arXiv:1906.04001v1 [math.DS] 10 Jun 2019

Finding extremal periodic orbits with polynomial optimisation, with application to a nine-mode model of shear flow

Mayur Lakshmi*[1], Giovanni Fantuzzi[1], Jesús Fernández-Caballero[2], Yongyun Hwang[1], and Sergei Chernyshenko[1]

[1]Department of Aeronautics, Imperial College London, London, SW7 2AZ, UK
[2]Mathematics Institute, University of Warwick, Coventry, CV4 7AL, UK

June 11, 2019

Abstract

Tobasco *et al.* [*Physics Letters A*, 382:382–386, 2018] recently suggested that trajectories of ODE systems which optimise the infinite-time average of a certain observable can be localised using sublevel sets of a function that arise when bounding such averages using so-called auxiliary functions. In this paper we demonstrate that this idea is viable and allows for the computation of extremal unstable periodic orbits (UPOs) for polynomial ODE systems. First, we prove that polynomial optimisation is guaranteed to produce auxiliary functions that yield near-sharp bounds on time averages, which is required in order to localise the extremal orbit accurately. Second, we show that points inside the relevant sublevel sets can be computed efficiently through direct nonlinear optimisation. Such points provide good initial conditions for UPO computations. We then combine these methods with a single-shooting Netwon–Raphson algorithm to study extremal UPOs for a nine-dimensional model of sinusoidally forced shear flow. We discover three previously unknown families of UPOs, one of which simultaneously minimises the mean energy dissipation rate and maximises the mean perturbation energy relative to the laminar state for Reynolds numbers approximately between 81.24 and 125.

1 Introduction

In the study of dynamical systems governed by nonlinear differential equations, the problem of finding unstable periodic orbits (UPOs) is of significant interest. The set of UPOs forms a skeleton for the chaotic dynamics of many nonlinear dissipative systems [1] and is sometimes dense in the chaotic attractor [2]. Therefore, the identification of UPOs is useful to understand the topology of the attractor. In addition, knowledge of UPOs and their respective stability eigenvalues can be leveraged to calculate the infinite-time average of any observable of interest over chaotic trajectories through a cycle expansion of the corresponding dynamical zeta function [3, 4]. A detailed treatment of cycle expansions can be found in [5]. Finally, UPOs also play a key role in the control of chaos [6].

One of the main challenges in the computation of UPOs is that the convergence of the Newton–Raphson algorithm applied to the flow map [5, Ch. 7], [7] relies on the availability of

*Email address for correspondence: mayur.lakshmi11@imperial.ac.uk

good initial guesses for both the UPO (or at least one point on it) and its period. One popular method to search for such initial guesses was introduced by Auerbach *et al.* [8], who suggested to look for near periodic sections of simulated chaotic trajectories, which signal shadowing of a UPO. Another approach exists for systems with at least two local attractors; see, for instance, [9–11]. The idea behind this approach is that initial conditions on the boundary of the local attraction basins converge to neither attractor, but rather to one of possibly many unstable "edge states", whose stable manifolds are embedded in the basin boundary. Using a simple bisection algorithm, one can obtain a point lying very close to the basin boundary, such that its forward trajectory approaches very closely an edge state—very often, a UPO. Both aforementioned methods have proved effective in practice and are applicable to very-high-dimensional systems. However, there is little control on which UPOs can be identified.

In this work, we consider the problem of computing UPOs that are *extremal*, in the sense that they maximise or minimise the infinite-time average of a certain observable. Such orbits are especially interesting for control purposes, since their knowledge allows to design control strategies to stabilise desired dynamics or suppress undesired ones. Observe also that every periodic orbit is extremal for at least one observable, such as its own indicator function (equal to one on the periodic orbit and zero elsewhere). Thus, studying extremal UPOs in not restrictive. In particular, although indicator functions can not be used in our approach, varying the observable whose time average is to be optimised still allows one to systematically probe the state space and potentially identify a large number of periodic orbits.

The identification of extremal UPOs is inherently linked to the study of infinite-time averages of scalar-valued observables, as these are typically maximised (or minimised) on periodic orbits [12]. For systems governed by ordinary differential equations (ODEs), extremal time averages can be bounded rigorously to arbitrary accuracy using so-called auxiliary functions of the system's state, which are similar to but distinct from the Lyapunov functions used in nonlinear stability analysis. This strategy, proposed by Chernyshenko *et al.* [13] and further considered in [14–16], generalises the "background method" developed by Doering & Constantin in fluid mechanics [17–20], which amounts to using a quadratic auxiliary function [21]. A dual formulation of the same approach, in the sense of convex duality, also exists and is based on the connection between infinite-time averages and invariant measures [22–24]. Auxiliary functions are attractive for two reasons. First, the optimisation of auxiliary functions is a convex problem, while directly optimising time averages over trajectories is a nonconvex and nonlinear problem. Second, for ODE systems with polynomial dynamics, polynomial auxiliary functions can be constructed and optimised numerically using methods for polynomial optimisation. Indeed, the constraint to be satisfied by any candidate auxiliary function becomes a polynomial inequality, and one can strengthen it into the condition that this polynomial admits a sum-of-squares (SOS) decomposition. Problems with SOS constraints can be posed as semidefinite programs (SDPs) and can be solved computationally using algorithms with polynomial-time complexity [25]. This connection has been exploited to study a number of properties of dynamical systems beyond time averages, including nonlinear stability analysis [13, 26–28], estimation of regions of attraction [29, 30] and nonlinear control [31–33].

The connection between auxiliary functions proving bounds on infinite-time averages and the corresponding extremal trajectories, which are often UPOs, has been recently established by Tobasco *et al.* [24]. Precisely, if an auxiliary function produces nearly sharp bounds, then it can be used to construct a function whose sublevel sets localise extremal and near-extremal trajectories in state space. Here we leverage this observation and propose a new approach to compute extremal UPOs for ODE systems with polynomial dynamics. The first step in this approach is to construct near-optimal auxiliary functions using SOS optimisation. We prove

that this is always possible with a small modification to the computational methods described in the previous works [13–15]. This follows from the argument in [24] using a standard argument in SOS optimisation, and is the dual version of Theorem 2 in [23]. The second step is to compute points in the ε-sublevel set of the function constructed using the near-optimal auxiliary function for small ε, and use these as initial conditions for converging to UPOs with existing algorithms. We demonstrate that while full sublevel sets of multivariate polynomials in high-dimensional spaces cannot be computed in practice, good initial conditions can be obtained with relatively inexpensive nonlinear optimisation methods such as quasi-Newton methods.

As a proof of concept, we apply our new method to a nine-dimensional model of sinusoidally forced shear flow [34]. We consider this particular system for two reasons. First, it is sufficiently low dimensional for the aforementioned SOS optimisation approach to be computationally feasible, but it is also sufficiently high dimensional that approximating extremal UPOs by computing full sublevel sets of polynomials as suggested in [24] is not possible. Second, lower bounds on the mean energy dissipation rate in the model were computed by optimising polynomial auxiliary functions [13]. For values of the Reynolds number Re larger than 40, these bounds appear to converge to a value strictly smaller than both the dissipation rate of the laminar flow state and the simulated dissipation rate of the turbulent flow state as the degree of the polynomial auxiliary function is raised. It was conjectured in [13] this behaviour is due to the existence of some periodic orbit which saturates the bounds. While a relatively large number of families of UPOs have already been found for this flow model [35], there remains the possibility that other UPOs have thus far eluded researchers.

Here we investigate whether optimising polynomial auxiliary functions of higher degree improves the bounds reported in [13], and whether for $Re > 40$ the mean energy dissipation rate is indeed minimised by UPOs. Using the methods described above, we obtain approximations to the UPOs which minimise the infinite-time average of the energy dissipation rate, as well as those maximising the infinite-time average of the energy of perturbations from the laminar flow state. We use our approximations to find three new families of UPOs which do not appear to be among those reported in [35].

The rest of this work is structured as follows. Section 2 reviews the auxiliary function method to bound extremal infinite-time averages. In section 3 we describe our approach to localising UPOs using auxiliary functions. Section 4 explains how to construct near-optimal auxiliary functions using SOS optimisation and includes theoretical results on the convergence of the numerical bounds to the extremal time average, as well as on the exploitation of symmetries in computations. Readers interested in computing UPOs may skip this section at first reading. Section 5 reports the results obtained when applying our methods to the nine-mode shear flow model from [34]. Finally, section 6 suggests possible future research directions and offers concluding remarks.

2 Infinite-time averages and auxiliary functions

Bounding infinite-time averages using auxiliary functions is the first step in our approach to computing extremal UPOs. Consider an autonomous dynamical system governed by the ODE

$$\frac{\mathrm{d}\mathbf{a}}{\mathrm{d}t} = \mathbf{f}(\mathbf{a}), \quad \mathbf{a}(0) = \mathbf{a}_0, \tag{1}$$

where $\mathbf{a} \in \mathbb{R}^n$ and $\mathbf{f} : \mathbb{R}^n \mapsto \mathbb{R}^n$ is continuously differentiable. Following [24], we assume that there exists a compact set Ω in which all solutions $\mathbf{a}(t\,;\mathbf{a}_0)$ of (1) eventually remain, irrespective of the initial conditions $\mathbf{a}_0$. Such a set may be found in a variety of ways, but one very common

approach is to let $\Omega = \{\mathbf{a} \in \mathbb{R}^n \mid W(\mathbf{a}) \leq C\}$, where $W : \mathbb{R}^n \to \mathbb{R}$ is a radially unbounded and continuously differentiable function such that

$$\lambda\, \mathbf{f}(\mathbf{a}) \cdot \nabla W(\mathbf{a}) \leq C - W(\mathbf{a}) \qquad \forall \mathbf{a} \in \mathbb{R}^n, \tag{2}$$

for some scalars C and $\lambda > 0$. While these conditions produce a globally absorbing set (this follows from Gronwall's inequality; see, e.g., [15, 16]), we need not assume that Ω is absorbing; trajectories may transiently leave Ω, provided that they eventually re-enter and remain in it. In particular, any closed ball containing a globally absorbing set is a suitable choice for Ω.

Given a continuous function $\Phi : \mathbb{R}^n \to \mathbb{R}$, which represents an observable of interest, we define its infinite-time average along the trajectory starting from $\mathbf{a}_0$ as

$$\overline{\Phi}\,(\mathbf{a}_0) := \limsup_{T\to\infty} \frac{1}{T} \int_0^T \Phi\left[\mathbf{a}\,(t\,;\mathbf{a}_0)\right] dt. \tag{3}$$

We use the lim sup in this definition because time averages need not converge. As already noted in [24], $\overline{\Phi}(\mathbf{a}_0)$ could alternatively be defined using the lim inf, with no effect on the results presented in this work. Conditions under which time averages do converge are discussed in [37].

We are interested in the maximal value of $\overline{\Phi}$ over all trajectories,

$$\overline{\Phi}^* := \max_{\mathbf{a}_0 \in \Omega} \overline{\Phi}(\mathbf{a}_0), \tag{4}$$

as well as the initial conditions and corresponding trajectories which achieve it. Observe that considering maximal time averages only is not restrictive because the minimal value of $\overline{\Phi}$ coincides with the maximum of $\overline{-\Phi}$. Additionally, there is no loss of generality in assuming that the initial condition belongs to Ω because every trajectory enters it after some finite time, and the part of the trajectory before this time gives no contribution to $\overline{\Phi}(\mathbf{a}_0)$.

Tobasco *et al.* [24] proved that (4) is a well-posed optimisation problem and that there exist optimal initial conditions $\mathbf{a}_0^*$ such that $\overline{\Phi}(\mathbf{a}_0^*) = \overline{\Phi}^*$. In fact, there are clearly infinitely many such optimal initial conditions because $\overline{\Phi}[\mathbf{a}(t\,;\mathbf{a}_0^*)] = \overline{\Phi}(\mathbf{a}_0^*)$ for any fixed time t. The same authors also proved that $\overline{\Phi}^*$ can be characterised equivalently as the optimal value of a minimisation problem, originally proposed in [13] and further studied in [14–16, 31], over continuously differentiable *auxiliary functions* $V : \mathbb{R}^n \to \mathbb{R}$. Precisely,

$$\overline{\Phi}^* = \inf_{V \in C^1(\Omega)} \max_{\mathbf{a}\in\Omega} \left\{\Phi(\mathbf{a}) + \mathbf{f}(\mathbf{a}) \cdot \nabla V(\mathbf{a})\right\}. \tag{5}$$

In fact, since Ω is compact by assumption, polynomials are dense in $C^1(\Omega)$ [38, Ch. 1, Theorem 1.1.2] and the minimisation on the right-hand side may be restricted to the space Π_n of n-variate polynomials:

$$\overline{\Phi}^* = \inf_{V \in \Pi_n} \max_{\mathbf{a}\in\Omega} \left\{\Phi(\mathbf{a}) + \mathbf{f}(\mathbf{a}) \cdot \nabla V(\mathbf{a})\right\}. \tag{6}$$

Replacing C^1 functions with polynomials may of course prevent the infimum over V in (6) from being attained even when that in (5) is, but this will not be important for our purposes. Whether the infimum in (5) is attained and under which conditions an optimal polynomial V exists are open theoretical questions that go beyond the scope of the present work.

The right-hand side of (6) can be understood as the minimisation of upper bounds on $\overline{\Phi}^*$ derived from two simple observations. The first is that

$$\overline{\Phi + \mathbf{f} \cdot \nabla V}(\mathbf{a}_0) \leq \max_{\mathbf{a}\in\Omega} \left\{\Phi(\mathbf{a}) + \mathbf{f}(\mathbf{a}) \cdot \nabla V(\mathbf{a})\right\} \tag{7}$$

irrespective of the initial condition $\mathbf{a}_0$ because infinite-time averages are determined only by the portion of trajectory that remains in Ω. The second is that the infinite-time averages of Φ and $\Phi + \mathbf{f} \cdot \nabla V$ are equal for any polynomial V because

$$\overline{\mathbf{f}[\mathbf{a}(t)] \cdot \nabla V[\mathbf{a}(t)]} = \overline{\frac{\mathrm{d}}{\mathrm{d}t} V[\mathbf{a}(t)]} = \lim_{T \to \infty} \frac{V[\mathbf{a}(T)] - V(\mathbf{a}_0)}{T} = 0. \tag{8}$$

The key feature of the minimax problems in (5) and (6) is that optimising auxiliary functions to obtain good upper bounds on $\overline{\Phi}^*$ does *not* require solving the ODE (1), but rather estimating the global maximum of $\Phi + \mathbf{f} \cdot \nabla V$ over Ω. For highly nonlinear or chaotic systems, therefore, finding nearly sharp bounds may be easier than a direct calculation of $\overline{\Phi}^*$ via extensive numerical simulations. In addition, while optimising fully general V is not easy, very good bounds can be obtained in practice by considering subsets of auxiliary functions that can be optimised numerically. This has already been clearly demonstrated by the results in [16, 36], where nearly optimal polynomial auxiliary functions for the Lorenz system and the Kuramoto–Sivashinsky equation were constructed numerically using SOS optimisation. In section 4 we show that if $\mathbf{f}$ and Φ are polynomials and Ω is a compact semialgebraic set subject to a mild technical condition, then arbitrarily sharp bounds on $\overline{\Phi}^*$ and the corresponding near-optimal auxiliary functions can be constructed numerically using a variation of the computational approach utilised in [13, 14, 16, 36].

3 Approximation of extremal trajectories

In addition to yielding arbitrarily sharp bounds on $\overline{\Phi}^*$, auxiliary functions can be used to localise the associated extremal trajectories in state space. To see this, suppose that a fully optimal auxiliary function V^* exists and is available. Then, the extremal trajectory $\mathbf{a}(t)$ must satisfy [14, 24]

$$\overline{\lambda - (\mathbf{f}[\mathbf{a}(t)] \cdot \nabla V^*[\mathbf{a}(t)] + \Phi[\mathbf{a}(t)])} = 0 \tag{9}$$

with $\lambda = \overline{\Phi}^*$. Since the quantity being averaged is nonnegative, the extremal trajectory must lie inside the set

$$\mathcal{S}_0 := \{\mathbf{a} \in \Omega : \lambda - [\mathbf{f}(\mathbf{a}) \cdot \nabla V^*(\mathbf{a}) + \Phi(\mathbf{a})] = 0\}. \tag{10}$$

While not all points in $\mathcal{S}_0$ necessarily belong to the extremal trajectory, they provide guidance to locate it.

Unfortunately, an optimal auxiliary function may not exist because the infimum in (6) need not be attained. When it does exist, moreover, it may not be computable. Nevertheless, equation (6) implies that for any $\delta > 0$ there exists a δ-suboptimal polynomial auxiliary function. More precisely, for all $\delta > 0$ there exists an integer d and $V \in \Pi_{n,d}$ which proves a bound $\overline{\Phi} \leq \overline{\Phi}^* + \delta$, where $\Pi_{n,d}$ denotes the set of n-variate polynomials of degree d or less. In addition, we show in the next section that such V can always be computed numerically, although doing so for high-dimensional systems might require prohibitively large computational resources. As described in [24], any suboptimal V constructed numerically can also be used to approximately localise the extremal trajectory, although the results are weaker compared to the case in which V is optimal. Specifically, for any pair (λ, V), where λ is a δ-suboptimal bound on $\overline{\Phi}^*$ and V is the corresponding δ-suboptimal auxiliary function, let $\mathcal{P}_{\lambda,V}$ denote the nonnegative polynomial

$$\mathcal{P}_{\lambda,V}(\mathbf{a}) := \lambda - [\mathbf{f}(\mathbf{a}) \cdot \nabla V(\mathbf{a}) + \Phi(\mathbf{a})] \tag{11}$$

and consider the set of all points where $\mathcal{P}_{\lambda,V}$ is no greater than some arbitrary $\varepsilon > 0$:

$$\mathcal{S}_\varepsilon = \{\mathbf{a} \in \Omega : \mathcal{P}_{\lambda,V} \leq \varepsilon\}. \tag{12}$$

For any δ-suboptimal V, the extremal trajectory is guaranteed to lie in $\mathcal{S}_\varepsilon$ for a fraction of time determined by the values of ε and δ. In particular, if the extremal trajectory is a periodic orbit, the fraction of its time period spent inside the set $\mathcal{S}_\varepsilon$ is no smaller than $F := 1 - \delta/\varepsilon$ [24, Section 3]. Since $\mathcal{P}_{\lambda,V}(\mathbf{a}) \geq 0$ on the compact set Ω, the volume of $\mathcal{S}_\varepsilon$ is small for small ε. When $\delta \ll 1$, we can have F close to 1 for ε not too large, so $\mathcal{S}_\varepsilon$ can be expected to approximate the location of extremal and near extremal trajectories. This has already been demonstrated in [24] for the Lorenz system [39].

Even when a polynomial auxiliary function V producing a near-optimal bound λ on $\overline{\Phi}^*$ is available, however, approximating extremal trajectories using the full set $\mathcal{S}_\varepsilon$ for a given ε is computationally intractable except for very-low-dimensional ODE systems. When simply "gridding" the state space and evaluating $\mathcal{P}_{\lambda,V}$ at the grid points to approximate $\mathcal{S}_\varepsilon$ is not viable, a simpler strategy to compute points in $\mathcal{S}_\varepsilon$ is to numerically look for points where $\mathcal{P}_{\lambda,V}(\mathbf{a}) \leq \varepsilon$ using any nonlinear minimisation algorithm, initialised using random initial conditions $\mathbf{a}_0 \in \Omega$. The polynomial $\mathcal{P}_{\lambda,V}$ is likely to be relatively "flat" in the immediate vicinity of the corresponding extremal trajectory, and steeper elsewhere. We therefore expect that the minimisation routine used would quickly descend to a flat region and then slowly approach a local minimum. In the process, one can obtain a sequence of points which are not local minima, but for which $\mathcal{P}_{\lambda,V}(\mathbf{a}) \leq \varepsilon$. In light of this, one should choose a minimisation algorithm which is unlikely to stall, but nevertheless converges slowly to local minima in order to produce many points in $\mathcal{S}_\varepsilon$. Typical algorithms which fit these requirements are variants of the quasi-Newton method.

An important complication is that only a finite number of local minima of $\mathcal{P}_{\lambda,V}$ may exist when V is suboptimal. Thus, in order to obtain a larger collection of points in $\mathcal{S}_\varepsilon$, one should not fully minimise $\mathcal{P}_{\lambda,V}$. We propose two simple strategies to avoid this issue and produce a better approximation to $\mathcal{S}_\varepsilon$. The first is to stop the minimisation routine prematurely by relaxing the tolerances. However, this prevents the computation of points where the functions used to define the stopping criteria take on values much smaller than their respective prescribed tolerances. Such points also lie in $\mathcal{S}_\varepsilon$. Therefore, one should repeat the computation for progressively tighter tolerance values. The second approach is to use tight tolerances and store the sequence of points generated by the minimisation procedure as it progresses. In theory, this allows one to obtain a more extensive description of $\mathcal{S}_\varepsilon$ than that obtained by not tracking the progress of the minimisation, leading to larger sections of the extremal trajectory being approximated. The two strategies are not exclusive, and the best results may be obtained when combining them. Whichever method is used, the process can be repeated by initialising the minimisation algorithm from different random initial conditions $\mathbf{a}_0 \in \Omega$, in order to generate a large collection of points in $\mathcal{S}_\varepsilon$. Observe that this process is amenable to a large degree of parallelisation, and can be scaled to high-dimensional systems by choosing inexpensive minimisation routines such as the BFGS quasi-Newton algorithm [40–43]. Our first approach was used in conjunction with the BFGS quasi-Newton algorithm to obtain the results presented in section 5.3. The MATLAB built-in function `fminunc` was used to implement the minimisation, with the step tolerance on $\mathbf{a}$ (relative lower bound on the size of a step) set to 10^{-6} for all sets of results, and the first-order optimality tolerance (lower bound on $||\nabla \mathcal{P}_{\lambda,V}||_\infty$) set to 10^{-6} for all results except for those presented in figure 7, where it was set to 5×10^{-7}.

Any points computed using this heuristic procedure provide educated initial guesses for algorithms that converge to UPOs by evolving the system's dynamics forward in time. However, it should be stressed that there are no theoretical guarantees that our heuristic procedure will produce sufficiently accurate approximations to the extremal orbit. First, there might exist points in $\mathcal{S}_\varepsilon$ that are not close to the extremal or near-extremal trajectories. Second, the analysis in [24] does not guarantee that the entire extremal trajectory is in $\mathcal{S}_\varepsilon$. Therefore, one may not be able

to approximate the full trajectory in practice. Nevertheless, the results presented in [24] and those in section 5.3 demonstrate that very good approximations of the extremal orbits are often obtained in practice when V is very close to optimal.

4 Numerical optimisation of auxiliary functions

The approach to approximating extremal UPOs described in the previous section relies on the availability of a near-optimal auxiliary function. The following subsections describe how suitable polynomial V and λ can be constructed computationally using SOS optimisation when Φ and (the entries of) $\mathbf{f}$ are polynomials. Readers who are primarily interested in the application of our methods can safely skip this section at first reading and proceed to section 5 for numerical results.

4.1 Optimising auxiliary functions with SOS optimisation

Upper bounds on $\overline{\Phi}^*$ can be optimised numerically when V is restricted to belong to the set $\Pi_{n,d}$. For each integer d, the best upper bound on $\overline{\Phi}^*$ available with $V \in \Pi_{n,d}$ is given by

$$\begin{aligned}\overline{\Phi}^* &\leq \inf_{V\in\Pi_{n,d}} \max_{\mathbf{a}\in\Omega} \{\Phi(\mathbf{a}) + \mathbf{f}(\mathbf{a})\cdot\nabla V(\mathbf{a})\} \\ &= \inf_{\substack{V\in\Pi_{n,d}\\ \lambda\in\mathbb{R}}} \{\lambda \mid \lambda - \Phi(\mathbf{a}) - \mathbf{f}(\mathbf{a})\cdot\nabla V(\mathbf{a}) \geq 0 \text{ on } \Omega\}\,. \end{aligned} \tag{13}$$

The right-hand side is a convex and finite-dimensional minimisation problem with a polynomial inequality constraint. The optimisation variables are λ and, in the most general case, the $\binom{n+d}{d}$ coefficients of V. Numerical solution of this problem to global optimality is difficult because polynomial inequalities are NP–hard except for a few special cases, such as univariate or quadratic polynomials [44]. To reduce the computational complexity, a common strategy is to replace nonnegativity with the sufficient condition that $\lambda - \Phi(\mathbf{a}) - \mathbf{f}(\mathbf{a})\cdot\nabla V(\mathbf{a})$ is representable as a sum of squares of other polynomials of degree no larger than $r(d)/2$, where

$$\begin{aligned} r(d) &= \deg\left(\lambda - \Phi - \mathbf{f}\cdot\nabla V\right) \\ &= \max\left\{\deg(\Phi), \deg(\mathbf{f}) + d - 1\right\}. \end{aligned} \tag{14}$$

While not all nonnegative polynomials admit an SOS representation, its existence (or lack thereof) can be established in polynomial time by solving an SDP [45, 46]. Additionally, a variety of open-source software packages are available to automatically reformulate optimisation problems with SOS constraints as SDPs and solve them. Upon strengthening the nonnegativity constraint in (13) with an SOS constraint, one obtains the computable upper bound

$$\overline{\Phi}^* \leq \inf_{\substack{V\in\Pi_{n,d}\\ \lambda\in\mathbb{R}}} \left\{\lambda \mid \lambda - \Phi - \mathbf{f}\cdot\nabla V \in \Sigma_{n,r(d)}\right\}, \tag{15}$$

where $\Sigma_{n,q}$ denotes the set of n-variate SOS polynomials of degree no larger than q. By increasing d, the degree of V, one obtains a sequence of nonincreasing bounds on $\overline{\Phi}^*$. This is the approach followed in [13, 14, 16, 36].

The SOS constraint in (15) enforces that $\lambda - \Phi(\mathbf{a}) - \mathbf{f}(\mathbf{a})\cdot\nabla V(\mathbf{a})$ is nonnegative everywhere on $\mathbb{R}^n$, not only on Ω. A weaker sufficient condition for nonnegativity on Ω, which also relies on SOS polynomials, can be formulated if Ω is a semialgebraic set. Precisely, assume that

$$\Omega = \{\mathbf{a}\in\mathbb{R}^n \mid g_1(\mathbf{a}) \geq 0, \ldots, g_m(\mathbf{a}) \geq 0\} \tag{16}$$

for some polynomials $g_1, \ldots, g_m$ of degree s or less. For each integer d such that $r(d) \geq s$, consider the set

$$\Lambda_d := \left\{ \sigma_0(\mathbf{a}) + \sum_{i=1}^{m} \sigma_i(\mathbf{a}) g_i(\mathbf{a}) \mid \sigma_0 \in \Sigma_{n,r(d)},\ \sigma_1, \ldots, \sigma_m \in \Sigma_{n,r(d)-s} \right\}. \tag{17}$$

In other words, polynomials in Λ_d are weighted sums of SOS polynomials, where the weights are exactly the polynomials defining Ω. Clearly, polynomials in Λ_d are non-negative over Ω and $\Sigma_{n,r(d)} \subset \Lambda_d$. Since weighted SOS constraints can also be transformed into SDPs[1], a computationally tractable upper bound on $\overline{\Phi}^*$ that improves on (15) is

$$\overline{\Phi}^* \leq \inf_{\substack{V \in \Pi_{n,d} \\ \lambda \in \mathbb{R}}} \left\{ \lambda \mid \lambda - \Phi - \mathbf{f} \cdot \nabla V \in \Lambda_d \right\}. \tag{18}$$

The real advantage of using weighted SOS constraints is that, if Ω satisfies an additional mild condition, then we can guarantee that the infimum on the right-hand side converges to $\overline{\Phi}^*$ as the degree d of V tends to infinity. This means that, at least in principle, arbitrarily sharp bounds on $\overline{\Phi}^*$ can be computed by optimising polynomial auxiliary functions of sufficiently high degree using SOS optimisation. Theorem 1 below formalises these observations and is one of the main contributions of this paper. Its proof, which we report in detail below for completeness, is a standard argument in SOS optimisation and uses a result due to Putinar [48, Lemma 4.1] on the existence of weighted SOS representations for strictly positive polynomials on a class of compact semialgebraic sets. For a detailed discussion of this result, known in the literature as Putinar's Positivstellensatz, we refer the reader to section 2.4 in [47].

Theorem 1. *Suppose that Ω is a compact semialgebraic set and that there exists L such that $L - \|\mathbf{a}\|^2 \in \Lambda_d$ for some integer d. Then,*

$$\overline{\Phi}^* = \lim_{d \to +\infty} \inf_{\substack{V \in \Pi_{n,d} \\ \lambda \in \mathbb{R}}} \left\{ \lambda \mid \lambda - \Phi - \mathbf{f} \cdot \nabla V \in \Lambda_d \right\}. \tag{19}$$

Proof. We only need to show that for every $\varepsilon > 0$ there exists an integer d and a polynomial $V \in \Pi_{n,d}$ such that $\overline{\Phi}^* + \varepsilon - \Phi - \mathbf{f} \cdot \nabla V \in \Lambda_d$. Equality (6) guarantees the existence of an integer d' and a polynomial $V \in \Pi_{n,d'}$ such that $\overline{\Phi}^* + \varepsilon - \Phi - \mathbf{f} \cdot \nabla V$ is *strictly* positive on Ω. Then, by virtue of our assumptions on Ω, Putinar's Positivstellensatz [48, Lemma 4.1] guarantees that $\overline{\Phi}^* + \varepsilon - \Phi - \mathbf{f} \cdot \nabla V$ belongs to the set of weighted SOS polynomials $\Lambda_{d''}$ for some integer d''. Setting $d = \max(d', d'')$ concludes the proof since $\Pi_{n,d'} \subset \Pi_{n,d}$ and $\Lambda_{d''} \subset \Lambda_d$. □

This result is dual to Theorem 2 in [23], since polynomial auxiliary functions are dual to sequences of approximations to the moments of the invariant measure supported on the trajectory achieving the optimal average $\overline{\Phi}^*$; see [47] for more details. In addition, observe that the assumption that $L - \|\mathbf{a}\|^2 \in \Lambda_d$ is mild and what is really needed is compactness. Indeed, any compact Ω is contained within some ball $L - \|\mathbf{a}\|^2 \geq 0$, so we can always ensure that $L - \|\mathbf{a}\|^2 \in \Lambda_d$ by adding $L - \|\mathbf{a}\|^2 \geq 0$ to the list of polynomial inequalities that define Ω.

[1] A weighted SOS constraint $p \in \Lambda_d$ is equivalent to the $m+1$ SOS constraints $p - \sum_{i=1}^{m} \sigma_i g_i \in \Sigma_{n,r(d)}$ and $\sigma_1, \ldots, \sigma_m \in \Sigma_{n,r(d)-s}$. This formulation of weighted SOS constraints, also referred to as the generalised S procedure [14, 30], is slightly suboptimal for computational purposes than the approach described in [47, Section 2.4], because it introduces more optimisation variables than necessary. However, it is convenient because most software packages for SOS optimisation do not natively support the method of [47, Section 2.4].

4.2 Exploiting symmetry in weighted SOS constraints

The computational cost of solving SOS optimisation problems increases quickly with problem size. As a result, solving the optimisation problem on the right hand side of (18) is practical only for small or medium sized problems, such that $\binom{n+d}{d}$ is $\mathcal{O}(100)$ or so. At the time of writing, the most computationally demanding part of our SOS approach to approximating extremal trajectories is the computation of near-optimal V and λ using SOS optimisation. This is due to the poor scalability of general-purpose algorithms for SDPs. To reduce computational cost, one can exploit information about the boundedness and symmetry of V that can be deduced a priori [16, Appendix A]. Theorem 2 below shows that symmetry can be exploited in weighted SOS constraints, too. A proof of this theorem is also presented below, where we adapt the argument of [16, Proposition 1].

Theorem 2. *Let $\mathcal{T} : \mathbb{R}^n \to \mathbb{R}^n$ be an invertible linear transformation that generates a finite symmetry group, meaning that $\mathcal{T}^K$ is the identity for some integer K. Suppose that the system, the function Φ and the set Ω are invariant under $\mathcal{T}$, meaning that $\mathbf{f}(\mathcal{T}\mathbf{a}) = \mathcal{T}\mathbf{f}(\mathbf{a})$, $\Phi(\mathcal{T}\mathbf{a}) = \Phi(\mathbf{a})$ and $g_i(\mathcal{T}\mathbf{a}) = g_i(\mathbf{a})$ for all $i \in \{1, \ldots, m\}$. If there exist $V, \sigma_0, \sigma_1, \ldots, \sigma_m$ that prove a bound via* (18)*, then there exist $\widehat{V}, \widehat{\sigma}_0, \widehat{\sigma}_1, \ldots, \widehat{\sigma}_m$ that are $\mathcal{T}-$ invariant and prove the same bound.*

Proof. Suppose that there exist $V, \sigma_0, \sigma_1, \ldots, \sigma_m$ which satisfy the weighted SOS constraint in (18):

$$\lambda - \Phi(\mathbf{a}) - \mathbf{f}(\mathbf{a}) \cdot \nabla V(\mathbf{a}) = \sigma_0(\mathbf{a}) + \sigma_1(\mathbf{a}) g_1(\mathbf{a}) + \cdots + \sigma_m(\mathbf{a}) g_m(\mathbf{a}). \tag{20}$$

Consider now symmetrised versions of $V, \sigma_0, \sigma_1, \ldots, \sigma_m$:

$$\widehat{V}(\mathbf{a}) := \frac{1}{K} \sum_{k=0}^{K-1} V(\mathcal{T}^k \mathbf{a}), \tag{21a}$$

$$\widehat{\sigma}_i(\mathbf{a}) := \frac{1}{K} \sum_{k=0}^{K-1} \sigma_i(\mathcal{T}^k \mathbf{a}) \quad \text{for} \quad i = 0, 1, \ldots, m. \tag{21b}$$

Note that $\widehat{V}(\mathcal{T}\mathbf{a}) = \widehat{V}(\mathbf{a})$ since $\mathcal{T}^K$ is the identity, and similarly for $\widehat{\sigma}_0, \widehat{\sigma}_1, \ldots, \widehat{\sigma}_m$. The claim is proven if:

$$\lambda - \Phi(\mathbf{a}) - \mathbf{f}(\mathbf{a}) \cdot \left[\frac{1}{K} \sum_{k=0}^{K-1} (\mathcal{T}^k)^T \nabla V(\mathcal{T}^k \mathbf{a}) \right] = \widehat{\sigma}_0(\mathbf{a}) + \widehat{\sigma}_1(\mathbf{a}) g_1(\mathbf{a}) + \cdots + \widehat{\sigma}_m(\mathbf{a}) g_m(\mathbf{a}). \tag{22}$$

To show that the above equality holds, evaluate (20) at $\mathcal{T}^k \mathbf{a}$ and use the symmetries of $\mathbf{f}$ and Φ to obtain the following:

$$\lambda - \Phi(\mathbf{a}) - \mathbf{f}(\mathbf{a}) \cdot [(\mathcal{T}^k)^T \nabla V(\mathcal{T}^k \mathbf{a})] = \sigma_0(\mathcal{T}^k \mathbf{a}) + \sigma_1(\mathcal{T}^k \mathbf{a}) g_1(\mathbf{a}) + \cdots + \sigma_m(\mathcal{T}^k \mathbf{a}) g_m(\mathbf{a}). \tag{23}$$

Averaging both sides of (23) for $k = 0, 1, \ldots, K-1$ gives (22), thus proving the claim. □

A similar argument also holds if one wishes to prove a globally absorbing set using an SOS relaxation of the sufficient condition (2). Specifically, if the ODE is invariant under a linear transformation $\mathcal{T}$, then there is no loss of generality in assuming that the function W in an SOS relaxation of the inequality (2) is also $\mathcal{T}-$ invariant.

5 Application to a model of shear flow

To demonstrate the techniques described so far, we apply them to study a nine-dimensional quadratic ODE system. The system was introduced by Moehlis *et al.* [34] as a model of sinusoidally forced shear flow in a periodic channel, and takes the form

$$\frac{da_i}{dt} = \lambda_1 \delta_{1i} - \frac{1}{Re}\lambda_j a_j + N_{ijk} a_j a_k, \qquad i,j,k = 1,\dots,9, \tag{24}$$

where summation over repeated indices is assumed and δ_{1i} is the usual Kronecker delta. The state $\mathbf{a} = (a_1, \dots, a_9)$ represents the amplitude of physically relevant flow modes, Re is the Reynolds number, and λ_j, N_{ijk} are numerical coefficients. All coefficients λ_j are strictly positive, so all modes are linearly damped, and $\lambda_1 \leq \lambda_j \leq \lambda_9$. The coefficients N_{ijk}, instead, are such that $N_{ijk} a_i a_j a_k = 0$, meaning that the quadratic terms conserve energy. Here we consider numerical values corresponding to the "NBC" configuration in [34, 35]. For this as well as all other possible configurations, system (24) is invariant under the two linear transformations

$$\mathcal{T}_1 \mathbf{a} = (a_1, a_2, a_3, -a_4, -a_5, -a_6, -a_7, -a_8, a_9)\,, \tag{25a}$$

$$\mathcal{T}_2 \mathbf{a} = (a_1, -a_2, -a_3, a_4, a_5, -a_6, -a_7, -a_8, a_9)\,. \tag{25b}$$

The linear transformations (25a,b) generate a four-element group $\{1, \mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_1\mathcal{T}_2\}$. As noted in [35], if one finds a periodic or fixed point solution to (24), there will be up to three other symmetry-related periodic orbits or fixed points obtained by the actions of the elements of the symmetry group.

At all values of Re, system (24) has a locally stable equilibrium point $\mathbf{a}_l = (1, 0, \dots, 0)$, which represents the laminar flow state. For our chosen configuration, unsteady and often chaotic solutions corresponding to turbulent flows are observed for $Re \geq 80.54$, and a large collection of UPOs have been computed [35]. The laminar state is expected to be globally asymptotically stable below this value of Re, and Lyapunov functions certifying global stability were constructed using SOS optimisation in [27] for $Re < 54.1$. For larger Re, Chernyshenko *et al.* [13] optimised lower bounds on the infinite-time average of the energy dissipation rate,

$$\mathcal{D} := \frac{\sum_{i=1}^{9} \lambda_i a_i^2}{Re}, \tag{26}$$

using polynomial auxiliary functions of degree up to 8. Here we repeat the computations using polynomials up to degree 10. We also compute bounds on the infinite-time average of the energy of perturbations from $\mathbf{a}_l$ (henceforth referred to as perturbation energy),

$$\mathcal{E} := (1 - a_1)^2 + \sum_{i=2}^{9} a_i^2. \tag{27}$$

In both cases, we use the auxiliary functions obtained with SOS optimisation to approximate the extremal trajectories as described in section 3. All our SOS computations were implemented using the MATLAB modelling toolbox YALMIP [49] and the SDP solver MOSEK [50]. We also exploited the symmetries defined by (25a,b) as described in section 4.2 in order to reduce computational cost.

It can be shown that the unit ball centred at the origin absorbs trajectories of (24). Theorem 1 thus guarantees the existence of near-optimal V for the weighted SOS problem (18). Although the existence of near-optimal V is not guaranteed for the standard SOS problem (15), preliminary investigations suggested that it also yields near-optimal V and good orbit approximations. All numerical results reported below were therefore obtained by solving (15), because it is computationally less expensive than (18).

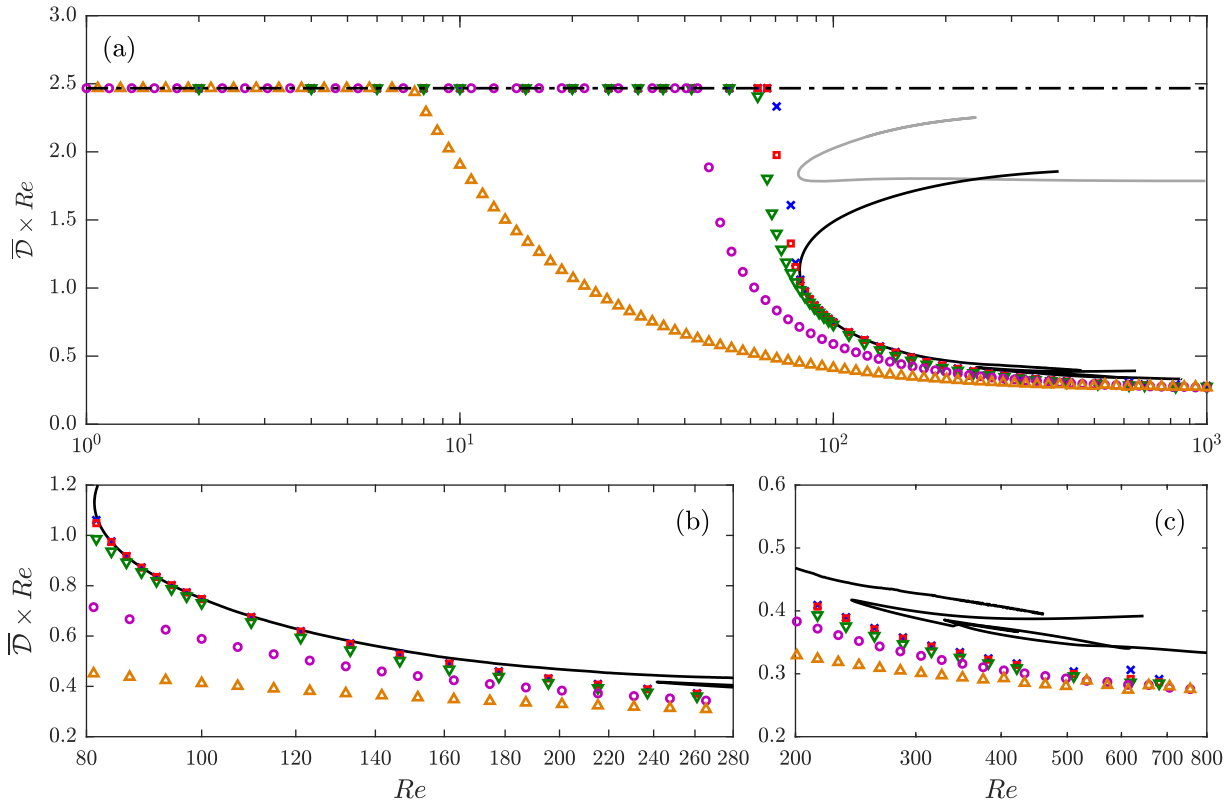

Figure 1: (a) Lower bounds on $\overline{\mathcal{D}}$ computed by optimising over polynomial auxiliary functions of degree 2 (△), 4 (○), 6 (▽), 8 (□) and 10 (×). Also shown is a family of UPOs computed by Moehlis *et al.* [35] (—), and three new families of UPOs (—). (b,c) Detailed views of the results in panel (a). Note the gap between the degree-10 bounds and the currently available UPO data at $Re \gtrapprox 125$.

5.1 Bounds on the infinite-time average of energy dissipation rate

Figure 1 illustrates lower bounds on $\overline{\mathcal{D}}$ computed by solving (15) with $\Phi = -\mathcal{D}$ for $d = 2, 4, 6, 8$ and 10. The polynomial auxiliary functions V used here are of more general ansätze, and some of higher degree, than those used to derive the bounds in [13], which results in tighter bounds. Also shown in the figure are one of the families of UPOs found in [35], born at $Re = 80.54$, and three branches discovered using the strategy outlined in section 3 (see section 5.3 for more details). These branches, which we refer to as PO1, PO2 and PO3, are born at $Re = 81.24$, $Re = 241.5$ and $Re = 330.2$ respectively, and are to the best of our knowledge new and not among those reported in [35].

In the range $82 \lessapprox Re \lessapprox 125$, as the degree of V increases, the lower bounds converge to the lower PO1 branch, which therefore represents a family of minimal orbits for $\overline{\mathcal{D}}$ for this range of Reynolds numbers. At $Re = 89$, for example, our numerically computed bound with degree-10 V is only 0.0494% less than the average over the numerically computed periodic orbit. This resolves the conjecture that the gap between the bounds on $\overline{\mathcal{D}}$ in [13, Figure 1] and values corresponding to the laminar solution $\mathbf{a}_l$, as well as that between the bounds and the simulated dissipation rate of the turbulent flow, was due to the existence of a branch of UPOs which saturates the bounds (at least in the aforementioned range of Re).

From $Re \approx 125$, the SOS bounds begin to deviate from the lowermost branch of PO1 orbits. This can be seen clearly in panel (b) of figure 1. Increasing the degree of V to 12 brings no

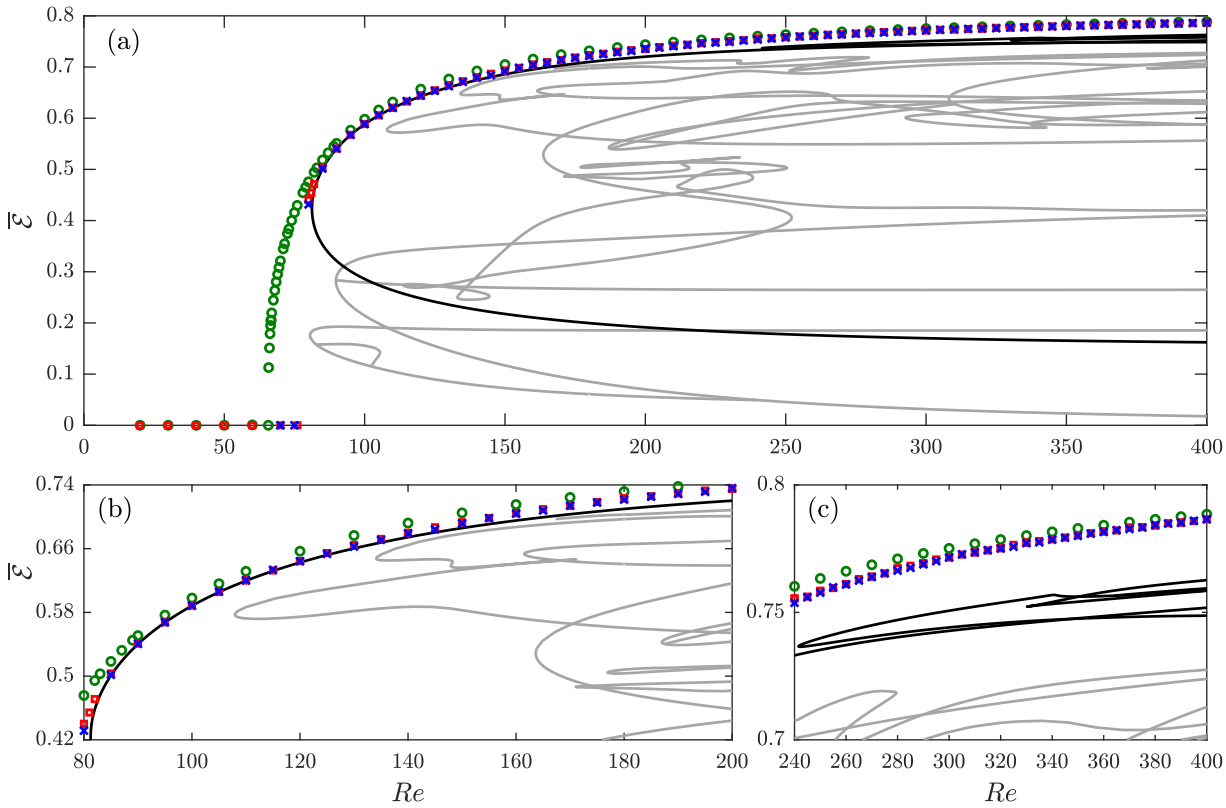

Figure 2: (a) Upper bounds on $\overline{\mathcal{E}}$ computed by optimising over polynomial auxiliary functions of degree 6 (○), 8 (□) and 10 (×). Also shown are families of UPOs computed by Moehlis *et al.* [35] (▬), and three new families of UPOs (▬). (b,c) Detailed views of the results in panel (a).

significant improvement in the bounds for $Re > 125$, suggesting that the degree-10 V used are almost optimal. We conjecture that there exists another family of UPOs with lower mean energy dissipation rate than all known UPOs, although attempts to find it have so far been unsuccessful. We will return to this issue in section 5.3.

Finally, for $Re \lessapprox 67$, the SOS bounds computed with degree-10 V deviate from the values corresponding to the laminar state by less than the numerical tolerance of the SDP solver. We believe that the laminar solution is globally stable up to $Re = 80.54$—the point at which the first family of orbits is born—meaning that our SOS bounds in the range $67 \lessapprox Re \lessapprox 80.54$ are far from sharp and V of very high degree is required to produce sharp bounds near criticality. We leave it to future work to determine possible reasons for this behaviour, and whether it is generic.

5.2 Bounds on the infinite-time average of perturbation energy

Figure 2 shows upper bounds on $\overline{\mathcal{E}}$ computed by solving (15) with $\Phi = \mathcal{E}$ for $d = 6, 8$ and 10. The figure also illustrates the mean perturbation energy of numerous families of UPOs discovered by Moehlis *et al.* [35], as well as that of the three new families PO1, PO2 and PO3. As the degree of V is raised, in the range $82 \lessapprox Re \lessapprox 125$ the upper bounds on $\overline{\mathcal{E}}$ converge to the same branch of UPOs in the family PO1 which minimises $\overline{\mathcal{D}}$. At $Re = 90$, for example, the bound computed with degree-10 V is larger than the average over the UPO by only 0.0775%. This branch of orbits is simultaneously maximal for $\overline{\mathcal{E}}$ and minimal for $\overline{\mathcal{D}}$. As for the case of the mean energy dissipation

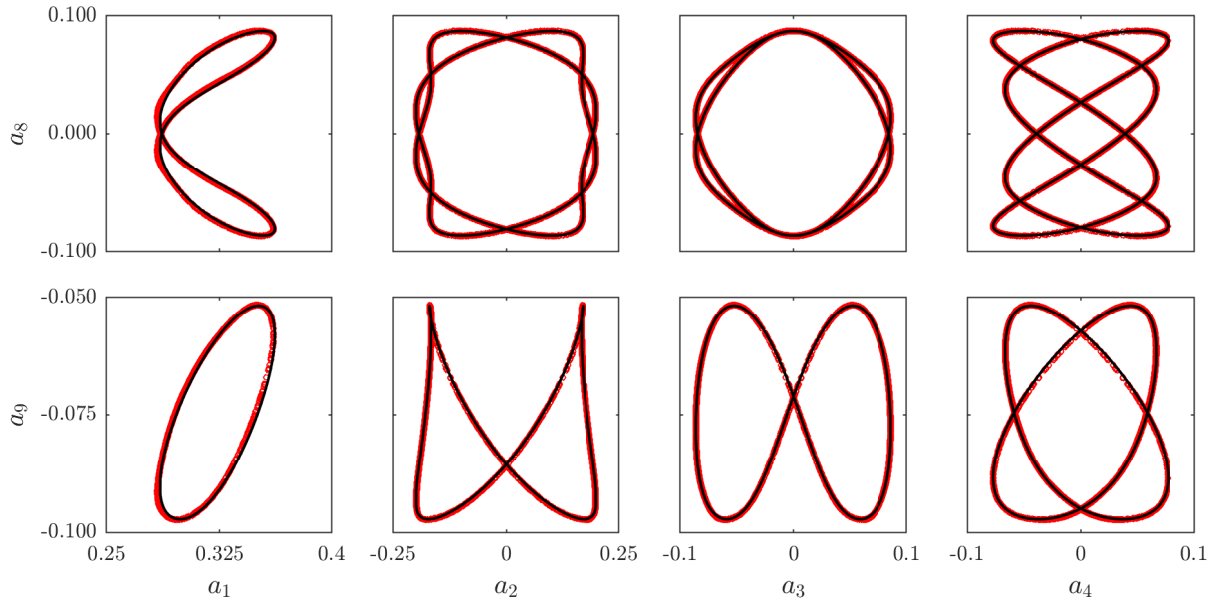

Figure 3: UPOs which maximise $\overline{\mathcal{E}}$ at $Re = 95$, projected onto various 2D subspaces (▬). Also shown are points in the set $\mathcal{S}_{2.55\text{e}-4}$ (○), with the auxiliary function used of the form $P_9(\mathbf{a}) + \|\mathbf{a}\|^{10}$.

rate, the SOS bounds on $\overline{\mathcal{E}}$ deviate from this branch for $Re \gtrapprox 125$. This can be seen clearly in panel (b) of figure 2. Furthermore, very little improvement in the bound is observed as the degree of V is increased beyond 10. Again, this suggests that the degree-10 V used are almost optimal and that another, yet undiscovered, branch of UPOs becomes maximal for $\overline{\mathcal{E}}$. Based on the results at low Re and the observations that solutions with lower dissipation rate have more available energy, we conjecture that this branch of UPOs will simultaneously maximise $\overline{\mathcal{E}}$ and minimise $\overline{\mathcal{D}}$.

The SOS bounds with V of degree 8 and 10 deviate from zero by less than the numerical tolerance of the SDP solver for $Re \lessapprox 76$. Since only the laminar solution $\mathbf{a}_l$ has a perturbation energy of zero, this confirms that it is globally stable in this range of Re. This statement could be made rigorous either by proving a zero upper bound on $\overline{\mathcal{E}}$ analytically, or by constructing more general Lyapunov functions than those considered in [27] using interval arithmetic.

5.3 Approximations to extremal periodic orbits

The same auxiliary functions optimised with SOS programming to bound $\overline{\mathcal{E}}$ and $\overline{\mathcal{D}}$ can be used to approximate the corresponding extremal trajectories as described in section 3. Fixing Φ to be either $-\mathcal{D}$ or $\mathcal{E}$, problem (15) was solved by optimising over polynomial auxiliary functions V of the form $P_9(\mathbf{a}) + \|\mathbf{a}\|^{10}$ or $P_{10}(\mathbf{a})$, where $P_k(\mathbf{a}) \in \Pi_{9,k}$. In each case, the polynomial $\mathcal{P}_{\lambda,V}$ defined as in (11) was built with numerically determined coefficients. We then proceeded by minimising $\mathcal{P}_{\lambda,V}$ over $\mathbb{R}^n$ from random initial conditions with the BFGS quasi-Newton algorithm [40–43] to obtain points that lie in the set $\mathcal{S}_\varepsilon$ for some small ε.

The red dots in figure 3 show the SOS approximation to the trajectory maximising $\overline{\mathcal{E}}$ at $Re = 95$. The results were obtained with $V = P_9(\mathbf{a}) + \|\mathbf{a}\|^{10}$ and $\varepsilon = 2.55 \times 10^{-4}$. These points resemble closed curves, suggesting that the extremal trajectory is a UPO. Indeed, using any of the points as the initial guess for a basic single-shooting Newton–Raphson algorithm [5, Ch. 7], [7], we found a UPO that fits our approximation extremely well. Other UPOs at $Re = 95$ which

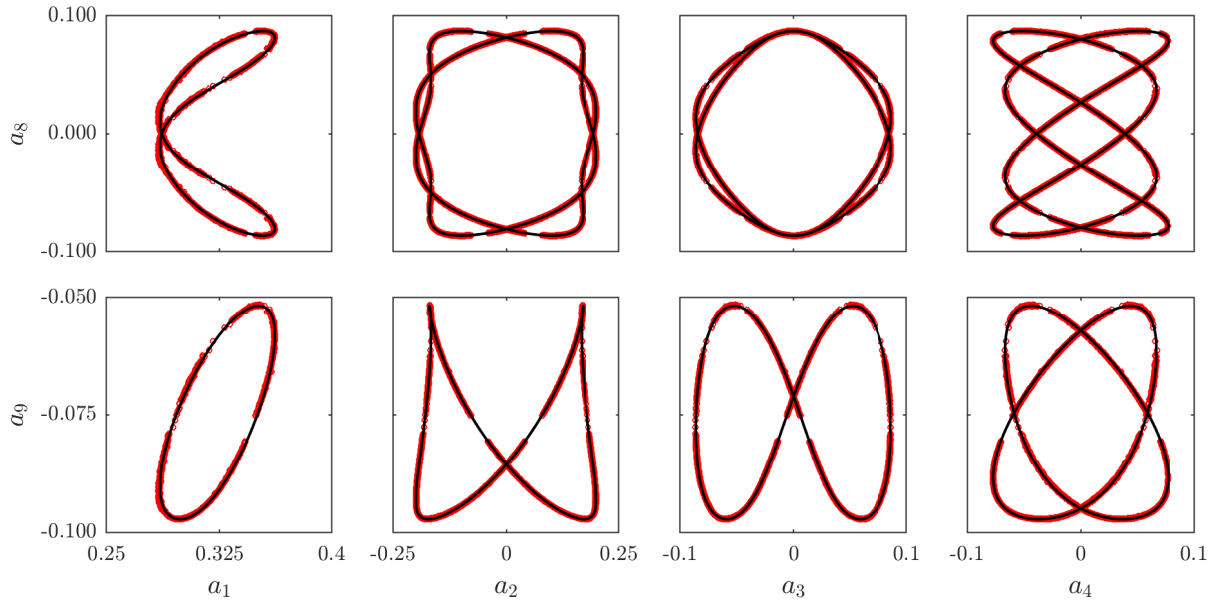

Figure 4: UPOs which minimise $\overline{\mathcal{D}}$ at $Re = 95$, projected onto various 2D subspaces (▬). Also shown are points in the set $\mathcal{S}_{1.23\mathrm{e}-7}$ (○), with the auxiliary function used of the form $P_9(\mathbf{a}) + \|\mathbf{a}\|^{10}$.

are simultaneously maximal for $\overline{\mathcal{E}}$ can be obtained from symmetry considerations, as described after (25b). All such UPOs are plotted as black lines in figure 3. Numerical continuation of one of these symmetry-related orbits using the package MatCont [51] yields the family of UPOs so far referred to as PO1. No bifurcations were detected using MatCont on the extremal PO1 branch (upper PO1 branch in figure 2). The PO1 family is not among the UPOs reported in [35], exists for $Re = 81.24$ and, as described in section 5.2, it maximises $\overline{\mathcal{E}}$ up to $Re \approx 125$.

The same UPOs shown in figure 3 are also plotted in figure 4, this time alongside our SOS approximation of the minimal orbits for $\overline{\mathcal{D}}$. This approximation was obtained with $V = P_9(\mathbf{a}) + \|\mathbf{a}\|^{10}$ and $\varepsilon = 1.23 \times 10^{-7}$; a different value for ε is used because near-optimal bounds on $\overline{\mathcal{D}}$ require a different V than near-optimal bounds on $\overline{\mathcal{E}}$. It is clear that the same branch of PO1 orbits simultaneously minimises the mean dissipation rate and maximises the mean perturbation energy. Although the approximation obtained when bounding the former is slightly worse, our Newton–Raphson solver converged robustly to the same UPO using initial guesses from either approximation. Surprisingly, for both choices of Φ specifying a more general ansatz for V results in a worse approximation of the extremal periodic orbits. Figure 5 shows our SOS approximation of the orbits maximizing $\overline{\mathcal{E}}$, obtained with a generic degree-10 V and $\varepsilon = 2.55 \times 10^{-4}$. The approximation is qualitatively correct but there is a marked offset of $\mathcal{O}(0.01)$ in the a_1 direction. Equally worse approximations are obtained with generic degree-10 V when approximating the orbits minimising $\overline{\mathcal{D}}$ (not shown for brevity). One possible explanation for this behaviour is that while multiple choices of V yielding the same bound on $\overline{\Phi}^*$ exist, not all provide a good approximation to the extremal trajectories. Optimising V over a larger class of polynomials might lead to the same bound on $\overline{\Phi}^*$ (within reasonable numerical tolerances) but a poorer approximation of the extremal trajectory, because this is not penalised. Another factor is that the optimisation problem (15) becomes increasingly ill-conditioned as the degree of V and the number of optimisation variables increases. This is a common problem for large SOS programs [52]. Restricting the form of V used in our computation counteracts both these facts, and we believe that this is

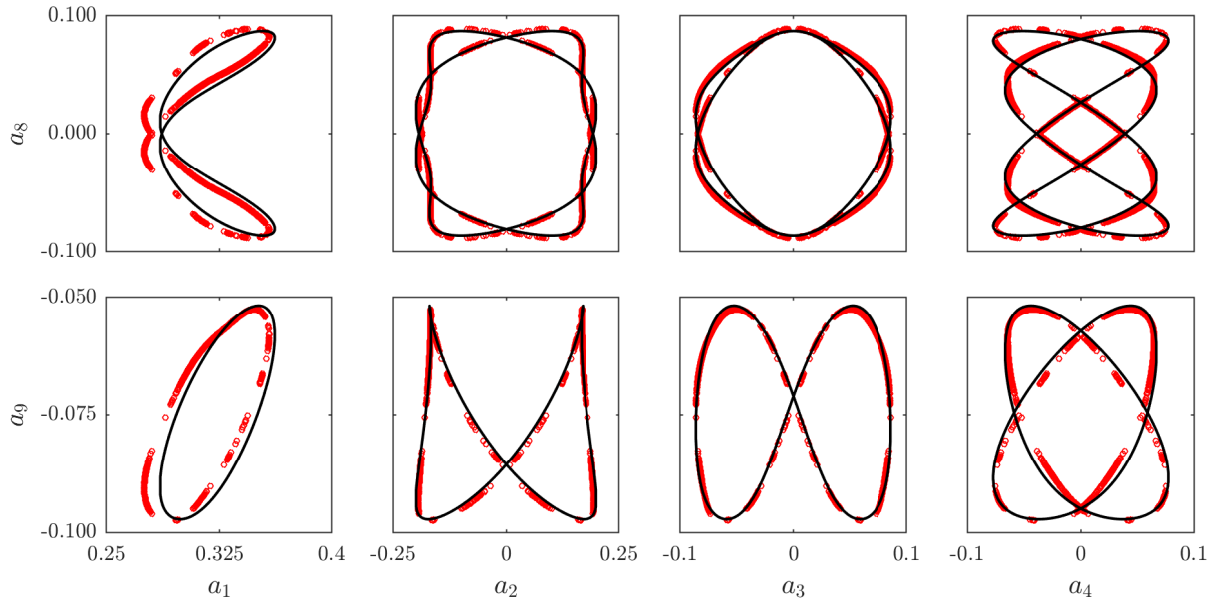

Figure 5: UPOs which maximise $\overline{\mathcal{E}}$ at $Re = 95$, projected onto various 2D subspaces (▬). Also shown are points in the set $\mathcal{S}_{2.55\mathrm{e}-4}$ (○), with the auxiliary function used of the form $P_{10}(\mathbf{a})$.

why it results in a slightly better orbit approximation. A more systematic investigation, both theoretical and computational, is left for future work.

An identical approximation of the maximal orbits for $\overline{\mathcal{E}}$ was performed at $Re = 105$. At this higher value of Re, the SOS bounds on $\overline{\mathcal{E}}$ and $\overline{\mathcal{D}}$ still agree very well with the averages over the most extremal UPOs of the PO1 family. Figure 6 shows these orbits, together with the SOS approximations obtained after optimising $V = P_9(\mathbf{a}) + \|\mathbf{a}\|^{10}$ to yield near-optimal bounds on $\overline{\mathcal{E}}$. The approximation points were computed with $\varepsilon = 3.93 \times 10^{-6}$. The results are just as convincing as at $Re = 95$, and equally good approximations are obtained when maximising lower bounds on $\overline{\mathcal{D}}$.

Finally, we repeated the analysis at $Re = 250$, for which there remains a gap between the available UPO data and our bounds on both $\overline{\mathcal{E}}$ and $\overline{\mathcal{D}}$. Figure 7 compares the most extremal known UPOs at $Re = 250$ to our SOS approximations of the extremal orbits, which consist of points in the set $\mathcal{S}_{4.71\mathrm{e}-6}$ obtained after optimising upper bounds on $\overline{\mathcal{E}}$ with $V = P_9(\mathbf{a}) + \|\mathbf{a}\|^{10}$. The two do not match and the SOS approximations do not look like closed orbits, although they do resemble sections of a UPO. Even worse results were obtained when maximising lower bounds on $\overline{\mathcal{D}}$. It should be stressed once again that the approximation procedure of section 3 is not guaranteed to localise the entire extremal UPO, but only regions of state space where it spends a large fraction of its time period, provided that V is very close to optimal. The small difference in the optimal bound on $\overline{\mathcal{E}}$ obtained with degree-8 and degree-10 V (cf. figure 2) suggests that the latter is close to optimal, and we conjecture that the extremal UPO at $Re = 250$ indeed passes near the approximating points shown in figure 7. We also expect that V of higher degree would produce a better approximation, but we could not test this due to both the increase in required computational resources and the aforementioned degradation in numerical conditioning.

Unfortunately, the approximation points in figure 7 did not provide sufficiently good initial conditions to converge to the extremal UPO at $Re = 250$ with our single-shooting Newton–Raphson algorithm. This is due in part to our degree-10 V not being sufficiently close to optimal, but also due to the UPOs of the system becoming more and more unstable as the Reynolds

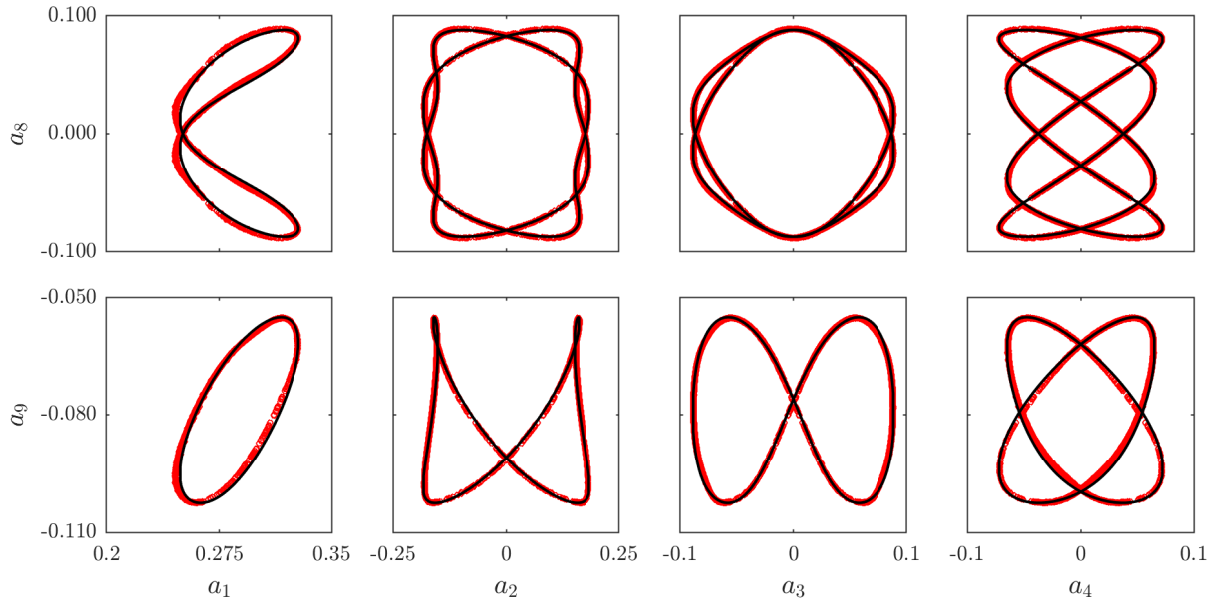

Figure 6: UPOs which maximise $\overline{\mathcal{E}}$ at $Re = 105$, projected onto various 2D subspaces (━). Also shown are points in the set $\mathcal{S}_{3.93\mathrm{e}-6}$ (○), with the auxiliary function used of the form $P_9(\mathbf{a}) + \|\mathbf{a}\|^{10}$.

number is raised. We believe that the latter issue is particularly relevant, and that it could be resolved with multiple-shooting or other, more sophisticated methods for converging UPOs beyond the authors' expertise.

For some initial points, however, our basic single-shooting Newton–Raphson algorithm did converge to a UPO not reported in [35], which is plotted in figure 7 together with three others obtained from symmetry considerations. This UPO was then numerically continued using MatCont to produce the family previously referred to as PO2. The orbits shown in figure 7 belong to the upper branch of this family in figure 2. A third family of UPOs, earlier referred to as PO3, bifurcates from this outer branch. Finally, although no bifurcations were detected on the outer extremal branch of the PO1 family using MatCont, the PO2 family in fact bifurcates from this outer branch. This can be seen clearly in panel (c) of figure 2.

6 Conclusions

In this paper we discussed the numerical implementation of an approach to approximate trajectories that maximise or minimise infinite-time averages of ODE systems governed by polynomial dynamics. The approach, originally proposed in [24], relies on the construction of an auxiliary function V that produces a near-sharp bound λ on the infinite-time average of a quantity Φ, followed by the computation of sublevel sets $\mathcal{S}_\varepsilon$ of the function $\lambda - \Phi - \mathbf{f} \cdot \nabla V$. Here, we have proved that arbitrarily optimal polynomial V and the corresponding bounds can always be constructed computationally when the optimisation problem (6) is implemented using weighted SOS constraints. Since such constraints can be reformulated as SDPs, the first step of the approximation procedure can always be carried out efficiently (precisely, in polynomial time) given enough computational resources. We have also proposed a simple and highly parallelisable scheme based on direct nonlinear minimisation of $\lambda - \Phi - \mathbf{f} \cdot \nabla V$ to compute points in the set $\mathcal{S}_\varepsilon$ for high-dimensional systems, for which finding the whole set $\mathcal{S}_\varepsilon$ is impossible in practice. These

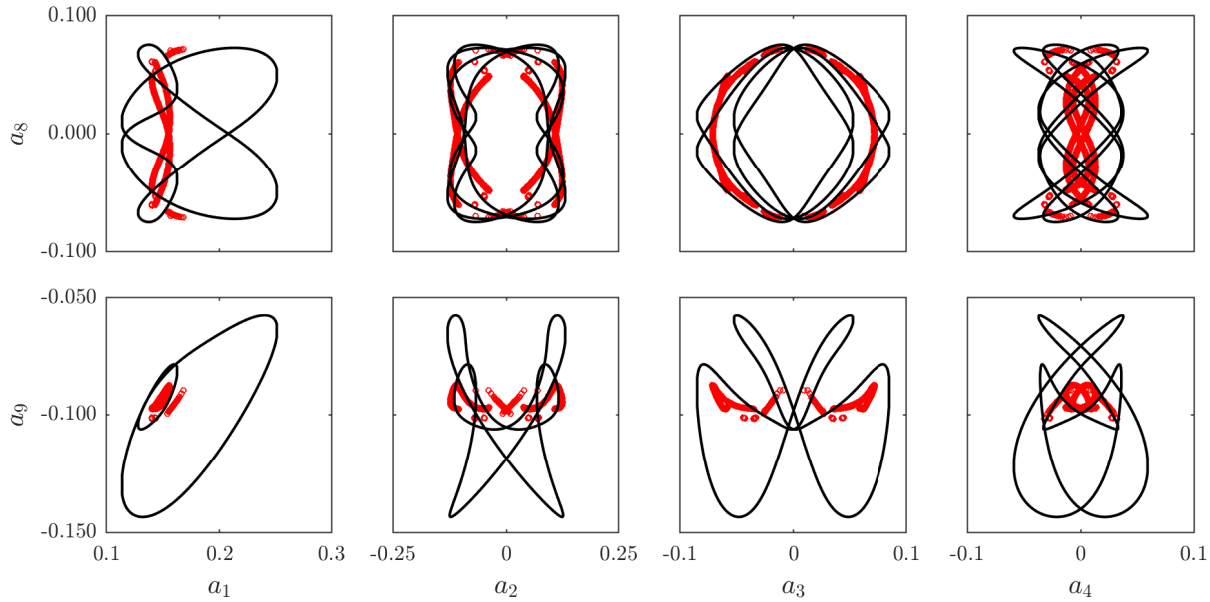

Figure 7: Most extremal known UPOs for $\overline{\mathcal{E}}$ at $Re = 250$, projected onto various 2D subspaces (▬). Also shown are points in the set $\mathcal{S}_{4.71\mathrm{e}-6}$ (○), with the auxiliary function used of the form $P_9(\mathbf{a}) + \|\mathbf{a}\|^{10}$.

points often approximate the extremal trajectories well, and provide good initial conditions to algorithms for converging UPOs.

The potential of these methods was demonstrated by applying them to a nine-dimensional model of shear flow [34]. We have calculated upper bounds on the mean energy of perturbations from the laminar state and lower bounds on the mean energy dissipation rate at a range of Re, and produced approximations to the corresponding extremal orbits. This has resulted in the discovery of three families of UPOs, born respectively at $Re = 81.24$, $Re = 241.5$ and $Re = 330.2$. These families are, to the best of our knowledge new, and not among those previously reported in [35]. The first family of UPOs is simultaneously minimal for $\overline{\mathcal{D}}$ and maximal for $\overline{\mathcal{E}}$ when $81.24 \lessapprox Re \lessapprox 125$, but cease to be extremal at higher Re. The other two families of UPOs appear not to be extremal at any Re. Since increasing the degree of V brings very little improvement in our bounds for both $\overline{\mathcal{D}}$ and $\overline{\mathcal{E}}$ when $Re \gtrapprox 125$, we conjecture the existence of another family of UPOs, which simultaneously optimises both quantities and attains values that almost saturate the corresponding bounds. Finding these extremal orbits remains a topic for future work, although figure 7 gives an idea of the regions of state space in which they lie.

While we have focused on finding UPOs that optimise $\overline{\mathcal{E}}$ and $\overline{\mathcal{D}}$, other orbits can be discovered by varying the choice of Φ. Indeed, any particular UPO can in principle be found by letting Φ be an indicator function, equal to unity on the UPO of interest and zero elsewhere, or a smooth polynomial approximation thereof. Of course, this cannot be done in practice unless a priori information on the location of a certain orbit is available. Nevertheless, varying Φ enables one to systematically search for periodic orbits guided by the SOS approximations. This much is true not only for the shear flow model considered here, but also for many other ODE systems. In addition, the methods can be extended to discrete-time and stochastic dynamics; see [23] for a discussion of this from the perspective of invariant measures.

In principle, the techniques discussed in this work are applicable to polynomial ODEs of arbitrarily high dimension. At the time of writing, however, constructing near-optimal auxiliary functions using off-the-shelf software becomes prohibitively expensive for ODEs with more

than $\mathcal{O}(10)$ states, because the computational resources required by standard interior-point SDP solvers grow quickly as the degree of V is raised, and the numerical conditioning degrades. To help alleviate these problems, Theorem 2 provides a method to reduce the number of decision variables in the related SDPs by invoking symmetries. Alternatives to computationally intensive general-purpose interior-point SDP solvers for optimisation with SOS constraints have been proposed recently [53–55] and further gains are possible if (scaled) diagonally dominant SOS constraints are used [56]. These are stronger than standard SOS conditions but can be implemented as linear or second-order cone programs instead of SDPs. We expect that the computational barrier to studying high-dimensional systems using SOS optimisation will be removed in the near future using a combination of any of these very promising techniques.

Finally, we remark that the algorithm we used in this work to numerically converge UPOs is a simple single-shooting Newton–Raphson method. Our SOS approximations to UPOs can be made even more useful if one employs more sophisticated techniques, which extract more information from the approximation than a single point used as the initial guess. One option is to initialise a multiple-shooting Newton–Raphson method [57] using a selection of the approximation points obtained from the SOS computation. Another interesting possibility is to leverage a variational method for finding UPOs introduced by Lan & Cvitanović [58] and further developed by Boghosian *et al.* [59]. This method finds a UPO by evolving an initial loop which approximates it. Connecting by smooth curves the approximation points obtained with the methods discussed in this work would generate such an initial loop, although it is not immediately clear how this connection should be performed. Whether the SOS approximation method presented here can successfully be utilised in conjunction with multiple-shooting or the variational approach of [58, 59] remains to be seen and should be investigated by future work. In principle, such a combined approach has the potential to become a powerful new method to tackle the long standing problem of searching for UPOs in differential dynamical systems.

Acknowledgements

The authors would like to thank Jeff Moehlis for providing data used to produce figure 2. During this work, one author (ML) was supported by EPSRC studentship award 2092930, and another author (GF) was supported by an EPSRC Doctoral Prize Fellowship and an Imperial College Research Fellowship.

References

[1] P. Cvitanović. Periodic orbits as the skeleton of classical and quantum chaos. *Physica D: Nonlinear Phenomena*, 51(1):138–151, 1991.

[2] J. P. Eckmann and D. Ruelle. Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics*, 57(3):617–656, 1985.

[3] R. Artuso, E. Aurell, and P. Cvitanovic. Recycling of strange sets: I. Cycle expansions. *Nonlinearity*, 3(2):325–359, 1990.

[4] P. Cvitanović. Dynamical averaging in terms of periodic orbits. *Physica D: Nonlinear Phenomena*, 83(1):109–123, 1995.

[5] P. Cvitanović, R. Artuso, R. Mainieri, G. Tanner, and G. Vattay. *Chaos: Classical and Quantum*. Niels Bohr Institute, 2017. http://www.chaosbook.org/ (version 15.9).

[6] E. Ott, C. Grebogi, and J. A. Yorke. Controlling chaos. *Physical Review Letters*, 64(11):1196–1199, 1990.

[7] C. T. Kelley. *Solving Nonlinear Equations with Newton's Method*. Fundamentals of Algorithms. SIAM, 2003.

[8] D. Auerbach, P. Cvitanović, J.-P. Eckmann, G. Gunaratne, and I. Procaccia. Exploring chaotic motion through periodic orbits. *Physical Review Letters*, 58(23):2387–2389, 1987.

[9] J. D. Skufca, J. A. Yorke, and B. Eckhardt. Edge of Chaos in a Parallel Shear Flow. *Physical Review Letters*, 96(17):174101, 2006.

[10] L. Kim and J. Moehlis. Characterizing the edge of chaos for a shear flow model. *Physical Review E*, 78(3):036315, 2008.

[11] M. Joglekar, U. Feudel, and J. A. Yorke. Geometry of the edge of chaos in a low-dimensional turbulent shear flow model. *Physical Review E*, 91(5):052903, 2015.

[12] T.-H. Yang, B. R. Hunt, and E. Ott. Optimal periodic orbits of continuous time chaotic systems. *Physical Review E*, 62(2):1950–1959, 2000.

[13] S. I. Chernyshenko, P. Goulart, D. Huang, and A. Papachristodoulou. Polynomial sum of squares in fluid dynamics: A review with a look ahead. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372(2020):20130350, 2014.

[14] G. Fantuzzi, D. Goluskin, D. Huang, and S. I. Chernyshenko. Bounds for Deterministic and Stochastic Dynamical Systems using Sum-of-Squares Optimization. *SIAM Journal on Applied Dynamical Systems*, 15(4):1962–1988, 2016.

[15] D. Goluskin. Bounding extreme values on attractors using sum-of-squares optimization, with application to the Lorenz attractor. arXiv:1807.09814v2 [math.DS], 2018.

[16] D. Goluskin and G. Fantuzzi. Bounds on mean energy in the Kuramoto–Sivashinsky equation computed using semidefinite programming. *Nonlinearity*, 32(5):1705–1730, 2019.

[17] C. R. Doering and P. Constantin. Energy dissipation in shear driven turbulence. *Physical Review Letters*, 69(11):1648–1651, 1992.

[18] C. R. Doering and P. Constantin. Variational bounds on energy dissipation in incompressible flows: Shear flow. *Physical Review E*, 49(5):4087–4099, 1994.

[19] P. Constantin and C. R. Doering. Variational bounds on energy dissipation in incompressible flows. II. Channel flow. *Physical Review E*, 51(4):3192–3198, 1995.

[20] C. R. Doering and P. Constantin. Variational bounds on energy dissipation in incompressible flows. III. Convection. *Physical Review E*, 53(6):5957–5981, 1996.

[21] S. Chernyshenko. Relationship between the methods of bounding time averages. arXiv:1704.02475 [physics.flu-dyn], 2017.

[22] D. Henrion. Semidefinite characterisation of invariant measures for one-dimensional discrete dynamical systems. *Kybernetika*, 48(6):1089–1099, 2012.

[23] M. Korda, D. Henrion, and I. Mezic. Convex computation of extremal invariant measures of nonlinear dynamical systems and Markov processes. arXiv:1807.08956 [math.OC], 2018.

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

[25] L. Vandenberghe and S. Boyd. Semidefinite Programming. *SIAM Review*, 38(1):49–95, 1996.

[26] A. Papachristodoulou and S. Prajna. A tutorial on sum of squares techniques for systems analysis. In *Proceedings of the 2005, American Control Conference, 2005.*, volume 4, pages 2686–2700, 2005.

[27] P. J. Goulart and S. Chernyshenko. Global stability analysis of fluid flows using sum-of-squares. *Physica D: Nonlinear Phenomena*, 241(6):692–704, 2012.

[28] D. Huang, S. Chernyshenko, P. Goulart, D. Lasagna, O. Tutty, and F. Fuentes. Sum-of-squares of polynomials approach to nonlinear stability of fluid flows: An example of application. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 471(2183):20150622, 2015.

[29] D. Henrion and M. Korda. Convex Computation of the Region of Attraction of Polynomial Control Systems. *IEEE Transactions on Automatic Control*, 59(2):297–312, 2014.

[30] W. Tan and A. Packard. Stability region analysis using sum of squares programming. In *2006 American Control Conference*, pages 2297–2302, 2006.

[31] D. Lasagna, D. Huang, O. R. Tutty, and S. Chernyshenko. Sum-of-squares approach to feedback control of laminar wake flows. *Journal of Fluid Mechanics*, 809:628–663, 2016.

[32] J. Lasserre, D. Henrion, C. Prieur, and E. Trélat. Nonlinear Optimal Control via Occupation Measures and LMI-Relaxations. *SIAM Journal on Control and Optimization*, 47(4):1643–1666, 2008.

[33] A. Majumdar, R. Vasudevan, M. M. Tobenkin, and R. Tedrake. Convex optimization of nonlinear feedback controllers via occupation measures. *The International Journal of Robotics Research*, 33(9):1209–1230, 2014.

[34] J. Moehlis, H. Faisst, and B. Eckhardt. A low-dimensional model for turbulent shear flows. *New Journal of Physics*, 6:56, 2004.

[35] J. Moehlis, H. Faisst, and B. Eckhardt. Periodic Orbits and Chaotic Sets in a Low-Dimensional Model for Shear Flows. *SIAM Journal on Applied Dynamical Systems*, 4(2):352–376, 2005.

[36] D. Goluskin. Bounding Averages Rigorously Using Semidefinite Programming: Mean Moments of the Lorenz System. *Journal of Nonlinear Science*, 28(2):621–651, 2018.

[37] Ö. Karabacak and P. Ashwin. On statistical attractors and the convergence of time averages. *Mathematical Proceedings of the Cambridge Philosophical Society*, 150(2):353–365, 2011.

[38] J. G. Llavona. *Approximation of Continuously Differentiable Functions*, volume 130 of *Mathematics Studies*. North-Holland, 1986.

[39] E. N. Lorenz. Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences*, 20(2):130–141, 1963.

[40] C. G. Broyden. The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations. *IMA Journal of Applied Mathematics*, 6(1):76–90, 1970.

[41] R. Fletcher. A new approach to variable metric algorithms. *The Computer Journal*, 13(3):317–322, 1970.

[42] D. Goldfarb. A Family of Variable-Metric Methods Derived by Variational Means. *Mathematics of Computation*, 24(109):23–26, 1970.

[43] D. F. Shanno. Conditioning of Quasi-Newton Methods for Function Minimization. *Mathematics of Computation*, 24(111):647–656, 1970.

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

[45] P. A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming, Series B*, 96(2):293–320, 2003.

[46] P. Parrilo. Polynomial Optimization, Sums of Squares, and Applications. In *Semidefinite Optimization and Convex Algebraic Geometry*, MOS-SIAM Series on Optimization, chapter 3, pages 47–157. SIAM, 2012.

[47] J. B. Lasserre. *An Introduction to Polynomial and Semi-Algebraic Optimization*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2015.

[48] M. Putinar. Positive Polynomials on Compact Semi-algebraic Sets. *Indiana University Mathematics Journal*, 42(3):969–984, 1993.

[49] J. Löfberg. YALMIP : A toolbox for modeling and optimization in MATLAB. In *2004 IEEE International Conference on Robotics and Automation (IEEE Cat. No.04CH37508)*, pages 284–289, 2004.

[50] MOSEK ApS. MOSEK Optimization Toolbox for MATLAB, version 8.1.0.72, 2018.

[51] A. Dhooge, W. Govaerts, Y. A. Kuznetsov, H. G. Meijer, and B. Sautois. New features of the software MatCont for bifurcation analysis of dynamical systems. *Mathematical and Computer Modelling of Dynamical Systems*, 14(2):147–175, 2008.

[52] J. Löfberg. Pre- and Post-Processing Sum-of-Squares Programs in Practice. *IEEE Transactions on Automatic Control*, 54(5):1007–1011, 2009.

[53] D. Papp and S. Yildiz. Sum-of-Squares Optimization without Semidefinite Programming. *SIAM Journal on Optimization*, 29(1):822–851, 2019.

[54] Y. Zheng, G. Fantuzzi, and A. Papachristodoulou. Exploiting Sparsity in the Coefficient Matching Conditions in Sum-of-Squares Programming Using ADMM. *IEEE Control Systems Letters*, 1(1):80–85, 2017.

[55] Y. Zheng, G. Fantuzzi, and A. Papachristodoulou. Fast ADMM for sum-of-squares programs using partial orthogonality. *IEEE Transactions on Automatic Control*, 2018. DOI: 10.1109/TAC.2018.2886170 (early access version).

[56] A. Ahmadi and A. Majumdar. DSOS and SDSOS Optimization: More Tractable Alternatives to Sum of Squares and Semidefinite Optimization. *SIAM Journal on Applied Algebra and Geometry*, 3(2):193–230, 2019.

[57] F. Christiansen. Multipoint shooting method. [5], chapter 16.2.

[58] Y. Lan and P. Cvitanović. Variational method for finding periodic orbits in a general flow. *Physical Review E*, 69(1):016217, 2004.

[59] B. M. Boghosian, L. M. Fazendeiro, J. Lätt, H. Tang, and P. V. Coveney. New variational principles for locating periodic orbits of differential equations. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369(1944):2211–2218, 2011.