Bounds for Deterministic and Stochastic Dynamical Systems Using Sum-of-Squares Optimisation

Giovanni Fantuzzi

October 15, 2015

1 Introduction

One of the challenges in studying physical systems that exhibit complex temporal (or spatio-temporal) dynamics is to obtain rigorous quantitative predictions of the system's behaviour. Given that rigorous, closed-form solution of the governing equations are not generally available, an interesting problem is to quantify the average properties of the system, described by an observable function φ of the system's state $\mathbf{x}$. Note that $\mathbf{x}$ may be infinite-dimensional and depend on position $\mathbf{s}$ and time t. Specifically, one is interested in estimating the value of

$$\langle \varphi(\mathbf{x}) \rangle = \lim_{T\to\infty} \frac{1}{T} \int_0^T \varphi[\mathbf{x}(t)]\, dt \tag{1}$$

or, for spatially-extended systems over a domain Ω with measure $\mu(\Omega)$,

$$\langle \varphi(\mathbf{x}) \rangle = \lim_{T\to\infty} \frac{1}{T}\frac{1}{\mu(\Omega)} \int_0^T \int_\Omega \varphi[\mathbf{x}(\mathbf{s},t)]\, d\mathbf{s}dt, \tag{2}$$

where we assume that the long-time limits exist. Assuming that the system has only one attracting set of dynamical interest (either a stable equilibrium, a periodic orbit or a strange attractor), this amounts to computing the time-average of φ as the system evolves on the attractor, irrespective of the specific initial condition.

Problems of this type have received increasing interest in recent years, focussing mainly on spatially extended systems described by partial differential equations (PDEs) and leading to the development of a variational technique known as the "background method" [4]. The method is based on the optimisation of a quadratic functional of the state variable, denoted by $V(\mathbf{x})$, that allows the derivation of rigorous bounds for $\langle\varphi\rangle$ (either upper or lower); in some cases, V can be interpreted as a Lyapunov function [2]. Classical applications of the background method include the estimation of the net turbulent heat transport in Rayleigh-Bénard convection [6, 7] and the computation of rigorous bounds on the energy dissipation in shear flows [5, 9]. Similar techniques can also be applied to finite-dimensional systems exhibiting chaotic behaviour, such as truncated low-order models of Rayleigh-Bénard convection including the well-known Lorenz system [14, 15].

An alternative approach to derive rigorous bounds for finite-dimensional systems with polynomial dynamics has been proposed recently [3]. The method, introduced in the context of fluid flows [3], is based on Sum-of-Squares (SOS) polynomial optimisation and allows

the construction of higher-than-quadratic Lyapunov-type functionals, generalising the background method.

Whilst the bounds obtained by these methods are rigorous and hold irrespectively of the system's initial condition, they do not generally represent the time averages observed in experiments or numerical simulations accurately. This is usually because the system possesses at least one unstable solution $\mathbf{x}_u(t)$— a fixed point or a periodic orbit — for which the rigorous bounds obtained with the aforementioned techniques are sharp; yet, $\mathbf{x}_u$ is never observed in practice since the real system is always subject to small perturbations. If a neighbourhood $\mathcal{U}$ of $\mathbf{x}_u$ is known to not belong to the system's attractor, a possible solution to this problem is to directly remove $\mathcal{U}$ from the analysis. Another solution, that in principle does not require any a priori knowledge of $\mathbf{x}_u$, is to model the physical disturbances by adding a stochastic forcing term of strength ε to the dynamical system [3].

In this work, we first review some of the ideas introduced in [3] in the context of a general dynamical system with polynomial dynamics. We then investigate how the influence of an unstable solution on the estimates of $\langle\varphi\rangle$ obtained with SOS optimisation can be removed. We will limit ourselves to unstable fixed points and develop the two approaches outlined above for systems with a repelling fixed point; we will not consider the case of saddle points or unstable limit cycles. We will also illustrate how these ideas work in practice by applying them to the well known Van der Pol oscillator.

2 Bounds Using SOS Optimisation: A Review

To make this work self-contained, we start by reviewing the ideas presented in [3]. Consider the dynamical system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^n \tag{3}$$

and assume that the trajectories $\mathbf{x}(t)$ are uniformly bounded as $t \to \infty$ regardless of the initial condition $\mathbf{x}_0$. Suppose there exists a function $V[\mathbf{x}(t)]$, continuous in $\mathbf{x}$, and a constant L such that

$$\dot{V} + \varphi - L \geq 0 \tag{4}$$

for all possible values of the state $\mathbf{x}$. Since any trajectory $x(t)$ is uniformly bounded as $t \to \infty$, so is V; hence, time averaging the last expression we obtain

$$\langle\varphi\rangle \geq L. \tag{5}$$

An upper bound U can be found in a similar way by reversing the inequality sign, and we summarise the above in the following:

Proposition 2.1. *Let $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ be a dynamical system whose trajectories are bounded at all times and let $\varphi(\mathbf{x})$ be an observable. If there exist continuous functions $V_u(\mathbf{x})$, $V_l(\mathbf{x})$, and constants U, L such that*

$$\mathcal{D}_u(\mathbf{x}) := \mathbf{f}\cdot\nabla V_u + \varphi - U \leq 0, \qquad \forall \mathbf{x} \in \mathbb{R}^n, \tag{6a}$$

$$\mathcal{D}_l(\mathbf{x}) := \mathbf{f}\cdot\nabla V_l + \varphi - L \geq 0, \qquad \forall \mathbf{x} \in \mathbb{R}^n, \tag{6b}$$

then

$$L \leq \langle\varphi\rangle \leq U. \tag{7}$$

The functions V_u and V_l (which we will occasionally refer to as *storage functions*) that achieve given bounds U and L may not be unique; yet, as one could expect, constructing them is generally a challenging task. However the problem is greatly simplified when $\mathbf{f}$ and φ are polynomials of the states x_i, $i \in \{1, ..., n\}$. In fact, if V_u and V_l are chosen to be polynomials, so are $\mathcal{D}_u$ and $\mathcal{D}_l$, hence (6a) and (6b) amount to verifying the non-negativity of a polynomial expression. Whilst this is an NP-hard problem, the computational complexity can be significantly reduced by replacing the conditions $\mathcal{D}_l(\mathbf{x}) \geq 0$ and $-\mathcal{D}_u(\mathbf{x}) \geq 0$ (note the minus sign) with the stronger conditions that $\mathcal{D}_l$ and $-\mathcal{D}_u$ admit a SOS decomposition, i.e. that there exists families of polynomials $\{p_i(\mathbf{x})\}_{i=1}^M$ and $\{q_i(\mathbf{x})\}_{i=1}^N$ such that

$$\begin{aligned} \mathcal{D}_u(\mathbf{x}) &= \sum_{i=1}^{M} p_i(\mathbf{x})^2, \\ \mathcal{D}_l(\mathbf{x}) &= \sum_{i=1}^{N} q_i(\mathbf{x})^2. \end{aligned} \tag{8}$$

These conditions can be formulated in terms of linear matrix inequality (LMI) constraints, a particular type of convex constraint; a brief explanation is given in Appendix A. Optimisation problems with LMI constraints, known as semidefinite programmes (SDPs), can in turn be solved efficiently with a number of software packages, e.g. YALMIP [10] and SOSTOOLS [12]. Consequently, polynomial storage functions and the corresponding bounds U and L may be constructed systematically by solving the SoS optimisation problems

$$\begin{aligned} &\min_{V_u, U} \quad U \\ &\text{such that} \quad U - \mathbf{f} \cdot \nabla V_u - \varphi \in \Sigma \end{aligned} \tag{9}$$

and

$$\begin{aligned} &\max_{V_l, L} \quad L \\ &\text{such that} \quad \mathbf{f} \cdot \nabla V_l + \varphi - L \in \Sigma \end{aligned} \tag{10}$$

where Σ denotes the set of SOS polynomials and the optimisation is over the coefficients of the polynomials V_u and V_l.

3 Improved Bounds for Deterministic Systems

As mentioned in the introduction and as noted in [3], the existence of unstable invariant trajectories $\mathbf{x}_u(t)$ (equilibria and/or limit cycles) poses a problem if one is interested in bounds that accurately describe the time averages measured in experiments. This is because the bounds U and L obtained from Proposition 2.1 must hold for any possible trajectory of the system, including unstable invariant solutions. To illustrate the idea, let $\mathbf{x}_u$ be an unstable equilibrium, such that $\mathbf{f}(\mathbf{x}_u) = 0$; for definiteness, assume that $\varphi(\mathbf{x}_u)$ is much lower than the observed time average $\langle\varphi\rangle$. Then, any lower bound L cannot coincide with $\langle\varphi\rangle$ since evaluating (6b) at $\mathbf{x}_u$ yields

$$\mathcal{D}_l(\mathbf{x}_u) = \varphi(\mathbf{x}_u) - L \geq 0, \tag{11}$$

implying that $L \leq \langle \varphi(\mathbf{x}_u) \rangle = \varphi(x_u) < \langle \varphi \rangle$. Similarly, integrating both sides of the inequality $\mathcal{D}_l(\mathbf{x}_u) \geq 0$ along an unstable limit cycle $\mathbf{x}_u(t)$, the term $\mathbf{f} \cdot \nabla V_l$ vanishes by periodicity and one obtains

$$L \leq \langle \varphi[\mathbf{x}_u(t)] \rangle \tag{12}$$

i.e. the bound is constrained by the time average of φ over the periodic orbit. The same problem arises for any upper bound U if $\varphi(\mathbf{x}_u) > \langle \varphi \rangle$.

One possible solution is to enforce (6a) and (6b) everywhere except for a neighbourhood $\mathcal{U}$ of an unstable invariant solution $\mathbf{x}_u$. This relaxation can indeed be carried out rigorously and implemented if $\mathbf{x}_u$ is known and if all trajectories starting at the perturbed position $\mathbf{x}_0 = \mathbf{x}_u + \delta\mathbf{x}$ permanently leave $\mathcal{U}$ after a finite time $\tau = \tau(\mathbf{x}_0)$.

Let us proceed formally, and assume that the dynamical system (3) has global attractor $\mathcal{A}$ and an unstable solution $\mathbf{x}_u$. Then, $\mathcal{B} = \mathbb{R}^n \setminus \{\mathbf{x}_u\}$ is the basin of attraction of $\mathcal{A}$. Since all trajectories leave $\mathcal{U}$, moreover, $\mathcal{T} = \mathbb{R}^n \setminus \mathcal{U}$ is an absorbing domain. Clearly, $\mathcal{A} \subseteq \mathcal{T} \subseteq \mathcal{B}$ and for any trajectory starting inside $\mathcal{B}$ one has

$$\langle \varphi \rangle = \lim_{T\to\infty} \frac{1}{T} \int_0^T \varphi[\mathbf{x}(t)]\, dt = \lim_{T\to\infty} \frac{1}{T} \int_\tau^T \varphi[\mathbf{x}(t)]\, dt \tag{13}$$

i.e. the time average of φ is completely determined by the dynamics inside $\mathcal{T}$. The same result applies if $\mathbf{x}_u$ is a saddle by letting $\mathcal{U}$ be a neighbourhood of the entire stable manifold $\mathcal{W}_s$ and $\mathcal{B} = \mathbb{R}^n \setminus \mathcal{W}_s$. A trivial extension of Proposition 2.1 is therefore

Proposition 3.1. *Let $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ be a dynamical system in $\mathbb{R}^n$, let $\mathcal{T}$ be a bounded absorbing domain containing an attractor $\mathcal{A}$ and let $\mathcal{B}$ be the basin of attraction of $\mathcal{A}$. If there exist continuous functions $V_u(\mathbf{x})$, $V_l(\mathbf{x})$, and constants U, L such that*

$$\mathcal{D}_u(\mathbf{x}) = \mathbf{f} \cdot \nabla V_u + \varphi - U \leq 0, \qquad \forall \mathbf{x} \in \mathcal{T}, \tag{14a}$$

$$\mathcal{D}_l(\mathbf{x}) = \mathbf{f} \cdot \nabla V_l + \varphi - L \geq 0, \qquad \forall \mathbf{x} \in \mathcal{T}, \tag{14b}$$

then for any initial condition $\mathbf{x}_0 \in \mathcal{B}$

$$L \leq \langle \varphi \rangle \leq U. \tag{15}$$

The problem of eliminating the influence of $\mathbf{x}_u$ on the bounds therefore reduces to that of finding a suitable absorbing domain for the attracting set. This is generally not a trivial task; however, when $\mathbf{f}(\mathbf{x})$ is polynomial an absorbing domain $\mathcal{T}$ may be constructed using SOS techniques [16]. Moreover, $\mathcal{T}$ is generally a semi-algebraic set; for clarity, let us assume that $\mathcal{T} = \{\mathbf{x} \,|\, g(\mathbf{x}) \geq 0\}$ for some polynomial g.

Equation (14b) then requires that $\mathcal{D}_l(\mathbf{x}) \geq 0$ when $g(\mathbf{x}) \geq 0$ (a similar argument holds for (14a) and will not be considered for brevity). It is easy to see that this condition is satisfied if there exists a non-negative polynomial $s(\mathbf{x})$ such that $\mathcal{D}_l(\mathbf{x}) - s(\mathbf{x})g(\mathbf{x}) \geq 0$. Whilst not necessary, this approximation — known as the generalised $\mathcal{S}$-procedure [16, Lemma 2.1] — allows the formulation of two SOS optimisation problems from Proposition 3.1 as

$$\begin{aligned} \min_{V_u, U, s} \quad & U \\ \text{such that} \quad & U - \mathbf{f} \cdot \nabla V_u - \varphi - s\, g \in \Sigma \\ & s \in \Sigma \end{aligned} \tag{16}$$

and

$$\begin{aligned} &\max_{V_l, L, s} \quad L \\ \text{such that} \quad &\mathbf{f} \cdot \nabla V_l + \varphi - L - s\, g \in \Sigma \\ &s \in \Sigma \end{aligned} \tag{17}$$

The $\mathcal{S}$-procedure generalises to more complicated semi-algebraic absorbing domains; more details and examples can be found in [16]. However, its applicability relies on the knowledge of an absorbing domain that does not contain unstable solutions and that is tractable using SOS techniques. This may not be the case for saddle point with a complicated stable manifold, or if the unstable trajectory cannot be separated from the attractor; an example combining both issues is the unstable saddle equilibrium at the origin in the well-known Lorenz system [17].

4 Bounds for Stochastically-Driven Systems

An alternative approach to eliminate the influence of unstable point on the bounds, proposed by Chernyshenko *et al.* [3], is to model the external disturbances that affect any real system with a small-amplitude stochastic forcing term. If the system is stochastically stable (in the sense of [18]), in fact, one can infer bounds for the original, unperturbed system by studying the vanishing-noise limit.

In Section 4.1, we extend the initial ideas of [3] by considering a stochastic dynamical system forced by finite-amplitude noise, and show how to determine bounds on its statistical properties using SOS programming. Our analysis applies not only in the small-noise limit, but to system which are inherently stochastic. In Section 4.2, we will then study the problem of computing rigorous bounds in the specific case of vanishing noise strength.

4.1 Bounds for system with finite noise

Consider the stochastic dynamical system driven by additive white noise

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \sqrt{2\varepsilon}\, \boldsymbol{\sigma} \boldsymbol{\xi}, \tag{18}$$

where $\boldsymbol{\xi}$ is a standard Wiener process and $\mathbf{x}, \boldsymbol{\xi} \in \mathbb{R}^n$. The constant matrix $\boldsymbol{\sigma} \in \mathbb{R}^{n \times n}$ describes the relative effect of each ξ_i on each state x_i, while the overall noise strength ε represents the balance between the deterministic and the stochastic dynamics.

This system can be interpreted as a stochastic perturbation of (3), and its state $\mathbf{x}$ (now a random variable) is described by the system's probability density function (PDF) ρ. We remind the reader that ρ must be a non-negative distribution such that $\|\rho\|_{L^1} = 1$. We assume that the system's trajectories remain bounded at all times, and that a statistical equilibrium is reached so the PDF satisfies the steady Fokker-Planck equation

$$\nabla \cdot (\varepsilon \mathbf{D} \nabla \rho - \mathbf{f} \rho) = 0, \tag{19}$$

where $\mathbf{D} = \boldsymbol{\sigma}^T \boldsymbol{\sigma}$. The stationary expectation of an observable $\varphi(\mathbf{x})$ can be computed as

$$\langle \varphi \rangle_\varepsilon = \int_{\mathbb{R}^n} \rho(\mathbf{x}) \varphi(\mathbf{x}) d\mathbf{x} \tag{20}$$

where we have introduced a subscript ε to indicate that the expectation depends on the overall noise strength. Clearly, L is a lower bound for $\langle\varphi\rangle_\varepsilon$ if

$$\int_{\mathbb{R}^n} \rho\,(\varphi - L)\,d\mathbf{x} \geq 0 \tag{21}$$

Enforcing (19) explicitly with a Lagrange multiplier function $V_l(\mathbf{x})$ and integrating by parts we obtain

$$\begin{aligned}&\int_{\mathbb{R}^n} \rho\,[\nabla\cdot(\varepsilon\mathbf{D}\nabla V_l) + \mathbf{f}\cdot\nabla V_l + \varphi - L]\,d\mathbf{x}\\ &+ \lim_{R\to\infty}\int_{\|\mathbf{x}\|=R} (\varepsilon\, V_l\,\mathbf{D}\nabla\rho - \varepsilon\,\rho\,\mathbf{D}\nabla V_l - \rho\, V_l\,\mathbf{f})\cdot\boldsymbol{\nu}(\mathbf{x})\,dS \geq 0\end{aligned} \tag{22}$$

where $\boldsymbol{\nu}(\mathbf{x})$ is the outwards unit normal to the sphere $\|\mathbf{x}\| = R$ and dS is the surface element. Since we have assumed that the system's trajectories are bounded when $\varepsilon = 0$, it is reasonable to expect that ρ decays exponentially at infinity, so that the boundary term vanishes if V_l does not grow too quickly. Thus, one is left with the condition

$$\int_{\mathbb{R}^n} \rho\,[\nabla\cdot(\varepsilon\mathbf{D}\nabla V_l) + \mathbf{f}\cdot\nabla V_l + \varphi - L]\,d\mathbf{x} \geq 0. \tag{23}$$

Since ρ is non-negative and V_l is arbitrary (up to some controlled-growth conditions at infinity), one could prove that L is a lower bound for $\langle\varphi\rangle_\varepsilon$ if there exists V_l such that the term in brackets is everywhere non-negative. Applying the same argument after reversing the inequality sign gives sufficient conditions for an upper bound U on $\langle\varphi\rangle_\varepsilon$, and we conclude the following:

Proposition 4.1. *Let $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \sqrt{2\varepsilon}\,\boldsymbol{\sigma}\boldsymbol{\xi}$, with $\mathbf{x}, \boldsymbol{\xi} \in \mathbb{R}^n$ and $\boldsymbol{\sigma} \in \mathbb{R}^{n\times n}$, be a stochastic system for which a steady PDF exists and let $\mathbf{D} = \boldsymbol{\sigma}^T\boldsymbol{\sigma}$. If there exist functions V_u and V_l such that*

$$\lim_{R\to\infty}\int_{\|\mathbf{x}\|=R} (\varepsilon V_u\mathbf{D}\nabla\rho - \varepsilon\rho\mathbf{D}\nabla V_u - \mathbf{f}\rho V_u)\cdot\boldsymbol{\nu}(\mathbf{x})\,dS(\mathbf{x}) = 0 \tag{24a}$$

$$\lim_{R\to\infty}\int_{\|\mathbf{x}\|=R} (\varepsilon V_l\mathbf{D}\nabla\rho - \varepsilon\rho\mathbf{D}\nabla V_l - \mathbf{f}\rho V_l)\cdot\boldsymbol{\nu}(\mathbf{x})\,dS(\mathbf{x}) = 0 \tag{24b}$$

and

$$\varepsilon\nabla\cdot(\mathbf{D}\nabla V_u) + \mathbf{f}\cdot\nabla V_u + \varphi - U \leq 0 \qquad \forall\mathbf{x}\in\mathbb{R}^n, \tag{25a}$$

$$\varepsilon\nabla\cdot(\mathbf{D}\nabla V_l) + \mathbf{f}\cdot\nabla V_l + \varphi - L \geq 0 \qquad \forall\mathbf{x}\in\mathbb{R}^n, \tag{25b}$$

then the stationary expectation of the random variable $\varphi(\mathbf{x})$ is bounded by

$$L \leq \langle\varphi\rangle_\varepsilon \leq U. \tag{26}$$

Note that the same result was derived in [3] using an alternative approach and fixing $\boldsymbol{\sigma}$ to be the identity matrix. For a given noise amplitude ε, a SOS relaxation of inequalities (25a) and (25b) yields the optimisation problems for the bounds

$$\begin{aligned}&\min_{V_u, U} \quad U\\ &\text{such that} \quad U - \varepsilon\nabla\cdot(\mathbf{D}\nabla V_u) - \mathbf{f}\cdot\nabla V_u - \varphi \in \Sigma\end{aligned} \tag{27}$$

and

$$\begin{aligned} &\max_{V_l, L} \quad L \\ &\text{such that} \quad \varepsilon\nabla\cdot(\mathbf{D}\nabla V_l) + \mathbf{f}\cdot\nabla V_l + \varphi - L \in \Sigma \end{aligned} \tag{28}$$

Note that we have assumed that the boundary terms vanish when V_u and V_l are polynomials, which is equivalent to the statement that all moments of the distribution ρ order up to the degree of the storage functions exist.

Furthermore, note that inequalities (25a) and (25b) are the same as (6a) and (6b), respectively, with the addition of a second order diffusive term. Consequently, a point $\mathbf{x}_u$ such that $\mathbf{f}(\mathbf{x}_u) = 0$ — corresponding to a fixed point of the deterministic system obtained for $\varepsilon = 0$ — does not constrain the bounds on $\langle\varphi\rangle_\varepsilon$ if V_u and V_l have large enough gradients.

4.2 Bounds for system with vanishing noise

Let us assume that a stochastic system is stable in the sense of [18]. Since stochastic bounds are not constrained by fixed points of the corresponding deterministic system (i.e.for $\varepsilon = 0$), the limit $\varepsilon \to 0$ can be studied to infer bounds on the corresponding deterministic system that are not affected by unstable equilibria.

Unfortunately, in practice polynomial storage functions and the SoS optimisation problems (27) and (28) give tight bounds only when the noise strength ε is relatively large. To illustrate the reason of this limitation, let us assume without loss of generality that $\mathbf{x}_u = 0$ and, for definiteness, consider $\varphi = \|\mathbf{x}\|^2$. To achieve a lower bound greater than the trivial result $L = 0$, say $L \sim \mathcal{O}(1)$, one needs

$$\varepsilon\nabla\cdot(\mathbf{D}\nabla V_l) \sim \mathcal{O}(1) \tag{29}$$

at least in a region near the origin, where $\mathbf{f}(\mathbf{x})$ and $\varphi(\mathbf{x})$ are almost negligible. Similar considerations can be made for V_u. When ε is small, large enough gradients can only be achieved if V_u and V_l are polynomials of very high degree, making the SOS problem numerically intractable.

One would therefore like a parametrisation of V_u and V_l with ε that satisfies (29) in a neighbourhood of the unstable states, and that is suitable for polynomial optimisation.

It turns out that an appropriate functional form to study the case $\varepsilon \to 0$ can be derived if the unstable solution is a repelling equilibrium. Henceforth, we will assume that the deterministic system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ has a repelling (focus or node) fixed point at $\mathbf{x} = 0$, i.e. all eigenvalues of the Jacobian $\mathbf{J}_0 \in \mathbb{R}^{n\times n}$ of $\mathbf{f}$ at the origin have positive real part. Without loss of generality, we will also assume that $\varphi(0) = 0$; this can always be achieved with an appropriate shift in φ. For definiteness, we will consider the problem of finding a lower bound when $\langle\varphi\rangle > 0$ and the bound is constrained by the unstable equilibrium at the origin; the analysis can be trivially extended to upper bounds.

Let us start by assuming that V_l is chosen so that $L = \langle\varphi\rangle_\varepsilon$ exactly. Repeating the derivation of Proposition 4.1 with an equality sign, we see that V_l satisfies

$$\varepsilon\,\nabla\cdot(\mathbf{D}\nabla V_l) + \mathbf{f}\cdot\nabla V_l + \varphi - \langle\varphi\rangle_\varepsilon = 0, \tag{30}$$

with boundary conditions described by (24b). When $\varepsilon \to 0$, this is a singularly perturbed boundary value problem, and the method of matched asymptotic expansions can be used

to study the behaviour of V_l. Rather than determining the solution V_l, however, we are interested in determining its scaling with ε and its approximate functional form near the origin, where we expect large gradients.

The appropriate "inner layer" coordinate stretching is $\mathbf{x} = \varepsilon^{1/2}\hat{\mathbf{x}}$, where $\hat{\mathbf{x}} \sim \mathcal{O}(1)$ as $\varepsilon \to 0$. We therefore expect that, near the origin, we can approximate $V_l = V_l(\varepsilon^{-1/2}\mathbf{x})$.

To determine a suitable functional form, instead, let us consider the intermediate region $\varepsilon^{1/2} \ll x_i \ll 1$, $i \in \{1, ..., n\}$, where, to leading order, equation (30) reduces to

$$\mathbf{x}^T \, \mathbf{J}_0^T \, \nabla V_l = \langle \varphi \rangle_\varepsilon. \tag{31}$$

Introducing a characteristic coordinate s such that

$$\frac{d\mathbf{x}}{ds} = \mathbf{J}_0 \, \mathbf{x}, \tag{32}$$

one has

$$V_l = \langle \varphi \rangle_\varepsilon \, s + const. \tag{33}$$

Moreover, if $\mathbf{v}_i$ and λ_i denote each of the n eigenvectors and eigenvalues of $\mathbf{J}_0$ (with $\Re\{\lambda_i\} > 0$ since the origin is a repellor), the solution of (32) can be written as

$$\mathbf{x} = \sum_{i=1}^{n} A_i \mathbf{v}_i e^{\lambda_i \, s} \tag{34}$$

for some constants A_i, suggesting that s should be some logarithmic function of $\mathbf{x}$.

Finally, since $\mathbf{x} = 0$ is a repelling point, it is not unreasonable to expect that there exists a linear coordinate transformation $\mathbf{x} \to \mathbf{u}(\mathbf{x})$ that makes the dynamics near the origin rotationally invariant; this situation is sketched in Figure 4.2 for a 2D system. In this case, one expects V_l to depend only on the (squared) radius $\zeta(\mathbf{x}) = \|\mathbf{u}(\mathbf{x})\|^2$. Note that ζ is a homogeneous, positive definite quadratic form of $\mathbf{x}$, i.e.

$$\zeta(\mathbf{x}) = \mathbf{x}^T \mathbf{Z} \mathbf{x}, \quad \mathbf{Z} \succ 0 \tag{35}$$

for a symmetric matrix $\mathbf{Z}$ to be chosen appropriately.

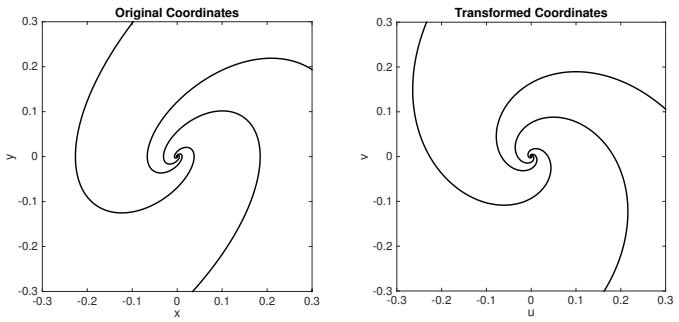

Figure 1: Sketch of trajectories for a 2D system in the original $\mathbf{x} = (x, y)$ coordinates (left) and the transformed $\mathbf{u}(x) = (u(\mathbf{x}), v(\mathbf{x}))$ coordinates.

Combining these heuristic arguments, we suggest that

$$V_l(\mathbf{x}) \approx \alpha \log[\zeta(\mathbf{x})] \tag{36}$$

in the intermediate layer, for some constant α. Consequently, we argue that an appropriate form for V_l is

$$V_l(\mathbf{x}) = \alpha \log[\varepsilon + \zeta(\mathbf{x})] + P_d(\mathbf{x}) \tag{37}$$

Here, ε has been added to the argument of the logarithm to regularise V_l at $\mathbf{x} = 0$ and maintain the correct balance of ε and $\mathbf{x}$ in the inner layer, while P_d is a polynomial of degree d that approximates the outer solution of (30). Moreover, this ansatz could be seen as the generalisation to multiple dimensions of the asymptotic results for one-dimensional systems, presented in Appendix B.

Despite not being polynomial, ansatz (37) is suitable for a SOS formulation. In fact we can substitute

$$\begin{aligned} \nabla V &= \frac{\alpha \nabla \zeta}{\varepsilon + \zeta} + \nabla P_d \\ \nabla \cdot (\mathbf{D} \nabla V) &= \alpha \frac{\nabla \cdot (\mathbf{D} \nabla \zeta)}{\varepsilon + \zeta} - \alpha \frac{\nabla \zeta \cdot (\mathbf{D}\, \nabla \zeta)}{(\varepsilon + \zeta)^2} + \nabla \cdot (\mathbf{D} \nabla P_d) \end{aligned} \tag{38}$$

into (30), multiply by $(\varepsilon + \zeta)^2$ and gather terms to obtain the polynomial inequality

$$\mathcal{L}(\mathbf{x}) := \mathcal{L}_0(\mathbf{x}) + \varepsilon \mathcal{L}_1(\mathbf{x}) + \varepsilon^2 \mathcal{L}_2(\mathbf{x}) + \varepsilon^3 \mathcal{L}_3(\mathbf{x}) \geq 0, \tag{39}$$

where

$$\begin{aligned} \mathcal{L}_0 &= \alpha\zeta\,(\mathbf{f} \cdot \nabla\zeta) + \zeta^2\,(\mathbf{f} \cdot \nabla P_d + \varphi - L)\,, \\ \mathcal{L}_1 &= \alpha\zeta\,\nabla \cdot (\mathbf{D}\nabla\zeta) - \alpha \nabla\zeta \cdot (\mathbf{D}\,\nabla\zeta) + \zeta^2\,\nabla \cdot (\mathbf{D}\nabla P_d) + \alpha \mathbf{f} \cdot \nabla\zeta + 2\zeta\,(\mathbf{f} \cdot \nabla P_d + \varphi - L)\,, \\ \mathcal{L}_2 &= \alpha \nabla \cdot (\mathbf{D}\nabla\zeta) + 2\zeta\,\nabla \cdot (\mathbf{D}\nabla P_d) + \mathbf{f} \cdot \nabla P_d + \varphi - L, \\ \mathcal{L}_3 &= \nabla \cdot (\mathbf{D}\nabla P_d). \end{aligned} \tag{40}$$

Consequently, a lower bound L on $\langle\varphi\rangle_\varepsilon$ can be calculated at a fixed, small ε by solving the optimisation problem

$$\begin{aligned} \max_{P_d, \alpha, L, \mathbf{Z}} \quad & L \\ \text{such that} \quad & \mathcal{L}(\mathbf{x}) \in \Sigma, \\ & \zeta(\mathbf{x}) = \mathbf{x}^T \mathbf{Z} \mathbf{x}, \quad \mathbf{Z} \succ 0. \end{aligned} \tag{41}$$

Note that when ε becomes small, the terms $\mathcal{L}_1$, $\mathcal{L}_2$ and $\mathcal{L}_3$ represent small perturbations of $\mathcal{L}_0$, which is an augmented version of the inequality constraint (25b) for the $\varepsilon = 0$ case. In particular, the term $\alpha\zeta\,(\mathbf{f} \cdot \nabla\zeta)$ represents the contribution of the large gradients and allows us to improve the bounds as $\varepsilon \to 0$.

In fact, we can further develop this idea and derive an optimisation problem which is *rigorous* in the limit $\varepsilon \to 0$, i.e. such that the computed L is a lower bound for $\lim_{\varepsilon \to 0} \langle\varphi\rangle_\varepsilon$. To show this, let us prove that in this limit (23) holds if $\mathcal{L}_0 \geq \gamma\zeta^2$ for any strictly positive

and arbitrarily small constant γ. In fact, since ζ is quadratic in $\mathbf{x}$, it can be verified that $\mathcal{L}_0$ is the dominant term in $\mathcal{L}$ for any fixed $\mathbf{x} \neq 0$. Moreover,

$$\lim_{\varepsilon \to 0} \varepsilon^n \int_{\|x\| \geq r} \rho(\mathbf{x}) \frac{\mathcal{L}_n(\mathbf{x})}{(\varepsilon + \zeta)^2} d\mathbf{x} = 0, \qquad n \in \{1, 2, 3\} \tag{42}$$

for any finite radius r since we have assumed that ρ decays faster than any polynomial. Consequently, if $\mathcal{L}_0 \geq \gamma \zeta^2$ then $\mathcal{L}(\mathbf{x})$ is positive when $\varepsilon \to 0$ at least outside a ball B_R of radius $R \sim \varepsilon^{1/2-\eta}$ with $0 < \eta < 1/2$. We conclude that

$$\int_{\mathbb{R}^n \setminus B_R} \rho \left[\varepsilon \nabla \cdot (\mathbf{D} \nabla V_l) + \mathbf{f} \cdot \nabla V_l + \varphi - L \right] d\mathbf{x} = \int_{\mathbb{R}^n \setminus B_R} \frac{\rho \mathcal{L}}{(\varepsilon + \zeta)^2} \, d\mathbf{x} \geq 0 \tag{43}$$

as $\varepsilon \to 0$. Moreover, although the integrand develops a singularity at $\mathbf{x} = 0$ when $\varepsilon \to 0$ it is possible to show (see Appendix C) that

$$\int_{B_R} \rho \left[\varepsilon \nabla \cdot (\mathbf{D} \nabla V_l) + \mathbf{f} \cdot \nabla V_l + \varphi - L \right] d\mathbf{x} \to 0 \text{ as } \varepsilon \to 0 \tag{44}$$

if ρ is bounded on B_R as $\varepsilon \to 0$ (a reasonable assumption since $\mathbf{x} = 0$ is an unstable equilibrium of the deterministic system). Therefore, a sufficient condition for L to be a valid lower bound on $\langle \varphi \rangle_\varepsilon$ in the limit of vanishing noise is that $\mathcal{L}_0 \geq \gamma \zeta^2$, i.e. (dropping a factor of ζ and rearranging)

$$\alpha \mathbf{f} \cdot \nabla \zeta + \zeta \left(\mathbf{f} \cdot \nabla P_d + \varphi \right) \geq (L + \gamma) \zeta. \tag{45}$$

Note that the role of γ is simply to decrease the tightest possible L by an arbitrarily small constant; consequently, we can drop it from the analysis and determine a rigorous bound L with the optimisation problem

$$\begin{aligned} &\max_{P_d, \alpha, L, \mathbf{Z}} \quad L \\ \text{such that} \quad &\alpha \mathbf{f} \cdot \nabla \zeta + \zeta \left(\mathbf{f} \cdot \nabla P_d + \varphi - L \right) \in \Sigma, \\ &\zeta(\mathbf{x}) = \mathbf{x}^T \mathbf{Z} \mathbf{x}, \quad \mathbf{Z} \succ 0. \end{aligned} \tag{46}$$

Finally, note that ζ is an unknown quadratic form, so the optimisation problem is bilinear. An optimal ζ (denoted by ζ^*) could be determined using a bilinear SDP solver. However, since the SOS constraint is homogeneous in ζ, any choice of $\zeta = \beta \zeta^*$ for $\beta > 0$ is optimal; this issue may be resolved by adding a constraint on the coefficients of ζ. A simpler solution is to fix ζ a priori according to the following observation. In a neighbourhood of the origin where $x, y \ll 1$, the SOS constraint becomes, to leading order in $\mathbf{x}$,

$$\alpha \tilde{\mathbf{f}} \cdot \nabla \zeta - L \zeta \geq 0 \tag{47}$$

where $\tilde{\mathbf{f}} = \mathbf{J}_0 \, \mathbf{x}$ denotes the linearised dynamics near the origin. Therefore, if L is to be positive we require

$$\alpha \tilde{\mathbf{f}} \cdot \nabla \zeta > 0, \tag{48}$$

i.e. that $\alpha \dot{\zeta}$ is positive near the unstable point.

A suitable ζ can be constructed if $\mathbf{J}_0$ can be diagonalised. Specifically, let $\mathbf{U}$ denote the matrix of eigenvectors of $\mathbf{J}_0$ and $\mathbf{\Lambda}$ be the usual diagonal matrix of eigenvalues, such that

$$\mathbf{U}^{-1}\mathbf{J}_0\mathbf{U} = \mathbf{\Lambda}, \tag{49}$$

and let $\mathbf{w} = \mathbf{U}^{-1}\mathbf{x}$. Then, an appropriate choice is

$$\zeta = \mathbf{w}^T\mathbf{w} = \mathbf{x}^T[\mathbf{U}^{-1}]^T\mathbf{U}^{-1}\mathbf{x}, \tag{50}$$

since near the origin we have

$$\begin{aligned}
\dot{\zeta} &= 2\dot{\mathbf{x}}^T[\mathbf{U}^{-1}]^T\mathbf{U}^{-1}\mathbf{x} \\
&= 2\mathbf{x}^T\mathbf{J}_0^T\,[\mathbf{U}^{-1}]^T\mathbf{U}^{-1}\mathbf{x} + \text{h.o.t.} \\
&= 2\mathbf{x}^T[\mathbf{U}\mathbf{U}^{-1}]^T\,\mathbf{J}_0^T\,[\mathbf{U}^{-1}]^T\mathbf{U}^{-1}\mathbf{x} + \text{h.o.t.} \\
&= 2\mathbf{w}^T[\mathbf{U}^{-1}\mathbf{J}_0\,\mathbf{U}]^T\mathbf{w} + \text{h.o.t.} \\
&= 2\mathbf{w}^T\mathbf{\Lambda}\mathbf{w} + \text{h.o.t.},
\end{aligned} \tag{51}$$

Neglecting the higher order terms near $\mathbf{x} = 0$ and recalling that $\mathbf{\Lambda}$ is positive definite since we are considering a repelling fixed point, we conclude that $\dot{\zeta} > 0$ and so (48) holds for $\alpha > 0$. Note, however, that this is not the only possible choice of ζ; see Appendix E for more examples.

4.3 Equivalence with the $\mathcal{S}$-procedure

Whilst the vanishing-noise formulation presented above and the $\mathcal{S}$-procedure of Section 3 have been obtained in completely separate ways, they are in fact related. To see this, consider the inequality

$$\alpha\,(\mathbf{f}\cdot\nabla\zeta) + \zeta\,(\mathbf{f}\cdot\nabla P_d + \varphi - L) \geq 0 \tag{52}$$

more carefully. The inequality is satisfied at $\mathbf{x} = 0$; for $\mathbf{x} \neq 0$, divide by ζ, rewrite the first term as a time-derivative and rearrange the terms to obtain

$$(\mathbf{f}\cdot\nabla P_d + \varphi - L) + \left(\frac{\alpha}{\zeta}\right)\dot{\zeta} \geq 0. \tag{53}$$

Having already noted that $\alpha > 0$ when ζ is as in (50), we recognise that this is a particular form of the more general $\mathcal{S}$-procedure

$$(\mathbf{f}\cdot\nabla P_d + \varphi - L) + s(\mathbf{x})\dot{\zeta} \geq 0, \tag{54}$$

i.e. we are imposing the inequality

$$\mathbf{f}\cdot\nabla P_d + \varphi - L \geq 0 \tag{55}$$

obtained from Proposition 2.1 *outside* a region $\mathcal{R}$ where $\dot{\zeta} > 0$. This represents a region of repulsion for the unstable fixed point, in which ζ acts as an "inverse" Lyapunov function. Consequently, (52) can be viewed as an application of Proposition 3.1 with $\mathcal{T} = \mathbb{R}^n \setminus \mathcal{R}$. Adding noise to a dynamical system with a repelling fixed point and using the logarithmic

ansatz (37) for the storage function V_l is therefore equivalent to carrying out an $\mathcal{S}$-procedure. Consequently, we expect that the larger the region of repulsion defined by the level sets of ζ, the better the lower bound L for a given degree of P_d in (37) — an observation that may assist the construction of a good ζ.

This equivalence could be expected since $\mathcal{L}_0$ in (39) does not inherit any noise-related terms from the full formulation. This is the result of the particular choice (37) of the form of V_l. If ε is taken as a small but finite value, however, the addition of noise is not equivalent to the $\mathcal{S}$-procedure. In this case, an alternative form of V_l has to be considered to keep the influence of noise when $\epsilon \to 0$.

5 Application to the Van der Pol Oscillator

Let us illustrate how the ideas presented so far can be applied in practice by considering the Van der Pol oscillator

$$\ddot{x} - \mu(1 - x^2)\dot{x} + x = 0, \tag{56}$$

or, in state-space representation,

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \quad \mathbf{x} := \begin{pmatrix} x \\ y \end{pmatrix}, \quad \mathbf{f}(\mathbf{x}) := \begin{pmatrix} y \\ \mu(1 - x^2)y - x \end{pmatrix}. \tag{57}$$

Here, $\mu > 0$ represents the strength of the nonlinear damping force. We are interested in finding upper and lower bounds for $\varphi = x^2 + \dot{x}^2 = x^2 + y^2 = \|\mathbf{x}\|^2$, a measure of the total (potential plus kinetic) energy in the system. As is well known, for any μ the equilibrium position $\mathbf{x} = 0$ is unstable, and the system settles into periodic oscillations for any initial perturbation (Figure 2).

5.1 Upper Bound for the deterministic oscillator

Upper bounds on $\langle\varphi\rangle$ were computed by solving the optimisation problem (9) for $0.1 \leq \mu \leq 5$ and for a range of polynomial degrees d. We used the SOS module of YALMIP [10] to transform (9) into a semidefinite program (SDP). Initial numerical experiments showed that the resulting SDP is ill-conditioned even for modest polynomial degrees, and cannot

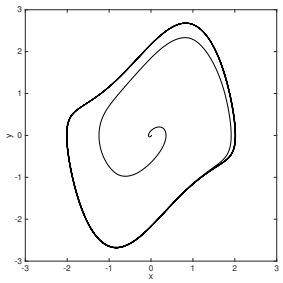

Figure 2: Sample state-space orbit of the Van der Pol oscillator for $\mu = 1$, starting near the unstable origin and converging to the periodic orbit.

be solved reliably by standard double-precision SDP solvers. We therefore resorted to the high-precision solver SDPA-GMP [8]; more details and comments on the numerical implementation can be found in Appendix D.

The upper bounds computed for $d = 6$, $d = 8$, $d = 10$ and $d = 12$ are shown in Figure 3, alongside the values of $\langle\varphi\rangle$ obtained by direct numerical integration of (56). As one would expect, the quality of the bound increases with d; the bounds are within approximately 5% of the simulated value for all values μ considered when $d = 10$, and almost sharp for $d = 12$. The contours of V_u, shown in Figure 4, suggest that better bounds are achieved when the storage function has negative peaks concentrated near the corners of the periodic orbit, where the system evolves slowly. This explains why higher polynomial degrees are necessary to achieve sharp bounds at large values μ, for which the periodic orbit becomes more elongated.

5.2 Lower bound for the deterministic oscillator

For the deterministic system, the trivial lower bound $\langle\varphi\rangle \geq 0$ cannot be improved using Proposition 2.1, since it is saturated by the equilibrium at the origin. In order to apply Proposition 3.1 and the $\mathcal{S}$-procedure to find a tight lower bound for trajectories attracted to the periodic orbit, we need to construct an absorbing domain that does not contain $\mathbf{x} = 0$. As already mentioned in Section 3, SOS optimisation can be used to construct absorbing domains that well approximate the periodic orbit [16]; however, this involves further complication and is beyond the scope of the present investigation. Instead, we will only consider the simple family of domains $\mathcal{T}_r = \{(x, y) \,|\, g(x, y) = x^2 + y^2 - r^2 \geq 0\}$ for $r \leq 1$. To show that these are indeed absorbing domains, let us reverse the direction of time in (57) and consider the energy $E = x^2 + y^2$. One has

$$\begin{aligned} \dot{E} &= 2x\dot{x} + \ y\dot{y} \\ &= -2xy - 2\mu y^2(1 - x^2) + 2xy \\ &= 2\mu y^2(x^2 - 1) \end{aligned} \tag{58}$$

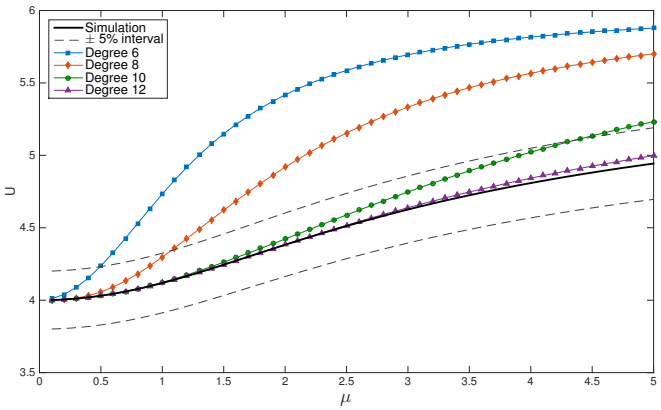

Figure 3: Upper bound for polynomial V_u of varying degree as a function of μ vs. the exact value of $\langle\varphi\rangle$ and the $\pm 5\%$ accuracy intervals.

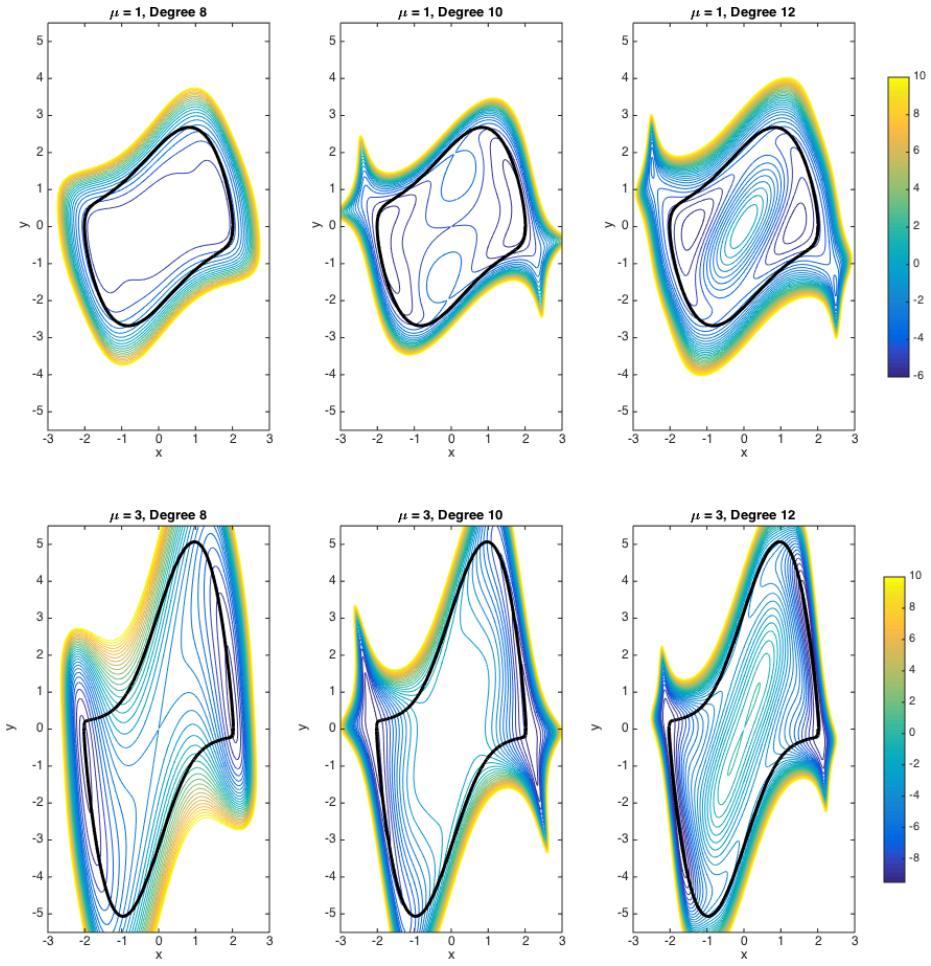

Figure 4: Contours of the optimal V_u for varying polynomial degree. Top: $\mu = 1$. Bottom: $\mu = 3$. The system's periodic orbit (thick black line) is also shown.

meaning that $E \leq 0$ when $|x| \leq 1$. Therefore, any contour of E which is contained in the strip $|x| \leq 1$ defines the boundary of a region of attraction of the origin for the time-reversed oscillator. One concludes that in the original system all orbits will leave the ball $x^2+y^2 < r^2$ if $r \leq 1$, i.e. $\mathcal{T}_r$ is an absorbing domain.

An immediate corollary of this simple proof is the lower bound $\langle\varphi\rangle \geq 1$. While this is already a step forward, more significant improvements and even sharp bounds on $\langle\varphi\rangle$ can be obtained by solving (17).

Initial numerical experiments revealed that it is sufficient to define V_l using monomials of even order only; this is a useful simplification, as it reduces the computational effort and improves the numerical conditioning of the SDP. In all cases, the degree of the $\mathcal{S}$-procedure multiplier s was fixed to be the same as V_l for simplicity.

Figure 5 illustrates the lower bounds computed for storage functions V_l of degree 8, 10 and 12 using the two different absorbing domains $\mathcal{T}_{0.5}$ and $\mathcal{T}_1$. Because $\mathcal{T}_1$ is a better

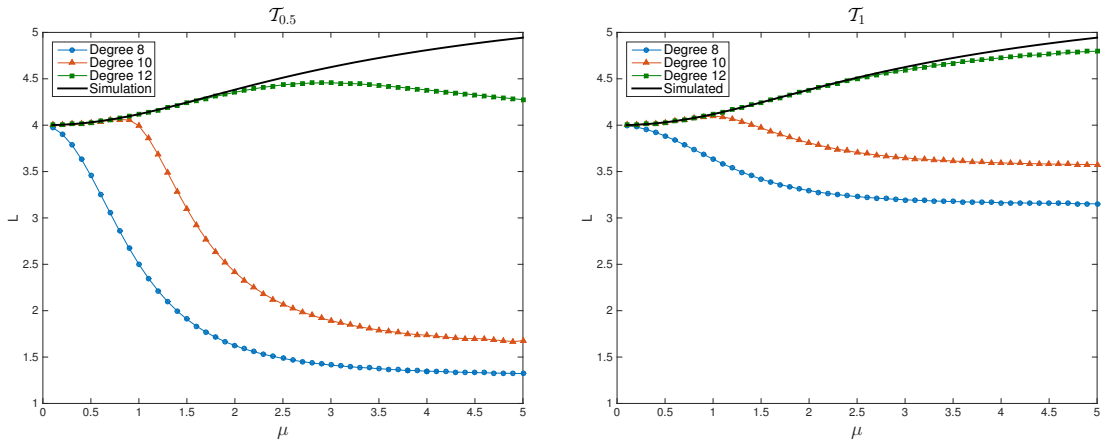

Figure 5: Lower bound computed with the $\mathcal{S}$-procedure for $d = 8, 10, 12$ using the absorbing domains $\mathcal{T}_{0.5}$ (left) $\mathcal{T}_1$ (right).

approximation to the periodic orbit than $\mathcal{T}_{0.5}$, it is not surprising that better bounds are obtained using the former for a given polynomial degree. For the same reason, the bounds worsen as μ increases since the periodic orbit becomes more elongated. We expect that if more sophisticated domain of attractions were computed, for example using the SOS techniques in [16], sharper bounds could be obtained for large μ. Alternatively, bounds of comparable accuracy could be obtained with a lower polynomial degree, reducing the cost of the optimisation.

5.3 Bounds for a stochastic Van der Pol oscillator

Let us now add a stochastic forcing term of strength $\sqrt{2\varepsilon}$ to the deterministic Van der Pol differential equation, i.e. consider

$$\ddot{x} - \mu(1 - x^2)\dot{x} + x = \sqrt{2\varepsilon}\,\xi. \tag{59}$$

where ξ denotes white noise. In state-space formulation, this becomes

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \sqrt{2\varepsilon}\begin{pmatrix} 0 \\ \xi \end{pmatrix} \tag{60}$$

Using the notation of Section 4, this corresponds to considering

$$\boldsymbol{\sigma} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{61}$$

First, let us try to compute upper and lower bounds assuming that V_u and V_l are polynomials. Figure 6 shows the results obtained after solving (27) and (28) for values of ε ranging from 10^{-6} to 1 and polynomial degrees 8, 10 and 12. All results were obtained for $\mu = 1$ and defining V_u and V_l using monomials of even order only. The Figure also shows some preliminary values for exact expectation, computed after solving the stationary

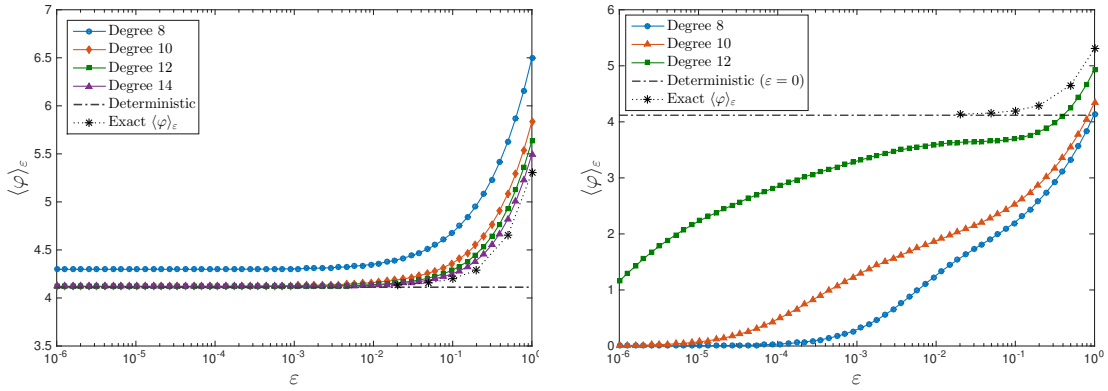

Figure 6: Bounds for the stochastic Van der Pol oscillator for fixed finite noise amplitude. The deterministic ($\varepsilon = 0$) value of $\langle\varphi\rangle$ is shown by the dashed line. Left: upper bound. Right: lower bound.

Fokker-Planck equation (19) using a low-order finite difference method; the steady state is achieved by time-stepping an initial distribution using an implicit Euler scheme, employing operator splitting between x and y derivatives (that is, half a time-step is taken ignoring any derivative in y, then another half time-step ignoring any derivative in x).

As far as the upper bounds are concerned, they are effectively indistinguishable from the value obtained for the deterministic oscillator in Section 5.1 for $\varepsilon < 10^{-3}$ approximately and for polynomial degrees larger than 8. This is not surprising, since (27) reduces to (9) as $\varepsilon \to 0$ when V_u is a polynomial. At larger ε, the bounds are reasonably accurate and capture the increase in $\langle\varphi\rangle_\varepsilon$. Such an increase is indeed consistent with the stronger effect the stochastic forcing has on the deterministic dynamics, suggesting that the well-defined, localised periodic orbit is "smeared" by the noise.

Regarding the lower bound, instead, the addition of noise is effective only when ε is relatively large; for all polynomial degrees, the bound L decreases to 0 as $\varepsilon \to 0$. This is consistent with the observation made in Section 4 that a polynomial V_l of fixed degree cannot have large enough gradients at $\mathbf{x} = 0$ to overcome the decrease in ε.

The numerical difficulties at small ε can be resolved if, instead of a polynomial, V_l is as in (37). For simplicity, rather than trying to determine an optimal ζ, we prescribed $\zeta = x^2 - xy + y^2$ so that the optimisation problem (41) is convex and can be solved using standard SDP solvers. It can be verified that this choice of ζ satisfies (48) for $\mu = 1$; more details can be found in Appendix E. Figure 7 shows the lower bounds computed using (41) for this choice of ζ, $\mu = 1$ and polynomials P_d of degree $d = 8$, $d = 10$ and $d = 12$. The improvement compared to the results obtained with a polynomial storage function are significant, and, for $d = 12$, L is indistinguishable from the deterministic bound when $\varepsilon < 10^{-3}$ approximately. Moreover, we expect that more accurate bounds could be obtained at large ε by increasing the degree of P_d.

Finally, we can compute lower bounds in the limit $\varepsilon \to 0$ by solving optimisation problem (46). We fixed the degree of P_d to 12, and considered different quadratic forms ζ, shown in Table 1. The quadratic forms ζ_2 and ζ_3 were constructed using the eigenvectors of the

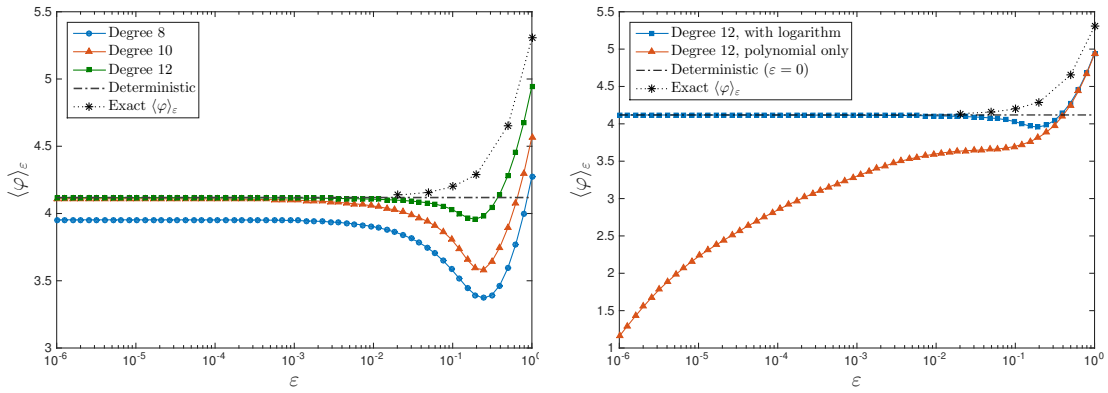

Figure 7: Lower bounds for the stochastic Van der Pol oscillator for fixed finite noise amplitude. Left: bounds for different degrees of P_d. Right: Comparison between the bounds computed from (28) and with (41) for $d = 12$.

Jacobian $\mathbf{J}_0$ (Appendix E); the different formulae for $\mu \leq 2$ and $\mu > 2$ are due to a change from complex to real eigenvectors. Instead, ζ_1 was obtained by arbitrarily fixing $\mu = 1$ in ζ_2. It can be verified that ζ_2 and ζ_3 satisfy (48) near the origin except when $\mu = 2$ (the "critical damping" condition), while ζ_1 satisfies (48) only for $4 - 2\sqrt{3} < \mu < 4 + 2\sqrt{3}$; see Appendix E for more details.

The lower bounds on $\langle\varphi\rangle$, computed as a function of μ, are shown in Figure 8. Overall,

Table 1: Choices of ζ for the Van der Pol oscillator.

	$\mu \leq 2$	$\mu > 2$
ζ_1	$x^2 - xy + y^2$	$x^2 - xy + y^2$
ζ_2	$x^2 - \mu xy + y^2$	$\mu x^2 - 4xy + \mu y^2$
ζ_3	$x^2 - \mu xy + y^2$	$(\mu^2 - 2)x^2 - 2\mu xy + 2y^2$

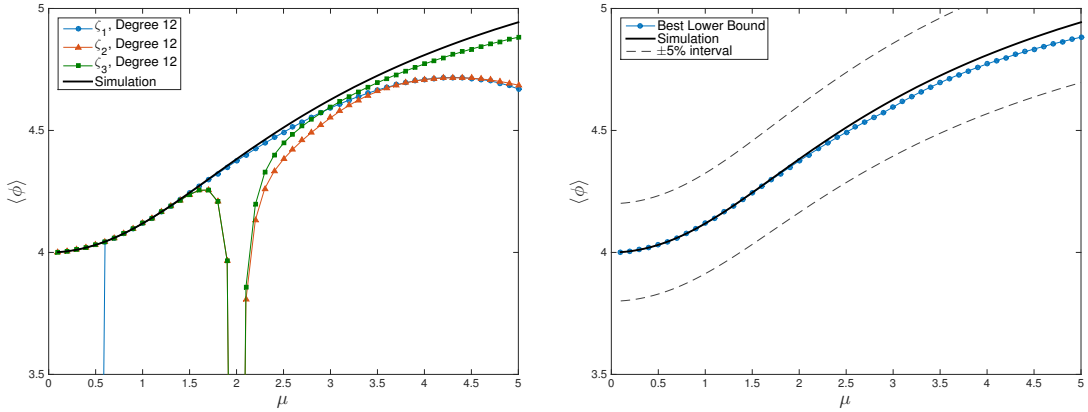

Figure 8: Left: lower bounds computed from (46) for the choices of ζ shown in Table 1 and $\deg(P_d) = 12$. Right: best lower bound compared to numerical integration of (57).

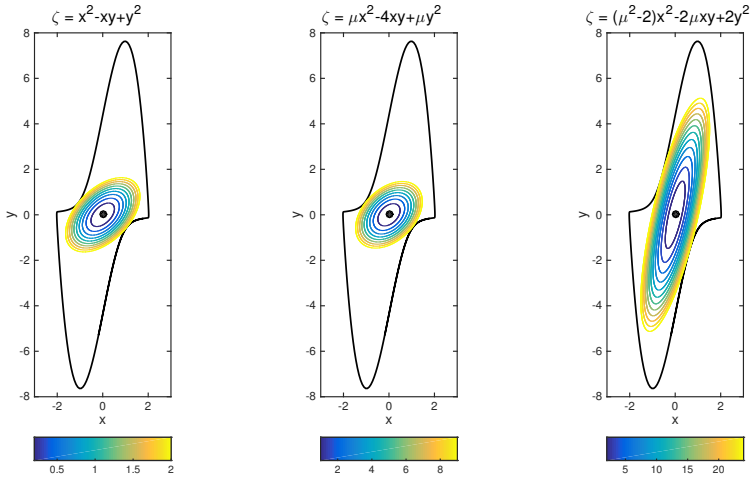

Figure 9: Contours of ζ_1 (left), ζ_2 (centre) and ζ_3 (right) for $\mu = 5$. The system's periodic orbit is also shown (thick black line).

our bounds are well within $\pm 5\%$ of the exact value of $\langle \varphi \rangle$; the poor performance of ζ_1 for $\mu \leq 4 - 2\sqrt{3} \approx 0.54$ and of ζ_2, ζ_3 near $\mu = 2$ was expected based on our previous comments. Moreover, ζ_3 significantly outperforms the other choices for $\mu > 3$ approximately. This can be understood by recalling the equivalence of the logarithmic ansatz and the $\mathcal{S}$-procedure: as shown in Figure 9, the contours of ζ_3 define a better region of attraction for the periodic orbit. Finally, the bounds worsen as μ increases for a fixed polynomial degree and for a fixed ζ, similarly to what was observed for the upper bounds of Section 5.1 and for the lower bounds obtained with the $\mathcal{S}$-procedure in Section 5.2. This could be resolved by increasing the degree of the polynomial P_n in (37) or by a more careful choice of ζ.

6 Further Comments

Although we have tried to keep our work as general as possible, we remark that the analysis of Section 4.2 is only appropriate to eliminate the influence of repelling fixed points; we have not considered the more common cases in which the bounds are constrained by saddle points or unstable limit cycles.

Unfortunately, many systems exhibiting interesting dynamics (such as the Lorenz system) possess unstable saddle points, and one cannot generally expect to successfully apply the techniques we have presented. For example, we expect that the logarithmic functional form we have proposed in Section 4 will not generally be suitable for systems with an unstable saddle point. The reason is that using the logarithmic ansatz is equivalent to implementing an $\mathcal{S}$-procedure, but a region of repulsion around a saddle point cannot generally be defined without including in it the entire stable manifold — normally, a convoluted set that cannot be easily approximated by polynomials. This was confirmed by a brief numerical investigation on the simple system

$$\begin{aligned} \dot{x} &= (x+y)(4-x^2-y^2) \\ \dot{y} &= y(2-x^2-y^2) \end{aligned} \tag{62}$$

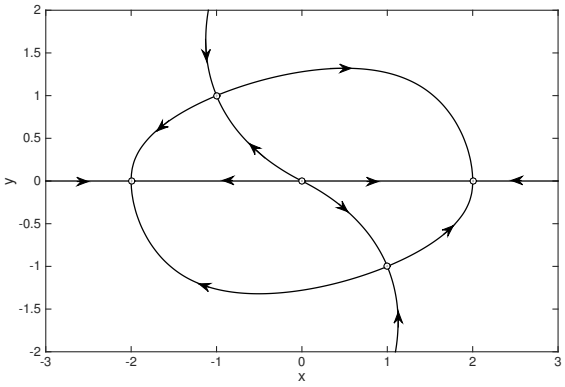

Figure 10: Fixed points of system (62) with their stable and unstable manifolds.

which has a repelling fixed point at the origin, two saddle points at $(\pm 1, \mp 1)$ and two stable equilibria at $(\pm 2, 0)$ as shown in Figure (10); tight bounds on $\varphi = y$ could not be obtained after adding noise to the system even with a logarithmic storage function.

Moreover, we expect that storage functions of impractically high degree will be needed even for noise of relatively large amplitude, making the SOS optimisation problem intractable. This is because the trajectories in the stable manifold will still try to approach the saddle point, forced by the deterministic component of the flow. Then, one expects that each of $\mathbf{f} \cdot \nabla V_u$ and $\mathbf{f} \cdot \nabla V_l$ in inequalities (25a) and (25b) will have opposite signs along the stable and unstable manifolds near the saddle point, unless ∇V_u and ∇V_l can change sign rapidly. This would require polynomial approximations of large degree.

Future work should therefore concentrate on determining an appropriate scaling and functional form for V_u and V_l for systems with saddle points and unstable periodic orbits.

Finally, we remark that the practical implementation of the SOS problems obtained throughout this work poses some technical challenges. Specifically, polynomials of high degree are required to compute relatively sharp bounds, significantly increasing the size of the SDP problems to be solved. This poses a limit on the dimension of the dynamical systems that one can study at a reasonable computational cost and time. Moreover, all SDP problems considered in this work were ill-conditioned, and the results we have presented could only be obtained using computationally expensive high-precision algorithms. Such numerical difficulties should be addressed more systematically in the future if, as it seems inevitable, the theoretical development of appropriate scaling arguments for the storage functions is to be reliably assisted by numerical investigations.

7 Conclusion

To summarise, we have demonstrated that bounds for long-time-averaged properties of systems with polynomial dynamics can be obtained by constructing suitable storage functions using SOS optimisation. Moreover, we have shown that the influence of unstable equilibria on the bounds can be removed via the $\mathcal{S}$-procedure (if a suitable absorbing domain can be defined), or, extending the ideas of [3], by adding noise to a system.

In particular, whilst the formulation of Section 4 holds for a general stochastic system with finite noise strength, a key development is the rigorous formulation of the optimisation problem in the vanishing noise limit for repelling fixed points. If the system is stochastically stable, rigorous bounds for a deterministic system can be inferred when Proposition 3.1 and/or the $\mathcal{S}$-procedure cannot be applied. In this context, we have demonstrated that simple polynomial storage functions are not appropriate to prove bounds that are insensitive to unstable solutions as $\varepsilon \to 0$, and a suitable asymptotic scaling of V_u and V_l with ε should be used.

Despite our successful application of the ideas we have presented to a simple example (the Van der Pol oscillator), some theoretical questions — such as whether it is possible to remove the influence of saddle points on the bounds — and some practical challenges in the implementation of the SOS optimisation remain unresolved. We anticipate that these issues will be the subject of future work, if rigorous bounds are to be obtained for physical systems of practical interest.

8 Acknowledgements

I would like to thank Dr. D. Goluskin for his thorough supervision and the constant support he provided. Moreover, I am grateful to Prof. G. Chini, Prof. S. I. Chernyshenko, Prof. C. R. Doering and Dr. A. Wynn for the numerous discussions and suggestions that helped to shape this work. Finally, I would like to thank all GFD fellows and staff for their patience and help throughout the duration of the programme.

A Introduction to SOS Optimisation

For simplicity, let us consider the problem of determining whether a polynomial $p(x)$ of degree $2N$, i.e.

$$p(x) = \sum_{n=0}^{2N} c_n \, x^n \tag{63}$$

is positive for any $x \in \mathbb{R}$ (the same argument can be generalised to multiple dimensions; for more details, see [13, 11, 1] and references therein). Clearly, a sufficient condition is that p admits a SOS decomposition, i.e. there exists a family of polynomials $\{g_i(x)\}_{i=0}^{M}$ such that

$$p(x) = g_0(x)^2 + g_1(x)^2 + \ldots + g_M(x)^2. \tag{64}$$

It can be shown that this is equivalent to the existence of a positive definite matrix $\mathbf{Q}$ (written as $\mathbf{Q} \succeq 0$) and a vector $\mathbf{z}(x)$ of monomials of x such that

$$p(x) = \mathbf{z}(x)^T \, \mathbf{Q} \, \mathbf{z}(x). \tag{65}$$

For example, if $p(x)$ has degree $2N$ one can take $\mathbf{z}(x)^T = (1, x, x^2, ..., x^N)$. Note that the matrix $\mathbf{Q}$ is generally not unique.

The problem of whether $p(x)$ admits a SOS decomposition can therefore be rewritten as the feasibility semidefinite problem

$$\begin{aligned} & \text{find} \quad \mathbf{Q} \\ \text{such that} \quad & p(x) - \mathbf{z}(x)^T \, \mathbf{Q} \, \mathbf{z}(x) = 0, \\ & \mathbf{Q} \succeq 0, \end{aligned} \tag{66}$$

where the equality constraint $p(x) - \mathbf{z}(x)^T \, \mathbf{Q} \, \mathbf{z}(x) = 0$ is interpreted as a set of equality constraints obtained by setting all coefficients of the difference $p(x) - \mathbf{z}(x)^T \, \mathbf{Q} \, \mathbf{z}(x)$ to zero. Note that these equality constraints are linear with respect to the coefficients c_n of $p(x)$, as well as with respect to the entries of $\mathbf{Q}$. Consequently, semidefinite programming can be used to find values of any unknown coefficients c_n such that $p(x)$ admits a SOS decomposition, or that minimise a linear function $f(c_0, c_1, \ldots, c_N)$ subject to $p(x)$ being a SOS polynomial. Again, more details can be found in [13, 11, 1].

B Asymptotic Analysis for 1D systems

Let us consider the one-dimensional dynamical system $\dot{x} = f(x) + \sqrt{2\varepsilon}\xi$, where $f(x)$ is a polynomial and ξ is a white noise process. Let $\rho_\varepsilon(x)$ be the stationary probability density function of the system, satisfying

$$\frac{\partial(f\rho_\varepsilon)}{\partial x} = \varepsilon \frac{\partial^2 \rho_\varepsilon}{\partial x^2}. \tag{67}$$

Integrating once we obtain

$$f\,\rho_\varepsilon = \varepsilon \frac{\partial \rho_\varepsilon}{\partial x}, \tag{68}$$

where the constant of integration has been set to zero since we must have $\rho \to 0$ as $|x| \to \infty$. This equation can be solved after letting

$$F(x) = \int f(x)\,dx \tag{69}$$

to find

$$\rho_\varepsilon(x) = \mathcal{N}\, e^{\frac{1}{\varepsilon}F(x)}, \qquad \mathcal{N} = \left[\int_{-\infty}^{+\infty} e^{\frac{1}{\varepsilon}F(x)}\,dx\right]^{-1}. \tag{70}$$

Consequently, the expectation of an observable φ can be computed as

$$\langle\varphi\rangle_\varepsilon = \int_{-\infty}^{+\infty} \varphi(x)\,\rho_\varepsilon(x)\,dx = \mathcal{N}\int_{-\infty}^{+\infty} \varphi(x)\, e^{\frac{1}{\varepsilon}F(x)}\,dx. \tag{71}$$

However, we are interested in computing $\langle\varphi\rangle_\varepsilon$ within the framework of Proposition 4.1, in the hope that we can gain some insight to tackle more complicated cases for which the Fokker-Planck equation cannot be solved as easily.

According to Proposition 4.1, $\langle\varphi\rangle_\varepsilon$ can be calculated by finding a function V what satisfies

$$\begin{aligned} \varepsilon\, V'' + f\, V' + \varphi - L_\varepsilon &= 0, \\ \lim_{|x|\to\infty} \left(\rho_\varepsilon\, V'\right) &= 0, \end{aligned} \tag{72}$$

where $(\cdot)'$ denotes differentiation with respect to x and the boundary term has been simplified using (68). As we will see in the following, a solution to this problem can only be found when $L_\varepsilon = \langle\varphi\rangle_\varepsilon$. Changing variable to $W = V'$, we can write an exact general solution to (72) for any value of ε as

$$W(x) = W_0 e^{-\frac{1}{\varepsilon}F(x)} + \frac{1}{\varepsilon}\, e^{-\frac{1}{\varepsilon}F(x)} \int_0^x \left[L_\varepsilon - \varphi(s)\right]\, e^{\frac{1}{\varepsilon}F(s)}\,ds. \tag{73}$$

The integration constant W_0 is determined by the boundary conditions, which using the expression for ρ_ε in (70) reduce to

$$W_0 + \frac{1}{\varepsilon}\int_0^{+\infty} \left[L_\varepsilon - \varphi(s)\right]\, e^{\frac{1}{\varepsilon}F(s)}\,ds = 0, \tag{74a}$$

$$W_0 - \frac{1}{\varepsilon}\int_{-\infty}^{0} \left[L_\varepsilon - \varphi(s)\right]\, e^{\frac{1}{\varepsilon}F(s)}\,ds = 0. \tag{74b}$$

Note that we have two boundary conditions for a first-order differential equation. In order to satisfy both, we let

$$L_\varepsilon = \langle \varphi \rangle_\varepsilon \tag{75}$$

such that, by definition of $\langle \varphi \rangle_\varepsilon$,

$$\int_{-\infty}^{+\infty} \left[L_\varepsilon - \varphi(x) \right] \, e^{\frac{1}{\varepsilon} F(x)} \, dx = 0. \tag{76}$$

Except from Section B.1 below, we will use (74a) and let

$$W_0 = \frac{1}{\varepsilon} \int_{-\infty}^{0} \left[L_\varepsilon - \varphi(s) \right] \, e^{\frac{1}{\varepsilon} F(s)} \, ds \tag{77}$$

so that W becomes

$$W(x) = \frac{1}{\varepsilon} \, e^{-\frac{1}{\varepsilon} F(x)} \int_{-\infty}^{x} \left(L_\varepsilon - s^2 \right) \, e^{\frac{1}{\varepsilon} F(s)} \, ds. \tag{78}$$

Note that the second boundary condition (74b) is also satisfied by virtue of (76). Alternatively, using (74b) to define W_0 one obtains the equivalent expression

$$W(x) = -\frac{1}{\varepsilon} \, e^{-\frac{1}{\varepsilon} F(x)} \int_{x}^{+\infty} \left(L_\varepsilon - s^2 \right) \, e^{\frac{1}{\varepsilon} F(s)} \, ds. \tag{79}$$

Unfortunately, neither (78) nor (79) give much information about the behaviour and scaling of W for a general dynamical system. We will therefore proceed by discussing some illustrative examples which allow us to draw some important conclusions about the applicability of SOS optimisation to determine sharp bounds for $\langle \varphi \rangle_\varepsilon$.

B.1 Case 1: $f(x) = x - x^3$, $\varphi(x) = x^2$

Let $f(x) = x - x^3$ and $\varphi(x) = x^2$ so that $F(x) = \frac{x^2}{2} - \frac{x^4}{4}$ is an even function with maxima at $x = \pm 1$ and a local minimum at $x = 0$, as shown in Figure 11. In this case, equation (72) can be solved using the method of matched asymptotic expansions (not show here) to find

$$W(x) \approx \underbrace{\frac{1}{x}}_{\text{outer sol.}} + \underbrace{\frac{1}{\sqrt{\varepsilon}} e^{-\frac{x^2}{2\varepsilon}} \int_0^{\frac{x}{\sqrt{\varepsilon}}} e^{\frac{s^2}{2}} \, ds}_{\text{inner sol.}} - \underbrace{\frac{1}{x}}_{\text{common part}}, \tag{80}$$

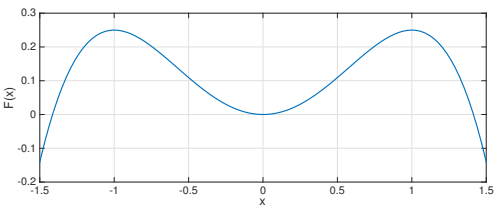

Figure 11: $F(x) = \frac{x^2}{2} - \frac{x^4}{4}$

where the inner solution is valid when $x \sim \sqrt{\varepsilon}\hat{x}$. In the intermediate layer $\sqrt{\varepsilon} \ll x \ll 1$ the solution reduces to the common part and one has

$$W(x) \sim \frac{1}{x} \Rightarrow V(x) = \int W(x)dx \sim \log|x| = \frac{1}{2}\log x^2. \tag{81}$$

This fact was used in Section 4 to justify the use of logarithmic ansatz for V in higher-dimensional systems.

The behaviour of W in the outer and inner regions could also be inferred from the expression for the exact solution. Rather than using (78) or (79), we note that φ is even and add the two conditions in (74) to deduce that $W_0 = 0$ is an appropriate choice for the integration constant in (73). Hence, we have

$$W(x) = \frac{1}{\varepsilon}\, e^{-\frac{1}{\varepsilon}F(x)} \int_0^x \left(L_\varepsilon - s^2\right)\, e^{\frac{1}{\varepsilon}F(s)}\, ds. \tag{82}$$

Moreover, subtracting (74b) with $W_0 = 0$ from this equation we find

$$W(x) = -\frac{1}{\varepsilon}\, e^{-\frac{1}{\varepsilon}F(x)} \int_x^{+\infty} \left(L_\varepsilon - s^2\right)\, e^{\frac{1}{\varepsilon}F(s)}\, ds, \tag{83}$$

which is the same as (79).

To study the asymptotic behaviour of W as $\varepsilon \to 0$, we note that W is an odd function so we restrict the attention to $x > 0$. First, we use Laplace's method to estimate

$$L_\varepsilon = \int_{-\infty}^{+\infty} x^2\, \rho_\varepsilon\, dx \sim 1 - \varepsilon + \mathcal{O}(\varepsilon^2). \tag{84}$$

When x is small, precisely $x = \varepsilon^{1/2}\hat{x}$ (where $\hat{x} \sim \mathcal{O}(1)$ and the scaling of ε is suggested by the term $\frac{1}{\varepsilon}F(x) \sim \frac{x^2}{2\varepsilon}$ when x is small) the leading order behaviour of L_ε can be used to estimate

$$W(x) \sim \frac{1}{\sqrt{\varepsilon}}\, e^{-\frac{x^2}{2\varepsilon}} \int_0^{\frac{x}{\sqrt{\varepsilon}}} e^{\frac{s^2}{2}}\, ds. \tag{85}$$

This is the same as the inner solution in (80).

When $x \sim \mathcal{O}(1)$ but $x < 1$, we can use Laplace's method to estimate the integral term in W, where the dominant contribution is given by the end-point x of the integration domain, since F is monotonically increasing over the interval $(0, 1)$ (cf. Figure 11). Recalling that $F' = f > 0$ for $x < 1$, we can show that

$$W(x) \sim \frac{1 - x^2}{x - x^3} = \frac{1}{x}, \tag{86}$$

which corresponds to the outer solution in (80). The same behaviour is found for $x > 1$ using (83).

Finally, we can estimate the behaviour at $x = 1$ to the leading order in ε as

$$\begin{aligned} W(1) &\sim \frac{1}{\varepsilon} e^{-\frac{1}{4\varepsilon}} \int_{-\infty}^{1} \left[1 - \varepsilon - 1 - 2(s-1)\right]\, e^{\left[\frac{1}{4\varepsilon} - \frac{1}{\varepsilon}(s-1)^2 + \ldots\right]}\, ds \\ &\sim 1 + \mathcal{O}(\sqrt{\varepsilon}), \end{aligned} \tag{87}$$

which is consistent with (86). A comparison between a direct numerical integration of (82) and its asymptotic expansion for $\varepsilon = 0.01$ is shown in Figure 12.

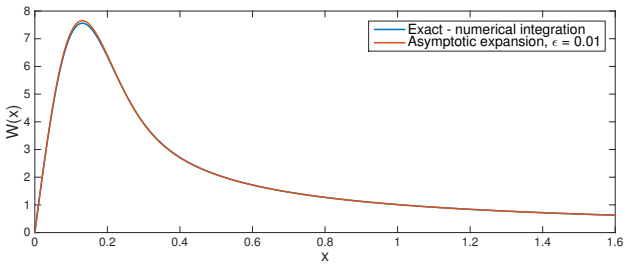

Figure 12: Comparison between the numerical integration and the asymptotic approximation of W, $\varepsilon = 0.01$.

B.2 Case 2: $f(x) = x - x^3$, $\varphi(x) = x$

A simple change from $\varphi = x^2$ to $\varphi = x$ implies a dramatic change in the behaviour of W. The functions $F(x)$ and $\rho_\varepsilon(x)$ are as in the previous example, and the symmetry of ρ_ε implies that

$$L_\varepsilon = \int_{-\infty}^{+\infty} x\, \rho_\varepsilon(x) dx = 0. \tag{88}$$

Equation (72) then becomes

$$\varepsilon W'(x) + f(x)W(x) + x = 0. \tag{89}$$

The method of matched asymptotic expansions fails in this case to produce an approximation to W if one assumes the usual outer solution

$$W_{\text{outer}} = \frac{L_\varepsilon - x}{x - x^3} = \frac{1}{x^2 - 1}, \tag{90}$$

since one cannot construct inner solutions at $x = \pm 1$ that satisfy the matching condition. Let us show this by trying to construct an inner solution at $x = 1$. The appropriate scaling for the inner variables is $x = 1 + \varepsilon^{1/2}\, \hat{y}$ and $W = \varepsilon^{-1/2}\hat{W}$, so the leading-order equation for $\hat{W}$ becomes

$$\hat{W}' - 2\hat{y}\hat{W} + 1 = 0, \quad \hat{y} = \frac{x-1}{\sqrt{\varepsilon}}. \tag{91}$$

Thus, we have

$$W = \frac{1}{\sqrt{\varepsilon}}\hat{W} = \frac{\sqrt{\pi}}{2\sqrt{\varepsilon}}\, e^{\hat{y}^2}\left[A - \operatorname{erf}(\hat{y})\right], \quad \hat{y} = \frac{x-1}{\sqrt{\varepsilon}}. \tag{92}$$

where erf is the standard error function and A is a constant of integration to be determined so as to match the assumed outer solution (90). Specifically, shifting coordinates $x = 1 + \sqrt{\varepsilon}\hat{y}$ in (90), we require

$$\lim_{\hat{y}\to+\infty} \frac{\sqrt{\pi}}{2\sqrt{\varepsilon}}\, e^{\hat{y}^2}\left[A - \operatorname{erf}(\hat{y})\right] \sim \frac{1}{2\sqrt{\varepsilon}\hat{y}}, \tag{93a}$$

$$\lim_{\hat{y}\to-\infty} \frac{\sqrt{\pi}}{2\sqrt{\varepsilon}}\, e^{\hat{y}^2}\left[A - \operatorname{erf}(\hat{y})\right] \sim \frac{1}{2\sqrt{\varepsilon}\hat{y}}. \tag{93b}$$

These conditions cannot be satisfied simultaneously, since the first one requires $A = 1$, while the second requires $A = -1$. Hence, matching is impossible.

This can be explained by considering the exact solution W, which according (78) and (79) can be written as

$$W(x) = \frac{1}{\varepsilon}\, e^{-\frac{1}{\varepsilon}F(x)} \int_{-\infty}^{x} s\, e^{\frac{1}{\varepsilon}F(s)}\, ds \tag{94a}$$

$$= -\frac{1}{\varepsilon}\, e^{-\frac{1}{\varepsilon}F(x)} \int_{x}^{+\infty} s\, e^{\frac{1}{\varepsilon}F(s)}\, ds. \tag{94b}$$

Note that this solution is even, so we only need to study its behaviour for $x > 0$.

When $x > 1$, W can be estimated from (94b) using Laplace's method, where the dominant contribution to the integral come from the end-point x of the domain of integration. We obtain

$$W(x) \sim \frac{1}{x^2 - 1}, \tag{95}$$

which corresponds to the usual outer solution (90). This, however, is not the appropriate outer solution when $0 < x < 1$; to show this, we again use Laplace's method on (94b), but this time the dominant contribution comes from $x = 1$. We obtain

$$W(x) \sim \sqrt{\frac{\pi}{\varepsilon}}\, e^{\frac{1}{4\varepsilon} - \frac{1}{\varepsilon}F(x)}, \tag{96}$$

and since $F(x) < \frac{1}{4}$ when $0 < x < 1$ (cf. Figure 11), this means that W behaves like a gaussian (cf. Figure 13).

This behaviour could have been derived from an asymptotic analysis of the original equation by applying the WKB method. If we assume that

$$W(x) = e^{\frac{1}{\varepsilon}\psi(x)} \left[W_0(x) + \varepsilon W_1(x) + ...\right], \tag{97}$$

equation (89) becomes

$$\left[\psi' W_0 + \varepsilon\psi' W_1 + \varepsilon W_0' + f\, W_0 + \varepsilon f\, W_1 + \mathcal{O}(\varepsilon^2)\right] e^{\frac{1}{\varepsilon}\psi} + x = 0. \tag{98}$$

Assuming that $\psi(x) \geq 0$, we can neglect the last term and impose

$$\mathcal{O}(\varepsilon^0): \qquad \psi' W_0 + f\, W_0 = 0 \tag{99}$$

$$\mathcal{O}(\varepsilon^1): \qquad \psi' W_1 + W_0' + f\, W_1 = 0 \tag{100}$$

In order to have a non-zero W_0, we must impose

$$\psi'(x) = -f(x) \implies \psi(x) = K - F(x). \tag{101}$$

Since $-F$ is bounded from below (cf. Figure 11), the integration constant K can indeed be chosen to satisfy $\psi \geq 0$ as originally assumed. It then follows from that (100) that W_0 =constant, and so

$$W(x) = W_0 e^{\frac{K}{\varepsilon}} e^{-\frac{F(x)}{e}} + \dots. \tag{102}$$

Incorporating the constant term $e^{\frac{K}{\varepsilon}}$ into W_0, the WKB outer solution for $0 < x < 1$ can be written to leading order as

$$W_{\text{outer}}(x) = W_0 e^{-\frac{1}{\varepsilon}F(x)}, \tag{103}$$

where W_0 has yet to be determined; note that we recover (96) if we choose

$$W_0 = \sqrt{\frac{\pi}{\varepsilon}}\, e^{\frac{1}{4\varepsilon}}. \tag{104}$$

This choice can indeed be motivated by matching the outer solution

$$W_{\text{outer}} = \begin{cases} W_0 e^{-\frac{1}{\varepsilon}F(x)}, & 0 < x < 1, \\ \dfrac{1}{x^2 - 1}, & x > 1. \end{cases} \tag{105}$$

with the inner solution near $x = 1$ given by (92). In particular, we choose $A = 1$ so that

$$W_{\text{inner}}(x) = \frac{\sqrt{\pi}}{2\sqrt{\varepsilon}}\, e^{\frac{(x-1)^2}{\varepsilon}} \left[1 - \operatorname{erf}\left(\frac{x-1}{\sqrt{\varepsilon}}\right)\right]. \tag{106}$$

It is easy to verify that the inner and outer solution match for $x > 1$, while matching is achieved for $x < 1$ if W_0 is as in (104). A comparison between the exact solution and its composite asymptotic expansion derived combining (105) and (106) is shown in Figure 13 for $\varepsilon = 0.02$ (the two curves are graphically indistinguishable).

B.3 General Third-Order Systems

From the examples above, we may conclude that an asymptotic solution of the equation

$$\varepsilon W'(x) + f(x)W(x) + \varphi(x) - L_\varepsilon = 0 \tag{107}$$

can only be achieved if one consider a more general outer solution than the "normal" outer solution

$$W_{\text{outer}} = \frac{L_\varepsilon - \varphi(x)}{f(x)}. \tag{108}$$

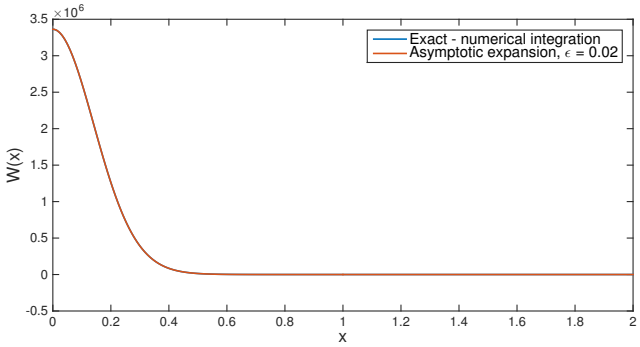

Figure 13: Comparison between the numerical integration and the asymptotic approximation of W, $\varepsilon = 0.02$ (the two curves are graphically indistinguishable).

To illustrate this concept, let us assume that f is a cubic polynomial with zeros at $x_{s,1} < x_u < x_{s,2}$ such that $x_{s,1}$ is the most stable fixed point of the deterministic system $\dot{x} = f(x)$. This means that F has a global maximum at $x_{s,1}$, a local minimum at x_u and a local maximum at $x_{s,2}$ (cf. Figure 14). In general, φ assumes different values at the fixed points and therefore one finds $L_\varepsilon = \varphi(x_{s,1}) + \mathcal{O}(\varepsilon)$. For example, one may consider $f(x) = 2x - x^2 - x^3$, $x_{s,1} = -2$ and $\varphi(x) = x^2$, in which case $L_\varepsilon = 4 + \mathcal{O}(\varepsilon)$.

Recall from (78) and (79) that if $L_\varepsilon = \langle \varphi \rangle_\varepsilon$ the solution of (107) can be written as

$$\begin{aligned} W(x) = & \quad \frac{1}{\varepsilon}\, e^{-\frac{1}{\varepsilon}F(x)} \int_{-\infty}^{x} [L_\varepsilon - \varphi(s)]\; e^{\frac{1}{\varepsilon}F(s)}\, ds \\ = & \quad -\frac{1}{\varepsilon}\, e^{-\frac{1}{\varepsilon}F(x)} \int_{x}^{+\infty} [L_\varepsilon - \varphi(s)]\; e^{\frac{1}{\varepsilon}F(s)}\, ds. \end{aligned} \tag{109}$$

Let us now study the behaviour of W for a different values of x. When $x \gg x_{s,2}$ an asymptotic analysis using Laplace's method shows that the "normal" outer solution (108) is an appropriate approximation to W (this is similar to our discussion in Section B.2). This approximation is valid as x decreases towards $x_{s,2}$.

Since F has a local maximum at $x_{s,2}$, when x is decreased past $x_{s,2}$ the behaviour of W changes; Laplace's method shows that

$$W(x) \sim \sqrt{\frac{2\pi}{|F''(x_{s,2})|\varepsilon}}\, [L_\varepsilon - \varphi(x_{s,2})]\; e^{\frac{1}{\varepsilon}[F(x_{s,2}) - F(x)]}, \tag{110}$$

i.e. W has an exponential behaviour. In particular, W reaches a maximum at x_u, when the difference $F(x_{s,2}) - F(x)$ is at its maximum (cf. Figure 14).

As x is decreased even further, one reaches a point x_0 at which $F(x_0) = F(x_{s,2})$. When $x < x_0$, the asymptotic behaviour changes again; one can show that, for $x < x_0$, W scales as in (108). This behaviour is maintained until $x = x_{s,1}$. Finally, under our assumption that $L_\varepsilon = \varphi(x_{s,1}) + \mathcal{O}(\varepsilon)$ it is can be shown that (108) holds for $x < x_{s,1}$ and, in fact, for $x = x_{s,1}$ (the analysis is analogous to that of Section B.1).

This analysis allows us to conclude that an asymptotic analysis of (107) should consider (108) as the outer solution for $x < x_0$ and $x > x_{s,2}$, while (110) should hold when

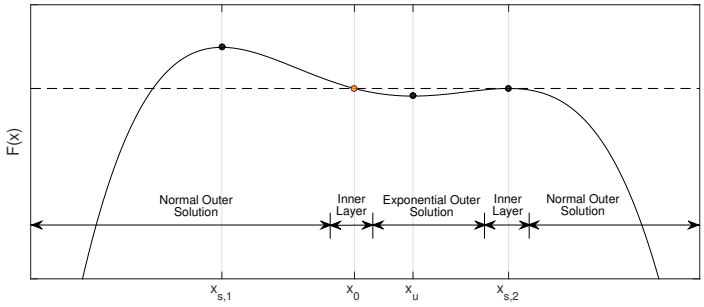

Figure 14: Validity region for different types of outer solutions for W. A typical profile for F satisfying our assumptions is shown. The location of inner layers where the behaviour transition smoothly is also sketched.

$x_0 < x < x_{s,2}$. Inner layers are required when $x \to x_0$ and $x \to x_{s,2}$ for a smooth transition between the different behaviours. The validity regions for each type of outer solutions are illustrated in Figure 14; the location of inner layers, where the behaviour smoothly transitions from one outer solution type to another, is also sketched.

The analysis for the inner layer near $x_{s,2}$ is identical to that carried out in Section B.2 for $x = x_{s,2} = 1$, so we will only consider the inner solution near $x \to x_0$. Letting $y = x - x_0$ and recalling our assumption that f is cubic, we may rewrite

$$f(x) = f(y + x_0) = -(a_0 + b_0 y + c_0 y^2 + d_0 y^3) \tag{111}$$

where the constants $a_0, ..., d_0$ depend on the value of x_0. In particular, $a_0 = -f(x_0) = -F'(x_0)$, and since we have assumed that $\mathbf{x}_{s,1} < x_0 < \mathbf{x}_u$, we conclude that $a_0 > 0$ (this can be seen from Figure 14). Equation (107) then becomes

$$\varepsilon W'(y) - \left(a_0 + b_0 y + c_0 y^2 + d_0 y^3\right) W(y) + \varphi(y + x_0) - L_\varepsilon = 0, \tag{112}$$

The appropriate coordinate stretching for this equation is $y = \varepsilon \hat{y}$, $W_{\text{inner}} = \hat{W}$; then, the inner solution satisfies the equation

$$\varepsilon W'_{\text{inner}} - a_0 W_{\text{inner}} + \varphi(x_0) - L_\varepsilon = 0 \tag{113}$$

and can be written explicitly as

$$W_{\text{inner}}(\hat{y}) = A\, e^{a_0 \hat{y}} - \frac{L_\varepsilon - \varphi(x_0)}{a_0}. \tag{114}$$

Recalling that $a_0 > 0$, it can be verified that this expression matches with (108) as $\hat{y} \to -\infty$ and with (110) as $\hat{y} \to +\infty$ if

$$A = \sqrt{\frac{2\pi}{|F''(x_{s,2})|\varepsilon}}\, [L_\varepsilon - \varphi(x_{s,2})]\; e^{-\frac{1}{\varepsilon}F(x_{s,2})}, \tag{115}$$

which is the required asymptotic behaviour (as illustrated in Figure 14).

B.4 Remarks & Implication for SOS Optimisation

In light of the examples discussed in the previous sections, we conclude that the appropriate form for the function W is highly dependent on both the system's dynamics $f(x)$ and the observable $\varphi(x)$. In general, there exist intervals in which W has an exponential growth of type $e^{\frac{1}{\varepsilon}(\cdot)}$. This behaviour is due to the existence of a stable fixed point x_s at which $L_\varepsilon - \varphi(x_s) \neq 0$ (to leading order in ε) such that standard outer solution (108) cannot be matched asymptotically to any appropriate inner solution. These considerations can be generalised to system with polynomial flows $f(x)$ of degree larger than 3.

The regions where W behaves exponentially, however, disappear when φ assumes the same value at *all* stable points (such as the example studied in Section B.1). In this special case, in fact, the "normal" outer solution does not become singular at any of the stable points, and no inner layers are required. An asymptotic solution similar to that proposed in Section B.1 can then be constructed.

For 1D system with bounded trajectories, one always has multiple stable points, hence these regions of exponential behaviour are generally unavoidable. Thus, one cannot usually approximate W with polynomials or rational functions that capture the correct scaling in the inner layers for arbitrary ε, as the degree of such polynomial approximations would have to be infinite.

The implication of these results is that, in general, one cannot hope to derive sharp lower bounds on $\langle\varphi\rangle_\varepsilon$ using SOS techniques that hold analytically as $\varepsilon \to 0$. Indeed, the computational effort and the order of polynomial approximations to the exact W increase so rapidly that SOS optimisation becomes impractical even when ε is fixed to a small value and not treated analytically.

C Vanishing Noise Limit: Proof of Negligible Contributions

In Section 4.2 we have let

$$V_l(\mathbf{x}) = \alpha \log[\varepsilon + \zeta(\mathbf{x})] + P_n(\mathbf{x})$$

and we have found a polynomial P_n such that

$$\int \rho \left[\varepsilon \nabla \cdot (\mathbf{D}\nabla V_l) + \mathbf{f} \cdot \nabla V_l + \varphi - L\right] d\mathbf{x} \geq 0$$

everywhere except from a ball B_R of radius $R \sim \varepsilon^{1/2-\eta}$ with $0 < \eta < 1/2$; for definiteness, let us write $R = C\varepsilon^{1/2-\eta}$ for some constant C. To complete the argument, we need to show that

$$\int_{B_R} \rho \left[\varepsilon \nabla \cdot (\mathbf{D}\nabla V_l) + \mathbf{f} \cdot \nabla V_l + \varphi - L\right] d\mathbf{x} \to 0 \text{ as } \varepsilon \to 0. \tag{116}$$

Upon substitution of the ansatz, the integral becomes

$$\int_{B_R} \rho \left[\varepsilon \frac{\alpha \nabla \cdot (\mathbf{D}\nabla\zeta)}{\varepsilon + \zeta} - \varepsilon \frac{\alpha \nabla\zeta \cdot (\mathbf{D}\nabla\zeta)}{(\varepsilon+\zeta)^2} + \varepsilon \nabla \cdot (\mathbf{D}\nabla P_n) + \frac{\alpha \mathbf{f} \cdot \nabla\zeta}{\varepsilon + \zeta} + \mathbf{f} \cdot \nabla P_n + \varphi - L\right] d\mathbf{x}.$$

Let us proceed term by term and let us assume that ρ is bounded in B_R uniformly as $\varepsilon \to 0$. Since P and φ are continuous,

$$\left| \int_{B_R} \rho \left[\varepsilon \nabla \cdot (\mathbf{D}\nabla P_n) + \mathbf{f} \cdot \nabla P_n + \varphi - L\right] d\mathbf{x} \right| \leq \frac{4\pi}{3} R^3 \max_{B_R} \left\{\rho |\varepsilon \nabla \cdot (\mathbf{D}\nabla P_n) + \mathbf{f} \cdot \nabla P + \varphi - L|\right\}$$

so

$$\boxed{\int_{B_R} \rho \left[\varepsilon \nabla \cdot (\mathbf{D}\nabla P_n) + \mathbf{f} \cdot \nabla P + \varphi - L\right] d\mathbf{x} \to 0 \text{ as } \varepsilon \to 0} \tag{117}$$

To study the other terms, we switch to polar coordinates, $(x_1, ..., x_n) \to (r, \theta_1, ..., \theta_{n-1})$ where $r \in [0, R]$, $\theta_1, ..., \theta_{n-2} \in [0, \pi]$, $\theta_{n-1} \in [0, 2\pi]$ and

$$d\mathbf{x} = r^{n-1} \sin^{n-2}(\theta_1) ... \sin(\theta_{n-2})\, dr d\theta_1 ... d\theta_{n-1}.$$

Since ζ is a homogeneous, positive definite quadratic form of $\mathbf{x}$ and $\mathbf{D}$ is positive semi-definite (recall that $\mathbf{D} = \boldsymbol{\sigma}^T \boldsymbol{\sigma}$) we can write

$$\zeta(\mathbf{x}) = r^2 F(\theta_1, ..., \theta_{n-1})$$
$$\nabla\zeta \cdot (\mathbf{D}\nabla\zeta) = r^2 G(\theta_1, ..., \theta_{n-1})$$

for some strictly positive function F and non-negative function G, while $\nabla \cdot (\mathbf{D}\nabla\zeta)$ is a real number. Moreover, let

$$F^* = \min_{\theta_1, ..., \theta_{n-1}} F(\theta_1, ..., \theta_{n-1}),$$
$$G^* = \max_{\theta_1, ..., \theta_{n-1}} G(\theta_1, ..., \theta_{n-1})$$

and

$$I = \int_{r=0}^{R} \int_{\theta_{n-1}=0}^{2\pi} \int_{\theta_{n-2}=0}^{\pi} ... \int_{\theta_1=0}^{\pi} \frac{r^{n-1} \sin^{n-2}(\theta_1) ... \sin(\theta_{n-2})}{\varepsilon + r^2 F(\theta_1, ..., \theta_{n-1})} \, dr d\theta_1 ... d\theta_{n-1}.$$

Then, we have

$$\left| \int_{B_R} \frac{\alpha\varepsilon \, \rho \, \nabla \cdot (\mathbf{D}\nabla\zeta)}{\varepsilon + \zeta} d\mathbf{x} \right| \leq \varepsilon \, |\alpha \nabla \cdot (\mathbf{D}\nabla\zeta)| \, \max_{B_R}(\rho) \, I$$

$$\leq \varepsilon \, |\alpha \nabla \cdot (\mathbf{D}\nabla\zeta)| \, \max_{B_R}(\rho) \, 2\,\pi^{n-1} \int_{r=0}^{R} \frac{r^{n-1}}{\varepsilon + r^2 F^*} dr$$

If $n = 2$, the last term can be integrated to give

$$\left| \int_{B_R} \frac{\alpha\varepsilon \, \rho \, \nabla \cdot (\mathbf{D}\nabla\zeta)}{\varepsilon + \zeta} d\mathbf{x} \right| \leq \varepsilon \left\{ |\alpha \nabla \cdot (\mathbf{D}\nabla\zeta)| \, \max_{B_R}(\rho) \, \frac{\pi^{n-1}}{F^*} \log\left(1 + C^2 \, F^* \varepsilon^{-2\eta}\right) \right\}$$

while when $n \geq 3$ we can estimate

$$\left| \int_{B_R} \frac{\alpha\varepsilon \, \rho \, \nabla \cdot (\mathbf{D}\nabla\zeta)}{\varepsilon + \zeta} d\mathbf{x} \right| \leq \varepsilon^{3/2-\eta} \left\{ 2C\pi^{n-1} |\alpha \nabla \cdot (\mathbf{D}\nabla\zeta)| \max_{B_R}(\rho) \max_{r\in[0,R]} \left(\frac{r^{n-1}}{\varepsilon + r^2 F^*} \right) \right\}.$$

It can be verified that the maximum of the last term is achieved at the endpoint $r = R = C\varepsilon^{1/2-\eta}$. Taking the limit shows that for all $n \geq 2$

$$\boxed{\int_{B_R} \frac{\alpha\varepsilon \, \rho \, \nabla \cdot (\mathbf{D}\nabla\zeta)}{\varepsilon + \zeta} d\mathbf{x} \to 0 \text{ as } \varepsilon \to 0} \tag{119}$$

Similarly, we can show

$$\left| \int_{B_R} \frac{\alpha\varepsilon\rho\nabla\zeta \cdot (\mathbf{D}\nabla\zeta)}{(\varepsilon + \zeta)^2} d\mathbf{x} \right| \leq \varepsilon \, |\alpha| 2\pi^{n-1} \max_{B_R}(\rho) \int_{r=0}^{R} \frac{r^{n+1} G^*}{\varepsilon^2 + r^4 F^{*2}} dr$$

$$\leq \begin{cases} \varepsilon \left\{ \dfrac{|\alpha|\pi^{n-1} G^*)}{2F^{*2}} \max_{B_R}(\rho) \, \log(1 + C^4 \, F^{*2} \varepsilon^{-4\eta}) \right\}, & n = 2 \\ \varepsilon^{3/2-\eta} \left\{ 2C|\alpha|\pi^{n-1} G^* \max_{B_R}(\rho) \max_{r\in[0,R]} \left(\dfrac{r^{n+1}}{\varepsilon^2 + r^4 F^{*2}} \right) \right\}, & n \geq 3 \end{cases}$$

where, again, the last maximum is achieved at $r = R = C\varepsilon^{1/2-\eta}$. Hence we deduce

$$\boxed{\int_{B_R} \frac{\alpha\varepsilon\rho\nabla\zeta \cdot (\mathbf{D}\nabla\zeta)}{(\varepsilon + \zeta)^2} d\mathbf{x} \to 0 \text{ as } \varepsilon \to 0} \tag{120}$$

Finally, since $\mathbf{f}(0) = 0$ and $\nabla\zeta$ is linear, the term $\mathbf{f} \cdot \nabla\zeta$ is a polynomial of $\mathbf{x}$ that only contains monomials of degree 2 and higher. Consequently, we can write

$$\mathbf{f} \cdot \nabla\zeta = \sum_{m=1}^{\deg \mathbf{f}} r^{1+m} H_m(\theta)$$

for some continuous functions H_m such that their standard L^∞ norm $\|H_m\|_\infty$ is finite. Each term in this series can be considered separately; for each m we have

$$\left| \int_{B_R} \rho \frac{\alpha r^{1+m} H_m(\theta)}{\varepsilon + \zeta} d\mathbf{x} \right| \leq 2\pi^{n-1} |\alpha| \, \|H_m\|_\infty \max_{B_R}(\rho) \int_0^R \frac{r^{n+m}}{\varepsilon + r^2 F^*} dr$$
$$\leq \varepsilon^{1/2-\eta} \left\{ 2C\pi^{n-1} |\alpha| \, \|H_m\|_\infty \max_{B_R}(\rho) \max_{r\in[0,R]} \left(\frac{r^{n+m}}{\varepsilon + r^2 F^*} \right) \right\},$$

which tends to 0 as $\varepsilon \to 0$ (since $n \geq 2$, $m \geq 1$, and the last maximum is obtained at $r = R = C\varepsilon^{1/2-\eta}$). We therefore conclude that

$$\boxed{\int_{B_R} \rho \frac{\alpha \mathbf{f} \cdot \nabla \zeta}{\varepsilon + \zeta} d\mathbf{x} \to 0 \text{ as } \varepsilon \to 0} \tag{121}$$

Combining (117)-(121) proves (116). Note that the proof presented for estimates (119) and (120) is valid for systems of dimension $n \geq 2$ or higher. For one-dimensional systems, instead, one needs to consider the behaviour of ρ explicitly (cf. Appendix B).

D Remarks on Technical Implementation of SOS Problems

Initial numerical experiments revealed that the SOS problems used to compute upper and lower bounds are particularly ill-conditioned and cannot generally be solved by standard double-precision SDP solvers. From experience, this seems especially true for lower bound problems, and in general for polynomial degrees higher than approximately 4 or 6. The results presented in this work were obtained by pre-processing the SOS problem using the MATLAB toolbox YALMIP [10], and solving the resulting SDP with SDPA-GMP [8] with the parameter `precision` set to `200`. The procedure outlined below was followed for each individual optimisation (we assume the reader is familiar with at least the basic commands of the SOS optimisation module in YALMIP, such as `sossolve`):

1. Set up the SOS problem in YALMIP using `sdpvar` variables and YALMIP's command `sos` to create SOS constraints. For example, a polynomial storage function `V` and the vector of its coefficients `Vcoeffs` are defined with the commands

   ```
   >> [V,Vcoeffs] = polynomial(x,degreeV);
   ```

   In an attempt to reduce the problem size, all SOS polynomials required by the $\mathcal{S}$-procedure were defined using the commands

   ```
   >> z = monolist(x,degreeS);
   >> Sm = sdpvar(size(z,1));
   >> S = z'*Sm*z;
   >> Constraints = [Sm>=0];
   ```

   rather that by using YALMIP's `polynomial` and `sos` commands. This was recommended in [16].

2. Export the SDP to SDPA-GMP using a modified version of YALMIP's `solvesos` function combined with the `export` command. This was done in order to exploit YALMIP's pre-processing capabilities (e.g. Newton polytope and symmetry reductions, [11]).

3. Solve the SDP problem with SDPA-GMP and import the solution back into MATLAB/YALMIP. This is necessary because the output from the solver cannot be easily interpreted in terms of the polynomials and `sdpvar` objects defined by the user in Step 1 above. We remark here that this is likely to introduce numerical errors, due to the different numerical precision used by the solver and by MATLAB.

4. Re-define the problem using a reduced monomial basis, then repeat Steps 1-3. The reduced monomial basis can be obtained after removing unused or zero coefficients from `Vcoeffs`. This is a basic attempt to reproduce YALMIP's post-processing routines [11] and should improve the numerical conditioning. Note that this should not change the problem's solution, simply improve the numerical conditioning and make any computation more reliable. Any significant change in the solution should be interpreted as a warning for numerical problems.

5. Repeat steps 1-4 until the monomial basis cannot be reduced any further.

6. Check the feasibility of the solution by fixing all variables to their optimal value and calling `solvesos` again.

7. [Optional] Check if a certificate of strict positivity exists using Theorem 4 in [11].

This procedure was implemented successfully for most SDPs related to upper bound problems. For lower bound prolems, instead, the numerical solution obtained with SDPA-GMP could not be post-processed reliably with YALMIP, and the feasibility test in step 6 above was almost never successful. However SDPA-GMP never reported infeasibility or numerical problems when solving the original SOS optimisation; in fact, it terminated with the optimal flag `pdOPT`, indicating that an optimal solution could be found. Consequently, the upper and lower bounds computed by SDPA-GMP were considered acceptable.

Further tests were carried out by disabling YALMIP's pre-processing routines; the results were unchanged but the computation time increased significantly. This suggests that the preliminary problem manipulation carried out by YALMIP is beneficial and does not constitute a large source of error. Yet, the solution of the SDP produced by YALMIP requires high-precision. Clearly, one would need to maintain high numerical precision to correctly interpret the data returned by the solver and recover the coefficients of the polynomials as defined originally by the user. Unfortunately, the "dictionary" used by YALMIP to operate this "translation" is not immediately available to the user, and the solver data must be returned to the SOS module in YALMIP if further processing/checking is needed. This, in our opinion, is the main source of numerical errors that prevented us to successfully carry out a feasibility test as in step 6 above.

Finally, a comment on the solver. Whilst SDPA-GMP allowed us to compute the results presented in this report, it does not allow multi-threading and, consequently, we expect that larger SDPs cannot be solved in a reasonable computation time. This is a fundamental issue if higher-dimensional systems are to be studied, as the SDP associated with the SOS formulation of the upper/lower bound problems become very large even for modest polynomial degrees.

Clearly, future work should address these issues more carefully, especially if further analysis of the optimal solution and certificates of true positivity are needed. As a first step, a high-precision parser for SOS problems should be implemented to carry out the necessary pre-processing and to interpret the solution from SDPA-GMP in a user-friendly format. As a long-term suggestion, however, we recommend the development of a high-precision SOS parser with pre- and post-processing capabilities similar to those of YALMIP and of a multi-threaded high-precision SDP solver.

E Constructing ζ from Eigenvector Analysis

Remark: we construct suitable quadratic forms ζ in the context of the Van der Pol Oscillator. However, the same ideas and derivations apply to a general system with a repelling fixed point.

In the vicinity of the origin, the Van der Pol oscillator

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= \mu(1-x^2)y - x \end{aligned} \tag{122}$$

can be expanded as

$$\dot{\mathbf{x}} = \mathbf{J}_0\mathbf{x} + \text{h.o.t.}, \qquad \mathbf{J}_0 = \begin{pmatrix} 0 & 1 \\ -1 & \mu \end{pmatrix}. \tag{123}$$

Letting $\mu = 2\nu$, the eigenvalues of $\mathbf{J}_0$ can be written as

$$\lambda_{1,2} = \nu \pm \sqrt{\nu^2 - 1} \tag{124}$$

and the matrix of normalised (i.e. unit norm) eigenvectors is

$$\mathbf{U} = \begin{pmatrix} A & B \\ A\nu + A\sqrt{\nu^2-1} & B\nu - B\sqrt{\nu^2-1} \end{pmatrix} \tag{125}$$

with

$$A = \left[2\nu^2 + 2\nu\sqrt{\nu^2-1}\right]^{-1/2}, \quad B = \left[2\nu^2 - 2\nu\sqrt{\nu^2-1}\right]^{-1/2}. \tag{126}$$

When $\mu = 2$ (i.e. $\nu = 1$), the two eigenvalues and eigenvectors coincide and $\mathbf{J_0}$ cannot be diagonalised; otherwise, we have

$$\mathbf{\Lambda} := \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \mathbf{U}^{-1}\,\mathbf{J}_0\,\mathbf{U}. \tag{127}$$

Note that $\mathbf{\Lambda}$ is a positive definite matrix.

E.1 Construction of ζ_1

We arbitrarily fix $\zeta_1 = x^2 - xy - y^2$, which is clearly positive definite. Moreover,

$$\begin{aligned} \dot{\zeta_1} &= 2x\dot{x} - \dot{x}y - x\dot{y} + 2y\dot{y} \\ &= x^2 - \mu xy + (2\mu - 1)y^2 + \text{h.o.t.} \end{aligned} \tag{128}$$

Neglecting the higher-order terms, the right hand side is positive in the vicinity of the origin if

$$\Delta = \mu^2 y^2 - 4(2\mu - 1)y^2 < 0, \tag{129}$$

which is satisfied for $4 - 2\sqrt{3} < \mu < 4 + 2\sqrt{3}$. Thus, ζ_1 satisfies the necessary condition (48) for $\alpha > 0$ if μ is in this range.

E.2 Construction of ζ_2

Let $\mathbf{w} = \mathbf{U}^{-1}\mathbf{x}$ and consider ζ_2 as

$$\zeta_2 \propto \|\mathbf{w}\|^2 = \mathbf{x}^T[\mathbf{U}^{-1}]^T\mathbf{U}^{-1}\mathbf{x}. \tag{130}$$

The (positive) proportionality constant can be chosen arbitrarily, since the constraint in (46) is homogeneous in ζ. In particular, one can chose such constants to write

$$\zeta_2 = \begin{cases} x^2 - \mu xy + y^2, & \mu < 2 \\ \mu x^2 - 4xy + \mu y^2, & \mu > 2 \end{cases} \tag{131}$$

Clearly, ζ_2 is positive definite. Moreover, it satisfies the necessary condition (48) near the origin. In fact,

$$\begin{aligned} \dot{\zeta_2} &\propto 2\mathbf{x}^T\mathbf{J}_0^T\,[\mathbf{U}^{-1}]^T\mathbf{U}^{-1}\mathbf{x} + \text{h.o.t.} \\ &= 2\mathbf{w}^T\mathbf{U}^T\mathbf{J}_0^T\,[\mathbf{U}^{-1}]^T\mathbf{w} + \text{h.o.t.} \\ &= 2\mathbf{w}^T[\mathbf{U}^{-1}\mathbf{J}_0\,\mathbf{U}]^T\mathbf{w} + \text{h.o.t.} \\ &= 2\mathbf{w}^T\mathbf{\Lambda}\mathbf{w} + \text{h.o.t.}. \end{aligned} \tag{132}$$

Neglecting the higher order terms in a neighbourhood of the origin and recalling that the matrix of eigenvalues is positive definite, we conclude that the last expression is positive. However, this expression cannot be used for $\mu = 2$, since $\mathbf{U}$ cannot be inverted.

E.3 Construction of ζ_3

In order to take into account the dynamics near the origin, and not only the geometric information contained in the eigenvectors, let

$$\zeta_3 \propto \mathbf{w}^T\mathbf{\Lambda}\mathbf{w}^T = \mathbf{x}^T[\mathbf{U}^{-1}]^T\mathbf{\Lambda}\mathbf{U}^{-1}\mathbf{x}. \tag{133}$$

Choosing appropriate proportionality constants to simplify the form of ζ_3, we can write

$$\zeta_3 = \begin{cases} x^2 - \mu xy + y^2, & \mu < 2, \\ \left(\mu^2 - 2\right)x^2 - 2xy + 2y^2, & \mu > 2. \end{cases} \tag{134}$$

Clearly, ζ_3 is positive definite. Moreover,

$$\begin{aligned} \dot{\zeta_3} &\propto 2\mathbf{x}^T\mathbf{J}_0^T\,[\mathbf{U}^{-1}]^T\mathbf{\Lambda}\mathbf{U}^{-1}\mathbf{x} + \text{h.o.t.} \\ &= 2\mathbf{w}^T\mathbf{U}^T\mathbf{J}_0^T\,[\mathbf{U}^{-1}]^T\mathbf{\Lambda}\mathbf{w} + \text{h.o.t.} \\ &= 2\mathbf{w}^T[\mathbf{U}^{-1}\mathbf{J}_0\,\mathbf{U}]^T\mathbf{\Lambda}\mathbf{w} + \text{h.o.t.} \\ &= 2\mathbf{w}^T\mathbf{\Lambda}^T\mathbf{\Lambda}\mathbf{w} + \text{h.o.t.} \\ &= 2\mathbf{w}^T\begin{pmatrix} |\lambda_1|^2 & 0 \\ 0 & |\lambda_2|^2 \end{pmatrix}\mathbf{w} + \text{h.o.t.} \end{aligned} \tag{135}$$

Neglecting the higher order terms in a neighbourhood of the origin, we conclude that the last expression is positive and therefore (48) is satisfied. However, this expression cannot be used for $\mu = 2$, since $\mathbf{U}$ cannot be inverted.

References

[1] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, 2004.

[2] J. C. Bronski and T. N. Gambill, *Uncertainty estimates and L_2 bounds for the Kuramoto-Sivashinsky equation*, Nonlinearity, 19 (2006), pp. 2023–2039.

[3] S. I. Chernyshenko, P. J. Goulart, D. Huang, and A. Papachristodoulou, *Polynomial sum of squares in fluid dynamics : a review with a look ahead*, Phylosophical Transactions of the Royal Society, A., 372 (2014).

[4] P. Constantin and C. R. Doering, *Variational bounds in dissipative systems*, Physica D: Nonlinear Phenomena, 82 (1995), pp. 221–228.

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

[6] ——, *Variational bounds on energy dissipation in incompressible flows. III. Convection*, Physical Review E, 53 (1996), pp. 5957–5981.

[7] C. R. Doering, F. Otto, and M. G. Reznikoff, *Bounds on vertical heat transport for infinite Prandtl number Rayleigh-Benard convection*, Journal of Fluid Mechanics, 560 (2006), pp. 229–241.

[8] K. Fujisawa, M. Fukuda, K. Kobayashi, M. Kojima, K. Nakata, M. Nakata, and M. Yamashita, *SDPA (SemiDefinite Programming Algorithm) and SDPA-GMP Users Manual Version 7.1.1*, tech. rep., Department of Mathematical and Computing Sciences Tokyo Institute of Technology, 2008.

[9] G. I. Hagstrom and C. R. Doering, *Bounds on Surface Stress-Driven Shear Flow*, Journal of Nonlinear Science, 24 (2014), pp. 185–199.

[10] J. Lofberg, *YALMIP: A toolbox for modeling and optimization in MATLAB*, in Proceedings of the CACSD Conference, Taipei, Taiwan, 2004.

[11] ——, *Pre-and post-processing sum-of-squares programs in practice*, IEEE Transactions on Automatic Control, 54 (2009), pp. 1007–1011.

[12] A. Papachristodoulou, J. Anderson, G. Valmorbida, S. Prajna, P. Seiler, and P. A. Parrilo, *SOSTOOLS: Sum of squares optimization toolbox for MATLAB*, `http://arxiv.org/abs/1310.4716`, 2013. Available from `http://www.eng.ox.ac.uk/control/sostools`.

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

[14] A. N. Souza and C. R. Doering, *Maximal transport in the Lorenz equations*, Physics Letters A, 379 (2014), pp. 1–13.

[15] A. N. Souza and C. R. Doering, *Transport bounds for a truncated model of Rayleigh-Bénard convection*, Physica D: Nonlinear Phenomena, 308 (2015), pp. 26–33.

[16] W. Tan, *Nonlinear Control Analysis and Synthesis using Sum-of-Squares Programming*, PhD Thesis, University of california, Berkeley, 2006.

[17] M. Viana, *Whats new on Lorenz strange attractors?*, The Mathematical Intelligencer, 22 (2000), pp. 6–19.

[18] L.-S. Young, *What are SRB measures, and which dynamical systems have them?*, Journal of Statistical Physics, 516 (2002), pp. 1–21.