addition, the computational methods described in section 4 can be applied with minor modifications to finite-dimensional polynomial maps.

Example 2.7. Recall Example 2.3, which shows that differentiable global auxiliary functions cannot give sharp bounds for the ODE (2.24) with Φ as in (2.25) and the single initial condition $X_0 = \{0\}$. For the auxiliary function

$$V(t,x) = \begin{cases} 0, & x \le 0, \\ 1, & 0 < x \le \frac{1}{2}, \\ \dfrac{4x}{1+4x^2}, & x > \frac{1}{2}, \end{cases} \tag{2.43}$$

which is discontinuous at $x = 0$, explicit ODE solutions confirm that V satisfies the nonincreasing condition (2.38). This V implies sharp bounds on Φ_∞^* for all sets X_0 of initial conditions, and in fact it is exactly the optimal V^* defined by (2.40). ∎

When trajectories are not known explicitly, the V^* defined by (2.40) cannot be used to find explicit bounds, but it can still be useful. For instance, in Appendix D we prove Theorem 2.5 by showing that V^* can be approximated with differentiable V. Moreover, V^* has arisen in various contexts. One field in which V^* arises is optimal control theory. Using ideas from dynamic programming for optimal stopping problems (see, e.g., section III.4.2 in [7]) one can show that if V^* is bounded and uniformly continuous on Ω, then it is exactly the so-called value function for problem (2.2) and is the unique viscosity solution to its corresponding Hamilton–Jacobi–Bellman complementarity system. This system consists of the auxiliary function constraints (2.5a,b) and the condition

$$\mathcal{L}V(t,x)[\Phi(t,x) - V(t,x)] = 0 \quad \forall (t,x) \in \Omega. \tag{2.44}$$

The auxiliary function framework studied in this work therefore can be seen as a relaxation of the Hamilton–Jacobi–Bellman system that results from dropping (2.44). A second connection between V^* and existing literature occurs in the particular case of linear dynamics on a Hilbert space, as explained in the following example.

Example 2.8. Let X be a Hilbert space with inner product $\langle \cdot, \cdot \rangle$. Consider the autonomous *linear* dynamical system $\dot{x} = Ax$ with initial condition $x(0) = x_0$, where A is a closed and densely defined linear operator, not necessarily bounded, that generates a strongly continuous semigroup $\{S_t\}_{t \ge 0}$. Trajectories satisfy $x(t) = S_t x_0$, so S_t is the flow map. Suppose S_t is compact for each $t > 0$. In various linear systems of this type, one is interested in the maximum possible amplification of the norm $\|x\| = \sqrt{\langle x, x \rangle}$, which in the present framework means that $\Phi(x) = \|x\|$ with the initial set $X_0 = \{x_0 \in X : \|x_0\| = 1\}$. In fluid mechanics, for instance, such problems have been studied to understand linear mechanisms by which perturbations are amplified (see, e.g., [72]). With the above choices, (2.40) and (2.41) reduce to the well-known result

$$\Phi_\infty^* = \sup_{\|x_0\|=1} \sup_{t \ge 0} \Phi(S_t x_0) = \sup_{t \ge 0} \sup_{\|x_0\|=1} \sqrt{\langle S_t x_0, S_t x_0 \rangle} = \sup_{t \ge 0} \sigma_{\max}(S_t), \tag{2.45}$$

where $\sigma_{\max}(S_t)$ denotes the maximum singular value of S_t. We stress, however, that the general bounding framework of subsection 2.1 does not require an explicit flow map and applies also to nonlinear systems. ∎

3. Optimal trajectories. So far we have presented a framework for bounding the magnitudes of extreme events without finding the extremal trajectories themselves. The latter is much harder in general, partly due to the non-convexity of searching over initial conditions. However, auxiliary functions producing bounds on Φ^* do give some information about optimal trajectories. Specifically, sublevel sets of any auxiliary function define regions of state space in which optimal and near-optimal trajectories must spend a certain fraction of time prior to the extreme event. A similar connection has been found between trajectories that maximize infinite-time averages and auxiliary functions that give bounds on these averages [71, 43]. The following discussion applies to both global and local auxiliary functions with either finite or infinite time horizons. The simpler case of exactly optimal auxiliary functions is addressed in subsection 3.1, followed by the general case in subsection 3.2.

3.1. Optimal auxiliary functions. Suppose for now that the optimal bound (2.8) is sharp and is attained by some V^*, in which case

$$\sup_{x_0 \in X_0} V^*(t_0, x_0) = \Phi^*. \tag{3.1}$$

Let $x_0^* \in X_0$ be an initial condition leading to an optimal trajectory, which attains the maximum value Φ^* at some time t^*. To determine the value of V^* on an optimal trajectory, note that the same reasoning leading to (2.8) yields

$$\begin{aligned}
\Phi^* &= \Phi[t, x(t^*; x_0^*)] \\
&\leq V^*(t_0, x_0^*) + \int_{t_0}^{t^*} \mathcal{L}V^*[\xi, x(\xi; t_0, x_0^*)] \,\mathrm{d}\xi \\
&\leq \sup_{x_0 \in X_0} V^*(t_0, x_0) + \int_{t_0}^{t^*} \mathcal{L}V^*[\xi, x(\xi; t_0, x_0^*)] \,\mathrm{d}\xi \\
&= \Phi^* + \int_{t_0}^{t^*} \mathcal{L}V^*[\xi, x(\xi; t_0, x_0^*)] \,\mathrm{d}\xi \\
&\leq \Phi^*
\end{aligned} \tag{3.2}$$

The above inequalities must be equalities and $\mathcal{L}V^* \leq 0$, so $\mathcal{L}V^* \equiv 0$ and $V^* \equiv \Phi^*$ along an optimal trajectory up to time t^*. These constant values of $\mathcal{L}V^*$ and V^* can be used to define sets in which optimal trajectories must lie:

$$\mathcal{R}_0 := \{(t, x) \in \Omega : \, \mathcal{L}V^*(t, x) = 0\}\,, \tag{3.3}$$

$$\mathcal{S}_0 := \left\{(t, x) \in \Omega : \, V^*(t, x) = \sup_{x_0 \in X_0} V^*(t_0, x_0)\right\}, \tag{3.4}$$

where we have used (3.1) in defining $\mathcal{S}_0$. The intersection $\mathcal{S}_0 \cap \mathcal{R}_0$ contains the graph of each optimal trajectory until the last time that trajectory attains the maximum value Φ^*. In general, $\mathcal{S}_0 \cap \mathcal{R}_0$ may also contain points not on any optimal trajectory.

3.2. General auxiliary functions. Consider an auxiliary function V and an initial condition x_0 that are a near-optimal pair, meaning that an upper bound on Φ^* implied by V and a lower bound implied by the trajectory starting from x_0 differ by no more than δ. That is, calling the upper bound λ,

$$\lambda - \delta \leq \sup_{t \in \mathcal{T}} \Phi[t, x(t; t_0, x_0)] \leq \Phi^* \leq \sup_{x_0 \in X_0} V(t_0, x_0) \leq \lambda. \tag{3.5}$$

The upper bound λ might be larger than $\sup_{x \in X_0} V(t_0, x)$ if the latter cannot be computed exactly, and the lower bound $\lambda - \delta$ might be smaller than $\sup_{t \in \mathcal{T}} \Phi[t, x(t; t_0, x_0)]$ if the trajectory starting from x_0 is only partly known.

Let t^* denote the latest time during the interval $\mathcal{T}$ when the trajectory starting at x_0 attains or exceeds the value $\lambda - \delta$. The constraints (2.5a,b) require V to decay along trajectories and bound Φ pointwise, so

$$\lambda - \delta \leq V[t^*, x(t^*; t_0, x_0)] \leq V[t, x(t; t_0, x_0)] \leq V(t_0, x_0) \leq \sup_{x \in X_0} V(t_0, x) \leq \lambda \tag{3.6}$$

for all $t \in [t_0, t^*]$. The above inequalities imply that the trajectory starting at x_0 satisfies

$$0 \leq \lambda - V[t, x(t; t_0, x_0)] \leq \delta \tag{3.7}$$

up to time t^*, so its graph must be contained in the set

$$\mathcal{S}_\delta := \{(t, x) \in \Omega : \, 0 \leq \lambda - V(t, x) \leq \delta\}, \tag{3.8}$$

which extends to suboptimal V the definition (3.4) of $\mathcal{S}_0$ for optimal V^*.

The definition (3.3) of $\mathcal{R}_0$ also can be extended to suboptimal V, but the resulting sets are guaranteed to contain optimal and near-optimal trajectories only for a certain amount of time. When V satisfies (3.5), an argument similar to (3.2) shows that

$$\Phi^* \leq \Phi^* + \delta + \int_0^{t^*} \mathcal{L}V[\xi, x(\xi; t_0, x_0)] \, \mathrm{d}\xi, \tag{3.9}$$

and therefore

$$-\int_{t_0}^{t^*} \mathcal{L}V[\xi, x(\xi; t_0, x_0)] \, \mathrm{d}\xi \leq \delta. \tag{3.10}$$

Since $\mathcal{L}V \leq 0$, the above condition can be combined with Chebyshev's inequality (cf. §VI.10 in [39]) to estimate, for any $\varepsilon > 0$, the total time during $[t_0, t^*]$ when $\mathcal{L}V \leq -\varepsilon$. Letting Θ_ε denote this total time and letting $\mathbb{1}_A$ denote the indicator function of a set A, we find

$$\Theta_\varepsilon := \int_{t_0}^{t^*} \mathbb{1}_{\{\xi : \, \mathcal{L}V[\xi, x(\xi; t_0, x_0)] < -\varepsilon\}} \, \mathrm{d}\xi \leq -\frac{1}{\varepsilon} \int_{t_0}^{t^*} \mathcal{L}V[\xi, x(\xi; t_0, x_0)] \, \mathrm{d}\xi \leq \frac{\delta}{\varepsilon}. \tag{3.11}$$

In other words, a trajectory on which $\Phi \geq \lambda - \delta$ at some time t^* cannot leave the set

$$\mathcal{R}_\varepsilon := \{(t, x) \in \Omega : \, -\varepsilon \leq \mathcal{L}V(t, x) \leq 0\} \tag{3.12}$$

for longer than δ/ε time units during the interval $[t_0, t^*]$. This statement is most useful when the upper bound $\Phi^* \leq \lambda$ implied by V is close to sharp, so there exist trajectories where Φ attains values $\lambda - \delta$ with small δ. Then one may take ε small enough for $\mathcal{R}_\varepsilon$ to exclude much of state space, while also having it be meaningful that near-optimal trajectories cannot leave $\mathcal{R}_\varepsilon$ for longer than δ/ε. The computational construction of $\mathcal{S}_\delta$ and $\mathcal{R}_\varepsilon$ for a polynomial ODE is illustrated by Example 4.3 in the next section.

4. Computing bounds for ODEs using SOS optimization. The optimization of auxiliary functions and their corresponding bounds is prohibitively difficult in many cases, even by numerical methods. However, computations often are tractable when the system (2.1) is an ODE with polynomial righthand side $F : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$, the observable Φ is polynomial, and the set of initial conditions X_0 is a basic semialgebraic set:

$$X_0 := \{x \in \mathbb{R}^n : \, f_1(x) \geq 0, \, \ldots, f_p(x) \geq 0, \, g_1(x) = 0, \, \ldots, g_q(x) = 0\} \tag{4.1}$$

for given polynomials $f_1, \ldots, f_p$ and $g_1, \ldots, g_q$. The set $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ in which the graphs of trajectories remain over the time interval $\mathcal{T}$ is assumed to be basic semialgebraic as well:

$$\Omega := \{(t, x) \in \mathbb{R} \times \mathbb{R}^n : \, h_1(t, x) \geq 0, \, \ldots, h_r(t, x) \geq 0, \, \ell_1(t, x) = 0, \, \ldots, \ell_s(t, x) = 0\} \tag{4.2}$$

for given polynomials $h_1, \ldots, h_r$ and $\ell_1, \ldots, \ell_s$. To construct global auxiliary functions with state space $\mathbb{R}^n$, the set Ω can be specified by a single inequality: $h_1(t, x) := t - t_0 \geq 0$ or $h_1(t, x) := (t - t_0)(T - t) \geq 0$ for infinite or finite time horizons, respectively. To construct local auxiliary functions, more inequalities or equalities must be added to define a smaller Ω.

For any integer d, let $\mathbb{R}_d[t, x]$ and $\mathbb{R}_d[x]$ denote the vector spaces of real polynomials of degree d or smaller in the variables (t, x) and x, respectively. Restricting the optimization over differentiable auxiliary functions in (2.6) to polynomials in $\mathbb{R}_d[t, x]$ gives

$$\Phi^* \leq \inf_{\substack{V \in \mathbb{R}_d[t,x] \\ \text{s.t. (2.5a,b)}}} \sup_{x_0 \in X_0} V(t_0, x_0). \tag{4.3}$$

Recalling that the supremum over X_0 is the smallest upper bound λ on that set, and substituting expression (2.4) for $\mathcal{L}V$ in the ODE case into (2.5a), we can express the righthand side of (4.3) as a constrained minimization over V and λ:

$$\begin{aligned} \Phi^* \leq \inf_{\substack{V \in \mathbb{R}_d[t,x] \\ \lambda \in \mathbb{R}}} \{\lambda : \, -\partial_t V(t, x) - F(t, x) \cdot \nabla_x V(t, x) &\geq 0 \text{ on } \Omega, \\ V(t, x) - \Phi(t, x) &\geq 0 \text{ on } \Omega, \\ \lambda - V(t_0, x) &\geq 0 \text{ on } X_0\}. \end{aligned} \tag{4.4}$$

Under the assumptions outlined above, the three constraints on V and λ are polynomial inequalities on basic semialgebraic sets. Checking such constraints is NP-hard in general [59], so a common strategy is to replace them with stronger but more tractable constraints. Here we require that the polynomials in (4.4) admit weighted sum-of-squares (WSOS) decompositions, which can be searched for computationally by solving SDPs. These WSOS constraints imply that the inequalities in (4.4) hold on Ω or X_0 but not necessarily outside these sets.

To define the relevant WSOS decompositions, let $\Sigma_\mu[t, x]$ and $\Sigma_\mu[x]$ be the cones of SOS polynomials of degrees up to μ in the variables (t, x) and x, respectively. That is, a polynomial $\sigma \in \mathbb{R}_\mu[x]$ belongs to $\Sigma_\mu[x]$ if and only if there exist a finite family of polynomials $q_1, \ldots, q_k \in \mathbb{R}_{\lfloor \mu/2 \rfloor}[x]$ such that $\sigma = \sum_{i=1}^k q_i^2$. For each integer μ that is no smaller than the highest polynomial degree appearing in the definition (4.1) of X_0, the set of degree-μ WSOS polynomials associated with X_0 is

$$\Lambda_\mu := \Big\{\sigma_0 + \sum_{i=1}^{p} f_i\sigma_i + \sum_{i=1}^{q} g_i\rho_i : \ \sigma_0 \in \Sigma_\mu[x], \quad \sigma_i \in \Sigma_{\mu-\deg(f_i)}[x],\ i = 1, \ldots, p \quad \rho_i \in \mathbb{R}_{\mu-\deg(g_i)}[x],\ i = 1, \ldots, q \Big\}. \tag{4.5}$$

In words, WSOS polynomials associated with X_0 can be written as a weighted sum of polynomials, where the weights are $\{1, f_1, \ldots, f_p, g_1, \ldots, g_q\}$ and the polynomials weighted by $\{1, f_1, \ldots, f_p\}$ are SOS. Every SOS polynomial is globally nonnegative, and it is WSOS with respect to any X_0 since all terms in the WSOS decomposition aside from σ_0 can be zero. On the other hand, WSOS polynomials need not be SOS.

Analogously to Λ_μ, the set of degree-μ WSOS polynomials associated with Ω is

$$\Gamma_\mu := \Big\{\sigma_0 + \sum_{i=1}^{r} h_i\sigma_i + \sum_{i=1}^{s} \ell_i\rho_i : \ \sigma_0 \in \Sigma_\mu[t,x], \quad \sigma_i \in \Sigma_{\mu-\deg(h_i)}[t,x],\ i = 1, \ldots, r \quad \rho_i \in \mathbb{R}_{\mu-\deg(\ell_i)}[t,x],\ i = 1, \ldots, s \Big\}. \tag{4.6}$$

If a polynomial belongs to Γ_μ or Λ_μ, then it is nonnegative on Ω or X_0, respectively. (The converse is false beyond a few special cases [34].) We can strengthen the inequality constraints on V in (4.4) by requiring WSOS representations instead of nonnegativity. This gives

$$\Phi^* \le \lambda_d^* := \inf_{\substack{V\in\mathbb{R}_d[t,x] \\ \lambda\in\mathbb{R}}} \{\lambda : \ -\partial_t V - F\cdot\nabla_x V \in \Gamma_{d-1+\deg(F)}, \quad V - \Phi \in \Gamma_d, \quad \lambda - V(t_0,\cdot) \in \Lambda_d\}. \tag{4.7}$$

For each integer d, the righthand side is a finite-dimensional optimization problem with WSOS constraints that are linear in the decision variables—the scalar λ and the coefficients of the polynomial V. It is well known that such problems can be reformulated as SDPs (e.g., Section 2.4 in [46]). Such SDPs can be solved numerically in polynomial time, barring problems with numerical conditioning. Open-source software is available to assist both with the reformulation of WSOS optimizations as SDPs and with the solution of the latter.[3] The SOS computations in Examples 2.1, 4.3, and 4.5, and in Appendix C, were set up in MATLAB using YALMIP [50, 51] or a customized version of SPOTLESS.[4] The resulting SDPs were solved with the interior-point solver MOSEK v.8 [58] except in Example 4.5, where the SDP was solved in multiple precision arithmetic with SDPA-GMP v.7.1.3 [24].

[3] Most modeling toolboxes for polynomial optimization, including the ones used in this work, do not natively support WSOS constraints. However, these can be implemented using standard SOS constraints. For instance, the WSOS constraint $P \in \Gamma_\mu$ can be implemented as the SOS constraint $P - \sum_{i=1}^{p} h_i\sigma_i - \sum_{i=1}^{q} \ell_i\rho_i \in \Sigma_\mu[t,x]$, along with the SOS constraints $\sigma_i \in \Sigma_{\mu-\deg(h_i)}[t,x]$ for $i = 1,\ldots,p$. This formulation, known as the generalized S-procedure [69, 20], introduces more decision variables than the direct WSOS approach of [46, Section 2.4]. The additional variables may lead to larger computations, but they can improve numerical conditioning by giving more freedom for the rescaling that is done within SDP solvers.

[4] https://github.com/aeroimperial-optimization/aeroimperial-spotless

The bounds λ_d^* found by solving (4.7) numerically form a nonincreasing sequence as the degree d of V is raised. These bounds appear to become sharp in various cases, including Example 2.1 above and Example 4.3 below. We cannot say whether such convergence occurs in all cases, even when auxiliary functions arbitrarily close to optimality are known to exist. This is due to our restriction to polynomial V and use of WSOS constraints, which are sufficient but not necessary for nonnegativity. However, if the sets X_0 and Ω are both compact and there exists a differentiable V attaining equality in (2.6), then the following theorem guarantees that bounds from SOS computations become sharp as the polynomial degree is raised. The proof is a standard argument in SOS optimization and relies on a result known as Putinar's Positivstellensatz [67, Lemma 4.1], which guarantees the existence of WSOS representations for strictly positive polynomials; details can be found in Section 2.4 of [46].

Theorem 4.1. *Let Ω and X_0 be compact semialgebraic sets. Assume the definitions of Ω and X_0 include inequalities $C_1 - t^2 - \|x\|_2^2 \geq 0$ and $C_2 - \|x\|_2^2 \geq 0$ for some C_1 and C_2, respectively, which can always be made true by adding inequalities that do not change the specified sets. Let λ_d^* be the bound from the optimization (4.7). If differentiable auxiliary functions give arbitrarily sharp bounds (2.33) on Φ_T^*, then $\lambda_d^* \to \Phi_T^*$ as $d \to \infty$.*

Proof. Assume that the semialgebraic definitions of Ω and X_0 include inequalities of the form $C_1 - t^2 - \|x\|_2^2 \geq 0$ and $C_2 - \|x\|_2^2 \geq 0$, respectively. If not, these inequalities can be added with C_1 and C_2 large enough to not change which points lie in Ω and X_0 since both sets are compact. Then, $C_1 - t^2 - \|x\|_2^2 \in \Gamma_\mu$ and $C_2 - \|x\|_2^2 \in \Lambda_\mu$ for all integers μ.[5]

To prove that $\lambda_d^* \to \Phi_T^*$ as $d \to \infty$, we establish the equivalent claim that, for each $\varepsilon > 0$, there exists an integer d such that $\lambda_d^* \leq \Phi_T^* + \varepsilon$. Choose $\gamma > 0$ such that

$$\gamma < \frac{2T\varepsilon}{5T - t_0}. \tag{4.8}$$

By assumption there exists an auxiliary function $W \in C^1(\Omega)$, not generally a polynomial, such that

$$W(t_0, x_0) \leq \Phi_T^* + \gamma \quad \text{on } X_0. \tag{4.9}$$

Since Ω is compact, polynomials are dense in $C^1(\Omega)$ (cf. Theorem 1.1.2 in [49]). That is, for each $\delta > 0$ there exists a polynomial P such that $\|W - P\|_{C^1(\Omega)} \leq \delta$, where $\|\cdot\|_{C^k(\Omega)}$ denotes the usual norm on $C^k(\Omega)$—the sum of the L^∞ norms of all derivatives up to order k. Fix such a P with

$$\delta < \frac{\gamma}{\max\left\{2, 2T, 2T\|F_1\|_{C^0(\Omega)}, \ldots, 2T\|F_n\|_{C^0(\Omega)}\right\}}. \tag{4.10}$$

By definition Ω contains the initial set $\{t_0\} \times X_0$, so $|W(t_0, \cdot) - P(t_0, \cdot)| < \delta$ uniformly on X_0. We define the polynomial auxiliary function

$$V(t, x) = P(t, x) + \gamma\left(1 - \frac{t}{2T}\right). \tag{4.11}$$

[5]Theorem 4.1 holds also when the semialgebraic definitions of Ω and X_0 satisfy Assumption 2.14 in [46, Section 2.4], which is a slightly weaker but more technical condition implying the inclusions $C_1 - t^2 - \|x\|_2^2 \in \Gamma_\mu$ and $C_2 - \|x\|_2^2 \in \Lambda_\mu$ for all sufficiently large integers μ.

With δ as in (4.10), γ as in (4.8), and W satisfying (4.9), elementary estimates show that

$$-\partial_t V - F \cdot \nabla_x V > 0 \quad \text{on } \Omega, \tag{4.12a}$$

$$V - \Phi > 0 \quad \text{on } \Omega, \tag{4.12b}$$

$$\Phi_T^* + \varepsilon - V(t_0, \cdot) > 0 \quad \text{on } X_0. \tag{4.12c}$$

The inequalities (4.12a–c) are strict. Since $C_1 - t^2 - \|x\|_2^2 \in \Gamma_\mu$ and $C_2 - \|x\|_2^2 \in \Lambda_\mu$ for all integers μ by assumption, a straightforward corollary of Putinar's Positivstellensatz [67, Lemma 4.1] guarantees that inequalities (4.12a–c) can be proved with WSOS certificates. Precisely, there exists an integer μ' such that the polynomials in (4.12a,b) belong to $\Gamma_{\mu'}$, and the polynomial in (4.12c) belongs to $\Lambda_{\mu'}$. We now set $d = \max\{\deg(V), \mu'\}$ and observe that V is feasible for the righthand problem in (4.7) with $\lambda = \Phi_T^* + \varepsilon$ because $\Gamma_{\mu'} \subseteq \Gamma_d$, $\Lambda_{\mu'} \subseteq \Lambda_d$, and $V \in \mathbb{R}_d[t, x]$. This proves the claim that $\lambda_d^* \le \Phi_T^* + \varepsilon$. ■

The computational cost of solving WSOS optimization problems grows quickly as d is raised. For instance, suppose the polynomials $f_1, \dots, f_p$ and $h_1, \dots, h_r$ all have the same degree ω, and let $d_F := d-1+\deg(F)$. Then, the time for standard primal-dual interior-point methods scales as $\mathcal{O}(L_1^{6.5} + (p+r)^{1.5} L_2^{6.5})$, where $L_1 = \binom{n+\lfloor d_F/2 \rfloor}{n}$ and $L_2 = \binom{n+\lfloor (d-\omega)/2 \rfloor}{n}$; see [63] and references therein for further details. Appendix C describes a way to improve bounds iteratively without raising d, but the improvement is small in the example tested. Poor computational scaling with increasing d can be partly mitigated if symmetries of optimal V can be anticipated and enforced in advance, leading to smaller SDPs. When the differential equations, the observable Φ, and the sets Ω and X_0 all are invariant under a symmetry transformation, then the optimal bound is unchanged if the symmetry is imposed also on V and the weights σ_i and ρ_i. The next proposition formalizes these observations; its proof is a straightforward adaptation of a similar result in Appendix A of [27], so we do not report it.

Proposition 4.2. *Let $A : \mathbb{R}^{n\times n}$ be an invertible matrix such that A^k is the identity for some integer k. Assume that $F(t, Ax) = AF(t, x)$, Φ is A-invariant in the sense that $\Phi(t, Ax) = \Phi(t, x)$, and all polynomials defining Ω and X_0 are A-invariant also. If $V \in \mathcal{V}(\Omega)$ gives a bound $\Phi^* \le \lambda$, then there exits $\widehat{V} \in \mathcal{V}(\Omega)$ that is A-invariant and proves the same bound. Moreover, if the pair (V, λ) satisfies the WSOS constraints in (4.7), then so does the pair $(\widehat{V}, \lambda)$ and there exist WSOS decompositions with A-invariant weights σ_i, ρ_i.*

We conclude this section with three computational examples. The first two demonstrate that SOS optimization can give extremely good bounds on both Φ_T^* and Φ_∞^* in practice, even when the assumptions of Theorems 2.5 and 4.1 do not hold. The first example also illustrates the approximation of optimal trajectories described in section 3. The third example, on the other hand, reveals a potential pitfall of SOS optimization applied to bounding Φ_∞^* for systems with periodic orbits: infeasible problems may appear to be solved successfully due to unavoidably finite tolerances in SDP solvers.

Example 4.3. Consider the nonlinear autonomous ODE system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0.2x_1 + x_2 - x_2(x_1^2 + x_2^2) \\ -0.4x_2 + x_1(x_1^2 + x_2^2) \end{bmatrix}, \tag{4.13}$$

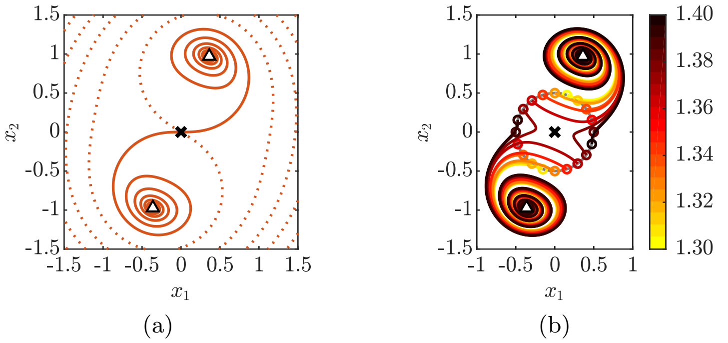

Figure 2. *(a) Phase portrait of the ODE (4.13) showing the attracting equilibria (△), the saddle (×), and the saddle's unstable (—) and stable (····) manifolds. (b) Sample trajectories starting from the circle $\|x\|_2^2 = 0.25$. Small circles mark the initial conditions. Colors indicate the maximum value of $\Phi = \|x\|_2^2$ along each trajectory.*

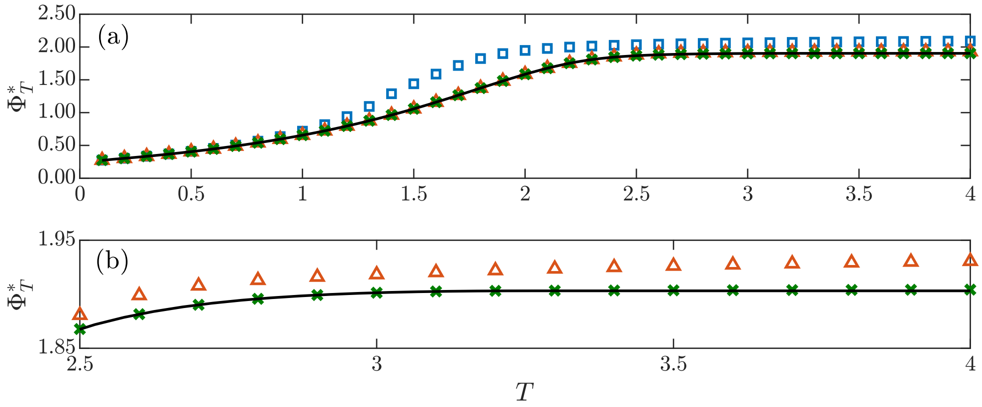

Figure 3. *(a) Upper bounds on Φ_T^* in Example 4.3 for various time horizons T, computed using auxiliary functions $V(t,x)$ with polynomial degrees 4 (□), 6 (△), and 8 (×). Lower bounds on Φ_T^* found by maximizing $\Phi[x(T;0,x_0)]$ over x_0 using adjoint optimization are also plotted (—). (b) Detailed view of part of panel (a).*

which is symmetric under $x \mapsto -x$. As shown in Figure 2(a), the system has a saddle point at the origin and a symmetry-related pair of attracting equilibria. Let $X_0 = \{x : \|x\|_2^2 = 0.25\}$. Aside from two points on the stable manifold of the origin, all points in X_0 produce trajectories that eventually spiral outwards towards the attractors, as shown in Figure 2(b).

Using SOS optimization, we have computed upper bounds on the value of $\Phi(x) = \|x\|_2^2$ among all trajectories starting from X_0, for both finite and infinite time horizons. For simplicity we considered only global auxiliary functions, meaning we used $\Omega = [0,T] \times \mathbb{R}^2$ and $\Omega = [0,\infty) \times \mathbb{R}^2$ to solve (4.7) in the finite- and infinite-time cases, respectively. Since both choices of Ω and the set of initial conditions $X_0 = \{x : \|x\|_2^2 = 0.25\}$ share the same symmetry as (4.13), we applied Proposition 4.2 to reduce the cost of solving (4.7). Our implementation used YALMIP to reformulate (4.7) into an SDP, which was solved with MOSEK.

Figure 3 shows upper bounds on Φ_T^* that we computed for a range of time horizons T by solving (4.7) with time-dependent polynomial V of degrees $d = 4$, 6, and 8. Also plotted in the figure are lower bounds on Φ_T^*, found by searching among initial conditions using adjoint optimization. The close agreement with our upper bounds shows that the degree-8 bounds

Table 2

Upper bounds on Φ_T^* and Φ_∞^* for Example 4.3, computed by solving (4.7). The bounds for Φ_T^* and Φ_∞^* were computed using time-dependent and time-independent V, respectively. Lower bounds are implied by the maximum of Φ on particular trajectories, whose initial conditions were found by adjoint optimization.

	deg(V)	$T = 2$	$T = 3$	$T = \infty$
Upper bounds	4	1.948016	2.062952	2.194343
	6	1.584910	1.918262	1.942396
	8	1.584055	1.901411	1.931330
	10	"	1.901409	1.916228
	12	"	"	1.903525
	14	"	"	1.903448
	16	"	"	1.903185
	18	"	"	1.903181
Lower bounds		1.584055	1.901409	1.903178

are very close to sharp, and that adjoint optimization likely has found the globally optimal initial conditions. We find that $\Phi_T^* = \Phi_\infty^* \approx 1.90318$ for all $T \geq 3.2604$, indicating that Φ attains its maximum over all time when $T \approx 3.2604$.

Table 2 reports upper bounds on Φ_T^* computed with time-dependent V up to degree 18 for $T = 2$ and $T = 3$, as well as upper bounds on Φ_∞^*. The infinite-time implementation was restricted to time-independent polynomial $V(x)$ because polynomial dependence on t gave no improvement in preliminary computations. This restriction lowers the computational cost because the first two WSOS constraints in (4.7) are independent of time and reduce to standard SOS constraints on $\mathbb{R}^2$. The resulting bounds are excellent for each T reported in Table 2. As the degree of V is raised, the upper bounds on Φ^* apparently converge to the lower bounds produced by adjoint optimization. Note that this convergence is not guaranteed by Theorems 2.5 and 4.1 because the domain Ω is not compact.

Finally, we illustrate how auxiliary functions can be used to localize optimal trajectories using the methods described in section 3. For a near-optimal V we take the time-independent degree-14 auxiliary function that gives the upper bound $\lambda = 1.903448$ reported in Table 2. Any trajectory that attains or exceeds a value $\lambda - \delta$ at some time t^* must spend the interval $[t_0, t^*]$ inside the set $\mathcal{S}_\delta$ defined by (3.8). In the present example, the lower bound $1.903178 \leq \Phi^*$ guarantees the existence of such trajectories for all $\delta \geq 0.00027$. In general a good lower bound on Φ^* may be lacking, in which case the sets $\mathcal{S}_\delta$ tell us where near-optimal trajectories must lie *if* they exist. With this general situation in mind, Figure 4(a,b) show $\mathcal{S}_\delta$ for $\delta = 0.01$ and 0.002, along with the exactly optimal trajectories. The $\mathcal{S}_\delta$ sets localize the optimal trajectories increasingly well as δ is lowered, although they contain other parts of state space also. Figure 4(c) shows the sets $\mathcal{R}_\varepsilon$, defined by (3.12), for $\varepsilon = 0.008$ and 0.004. Each trajectory coming within $\delta = 0.002$ of the upper bound, for example, cannot leave these $\mathcal{R}_\varepsilon$ for longer than $\delta/\varepsilon = 0.25$ and 0.5 time units, respectively, prior to any time at which $\Phi \geq \lambda - \delta$. The same is true of the intersections of these sets with $\mathcal{S}_\delta$, which are shown in Figure 4(d). ■

Example 4.4. Here we consider a 16-dimensional ODE model obtained by projecting the Burgers equation (2.14) with ordinary diffusion ($\alpha = 1$) onto modes $u_n(x) = \sqrt{2}\sin(2n\pi x)$,

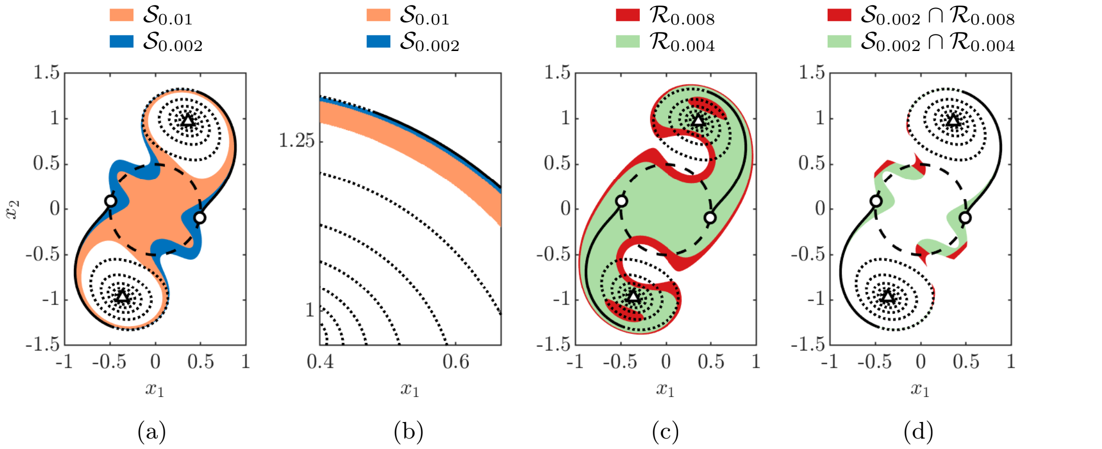

Figure 4. *Sets approximating the trajectories that attain Φ_∞^* for Example 4.3: (a) $\mathcal{S}_{0.01}$ and $\mathcal{S}_{0.002}$. (b) Detail view of part of panel (a). (c) $\mathcal{R}_{0.008}$ and $\mathcal{R}_{0.004}$. (d) $\mathcal{S}_{0.002} \cap \mathcal{R}_{0.008}$ and $\mathcal{S}_{0.002} \cap \mathcal{R}_{0.004}$. All sets were computed using the same degree-14 polynomial $V(x)$ that yields the nearly sharp bounds in Table 2. Also plotted are the attracting equilibria (△), the set of initial conditions X_0 (- - -), the optimal initial conditions (○), and the optimal trajectories before (—) and after (·····) the point at which Φ_∞^* is attained.*

$n = 1, \ldots, 16$. In other words, we substitute the expansion $u(x,t) = \sum_{m=1}^{16} a_m(t) u_m(x)$ into (2.14) with $\alpha = 1$ and integrate the result against each $u_n(x)$ to derive 16 nonlinear coupled ODEs for the amplitudes $a_1(t), \ldots, a_{16}(t)$. This gives

$$\dot{a}_n = -(2\pi n)^2 a_n + \sqrt{2}\pi n \left[\sum_{m=1}^{16-n} a_m a_{m+n} - \frac{1}{2} \sum_{m=1}^{n-1} a_m a_{n-m} \right], \qquad n = 1, \ldots, 16. \tag{4.14}$$

Let $a = (a_1, \ldots, a_{16})$ denote the state vector. Similarly to what is done for the PDE in Example 2.2, we bound the projected enstrophy $\Phi(a) := 2\pi^2 \sum_{n=1}^{16} n^2 a_n^2$ along trajectories with initial conditions in the set $X_0 = \{a \in \mathbb{R}^{16} : \Phi(a) = \Phi_0\}$, and we consider various values Φ_0 of the initial enstrophy. We construct time-independent degree-d polynomial V of the form

$$V(a) = c\|a\|_2^d + P_{d-1}(a), \tag{4.15}$$

where d is even, c is a tunable constant, and $P_{d-1}(a)$ is a tunable polynomial of degree $d-1$. Since the nonlinear terms in (4.14) conserve the leading $\|a\|_2^d$ term, $\mathcal{L}V$ has the same even leading degree as V, which is necessary for (2.5a,b) to hold over the global spacetime set $\Omega = [0,\infty) \times \mathbb{R}^{16}$. We also construct local V of the form (4.15) by imposing (2.5a,b) only on the smaller spacetime set $\Omega = [0,\infty) \times X$ with

$$X := \left\{ a \in \mathbb{R}^{16} : \|a\|_2^2 \le \frac{\Phi_0}{2\pi^2} \right\}. \tag{4.16}$$

All trajectories starting from X_0 remain in X because (4.14) implies $\frac{\mathrm{d}}{\mathrm{d}t}\|a\|_2^2 = -4\Phi(a) \le 0$, so $\|a\|_2^2$ is bounded by its initial value, and $\|a\|_2^2 \le \frac{1}{2\pi^2}\Phi(a)$ pointwise.

Figure 5 shows upper bounds on Φ_∞^* computed for Φ_0 values spanning four orders of magnitude using both global and local V of degrees 4 and 6. Also shown are lower bounds

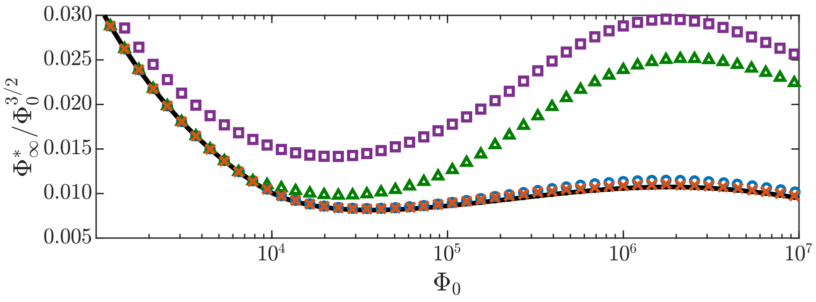

Figure 5. *Bounds on Φ^*_∞ for (4.14) computed with both global and local polynomial auxiliary functions V of the form (4.15) for $d = 4$ (□ global, △ local) and $d = 6$ (○ global, × local). Also plotted are lower bounds on Φ^*_∞ obtained with adjoint optimization (—). All results are normalized by $\Phi_0^{3/2}$, the expected scaling at large Φ_0 [5].*

obtained using adjoint optimization. (Note that the 16-mode truncation (4.15) accurately resolves Burgers equation only in cases with $\Phi_0 \lesssim 2 \cdot 10^5$.) We used SPOTLESS and MOSEK to solve (4.7) and applied Proposition 4.2 to exploit symmetry under the transformation $a_n \mapsto (-1)^n a_n$. At each Φ_0 value, constructing quartic V required approximately 60 seconds on 4 cores with 16GB of memory. Local quartic V produce better bounds than global ones, the results obtained with the former being within 1% of the lower bounds from adjoint optimization for $\Phi_0 \lesssim 8000$. The results improve significantly with sextic V: for all tested Φ_0, the upper bounds produced by global and local sextic V are within 9% and 5% of the adjoint optimization results, respectively. Constructing sextic V at a single Φ_0 value required 16 hours on a 12-core workstation with 48GB of memory, which is significantly more expensive than adjoint optimization. However, we stress that auxiliary functions yield *upper* bounds on Φ^*_∞, while adjoint optimization gives only *lower* bounds on Φ^*_∞, so the two approaches give different and complementary results. ■

It is evident that SOS optimization can produce excellent bounds on extreme events given enough computational resources, but care must be taken to assess whether numerical results can be trusted. As observed already in the context of SOS optimization [82], numerical SDP solvers can return solutions that appear to be correct but are provably not so. The next example shows that this issue can arise when bounding Φ^*_∞ in systems with periodic orbits.

Example 4.5. Consider a scaled version of the van der Pol oscillator [77],

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix} = \begin{bmatrix} x_2 \\ (1-9x_1^2)x_2 - x_1\end{bmatrix}, \tag{4.17}$$

which has a limit cycle attracting all trajectories except the unstable equilibrium at the origin (see Figure 6). Let $\Phi = \|x\|_2^2$ be the observable of interest. We seek bounds on Φ^*_∞ along trajectories starting from the circle $\|x\|_2^2 = 0.04$. All such trajectories approach the limit cycle from the inside, so Φ^*_∞ coincides with the pointwise maximum of Φ on the limit cycle. Maximizing Φ numerically along the limit cycle yields $\Phi^*_\infty \approx 0.889856$.

We implemented (4.7) with YALMIP using a time-independent polynomial auxiliary function $V(x)$ of degree 22. To confirm that difficulties were not easily avoided by increasing precision, we solved the resulting SDP in multiple precision arithmetic using the solver SDPA-GMP

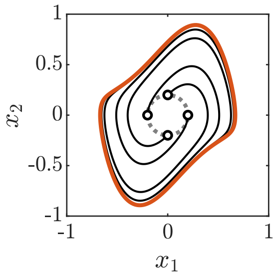

Figure 6. *Limit cycle (—) for the scaled van der Pol oscillator (4.17). Also plotted are trajectories (—) with initial conditions (○) on the circle $\|x\|_2^2 = 0.04$ (- - -).*

Table 3

Parameters for SDPA-GMP used in Example 4.5 to produce an invalid degree-22 auxiliary function for the scaled van der Pol oscillator. A description of each parameter can be found in [24].

epsilonStar	10^{-25}	betaStar	0.1	lowerBound	-10^{25}	maxIteration	200
epsilonDash	10^{-25}	betaBar	0.3	upperBound	10^{25}	precision	200
lambdaStar	10^{4}	gammaStar	0.7	omegaStar	2		

v.7.1.3. The solver parameters we used are listed in Table 3 in order to ensure that our results are reproducible; see [24] for the meaning of each parameter. The solver terminated successfully after 95 iterations, reporting no error and returning the upper bound $\Phi_\infty^* \leq 0.956911$. Although this bound is true, it reflects an invalid SOS solution because no time-independent polynomial V of any degree can satisfy (2.5a). To see this, suppose that (2.5a) holds, so V cannot increase along trajectories of (4.17). In particular, if $x(t)$ lies on the limit cycle and τ is the period, then for all $\alpha \in (0, 1)$,

$$V[x(t)] \geq V[x(t + \alpha\tau)] \geq V[x(t + \tau)] = V[x(t)]. \tag{4.18}$$

Thus, time-independent V giving finite bounds on Φ_∞^* must be constant on the limit cycle. This is impossible if V is polynomial because the limit cycle is not an algebraic curve [61].

There are two possible reasons why the SDP solver does not detect that the problem is infeasible despite the use of multiple precision. The first is that inevitable roundoff errors mean that our bound does not apply to (4.17), but to a slightly perturbed system whose limit cycle *is* an algebraic curve. The second possibility, which seems more likely, is that although no time-independent polynomial V is feasible, there exists a feasible nonpolynomial V that can be approximated accurately near the limit cycle by a degree-22 polynomial. In particular, the approximation error is smaller than the termination tolerances used by the solver, which therefore returns a solution that is not feasible but very nearly so. This interpretation is supported by the fact that SDPA-GMP issues a warning of infeasibility when its tolerances are tightened by lowering the values of parameters epsilonDash and epsilonStar to 10^{-30}. ∎

5. Extensions.

The framework for bounding extreme events presented in section 2 can be extended in several ways. Here we briefly summarize two extensions. Both are covered by the measure-theoretic approach of [81, 80, 48, 79], but we give a more direct derivation.

The first extension applies when upper bounds are sought on the maximum of Φ at a fixed finite time T, rather than its maximum over the time interval $[0,T]$. Such bounds can be proved by relaxing inequality (2.5b) to require that V bounds Φ only at time T.

A second extension lets extreme events be defined using integrals over trajectories in addition to instantaneous values. Precisely, suppose the quantity we want to bound from above is

$$\sup_{\substack{x_0\in X_0\\ t\in\mathcal{T}}} \left\{ \Phi[t,x(t;t_0,x_0)] + \int_{t_0}^{t} \Psi[\xi,x(\xi;t_0,x_0)]\,\mathrm{d}\xi \right\} \tag{5.1}$$

with chosen Φ and Ψ. One way to proceed is to augment the original dynamical system (2.1) with the scalar ODE $\dot{z} = \Psi(t,x)$, $z(t_0) = 0$. Bounding (5.1) along trajectories of the original system is equivalent to bounding the maximum of $\Phi(t,x) + z$ pointwise in time along trajectories of the augmented system, and this can be done with the methods described in the previous sections. Another way to bound (5.1), without introducing an extra ODE, is to replace condition (2.5a) with

$$\mathcal{L}V(t,x) + \Psi(t,x) \le 0 \quad \forall (t,x)\in\Omega. \tag{5.2}$$

Minor modification to the argument leading to (2.6) proves that

$$\sup_{\substack{x_0\in X_0\\ t\in\mathcal{T}}} \left\{ \Phi[t,x(t;t_0,x_0)] + \int_{t_0}^{t} \Psi[\xi,x(\xi;t_0,x_0)]\,\mathrm{d}\xi \right\} \le \inf_{\substack{V:\,(2.5\mathrm{b})\\ (5.2)}} \sup_{x_0\in X_0} V(t_0,x_0). \tag{5.3}$$

As in (2.6), the righthand minimization is a convex problem and can be tackled computationally using SOS optimization for polynomial ODEs when Φ and Ψ are polynomial. Analogues of Theorems 2.5 and 4.1 for (5.3) hold if Ψ is continuous.

6. Conclusions.

We have discussed a convex framework for constructing *a priori* bounds on extreme events in nonlinear dynamical systems governed by ODEs or PDEs. Precisely, we have described how to bound from above the maximum value Φ^* of an observable $\Phi(t,x)$ over a given finite or infinite time interval, among all trajectories that start from a given initial set. This approach, which is a particular case of general relaxation frameworks for optimal control and optimal stopping problems [48, 11], relies on the construction of auxiliary functions $V(t,x)$ that decay along trajectories and bound Φ pointwise from above. These constraints amount to the pointwise inequalities (2.5a,b) in time and state space, which can be either imposed globally or imposed locally on any spacetime set that contains all trajectories of interest. Suitable global or local V can be constructed without knowing any system trajectories, so Φ^* can be bounded above even when trajectories are very complicated. We have given a range of ODE examples in which analytical or computational constructions give very good and sometimes sharp bounds. As a PDE example, we have proved analytical upper bounds on a quantity called fractional enstrophy for solutions to the one-dimensional Burgers equation with fractional diffusion.

The convex minimization of upper bounds on Φ^* over global or local auxiliary functions is dual to the non-convex maximization of Φ along trajectories. In the case of ODEs and local auxiliary functions, Theorem 2.5, which is a corollary of Theorem 2.1 and equation (5.3) in [48], guarantees that this duality is strong when the time interval is finite and the ODE satisfies certain continuity and compactness assumptions. This means that the infimum over bounds is equal to the maximum over trajectories, so there exist V proving arbitrarily sharp bounds on Φ^*. Further, strong duality holds in several of our ODE examples to which the assumptions of Theorem 2.5 do not apply, including formulations with global V or infinite time horizons. However, neither the proofs in [48] nor our alternative proof in Appendix D can be easily extended to these cases because they rely on compactness, and we have given counterexamples to strong duality with infinite time horizon even when trajectories remain in a compact set. Better characterizing the dynamical systems for which strong duality holds remains an open challenge.

Regardless of whether duality is weak or strong for a given dynamical system, constructing auxiliary functions that yield good bounds often demands ingenuity. Fortunately, as described in section 4, computational methods of sum-of-squares (SOS) optimization can be applied in the case of polynomial ODEs with polynomial Φ. Moreover, Theorem 4.1 guarantees that if strong duality and mild compactness assumptions hold, then bounds computed by solving the SOS optimization problem (4.7) become sharp as the polynomial degree of the auxiliary function V is raised. In practice, computational cost can become prohibitive as either the dimension of the ODE system or the polynomial degree of V increases, at least with the standard approach to SOS optimization wherein generic semidefinite programs are solved by second-order symmetric interior-point algorithms. For instance, given a 10-dimensional ODE system with no symmetries to exploit, the degree of V is currently limited to about 12 on a large-memory computer. Larger problems may be tackled using specialized nonsymmetric interior-point [63] or first-order algorithms [86, 87]. One also could replace the weighted SOS constraints in (4.7) with stronger constraints that may give more conservative bounds at less computational expense [1, 2].

In the case of PDEs, the bounding framework of section 2 can produce valuable bounds, as in Example 2.2, but theoretical results and computational tools are lacking. Theorem 2.5, which guarantees arbitrarily sharp bounds for many ODEs, does not apply to PDEs, nor can we directly apply the computational methods of section 4 that work well for polynomial ODEs. On the theoretical side, guarantees that feasible auxiliary functions exist for PDEs would be of great interest, not least because bounds on certain extreme events can preclude loss of regularity. Statements formally dual to results in [11] for optimal stopping problems would imply that near-optimal auxiliary functions exist for autonomous PDEs, at least when extreme events occur at finite time, but such statements have not yet been proved. On the computational side, constructions of optimal V for PDEs would be very valuable, both to guide rigorous analysis and to improve on conservative bounds proved by hand. Methods of SOS optimization can be applied to PDEs in two ways. The first is to approximate the PDE as an ODE system and bound the error this incurs, obtaining an "uncertain" ODE system to which standard SOS techniques can be applied [28, 10, 35, 27]. The second approach is to work directly with the PDE using either the integral inequality methods of [74, 76, 73] or the moment relaxation techniques of [42, 57]. These strategies have been used to study PDE stability, time averages, and optimal control, but they are in relatively early development. They have not yet been applied to extreme events as studied here, although the method in [42] applies to extreme behavior at a fixed time and could be extended to time intervals. It remains to be seen whether any of these strategies can numerically optimize auxiliary functions for PDEs of interest at reasonable computational cost, but recent advances in optimization-based formulations and corresponding numerical algorithms give us hope that this will be possible in the near future.

Acknowledgments. We are indebted to Andrew Wynn, Sergei Chernyshenko, Ian Tobasco, and Charles Doering, who offered many insightful comments on this work. We also thank the anonymous referees for comments that considerably improved the original version of this work.

Appendix A. Optimality of the quadratic V in Example 2.1. The V given by (2.10) is optimal among all quadratic global auxiliary functions that produce upper bounds on $\Phi = x_1$ along the trajectory starting from the point $(0, 1)$. To prove this, consider a general quadratic global auxiliary function,

$$\begin{aligned} V(t, x_1, x_2) = C_0 + C_1 x_1 + C_2 x_2 + C_3 t \\ + C_4 x_1^2 + C_5 x_2^2 + C_6 t^2 + 2C_7 x_1 x_2 + 2C_8 t x_1 + 2C_9 t x_2. \end{aligned} \tag{A.1}$$

The coefficients $C_0, \ldots, C_9$ must be chosen to minimize the bound $\Phi^* \le V(0, 0, 1)$ implied by (2.8), subject to the inequality constraints (2.5a,b). Differentiating V along solutions of (2.9) yields

$$\begin{aligned} \mathcal{L}V(t, x_1, x_2) = C_3 + (2C_9 - C_2)x_2 + (2C_8 - 0.1C_1)x_1 + 2C_6 t + (C_2 - 0.2C_4)x_1^2 \\ - (2.2C_7 + C_1)x_1 x_2 - 2C_5 x_2^2 + (C_1 - 2C_9)t x_2 - (C_2 + 0.2C_8)t x_1 \\ + 2C_7 x_1^3 - 2C_7 x_1 x_2^2 + 2(C_5 - C_4)x_1^2 x_2 + 2C_7 t x_2^2 \\ + 2(C_4 - C_5 - C_8)t x_1 x_2 + 2(C_9 - C_7)t x_1^2 - 2C_9 t^2 x_1 + 2C_8 t^2 x_2. \end{aligned} \tag{A.2}$$

In order for this expression to be nonpositive for all $(x_1, x_2) \in \mathbb{R}^2$ and $t \ge 0$, as required by (2.5a), the indefinite cubic terms and the quadratic terms proportional to t must vanish. This forces us to set $C_1, C_2, C_7, C_8, C_9 = 0$ and $C_4 = C_5$, so the expressions for V and $\mathcal{L}V$ reduce to

$$V(t, x_1, x_2) = C_0 + C_3 t + C_6 t^2 + C_5 \left(x_1^2 + x_2^2\right), \tag{A.3a}$$

$$\mathcal{L}V(t, x_1, x_2) = C_3 + 2C_6 t - 0.2C_5 x_1^2 - 2C_5 x_2^2. \tag{A.3b}$$

Condition (2.5a), which requires $\mathcal{L}V \le 0$, is satisfied only if $C_3, C_6 \le 0$ and $C_5 \ge 0$. With $\Phi = x_1$ condition (2.5b) becomes $C_0 - x_1 + C_5 x_1^2 + C_3 t + C_6 t^2 + C_5 x_2^2 \ge 0$, which in turn requires $4C_0 C_5 \ge 1$. Minimizing the bound $\Phi^* \le V(0, 0, 1) = C_0 + C_5$ under these constraints yields $C_0, C_5 = \frac{1}{2}$, and we are free to choose any $C_3, C_6 \le 0$. Any such V is optimal, including (2.10) which results from choosing $C_3, C_6 = 0$.

Appendix B. Sharp bounds for nonzero initial conditions in Example 2.3. Auxiliary functions that give sharp bounds on $\Phi = 4x/(1 + 4x^2)$ along single trajectories of the ODE (2.24) exist for every nonzero initial condition x_0. Here we give global V, which also are local V on any Ω in which trajectories remain. In the $x_0 > 0$ case, a global V giving sharp upper bounds on Φ^*_∞ is

$$V(t,x) = \begin{cases} 1, & x \le \frac{1}{2}, \\ \dfrac{4x}{1+4x^2}, & x > \frac{1}{2}. \end{cases} \tag{B.1}$$

This function is continuously differentiable and satisfies (2.5a,b). It is optimal because the bound on Φ^*_∞ implied by (2.6) with $X_0 = \{x_0\}$ is

$$\Phi^*_\infty \le V(0,x_0) = \begin{cases} 1, & 0 < x_0 \le \frac{1}{2}, \\ \dfrac{4x_0}{1+4x_0^2}, & x_0 > \frac{1}{2}, \end{cases} \tag{B.2}$$

which coincides with the expression (2.26) for Φ^*_∞.

The $x_0 < 0$ case requires a more complicated construction. An argument similar to that in Example 2.3 shows that any global optimal V providing the sharp bound $\Phi^*_\infty \le 0$ must be time-dependent. The same is true for local V unless $\Omega \subseteq [0,\infty) \times (-\infty, 0]$, in which case $V = 0$ is optimal. To construct a time-dependent global V that is optimal for $X_0 = \{x_0\}$ with x_0 negative, we note that $\beta(t) = x_0/(1 - x_0 t)$ solves the ODE (2.24) with initial condition $x(0) = x_0$. Observe that $\beta(0) = x_0$, $\beta(t) < 0$, and $\beta'(t) = \beta(t)^2$. Consider

$$\rho(x) = \begin{cases} \exp\left(1 - \dfrac{1}{1-x^2}\right), & |x| < 1, \\ 0, & |x| \ge 1, \end{cases} \tag{B.3}$$

which is a smooth nonnegative function. We claim that

$$V(t,x) := \begin{cases} \rho\left(\dfrac{x}{\beta(t)}\right), & x \le 0, \\ 1, & x > 0 \end{cases} \tag{B.4}$$

is an optimal global auxiliary function. This V implies the sharp bound $\Phi^*_\infty \le V(0,x_0) = 0$ since $\rho(1) = 0$, so it remains only to check (2.5a,b). Inequality (2.5b) holds because Φ is nonpositive for $x \le 0$ and is bounded above by 1 pointwise. To verify (2.5a), we consider positive and nonpositive x separately. The $x > 0$ case is immediate because $\mathcal{L}V(t,x) = 0$. For $x \le 0$, a straightforward calculation using $\beta'(t) = \beta(t)^2$ gives

$$\mathcal{L}V(t,x) = \partial_t V + x^2\,\partial_x V = \frac{x}{\beta(t)}\,[x - \beta(t)]\,\rho'\left(\frac{x}{\beta(t)}\right). \tag{B.5}$$

Observe that $\rho'(s)$ vanishes if $s = 0$ or $|s| \ge 1$, so $\mathcal{L}V = 0$ if $x \le \beta(t)$ or $x = 0$. When $\beta(t) < x < 0$ instead, $\mathcal{L}V < 0$ because the first two factors in (B.5) are positive, while $\rho'(s)$ is negative for $0 < s < 1$. Combining these observations shows that $\mathcal{L}V \le 0$ for all times if $x \le 0$. Figure 7 illustrates the behavior of V and $\mathcal{L}V$ when $x_0 = -\frac{3}{4}$.

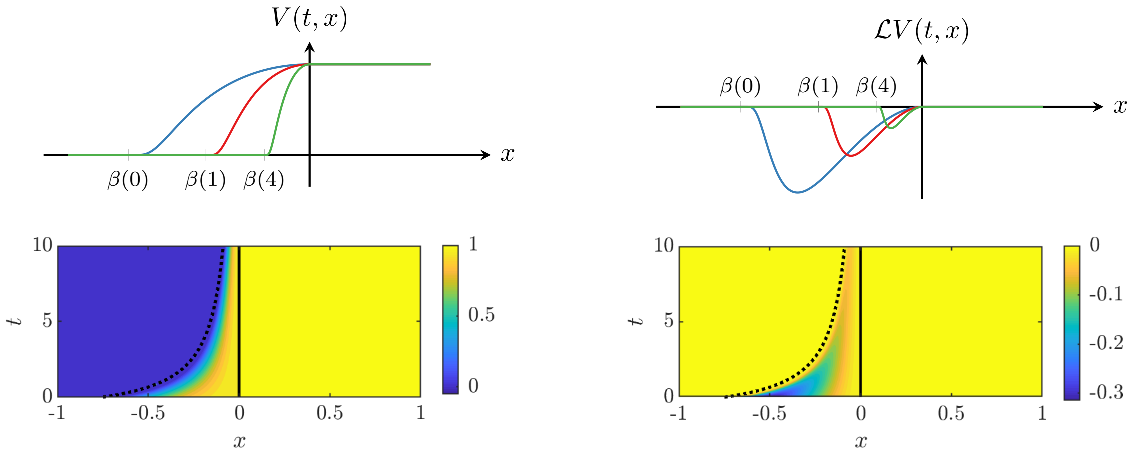

Figure 7. *Top row: Profiles of the auxiliary function $V(t,x)$ in (B.4) and its derivative along trajectories $\mathcal{L}V(t,x)$, plotted as a function of x for $t=0$, $t=1$, and $t=4$. Bottom row: Contours of V (left) and $\mathcal{L}V$ (right). Lines mark the trajectory $x=\beta(t)$ (···) and the semistable equilibrium $x=0$ (—). Outside the region between these two lines, $\mathcal{L}V=0$. All plots are for $x_0=-\frac{3}{4}$.*

Table 4

Upper bounds on Φ^*_∞ for Example 4.3, computed using time-independent polynomial auxiliary functions $V(x)$ of degree d by the iterative procedure described in Appendix C.

Iteration	$d=4$	$d=6$	$d=8$	$d=10$	$d=12$	$d=14$
1	2.194343	1.942396	1.931330	1.916228	1.903525	1.903448
2	2.194343	1.934692	1.926088	1.913889	1.903346	1.903307
3	2.194343	1.934643	1.926088	1.913817	1.903280	1.903250
4	2.194342	1.934642	1.926086	1.913815	1.903260	1.903222
5	2.194342	1.934642	1.926086	1.913814	1.903249	1.903207

Appendix C. Improving bounds iteratively with polynomial V of fixed degree. Bounds computed with (4.7) can be improved without increasing the degree d by using an iterative procedure. First, solve (4.7) to obtain an upper bound $\Phi^* \le \lambda^*_{d,0}$, which implies $\Phi(t,x) \le \lambda^*_{d,0}$ along trajectories of interest. Then, replace the original set Ω in which trajectories remain with its subset $\Omega_1 := \Omega \cap \{(t,x) : \Phi(t,x) \le \lambda^*_{d,0}\}$. Since $\Omega_1 \subseteq \Omega$ is still basic semialgebraic, one can solve (4.7) again, but with the WSOS constraints defined on Ω_1 rather than Ω. This produces a new bound, $\Phi^* \le \lambda^*_{d,1} \le \lambda^*_{d,0}$. The process can be iterated by taking $\Omega_{i+1} = \Omega \cap \{(t,x) : \Phi(t,x) \le \lambda^*_{d,i}\}$, $i = 1, 2, \ldots$, until the bound on Φ^* stops improving. The WSOS optimization problem to be solved for each i has constant computational cost, which is higher than the original one but typically much smaller than solving (4.7) with larger d.

Table 4 reports bounds on Φ^*_∞ obtained with this iterative procedure for the problem described in Example 4.3, using polynomial V of degrees up to 14. Each iteration lowers the bound as expected. The improvement with each iteration is small in this example, especially with lower-degree V. Raising d by 2 offers much more improvement except when the bound is nearly sharp already. It remains to be tested whether the iterative scheme brings more gains for other problems.

Appendix D. An elementary proof of Theorem 2.5.

Under the assumptions of Theorem 2.5, differentiable auxiliary functions that produce arbitrarily sharp bounds on Φ_T^* can be constructed by approximating the optimal but generally discontinuous V^* defined in section 2.3.2. This construction, which resembles the argument in [33], yields Theorem 2.5 without the measure theory or convex analysis used in the proofs of [48].

D.1. Construction of near-optimal V.

Let $\delta > 0$. We must show that there exists a C^1 function V on $\Omega = [t_0, T] \times X$ that satisfies (2.5a,b) and

$$\sup_{x_0 \in X_0} V(t_0, x_0) \le \Phi_T^* + \delta. \tag{D.1}$$

To do this we construct $W \in C^1(\Omega)$ such that

$$\mathcal{L}W(t,x) \le \frac{\delta}{5(T-t_0)} \qquad \text{on } \Omega, \tag{D.2a}$$

$$\Phi(t,x) \le W(t,x) + \frac{2}{5}\delta \qquad \text{on } \Omega, \tag{D.2b}$$

$$\sup_{x_0 \in X_0} W(t_0, x_0) \le \Phi_T^* + \frac{2}{5}\delta. \tag{D.2c}$$

Then, (2.5a,b) and (D.1) are satisfied by the continuously differentiable function

$$V(t,x) := W(t,x) + \frac{2}{5}\delta + \frac{(T-t)\delta}{5(T-t_0)}. \tag{D.3}$$

Our construction of W uses the flow map $S_{(s,t)} : Y \to \mathbb{R}^n$, defined for any two fixed time instants s and t such that $t_0 \le s \le t \le t_1$ as $S_{(s,t)}y = x(t; s, y)$. In other words, $S_{(s,t)}y$ is the point at time t on the trajectory of the ODE $\dot{x} = F(\xi, x)$ that passed through y at time s. An explicit expression for the flow map is generally not available. Nonetheless, under the assumptions of Theorem 2.5, the flow map is well defined and satisfies

$$S_{(s,t)}y = y + \int_s^t F[\xi, S_{(s,\xi)}y] \,\mathrm{d}\xi, \tag{D.4a}$$

$$S_{(s,t)} \circ S_{(r,s)} = S_{(r,t)} \qquad \forall r,t,s :\ t_0 \le r \le s \le t. \tag{D.4b}$$

The function $(t, s, y) \mapsto S_{(s,t)}y$ is uniformly continuous with respect to both s and y for t in compact time intervals; see, for instance, [30, Chapter V, Theorem 2.1]. It also is locally Lipschitz in the sense of the following Lemma, which is proved in Appendix D.2.

Lemma D.1. *Suppose the assumptions of Theorem 2.5 hold and let $[a, b] \times K$ be a compact subset of $[t_0, t_1] \times Y$. There exist positive constants C_1 and C_2, dependent only on a, b, K, t_0 and t_1, such that:*

(i) $\|S_{(t,\xi)}x - S_{(t,\xi)}y\| \le C_1 \|x - y\|$ *for all $x, y \in K$, all $t \in [a, b]$, and all $\xi \in [t, t_1]$.*

(ii) $\|S_{(t,\xi)}x - S_{(s,\xi)}x\| \le C_2\, |t - s|$ *for all $x \in K$, all $t, s \in [a, b]$, and all $\xi \in [\max(t, s), t_1]$.*

We also need the following Lemma, proved in Appendix D.3, which states that the optimal but possibly discontinuous auxiliary function defined by (2.40) can be approximated by a locally Lipschitz function.

Lemma D.2. *There exist* $t_2 \in (T, t_1)$ *and a locally Lipschitz function* $U : [t_0, t_2] \times Y \to \mathbb{R}$ *that satisfies*

$$\Phi(t,x) \le U(t,x) + \frac{\delta}{5} \quad \text{on } \Omega, \tag{D.5a}$$

$$\sup_{x_0 \in X_0} U(t_0, x_0) \le \Phi_T^* + \frac{\delta}{5}, \tag{D.5b}$$

and, for each fixed $(t,x) \in [t_0, t_2) \times Y$,

$$U(t+\varepsilon, S_{(t,t+\varepsilon)}x) \le U(t,x) \qquad \forall \varepsilon \in (0, t_2 - t). \tag{D.5c}$$

A function $W \in C^1(\Omega)$ that satisfies (D.2a,b,c) can be constructed by mollifying U "forward in time" on Ω. Precisely, fix any nonnegative differentiable mollifier $\rho(t,x)$ that is supported on the closed unit ball of $\mathbb{R} \times \mathbb{R}^n$ and has unit integral. For each $k \ge 1$ define

$$\rho_k(t,x) := k^{n+1} \rho(kt+1, kx). \tag{D.6}$$

Observe that ρ_k is supported on $R_k = [-2k^{-1}, 0] \times B_n(0, k^{-1})$, where $B_n(0,r)$ denotes the closed n-dimensional ball of radius r centered at the origin, and has unit integral. Let k be large enough that $[t_0, t_2] \times Y$ contains the compact set

$$\mathcal{N} = \{(t-s, x-y) : (t,x) \in \Omega,\ (s,y) \in R_k\}. \tag{D.7}$$

Note that $\Omega \subset \mathcal{N}$. For each $(t,x) \in \Omega$, define

$$W(t,x) := (\rho_k * U)(t,x) = \int_{R_k} \rho_k(s,y) U(t-s, x-y)\, \mathrm{d}s\, \mathrm{d}^n y. \tag{D.8}$$

Since R_k contains only nonpositive times $s \le 0$, W is a forward-in-time mollification of U. Standard arguments [19, Appendix C.4] show that W is continuously differentiable on Ω. Because Ω is compact and U is continuous, $W \to U$ uniformly on Ω as $k \to \infty$. Thus we can choose k large enough to ensure

$$\|U - W\|_{C^0(\Omega)} \le \frac{\delta}{5}, \tag{D.9}$$

To see that W satisfies (D.2c), combine (D.9) with (D.5b) to estimate

$$\sup_{x_0 \in X_0} W(t_0, x_0) \le \sup_{x_0 \in X_0} U(t_0, x_0) + \|U - W\|_{C^0(\Omega)} \le \Phi_T^* + \frac{2}{5}\delta. \tag{D.10}$$

We similarly obtain (D.2b) by estimating the righthand side of (D.5a) as

$$\Phi(t,x) \le U(t,x) + \frac{\delta}{5} \le W(t,x) + \|U - W\|_{C^0(\Omega)} + \frac{\delta}{5} \le W(t,x) + \frac{2}{5}\delta. \tag{D.11}$$

To prove (D.2a), fix $(t, x) \in \Omega$ and bound

$$
\begin{aligned}
\text{(D.12)} \quad \mathcal{L}W(t,x) &= \lim_{\varepsilon\to 0} \frac{W(t+\varepsilon, S_{(t,t+\varepsilon)}x) - W(t,x)}{\varepsilon} \\
&= \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \int_{R_k} \rho_k(s,y) \left[U(t+\varepsilon-s, S_{(t,t+\varepsilon)}x - y) - U(t-s, x-y)\right] \mathrm{d}s\, \mathrm{d}^n y \\
&\le \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \int_{R_k} \rho_k(s,y) \big\{U(t+\varepsilon-s, S_{(t,t+\varepsilon)}x - y) \\
&\qquad\qquad -U[t+\varepsilon-s, S_{(t-s,t-s+\varepsilon)}(x-y)]\big\} \mathrm{d}s\, \mathrm{d}^n y \\
&\le \lim_{\varepsilon\to 0} \frac{C}{\varepsilon} \int_{R_k} \rho_k(s,y) \left\|S_{(t,t+\varepsilon)}x - y - S_{(t-s,t-s+\varepsilon)}(x-y)\right\| \mathrm{d}s\, \mathrm{d}^n y,
\end{aligned}
$$

where C is a positive constant independent of t and x. The two inequalities above follow, respectively, from (D.5c) and the uniform Lipschitz continuity of U on compact sets.

Since $t \le T < t_2$, forward-in-time trajectories are well defined for sufficiently small ε. Moreover, reasoning as in the proof of Lemma D.1 in Appendix D.2 shows that trajectories starting from the compact neighborhood $\mathcal{N}$ of Ω defined in (D.7) are uniformly bounded up to time t_2. Thus the rightmost integrand in (D.12) is uniformly bounded and, by the dominated convergence theorem, we can exchange the limit and the integral. Then, we can further estimate $\mathcal{L}W$ using the fact that ρ_k has unit integral over R_k, the relation (D.4a), and the mean value theorem:

$$
\begin{aligned}
\text{(D.13)} \quad \mathcal{L}W(t,x) &\le C \max_{(s,y)\in R_k} \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \left\|S_{(t,t+\varepsilon)}x - y - S_{(t-s,t-s+\varepsilon)}(x-y)\right\| \\
&= C \max_{(s,y)\in R_k} \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \left\| \int_t^{t+\varepsilon} F(\xi, S_{(t,\xi)}x)\, \mathrm{d}\xi - \int_{t-s}^{t-s+\varepsilon} F[\xi, S_{(t-s,\xi)}(x-y)]\, \mathrm{d}\xi \right\| \\
&= C \max_{(s,y)\in R_k} \left\|F(t,x) - F(t-s, x-y)\right\|.
\end{aligned}
$$

Both (t, x) and $(t-s, x-y)$ lie in the compact set $\mathcal{N}$. Since F is locally Lipschitz by assumption, it is uniformly Lipschitz on $\mathcal{N}$. Consequently, there exist a constant C', independent of t and x, and a k sufficiently large such that

$$
\text{(D.14)} \qquad \mathcal{L}W(t,x) \le C' \max_{(s,y)\in R_k} (|s| + \|y\|) = \frac{3C'}{k} \le \frac{\delta}{5(T-t_0)},
$$

meaning that W satisfies (D.2a) as claimed. This concludes the proof of Theorem 2.5.

Remark D.3. Defining ρ_k such that the mollification (D.8) is forward in time, so $s \le 0$ on R_k, is key to prove (D.14) for all $(t, x) \in \Omega = [t_0, T] \times X$. If $s > 0$ anywhere on R_k, given any finite k we would have $t - s < t_0$ for all $t \in [t_0, t_k]$ and some $t_k > t_0$. In this case, we would not have the first inequality in (D.12) for all $(t, x) \in \Omega$ because (D.5c) holds only after time t_0.

D.2. Proof of Lemma D.1. To establish part *(i)* of Lemma D.1, observe that assumption (A.2) in Theorem 2.5 guarantees that the trajectory starting from any $x \in K$ at any time $t \in [a, b]$ exists up to time t_1, so in particular $\|S_{(t,\xi)}x\|$ is bounded for all $\xi \in [t, t_1]$. Combining the compactness of $[a, b] \times K$ with the continuity of trajectories with respect to both the initial point and the initial time [30, Chapter V, Theorem 2.1] shows that trajectories are uniformly bounded in norm. Precisely, there exists a constant M, depending only on a, b, K and t_1, such that $\|S_{(t,\xi)}x\| \leq M$ for all $(t, x) \in [a, b] \times K$ and all $\xi \in [t, t_1]$. We therefore can apply Lemma 2.9 from [68] and the local Lipschitz continuity of $F(\cdot, \cdot)$ to find a constant Λ_1, dependent only on a, b and K, such that

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\|S_{(t,\xi)}x - S_{(t,\xi)}y\| \leq \|F(\xi, S_{(t,\xi)}x) - F(\xi, S_{(t,\xi)}y)\| \leq \Lambda_1\|S_{(t,\xi)}x - S_{(t,\xi)}y\| \tag{D.15}$$

for all $x, y \in K$, all $t \in [a, b]$, and all $\xi \in [t, t_1]$. Assertion *(i)* then follows with $C_1 = \mathrm{e}^{\Lambda_1 t_1}$ after applying Gronwall's inequality to bound

$$\|S_{(t,\xi)}x - S_{(t,\xi)}y\| \leq \mathrm{e}^{\Lambda_1 \xi}\|x - y\| \leq \mathrm{e}^{\Lambda_1 t_1}\|x - y\|. \tag{D.16}$$

To prove part *(ii)* of Lemma D.1, assume without loss of generality that $s < t$. For all $\xi \in [t, t_1]$, identity (D.4b) gives $\|S_{(t,\xi)}x - S_{(s,\xi)}x\| = \|S_{(t,\xi)}x - S_{(t,\xi)}S_{(s,t)}x\|$. Proceeding as above with $y = S_{(s,t)}x$ shows that

$$\|S_{(t,\xi)}x - S_{(s,\xi)}x\| \leq \Lambda_2 \left\|x - S_{(s,t)}x\right\| \tag{D.17}$$

for some positive constant Λ_2. Moreover, we can use (D.4a) to estimate

$$\left\|S_{(s,t)}x - x\right\| = \left\|\int_s^t F(\xi, S_{(s,\xi)}x)\,\mathrm{d}\xi\right\| \leq \sqrt{n}\int_s^t \|F(\xi, S_{(s,\xi)}x)\|\,\mathrm{d}\xi. \tag{D.18}$$

Since F is continuous and, as noted above, $\|S_{(s,\xi)}x\| \leq M$ for all $(s, x) \in [a, b] \times K$ and all $\xi \in [s, t_1] \subset [a, t_1]$,

$$\left\|S_{(s,t)}x - x\right\| \leq \sqrt{n} \max_{\substack{\xi\in[a,t_1]\\ \|y\|\leq M}} \|F(\xi, y)\|\;|t - s|\,. \tag{D.19}$$

Combining this with (D.17) proves the claim for a suitable choice of C_2.

D.3. Proof of Lemma D.2. Fix $t_2 = T + \gamma$ for some $\gamma > 0$ sufficiently small and to be determined later. Arguing as in the proof of Lemma D.1 *(i)*, trajectories starting from $x_0 \in X_0$ remain bounded uniformly in the initial condition and time. Precisely, there exists a constant M such that $\|S_{(t_0,t)}x_0\| \leq M$ for all $x_0 \in X_0$ and $t \in [t_0, t_2]$. If $\mathcal{B}$ denotes the n-dimensional ball of radius M centered at the origin, we conclude that the compact set $[t_0, t_2] \times \mathcal{B}$ contains $\Omega = [t_0, T] \times X$, the spacetime set in which trajectories starting from $x_0 \in X_0$ at time t_0 remain up to time T.

Let $\Psi : \mathbb{R} \times \mathbb{R}^n \times Y \to \mathbb{R}$ be a Lipschitz approximation of Φ satisfying

$$\|\Phi - \Psi\|_{C^0([t_0,t_2]\times\mathcal{B})} \leq \frac{\delta}{10}. \tag{D.20}$$

Such Ψ may be constructed in a number of ways, for instance by using the Stone–Weierstrass theorem to approximate Φ uniformly on the compact set $[t_0, t_2] \times \mathcal{B}$ by a polynomial, and extending such polynomial to a Lipschitz function on $\mathbb{R} \times \mathbb{R}^n$. We claim that t_2 can be chosen such that the function $U : [t_0, t_2] \times Y \to \mathbb{R}$ defined as

$$U(t,x) := \sup_{\tau \in [t,t_2]} \Psi[\tau, S_{(t,\tau)}x] \tag{D.21}$$

satisfies (D.5a–c). This U cannot be computed in practice but is well defined. Note that if Φ is Lipschitz we can choose $\Psi = \Phi$ and the restriction of U to Ω tends to the optimal but possibly discontinuous auxiliary function defined in (2.40) as $\gamma = t_2 - T$ tends to zero. If γ is finite but small, then U approximates this optimal auxiliary function. The same is true when Ψ only approximates Φ.

To see that (D.5a) holds, note that $U(t,x) \geq \Psi(t,x)$. Since $\Omega \subset [t_0, t_2] \times \mathcal{B}$ we conclude from (D.20) that, for all $(t,x) \in \Omega$,

$$\Phi(t,x) \leq \Psi(t,x) + \|\Phi - \Psi\|_{C^0([t_0,t_2]\times\mathcal{B})} \leq U(t,x) + \frac{\delta}{10} < U(t,x) + \frac{\delta}{5}. \tag{D.22}$$

To prove (D.5b), we will choose $\gamma = t_2 - T$ such that

$$U(t_0, x_0) = \sup_{\tau \in [t_0,t_2]} \Psi[\tau, S_{(t_0,\tau)}x_0] \leq \Phi_T^* + \frac{\delta}{5} \tag{D.23}$$

uniformly in the initial condition $x_0 \in X_0$. To do this, fix $x_0 \in X_0$ and observe that the supremum over $\tau \in [t_0, t_2]$ must be attained because the function $\tau \mapsto \Psi[\tau, S_{(t_0,\tau)}x_0]$ is continuous. If the supremum is attained on the interval $[t_0, T]$, then

$$\begin{aligned} \sup_{\tau \in [t_0,t_2]} \Psi[\tau, S_{(t_0,\tau)}x_0] &= \sup_{\tau \in [t_0,T]} \Psi[\tau, S_{(t_0,\tau)}x_0] \\ &\leq \sup_{\tau \in [t_0,T]} \Phi[\tau, S_{(t_0,\tau)}x_0] + \|\Phi - \Psi\|_{C^0([t_0,t_2]\times\mathcal{B})} \\ &\leq \Phi_T^* + \frac{\delta}{10}. \end{aligned} \tag{D.24}$$

Instead, if the supremum is attained at time $t^* \in [T, t_2]$, then we can use the Lipschitz continuity of Ψ, the group property (D.4b) of the flow map, and Lemma D.1*(ii)* to find constants C and C', dependent on t_0, t_1 and the set X_0 but not on the choice of $x_0 \in X_0$, such that

$$\begin{aligned} \sup_{\tau \in [t_0,t_2]} \Psi[\tau, S_{(t_0,\tau)}x_0] &= \Psi[t^*, S_{(t_0,t^*)}x_0] \\ &\leq \Psi[T, S_{(t_0,T)}x_0] + \left|\Psi[t^*, S_{(t_0,t^*)}x_0] - \Psi[T, S_{(t_0,T)}x_0]\right| \\ &\leq \Psi[T, S_{(t_0,T)}x_0] + C\,|t^* - T| + C\|S_{(T,t^*)}S_{(t_0,T)}x_0 - S_{(t_0,T)}x_0\| \\ &\leq \Phi[T, S_{(t_0,T)}x_0] + \|\Phi - \Psi\|_{C^0([t_0,t_2]\times\mathcal{B})} + (C + C')\,|t^* - T| \\ &\leq \Phi_T^* + \frac{\delta}{10} + (C + C')\gamma. \end{aligned} \tag{D.25}$$

Upon setting $\gamma = \delta/[10(C+C')]$, (D.24) and (D.25) prove that (D.23) holds uniformly in the initial condition x_0 irrespective of whether the sup over τ is attained before or after time T.

Finally, to obtain (D.5c), fix $(t,x) \in [t_0,t_2) \times Y$ and observe that, for all $\varepsilon \in (0, t_2-t)$,

$$\begin{aligned} U(t+\varepsilon, S_{(t,t+\varepsilon)}x) &= \sup_{\tau\in[t+\varepsilon,t_2]} \Psi[\tau, S_{(t+\varepsilon,\tau)}S_{(t,t+\varepsilon)}x] \\ &= \sup_{\tau\in[t+\varepsilon,t_2]} \Psi[\tau, S_{(t,\tau)}x] \\ &\le \sup_{\tau\in[t,t_2]} \Psi[\tau, S_{(t,\tau)}x] \\ &= U(t,x). \end{aligned} \tag{D.26}$$

To conclude the proof of Lemma D.2, we must prove that U is locally Lipschitz on $[t_0,t_2]\times Y$, meaning that for each compact subset $[a,b]\times K$ of $[t_0,t_2]\times Y$ there exists a constant C (dependent only on a, b, K, t_0, and t_2) such that

$$|U(t,x) - U(s,y)| \le C\left(|s-t| + \|x-y\|\right) \qquad \forall\,(t,x),(s,y) \in [a,b]\times K. \tag{D.27}$$

Clearly, it suffices to find constants C' and C'' such that

$$U(t,x) - U(s,y) \le C'\left(|t-s| + \|x-y\|\right), \tag{D.28a}$$

$$U(s,y) - U(t,x) \le C''\left(|t-s| + \|x-y\|\right), \tag{D.28b}$$

To simplify the presentation below, we let C to denote any absolute constant; its value may vary from line to line. We also assume without loss of generality that $s \le t$.

To prove (D.28a) observe that, since $s \le t$,

$$\begin{aligned} U(t,x) - U(s,y) &= \sup_{\tau\in[t,t_2]} \Psi[\tau, S_{(t,\tau)}x] - \sup_{\tau\in[s,t_2]} \Psi[\tau, S_{(s,\tau)}y] \\ &\le \sup_{\tau\in[t,t_2]} \Psi[\tau, S_{(t,\tau)}x] - \sup_{\tau\in[t,t_2]} \Psi[\tau, S_{(s,\tau)}y] \\ &\le \sup_{\tau\in[t,t_2]} \left\{ \Psi[\tau, S_{(t,\tau)}x] - \Psi[\tau, S_{(s,\tau)}y] \right\}. \end{aligned} \tag{D.29}$$

The term inside the last supremum can be bounded uniformly in τ. The Lipschitz continuity of Ψ and Lemma D.1 imply

$$\begin{aligned} \Psi[\tau, S_{(t,\tau)}x] - \Psi[\tau, S_{(s,\tau)}y] &\le C\|S_{(t,\tau)}x - S_{(s,\tau)}y\| \\ &\le C\|S_{(t,\tau)}x - S_{(t,\tau)}y\| + C\|S_{(t,\tau)}y - S_{(s,\tau)}y\| \\ &\le C\left(\|x-y\| + |t-s|\right). \end{aligned} \tag{D.30}$$

Combining this estimate with (D.29) yields (D.28a).

To show (D.28b) we seek an upper bound on

$$U(s,y) - U(t,x) = \sup_{\tau\in[s,t_2]} \Psi[\tau, S_{(s,\tau)}y] - \sup_{\tau\in[t,t_2]} \Psi[\tau, S_{(t,\tau)}x]. \tag{D.31}$$

If the first supremum can be restricted to $[t, t_2]$ without affecting its value, then we proceed as before. Otherwise, we restrict the supremum to $[s, t]$ and estimate

$$U(s, y) - U(t, x) \leq \sup_{\tau \in [s,t]} \Psi[\tau, S_{(s,\tau)} y] - \Psi(t, x) = \sup_{\tau \in [s,t]} \left\{ \Psi[\tau, S_{(s,\tau)} y] - \Psi(t, x) \right\}. \tag{D.32}$$

As before, the term inside the supremum can be bounded uniformly in τ using Lipschitz continuity and Lemma D.1. Precisely, since $\tau \leq t$ and $S_{(\tau,\tau)} y = y$,

$$\begin{aligned} \Psi(\tau, S_{(s,\tau)} y) - \Psi(t, x) &\leq C \left(|\tau - t| + \|S_{(s,\tau)} y - x\| \right) \\ &\leq C \left(|t - s| + \|S_{(s,\tau)} y - S_{(\tau,\tau)} y\| + \|y - x\| \right) \\ &\leq C \left(|t - s| + \|x - y\| \right). \end{aligned} \tag{D.33}$$

Combining these estimates with (D.32) yields (D.28b).

REFERENCES

[1] A. A. Ahmadi and G. Hall, *Sum of squares basis pursuit with linear and second order cone programming*, in Algebraic and Geometric Methods in Discrete Mathematics, H. A. Harrington, M. Omar, and M. Wright, eds., vol. 685 of Contemporary Mathematics, AMS, 2015, pp. 27–53, https://doi.org/10.1090/conm/685/13712.

[2] A. A. Ahmadi and A. Majumdar, *DSOS and SDSOS optimization: More tractable alternatives to sum of squares and semidefinite optimization*, SIAM J. Appl. Algebra Geometry, 3 (2019), pp. 1–8, https://doi.org/10.1109/CISS.2014.6814141.

[3] M. Ahmadi, G. Valmorbida, and A. Papachristodoulou, *Dissipation inequalities for the analysis of a class of PDEs*, Automatica, 66 (2016), pp. 163–171, https://doi.org/10.1016/j.automatica.2015.12.010.

[4] M. Ahmadi, G. Valmorbida, and A. Papachristodoulou, *Safety verification for distributed parameter systems using barrier functionals*, Syst. Control Lett., 108 (2017), pp. 33–39, https://doi.org/10.1016/j.sysconle.2017.08.002.

[5] D. Ayala and B. Protas, *On maximum enstrophy growth in a hydrodynamic system*, Phys. D, 240 (2011), pp. 1553–1563, https://doi.org/10.1016/j.physd.2011.07.003.

[6] D. Ayala and B. Protas, *Maximum palinstrophy growth in 2D incompressible flows*, J. Fluid Mech., 742 (2014), pp. 340–367, https://doi.org/10.1017/jfm.2013.685.

[7] M. Bardi and I. Capuzzo-Dolcetta, *Optimal control and viscosity solutions of Hamilton–Jacobi–Bellman equations*, Birkhäuser Boston, 1997, https://doi.org/10.1007/978-0-8176-4755-1.

[8] J. Carlson, A. Jaffe, and A. Wiles, eds., *The Millennium Prize Problems*, AMS, Providence, R.I. (USA), 2006.

[9] S. I. Chernyshenko, *Relationship between the methods of bounding time averages.* arXiv:1704.02475 [physics.flu-dyn], 2017, https://arxiv.org/abs/1704.02475.

[10] S. I. Chernyshenko, P. J. Goulart, D. Huang, and A. Papachristodoulou, *Polynomial sum of squares in fluid dynamics: a review with a look ahead*, Philos. Trans. Roy. Soc. A, 372 (2014), p. 20130350, https://doi.org/10.1098/rsta.2013.0350.

[11] M. J. Cho and R. H. Stockbridge, *Linear programming formulation for optimal stopping problems*, SIAM J. Control Optim., 40 (2002), pp. 1965–1982, https://doi.org/10.1137/S0363012900377663.

[12] J. Cyranka, M. A. Islam, G. Byrne, P. Jones, S. A. Smolka, and R. Grosu, *Lagrangian reachability*, in Computer Aided Verification, R. Majumdar and V. Kunčak, eds., Cham, 2017, Springer International Publishing, pp. 379–400, https://doi.org/10.1007/978-3-319-63387-9_19.

[13] S. Dashkovskiy and A. Mironchenko, *Input-to-state stability of infinite-dimensional control systems*, Math. Control Signals Systems, 25 (2013), pp. 1–35, https://doi.org/10.1007/s00498-012-0090-2.

[14] R. Datko, *Extending a theorem of A. M. Liapunov to Hilbert Space*, J. Math. Anal. Appl., 32 (1970), pp. 610–616, https://doi.org/10.1016/0022-247X(70)90283-0.

[15] C. R. Doering, *The 3D Navier–Stokes Problem*, Annu. Rev. Fluid Mech., 41 (2009), pp. 109–128, https://doi.org/10.1146/annurev.fluid.010908.165218.

[16] C. R. Doering and P. Constantin, *Energy dissipation in shear driven turbulence*, Phys. Rev. Lett., 69 (1992), pp. 1648–1651, https://doi.org/10.1103/PhysRevLett.69.1648.

[17] C. R. Doering and E. S. Titi, *Exponential decay rate of the power spectrum for solutions of the Navier–Stokes equations*, Phys. Fluids, 7 (1995), pp. 1384–1390, https://doi.org/10.1063/1.868526.

[18] L. Escauriaza, G. A. Seregin, and V. Šverák, *$L_{3,\infty}$-solutions of the Navier–Stokes equations and backward uniqueness*, Russian Math. Surveys, 58 (2003), pp. 211–250, https://doi.org/10.1070/rm2003v058n02abeh000609.

[19] L. C. Evans, *Partial Differential Equations*, AMS, Providence, R.I. (USA), 1998.

[20] G. Fantuzzi, D. Goluskin, D. Huang, and S. I. Chernyshenko, *Bounds for deterministic and stochastic dynamical systems using sum-of-squares optimization*, SIAM J. Appl. Dyn. Syst., 15 (2016), pp. 1962–1988, https://doi.org/10.1137/15M1053347.

[21] M. Farazmand and T. P. Sapsis, *A variational approach to probing extreme events in turbulent dynamical systems*, Science Advances, 3 (2017), pp. e1701533(1–7), https://doi.org/10.1126/sciadv.1701533.

[22] C. Foias and R. Temam, *Gevrey class regularity for the solutions of the Navier–Stokes equations*, J. Funct. Anal., 87 (1989), pp. 359–369, https://doi.org/10.1016/0022-1236(89)90015-3.

[23] D. P. G. Foures, C. P. Caulfield, and P. J. Schmid, *Optimal mixing in two-dimensional plane Poiseuille flow at finite Péclet number*, J. Fluid Mech., 748 (2014), pp. 241–277, https://doi.org/10.1017/jfm.2014.182.

[24] K. Fujisawa, M. Fukuda, K. Kobayashi, M. Kojima, K. Nakata, M. Nakata, and M. Yamashita, *SDPA (SemiDefinite Programming Algorithm) and SDPA-GMP User's Manual – Version 7.1.1*, tech. report, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Japan, 2008.

[25] D. Goluskin, *Bounding averages rigorously using semidefinite programming: Mean moments of the Lorenz system*, J. Nonlinear Sci., 28 (2016), pp. 621–651, https://doi.org/10.1007/s00332-017-9421-2.

[26] D. Goluskin, *Bounding extrema over global attractors using polynomial optimization*, Nonlinearity, In press (2019), http://arxiv.org/abs/1807.09814.

[27] D. Goluskin and G. Fantuzzi, *Bounds on mean energy in the Kuramoto–Sivashinsky equation computed using semidefinite programming*, Nonlinearity, 32 (2019), pp. 1705–1730, https://doi.org/10.1088/1361-6544/ab018b.

[28] P. J. Goulart and S. I. Chernyshenko, *Global stability analysis of fluid flows using sum-of-squares*, Phys. D, 241 (2012), pp. 692–704, https://doi.org/10.1016/j.physd.2011.12.008.

[29] M. D. Gunzburger, *Perspectives in Flow Control and Optimization*, SIAM, Philadelphia, PA, USA, 2003, https://doi.org/10.1137/1.9780898718720.

[30] P. Hartman, *Ordinary differential equations*, Classics in Applied Mathematics, vol. 38, SIAM, Philadelphia, 2002, https://doi.org/10.1137/1.9780898719222.

[31] D. Henrion and M. Korda, *Convex computation of the region of attraction of polynomial control systems*, IEEE Trans. Automat. Control, 59 (2014), pp. 297–312, https://doi.org/10.1109/TAC.2013.2283095.

[32] D. Henrion, J. B. Lasserre, and C. Savorgnan, *Nonlinear optimal control synthesis via occupation measures*, in Proc. IEEE Conf. Decis. Control, Cancun, Mexico, 2008, IEEE, pp. 4749–4754, https://doi.org/10.1109/CDC.2008.4739136.

[33] D. Hernández-Hernández, O. Hernández-Lerma, and M. Taksar, *The linear programming approach to deterministic optimal control problems*, Appl. Math. (Warsaw), 24 (1996), pp. 17–33, https://doi.org/10.4064/am-24-1-17-33.

[34] D. Hilbert, *Ueber die Darstellung definiter Formen als Summe von Formenquadraten*, Math. Ann., 32 (1888), pp. 342–350, https://doi.org/10.1007/BF01443605, https://doi.org/10.1007/BF01443605.

[35] D. Huang, S. I. Chernyshenko, P. J. Goulart, D. Lasagna, O. R. Tutty, and F. Fuentes, *Sum-of-squares of polynomials approach to nonlinear stability of fluid flows: an example of application*, Proc. Roy. Soc. A, 471 (2015), p. 20150622, https://doi.org/10.1098/rspa.2015.0622.

[36] M. Jones and M. M. Peet, *Using SOS and sublevel set volume minimization for estimation of forward reachable sets*. arXiv:1901.11174 [math.ds], 2019, https://arxiv.org/abs/1901.11174.

[37] R. R. Kerswell, C. C. T. Pringle, and A. P. Willis, *An optimization approach for analysing nonlinear stability with transition to turbulence in fluids as an exemplar*, Rep. Progr. Phys., 77 (2014),

p. 85901, https://doi.org/10.1088/0034-4885/77/8/085901.
[38] A. Kiselev, F. Nazarov, and R. Shterenberg, *Blow up and regularity for fractal burgers equation*, Dyn. Partial Differ. Equ., 5 (2008), pp. 211–240, https://doi.org/10.4310/DPDE.2008.v5.n3.a2.
[39] A. W. Knapp, *Basic Real Analysis*, Birkhäuser Boston, 2005, https://doi.org/10.1007/0-8176-4441-5.
[40] M. Korda, D. Henrion, and C. N. Jones, *Inner approximations of the region of attraction for polynomial dynamical systems*, IFAC Proceedings Volumes (IFAC-PapersOnline), 43 (2013), pp. 534–539, https://doi.org/10.3182/20130904-3-FR-2041.00002.
[41] M. Korda, D. Henrion, and C. N. Jones, *Controller design and value function approximation for nonlinear dynamical systems*, Automatica, 67 (2016), pp. 54–66, https://doi.org/10.1016/j.automatica.2016.01.022.
[42] M. Korda, D. Henrion, and J. B. Lasserre, *Moments and convex optimization for analysis and control of nonlinear partial differential equations.* arXiv:1804.07565 [math.OC], 2018, http://arxiv.org/abs/1804.07565.
[43] M. Korda, D. Henrion, and I. Mezić, *Convex computation of extremal invariant measures of nonlinear dynamical systems and Markov processes.* arXiv:1807.08956 [math.OC], 2018, http://arxiv.org/abs/1807.08956.
[44] J. Kuntz, M. Ottobre, G.-B. Stan, and M. Barahona, *Bounding stationary averages of polynomial diffusions via semidefinite programming*, SIAM J. Sci. Comput., 38 (2016), pp. A3891–A3920, https://doi.org/10.1137/16M107801X.
[45] J. B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim., 11 (2001), pp. 796–817, https://doi.org/10.1137/S1052623400366802, http://epubs.siam.org/doi/10.1137/S1052623400366802.
[46] J. B. Lasserre, *An introduction to polynomial and semi-algebraic optimization*, Cambridge University Press, 2015, https://doi.org/10.1017/CBO9781107447226.
[47] J. B. Lasserre, D. Henrion, C. Prieur, and E. Trélat, *Nonlinear optimal control via occupation measures and LMI-relaxations*, SIAM J. Control Optim., 47 (2008), pp. 1643–1666, https://doi.org/10.1137/070685051.
[48] R. M. Lewis and R. B. Vinter, *Relaxation of optimal control problems to equivalent convex programs*, J. Math. Anal. Appl., 74 (1980), pp. 475–493, https://doi.org/10.1016/0022-247X(80)90143-2.
[49] J. G. Llavona, *Approximation of smooth functions on manifolds*, in Approximation of Continuously Differentiable Functions, J. G. Llavona, ed., North–Holland, 1986, pp. 23–51, https://doi.org/10.1016/S0304-0208(08)71442-7.
[50] J. Löfberg, *YALMIP: A toolbox for modeling and optimization in MATLAB*, in IEEE International Symposium on Computer Aided Control Systems Design, Taipei, 2004, IEEE, pp. 284–289, https://doi.org/10.1109/CACSD.2004.1393890.
[51] J. Löfberg, *Pre-and post-processing sum-of-squares programs in practice*, IEEE Trans. Automat. Control, 54 (2009), pp. 1007–1011, https://doi.org/10.1109/TAC.2009.2017144.
[52] L. Lu and C. R. Doering, *Limits on enstrophy growth for solutions of the three-dimensional Navier–Stokes equations*, Indiana Univ. Math. J., 57 (2008), pp. 2693–2727, https://doi.org/10.1512/iumj.2008.57.3716.
[53] A. M. Lyapunov, *Stability of motion: General problem*, Internat. J. Control, 55 (1992), pp. 539–589, https://doi.org/10.1080/00207179208934254. Translated by A. T. Fuller from a French translation of Lyapunov's original 1892 dissertation.
[54] V. Magron, P.-L. Garoche, D. Henrion, and X. Thirioux, *Semidefinite approximations of reachable sets for discrete-time polynomial systems*, SIAM J. Control Optim., 57 (2019), pp. 2799–2820, https://doi.org/10.1137/17M1121044.
[55] A. Majumdar, R. Vasudevan, M. M. Tobenkin, and R. Tedrake, *Convex optimization of nonlinear feedback controllers via occupation measures*, Int. J. Robot. Res., 33 (2014), pp. 1209–1230, https://doi.org/10.1177/0278364914528059.
[56] F. Marcotte and C. P. Caulfield, *Optimal mixing in two-dimensional stratified plane Poiseuille flow at finite Péclet and Richardson numbers*, J. Fluid Mech., 853 (2018), pp. 359–385, https://doi.org/10.1017/jfm.2018.565.
[57] S. Marx, T. Weisser, D. Henrion, and J. B. Lasserre, *A moment approach for entropy solutions to nonlinear hyperbolic PDEs.* arXiv:1807.02306 [math-AP], 2018, http://arxiv.org/abs/1807.02306.

[58] MOSEK ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 8.1*, 2017, http://docs.mosek.com/8.1/toolbox/index.html.
[59] K. G. MURTY AND S. N. KABADI, *Some NP-complete problems in quadratic and nonlinear programming*, Math. Program., 39 (1987), pp. 117–129, https://doi.org/10.1007%2FBF02592948.
[60] Y. NESTEROV, *Squared functional systems and optimization problems*, in High performance optimization, H. Frenk, K. Roos, T. Terlaky, and S. Zhang, eds., Springer, 2000, pp. 405–440, https://doi.org/10.1007/978-1-4757-3216-0_17, http://link.springer.com/chapter/10.1007/978-1-4757-3216-0{_}17.
[61] K. ODANI, *The limit cycle of the van der Pol equation is not algebraic*, J. Differential Equations, 115 (1995), pp. 146–152, https://doi.org/10.1006/jdeq.1995.1008.
[62] M. ONORATO, S. RESIDORI, U. BORTOLOZZO, A. MONTINA, AND F. T. ARECCHI, *Rogue waves and their generating mechanisms in different physical contexts*, Phys. Rep., 528 (2013), pp. 47–89, https://doi.org/10.1016/j.physrep.2013.03.001.
[63] D. PAPP AND S. YILDIZ, *Sum-of-squares optimization without semidefinite programming*, SIAM J. Optim., 29 (2019), pp. 822–851, https://doi.org/10.1137/17M1160124.
[64] P. A. PARRILO, *Semidefinite programming relaxations for semialgebraic problems*, Math. Program. B, 96 (2003), pp. 293–320, https://doi.org/10.1007/s10107-003-0387-5.
[65] D. PELINOVSKY, *Sharp bounds on enstrophy growth in the viscous Burgers equation*, Proc. Roy. Soc. A, 468 (2012), pp. 3636–3648, https://doi.org/10.1098/rspa.2012.0200.
[66] S. PRAJNA, A. JADBABAIE, AND G. J. PAPPAS, *A framework for worst-case and stochastic safety verification using barrier certificates*, IEEE Trans. Automat. Control, 52 (2007), pp. 1415–1428, https://doi.org/10.1109/TAC.2007.902736.
[67] M. PUTINAR, *Positive polynomials on compact semi-algebraic sets*, Indiana Univ. Math. J., 42 (1993), pp. 969–984.
[68] J. C. ROBINSON, *Infinite-dimensional dynamical systems: An introduction to dissipative parabolic PDEs and the theory of global attractors*, Cambridge University Press, Cambridge, 2001.
[69] W. TAN AND A. PACKARD, *Stability region analysis using sum of squares programming*, in Proc. Amer. Control Conf., Minneapolis, MN, USA, June 14-16, 2006, IEEE, pp. 2297–2302, https://doi.org/10.1109/ACC.2006.1656562.
[70] R. TEMAM, *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, Springer-Verlag New York, 2 ed., 1997, https://doi.org/10.1007/978-1-4612-0645-3.
[71] I. TOBASCO, D. GOLUSKIN, AND C. R. DOERING, *Optimal bounds and extremal trajectories for time averages in nonlinear dynamical systems*, Phys. Lett. A, 382 (2018), pp. 382–386, https://doi.org/10.1016/j.physleta.2017.12.023.
[72] L. N. TREFETHEN, A. E. TREFETHEN, S. C. REDDY, AND T. A. DRISCOLL, *Hydrodynamic stability without eigenvalues*, Science, 261 (1993), pp. 578–584, https://doi.org/10.1126/science.261.5121.578.
[73] G. VALMORBIDA, M. AHMADI, AND A. PAPACHRISTODOULOU, *Convex solutions to integral inequalities in two-dimensional domains*, in Proc. IEEE Conf. Decis. Control, Osaka, Japan, 2015, IEEE, pp. 7268–7273, https://doi.org/10.1109/CDC.2015.7403366.
[74] G. VALMORBIDA, M. AHMADI, AND A. PAPACHRISTODOULOU, *Stability analysis for a class of partial differential equations via semidefinite programming*, IEEE Trans. Automat. Control, 61 (2016), pp. 1649–1654, https://doi.org/10.1109/TAC.2015.2479135.
[75] G. VALMORBIDA AND J. ANDERSON, *Region of attraction estimation using invariant sets and rational Lyapunov functions*, Automatica, 75 (2017), pp. 37–45, https://doi.org/10.1016/j.automatica.2016.09.003.
[76] G. VALMORBIDA AND A. PAPACHRISTODOULOU, *Introducing INTSOSTOOLS : A SOSTOOLS plug-in for integral inequalities*, in Proc. of the 2015 Eur. Control Conf., Linz, Austria, 2015, IEEE, pp. 1231–1236, https://doi.org/10.1109/ECC.2015.7330708.
[77] B. VAN DER POL, *On relaxation-oscillations*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 2 (1926), pp. 978–992, https://doi.org/10.1080/14786442608564127.
[78] L. VANDENBERGHE AND S. BOYD, *Semidefinite Programming*, SIAM Rev., 38 (1996), pp. 49–95, https://doi.org/10.1137/1038003.
[79] R. B. VINTER, *Convex duality and nonlinear optimal control*, SIAM J. Control Optim., 31 (1993), pp. 518–538, https://doi.org/10.1137/0331024.
[80] R. B. VINTER AND R. M. LEWIS, *A necessary and sufficient condition for optimality of dynamic pro-*

gramming type, making no a priori assumptions on the controls, SIAM J. Control Optim., 16 (1978), pp. 571–583, https://doi.org/10.1137/0316038.

[81] R. B. Vinter and R. M. Lewis, *The equivalence of strong and weak formulations for certain problems in optimal control*, SIAM J. Control Optim., 16 (1978), pp. 546–570, https://doi.org/10.1137/0316037.

[82] H. Waki, M. Nakata, and M. Muramatsu, *Strange behaviors of interior-point methods for solving semidefinite programming problems in polynomial optimization*, Comput. Optim. Appl., 53 (2012), pp. 823–844, https://doi.org/10.1007/s10589-011-9437-8.

[83] J. C. Willems, *Dissipative dynamical systems part I: General theory*, Arch. Ration. Mech. Anal., 45 (1972), pp. 311–314, https://doi.org/10.1007%2FBF00276493.

[84] D. Yun and B. Protas, *Maximum rate of growth of enstrophy in solutions of the fractional Burgers equation*, J. Nonlinear Sci., 28 (2018), pp. 395–422, https://doi.org/10.1007/s00332-017-9412-3.

[85] P. Zhao, S. Mohan, and R. Vasudevan, *Control synthesis for nonlinear optimal control via convex relaxations*, in Proc. Amer. Control Conf., Seattle, WA, USA, 2017, IEEE, pp. 2654–2661, https://doi.org/10.23919/ACC.2017.7963353.

[86] Y. Zheng, G. Fantuzzi, and A. Papachristodoulou, *Exploiting sparsity in the coefficient matching conditions in sum-of-squares programming using ADMM*, IEEE Control Syst. Lett., 1 (2017), pp. 80–85, https://doi.org/10.1109/LCSYS.2017.2706941.

[87] Y. Zheng, G. Fantuzzi, and A. Papachristodoulou, *Fast ADMM for sum-of-squares programs using partial orthogonality*, IEEE Trans. Automat. Control (Early Access), PP (2018), pp. 1–8, https://doi.org/10.1109/TAC.2018.2886170.