

MIT Open Access Articles

*Least squares shadowing sensitivity analysis
of a modified Kuramoto-Sivashinsky equation*

The MIT Faculty has made this article openly available. **Please share**
how this access benefits you. Your story matters.

Citation: Blonigan, Patrick J., and Qiqi Wang. "Least squares shadowing sensitivity analysis of a modified Kuramoto-Sivashinsky equation." *Chaos, Solitons & Fractals* 64 (July 2014), pp.16-25.

As Published: <http://dx.doi.org/10.1016/j.chaos.2014.03.005>

Publisher: Elsevier

Persistent URL: <http://hdl.handle.net/1721.1/103866>

Version: Original manuscript: author's manuscript prior to formal peer review

Terms of use: Creative Commons Attribution-NonCommercial-NoDerivs License



Least Squares Shadowing Sensitivity Analysis of a Modified Kuramoto-Sivashinsky Equation

Patrick J. Blonigan^a, Qiqi Wang^a

^a*Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139, United States*

Abstract

Computational methods for sensitivity analysis are invaluable tools for scientists and engineers investigating a wide range of physical phenomena. However, many of these methods fail when applied to chaotic systems, such as the Kuramoto-Sivashinsky (K-S) equation, which models a number of different chaotic systems found in nature. The following paper discusses the application of a new sensitivity analysis method developed by the authors to a modified K-S equation. We find that least squares shadowing sensitivity analysis computes accurate gradients for solutions corresponding to a wide range of system parameters.

Keywords: Sensitivity Analysis, Kuramoto-Sivashinsky Equation

1. Introduction

Sensitivity analysis is of great importance to scientists and engineers. It is used to compute sensitivity derivatives of key quantities of interest to parameters that influence a system which is governed by some ordinary differential equation (ODE) or partial differential equation (PDE). These sensitivity derivatives can be used for design optimization, inverse problems, data assimilation, and uncertainty quantification.

Sensitivity analysis for chaotic dynamical systems is important because of the prevalence of chaos in many scientific and engineering fields. A relatively simple PDE that can be chaotic, the Kuramoto-Sivashinsky (K-S) equation, (1), is a great illustration of how widespread chaos is in nature, as it has been found to model a wide range of physical phenomena.

$$\frac{\partial u}{\partial t} = -u \frac{\partial u}{\partial x} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^4 u}{\partial x^4} \quad (1)$$

Kuramoto derived the equation for angular-phase turbulence for a system of reaction-diffusion equations modeling the Belousov-Zhabotinskii reaction in three spatial dimensions [1, 2]. Sivashinsky also derived the

Email addresses: blonigan@mit.edu (Patrick J. Blonigan), qiqi@mit.edu (Qiqi Wang)

equation to model the evolution of instabilities in a distributed plane flame front [3, 4]. In addition the K-S equation has also been shown to be a model of Poiseuille flow of a film layer on an inclined plane [5].

In the study of periodic and chaotic systems, long time averaged quantities, such as mean temperature and mean aerodynamic forces of turbulent fluid flows, are of interest. However, many sensitivity analysis methods fail when applied to long time averaged quantities in chaotic dynamical systems. The sensitivity gradient predicted by methods including the adjoint method are observed to diverge as simulation time is increased [6].

A recently developed method, the least squares shadowing (LSS) method, can compute accurate gradients for ergodic chaotic systems [7, 8]. The LSS method finds a perturbed trajectory (or solution) that does not diverge exponentially from some trajectory in phase space. This non-diverging trajectory, called a “shadow trajectory”, has its existence guaranteed by the shadowing lemma [9] for a large number of chaotic systems and can be used to compute sensitivities.

This paper discusses the application of the LSS method to a one-dimensional modified K-S equation:

$$\begin{aligned} \frac{\partial u}{\partial t} &= -(u + c) \frac{\partial u}{\partial x} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^4 u}{\partial x^4} & (2) \\ x &\in [0, L], t \in [0, \infty) \\ u(0, t) &= u(L, t) = 0 \\ \frac{\partial u}{\partial x} \Big|_{x=0} &= \frac{\partial u}{\partial x} \Big|_{x=L} = 0 \\ u(x, 0) &= u_0(x) \end{aligned}$$

Homogenous Dirichlet and Neumann boundary conditions are used to make the system ergodic. The parameter c is added to demonstrate LSS. Sensitivity of the quantity of interests $\langle \bar{u} \rangle$ and $\langle \bar{u}^2 \rangle$ to c will be investigated, where:

$$\langle \bar{u} \rangle \equiv \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \bar{u} dt, \quad \bar{u} \equiv \frac{1}{L} \int_0^L u dx, \quad \bar{u}^2 \equiv \frac{1}{L} \int_0^L u^2 dx$$

We have chosen $L = 128$ to ensure the solution is chaotic [10].

The remainder of this paper is organized as follows: first, section 2 discusses the numerical simulation used to compute solutions of the modified K-S equation. Next, the reasons for the modifications to the K-S equation are discussed in more detail in section 3. Thirdly, a brief summary of the LSS method will be given

in section 4, followed by a presentation and a discussion of the gradients computed using LSS in section 5. Section 6 offers some concluding remarks and a discussion of future work.

2. Numerical Simulation

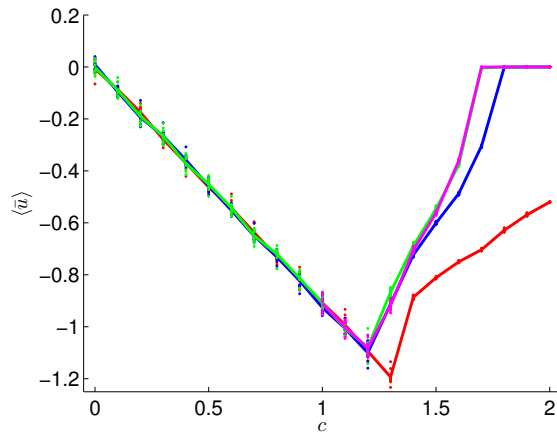
The modified K-S equation was discretized with a 2nd order finite difference scheme. We number the nodes $i = 0, 1, 2, \dots, n, n + 1$, where $i = 0$ and $i = n + 1$ are the boundary nodes and $i = 1, 2, \dots, n$ denote the interior nodes. Define $x_i = i\Delta x$, where $\Delta x = L/(n + 1)$ is the spacing between each node, and $u_i = u(x_i)$. The terms of the modified K-S equation were approximated as follows on the interior nodes:

$$\begin{aligned} \left. \frac{\partial u}{\partial x} \right|_i &\approx \frac{u_{i+1} - u_{i-1}}{2\Delta x}, \quad i = 1, 2, \dots, n \\ u \left. \frac{\partial u}{\partial x} \right|_i &= \frac{1}{2} \left. \frac{\partial u^2}{\partial x} \right|_{x_i} \approx \frac{u_{i+1}^2 - u_{i-1}^2}{4\Delta x}, \quad i = 1, 2, \dots, n \\ \left. \frac{\partial^2 u}{\partial x^2} \right|_i &\approx \frac{u_{i+1} - 2u_i + u_{i-1}}{\Delta x^2}, \quad i = 1, 2, \dots, n \\ \left. \frac{\partial^4 u}{\partial x^4} \right|_i &\approx \frac{u_{i-2} - 4u_{i-1} + 6u_i - 4u_{i+1} + u_{i+2}}{\Delta x^4}, \quad i = 2, 3, \dots, n-1 \end{aligned}$$

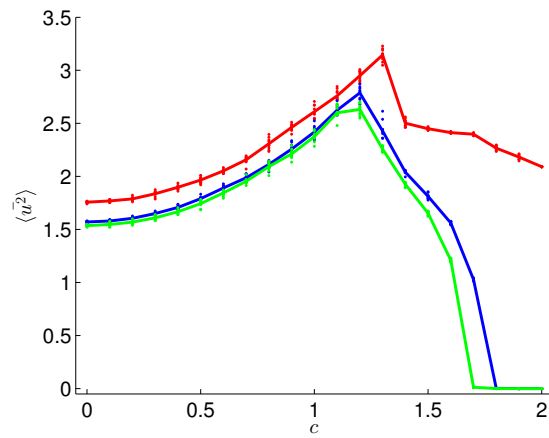
To enforce the homogeneous Dirichlet boundary conditions we set $u_0 = u_{n+1} = 0$. To enforce the homogeneous Neumann boundary conditions, we make use of ghost nodes, setting $u_{-1} = u_1$ and $u_{n+2} = u_n$. This ensures that the central difference approximation of $\frac{\partial u}{\partial x}$ is zero at nodes 0 and n , which correspond to $x = 0$ and $x = L$, respectively. Therefore:

$$\begin{aligned} \left. \frac{\partial u}{\partial x} \right|_1 &\approx \frac{u_2}{2\Delta x}, & \left. \frac{\partial u}{\partial x} \right|_n &\approx -\frac{u_{n-1}}{2\Delta x} \\ u \left. \frac{\partial u}{\partial x} \right|_1 &\approx \frac{u_2^2}{4\Delta x}, & u \left. \frac{\partial u}{\partial x} \right|_n &\approx -\frac{u_{n-1}^2}{4\Delta x} \\ \left. \frac{\partial^2 u}{\partial x^2} \right|_1 &\approx \frac{u_2 - 2u_1}{\Delta x^2}, & \left. \frac{\partial^2 u}{\partial x^2} \right|_n &\approx \frac{u_{n-1} - 2u_n}{\Delta x^2} \\ \left. \frac{\partial^4 u}{\partial x^4} \right|_1 &\approx \frac{7u_1 - 4u_2 + u_3}{\Delta x^4}, & \left. \frac{\partial^4 u}{\partial x^4} \right|_n &\approx \frac{7u_n - 4u_{n-1} + u_{n-2}}{\Delta x^4} \\ \left. \frac{\partial^4 u}{\partial x^4} \right|_2 &\approx \frac{-4u_1 + 6u_2 - 4u_3 + u_4}{\Delta x^4}, & \left. \frac{\partial^4 u}{\partial x^4} \right|_{n-1} &\approx \frac{-4u_n + 6u_{n-1} - 4u_{n-2} + u_{n-3}}{\Delta x^4} \end{aligned}$$

The coarsest mesh we used to solve the modified K-S equation is a uniform mesh with 127 interior nodes,



(a) $J = \bar{u}$



(b) $J = \bar{u}^2$

Figure 1: Two different objective functions $\langle J \rangle$ versus c . Dots represent each realization, lines represent the average of 10 realizations. Different colors correspond to different spatial discretizations, with red for $n = 127$, blue for $n = 255$, green for $n = 511$, and maroon for $n = 700$, where n is the number of nodes. Each realization was run for 1000 time units before averaging was begun. Averages were taken over 2000 time unit intervals.

0	0			
1/3	0	1/3		
1	0	1/2	1/2	
1	0	3/4	-1/4	1/2
	0	3/4	-1/4	1/2

0	0			
1/3	1/3	0		
1	0	1	0	
1	0	3/4	1/4	0
	0	3/4	-1/4	1/2

Table 1: Butcher Tableau for the Diagonally Implicit Runge-Kutta (DIRK) scheme (left) and the Explicit Runge-Kutta (ERK) scheme (right) used by IMEXRK34S[2R]L α . The scheme was designed to maximize the accuracy of the explicit portion of the solver [12]. In the case of this paper, the explicit portion is the convective terms of (2), $-(u + c)\frac{\partial u}{\partial x}$.

as in Brummitt and Sprott [11]. From figure 1, we can see that while this mesh is accurate for $0 \leq c \leq 1.2$, where the mean values of $\langle \bar{u} \rangle$ never vary more than 5%. However. The approximation is poor for larger magnitude values of c and for computing $\langle \bar{u}^2 \rangle$. A mesh with $n = 511$ interior nodes and $\Delta x = 0.25$ is used for these computations, as this grid resolution appears to offer sufficient accuracy in figure 1.

Time integration was mainly conducted with a 3rd order accurate implicit-explicit Runge-Kutta time integration scheme, IMEXRK34S[2R]L α [12], with the Butcher tableau shown in table 1. The 2nd and 4th order derivatives in equation (2) were integrated with the Diagonally Implicit Runge-Kutta (DIRK) scheme, and the convective terms, $-(u + c)\frac{\partial u}{\partial x}$, were integrated with the Explicit Runge-Kutta (ERK) scheme. For the results presented in this paper, a time step of $\Delta t = 0.1$ was used unless otherwise stated.

Some of the ensemble average computations (see section 3) were conducted using MATLAB's ODE45 hybrid 4th/5th order Runge-Kutta scheme and a time step size of $\Delta t = 0.2$ to ensure stability. Solutions to the modified K-S equation used for the work presented in this paper range from 100 to 4000 time units.

Finally, unless otherwise stated, the initial condition $u_0(x)$ was formed by randomly selecting numbers at each spatial node from a uniform distribution between $u = -0.5$ and $u = 0.5$.

3. The Modified Kuramoto-Sivashinsky Equation

3.1. Boundary Conditions

For LSS to compute accurate gradients, the system being analyzed must be ergodic; the long time behavior of the system is independent of initial conditions. The K-S equation with periodic boundary conditions is not ergodic, and to show this we consider the spatial average of equation (2):

$$\frac{1}{L} \int_0^L \left(\frac{\partial u}{\partial t} \right) dx = \frac{1}{L} \int_0^L \left(-(u + c) \frac{\partial u}{\partial x} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^4 u}{\partial x^4} \right) dx$$

Taking the time derivative outside of the integral, using the notation $\bar{u} = \frac{1}{L} \int_0^L u dx$, and multiplying both sides of the equation by L , we obtain:

$$L \frac{\partial \bar{u}}{\partial t} = -\frac{1}{2} u^2 \Big|_0^L - c u \Big|_0^L - \frac{\partial u}{\partial x} \Big|_0^L - \frac{\partial^3 u}{\partial x^3} \Big|_0^L \quad (3)$$

For periodic boundary conditions, u and all of its derivatives are equal at $x = 0$ and $x = L$ for all time. Therefore equation (3) becomes:

$$\frac{\partial \bar{u}}{\partial t} = 0$$

This means that \bar{u} is always equal to the average of the initial condition $u_0(x)$. Therefore, our choice of the initial condition dictates the long time behavior of our quantity of interest \bar{u} and the system is not ergodic. On the other hand if homogeneous Dirichlet and Neumann boundary conditions are used, ergodic behavior can be shown from numerical solutions. Ergodicity can be observed in the behavior of the ensemble averaged time averaged solution, defined as

$$\langle \widetilde{u} \rangle \equiv \frac{1}{K} \sum_{k=1}^K \langle u \rangle$$

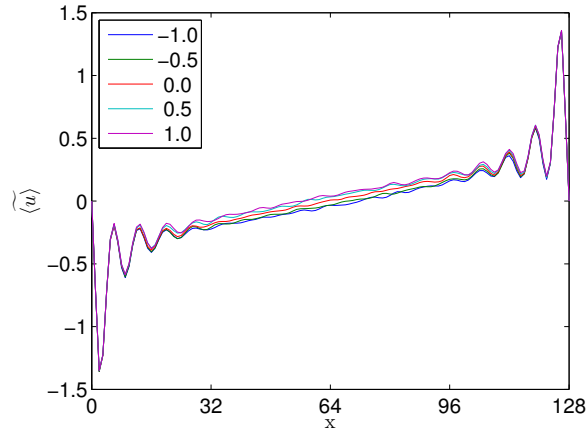
where $\langle u \rangle$ is the time average of $u(x, t)$. The ensemble average is conducted over K realizations, each starting from a different initial condition, comprised of the random noise described in section 2 shifted to some non-zero mean. Figure 2 shows that $\langle \widetilde{u} \rangle$ converges to a single function for initial conditions with different means. The initial condition also contained random noise like that described in section 2. The random noise component of the initial condition is comprised of many different functions, which means that the convergence shown in figure 2 strongly implies that the system is ergodic. Note that similar convergence of the solutions was observed at a number of different values of c , suggesting that the modified K-S equation is ergodic over a wide range of values of c .

Finally, note that the initial condition $u_0(x) = 0$ results in a trivial solution $u(x, t) = 0$. However, this trivial solution is unstable; any small perturbation to it will lead to a non-trivial, chaotic solution [10]. One can think of the trivial solution to the modified K-S equation as an unstable fixed point, like $(x, y, z) = (0, 0, 0)$ in the Lorenz System.

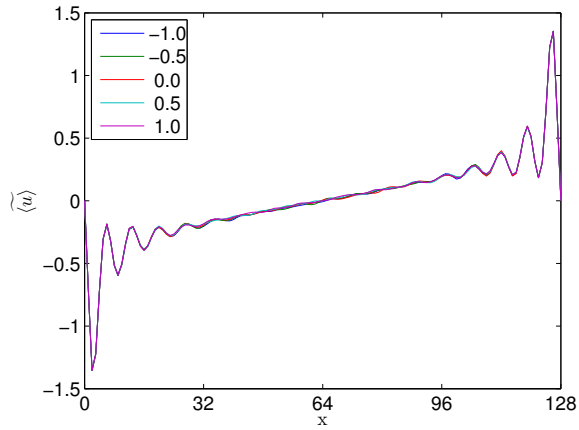
3.2. Additional Linear Convection Parameter c

To see the effect of the parameter c , we again consider ensemble averaged time averaged solutions $\langle \widetilde{u} \rangle$. Figure 3 shows that as c is increased, $\langle \widetilde{u} \rangle$ is decreased at all values of x other than the boundaries, where the Dirichlet boundary conditions are imposed.

This decrease in $\langle \widetilde{u} \rangle$ occurs because increasing c leads to increased linear convection in the positive x

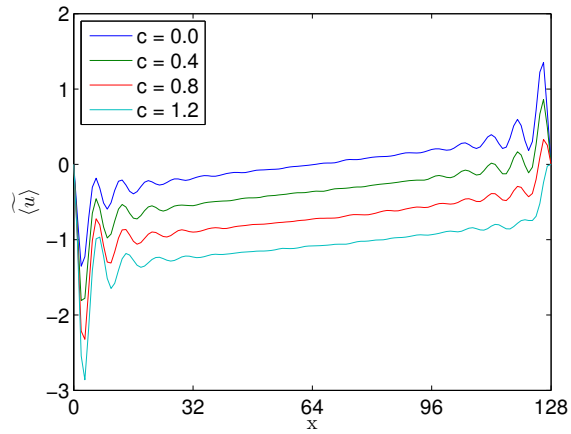


(a) $T_0 = 500$

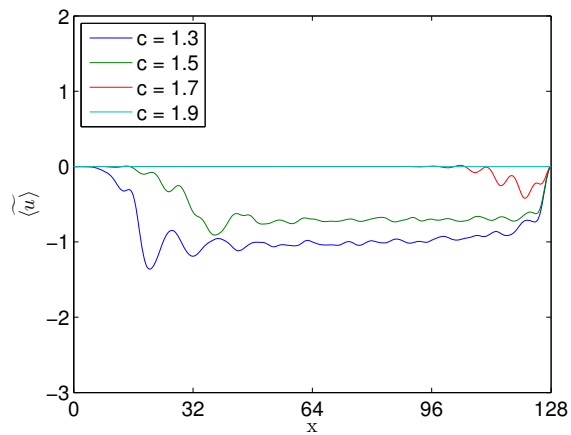


(b) $T_0 = 1000$

Figure 2: The ensemble averaged time averaged solution, $\langle \tilde{u} \rangle$, for $c = 0$ for initial conditions with five different means \bar{u}_0 . Both solutions were run for T_0 time units before the averaging was started. The averaging interval was 100 time units long and 4000 realizations were used for both ensemble averages. All five initial conditions were formed by randomly selecting numbers at each grid point from a uniform distribution between $u = -0.5$ and $u = 0.5$ and then changing the mean.



(a) $c < 1.3$



(b) $c > 1.3$

Figure 3: The ensemble averaged, time averaged solution, $\langle u \rangle$ for different values of c . The averaging interval was 100 time units long and 4000 realizations were used for both ensemble averages.

direction, transporting negative u from one side of the domain to the other and transporting positive u out of the domain.

This trend can be seen in figure 1, which shows how time and space averaged u , $\langle \bar{u} \rangle$, varies with c . From $c = 0$ to $c \approx 1.2$, we see the behavior shown in figure 3 and $\langle \bar{u} \rangle$ decreases linearly with c . Note that because the the boundary conditions are symmetric, $\langle \bar{u} \rangle$ as a function of c is anti-symmetric. Decreasing c from $c = 0$ results in a linear increase in $\langle \bar{u} \rangle$, until $c \approx -1.2$.

For $c > 1.2$ (or $c < -1.2$) we see a change in the slope of $\langle \bar{u} \rangle$. Another change occurs around $c = 1.8$. These changes in the trend of $\langle \bar{u} \rangle$ are a result of changes in the dynamics of the modified K-S equation. Example of solutions $u(x, t)$ in these regimes, $0 \leq c \leq 1.2$, $1.2 \leq c \leq 1.8$, and $c > 1.8$ are shown in figure 4.

In the first regime, $0 \leq c \leq 1.2$, the “light turbulence dominated regime”, we see the chaotic spatio-temporal structures referred to in many past studies of the K-S. In the left-most plot of figure 4 we see that some of the structures between roughly $x = 64$ and $x = 128$ are convecting towards $x = 128$, due to the linear convection term.

As c is increased, convection in the positive x -direction is increased. This can be seen visually from the tilt or slope of the spatio-temporal structures in all three $x - t$ diagrams in figure 4. The smaller the slope of the structure, the faster the structure is being convected.

In the second regime, $1.2 \leq c \leq 1.8$, the “convection dominated regime”, All spatio-temporal structures are convecting towards $x = 128$, as seen in the center plot of figure 4. Additionally, a region of $u = 0$ now exists near $x = 0$. As x increases, the spatio-temporal structures grow, until $x \approx 30$ for $c = 1.4$ (figure 4). The size of this region grows as c is increased.

In the final regime, $c > 1.8$, the “steady regime”, the chaotic structures are convected out of the domain, as in the right-most plot in figure 4. As c is increased, the structures are convected out of the spatial domain faster. The trivial solution has changed from an unstable fixed point to a stable fixed point, and the solution becomes $u(x, t) = 0$ after some time. This is different from solutions in the first two regimes, both of which are on chaotic attractors, as indicated by ergodicity and presence of spatio-temporal chaos.

4. The Least Squares Shadowing Method

Say we are interested in the sensitivity of our long-time averaged quantity $\langle \bar{u} \rangle$ to the parameter c :

$$\frac{d\langle \bar{u} \rangle}{dc} = \frac{d}{dc} \left(\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \bar{u} dt \right)$$

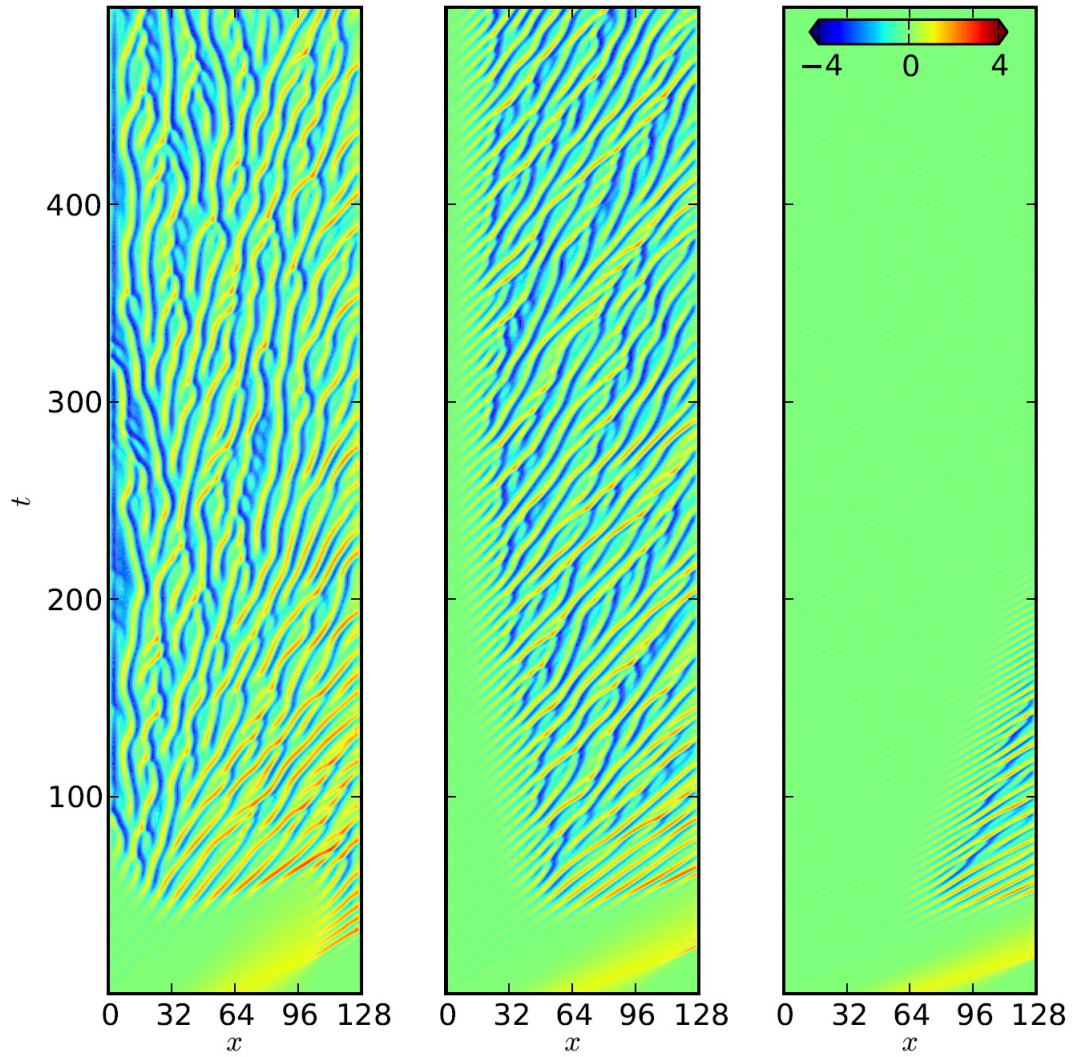


Figure 4: From left to right: x vs. t plots for $u(x,t)$ for $c = 0.8$ (light turbulence dominated regime), $c = 1.4$ (convection dominated regime), and $c = 2.0$ (steady regime). All three histories were computed with $\Delta x = 0.25$. The initial condition for all three histories was $u_0(x) = e^{-\frac{(x-64)^2}{512}}$.

For non-chaotic solutions of the K-S equation solved from the **initial value problem**, we can exchange the derivative and the time limit:

$$\frac{d\langle \bar{u} \rangle}{dc} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \frac{\partial \bar{u}}{\partial u} \frac{\partial u}{\partial c} dt \quad (4)$$

Where

$$\frac{\partial u}{\partial c} = \lim_{\varepsilon \rightarrow 0} \frac{u(t; c + \varepsilon) - u(t; c)}{\varepsilon} \quad (5)$$

We can solve for the *tangent solution*, $v \equiv \frac{\partial u}{\partial c}$, with the linearization of equation (2), also called the *tangent equation*. However, if $u(t; c)$ is chaotic, then equations (4) and (5) do not hold and the tangent solution, v , diverges exponentially as the simulation time T is increased [6, 7].

For a steady or periodic solution, a slight change ε in c results in a slightly different solution $u(t; c + \varepsilon)$ for all time and the tangent solution, v , can be computed accurately. However, the positive Lyapunov exponent(s) present on strange attractors ensure that the solutions $u(t; c + \varepsilon)$ and $u(t; c)$ will be very different after some time, as illustrated in figure 5. We see that if the perturbed solution, or phase space trajectory, has the same initial condition as the unperturbed trajectory, the two trajectories diverge exponentially. This exponential divergence of the two trajectories causes the tangent solution, v , to diverge, resulting in the issues with traditional sensitivity analysis mentioned in the introduction.

However, the assumption of ergodicity means that it is not necessary to compare a perturbed and an unperturbed trajectory with the same initial condition if the quantities of interest are statistics of the system such as long time averages. Therefore, an initial condition can be chosen such that the perturbed and unperturbed trajectories do not diverge, resulting in the blue trajectory in figure 5. The existence of this trajectory, called a “shadow trajectory”, follows from the shadowing lemma [9]:

For any $\delta > 0$ there exists $\varepsilon > 0$, such that for every “ ε -pseudo-solution” u_ε that satisfies $\|du_\varepsilon/dt - \mathcal{R}(u_\varepsilon)\| < \varepsilon$, $0 \leq t \leq T$, there exists a true solution u and a time transformation $\tau(t)$, such that $\|u(\tau(t)) - u_\varepsilon(t)\| < \delta$, $|1 - d\tau/dt| < \delta$ and $du/d\tau - \mathcal{R}(u) = 0$, $0 \leq \tau \leq \mathcal{T}$.

Where the norm $\|\cdot\|$ refers to distance in phase space. For the K-S equation, the operator \mathcal{R} is comprised of the spatial derivative operators on the right hand side of equation (2).

The time transformation alluded to in the shadowing lemma is required to deal with the zero (neutrally stable) Lyapunov exponent on the strange attractor. The need for this transformation is clarified in figure

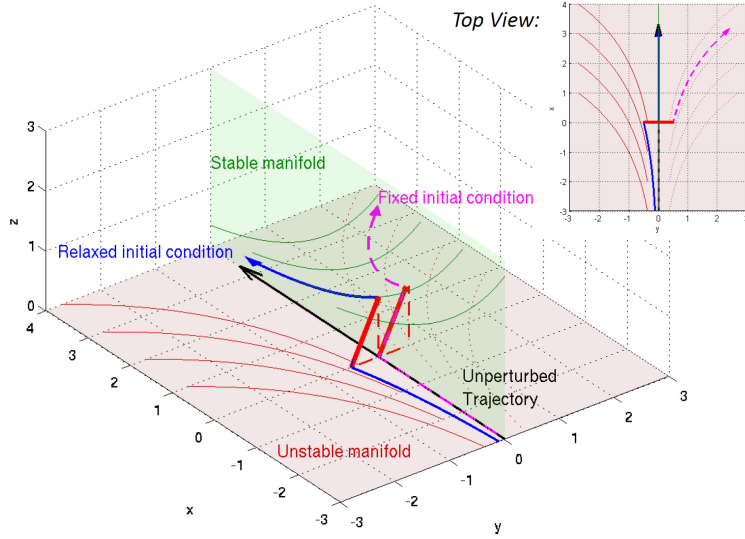


Figure 5: Phase space trajectory of a chaotic dynamical system. The unstable manifold, in red, is the space of all Lyapunov covariant vectors corresponding to positive exponents. The stable manifold, in green, corresponds to the space of all covariant vectors associated with negative exponents. A perturbation to the system (in red) has components in both manifolds, and the unstable component causes the perturbed trajectory (pink) to diverge exponentially from the unperturbed trajectory (in black). LSS chooses a perturbed trajectory with a different initial condition (in blue) that does not diverge from the unperturbed trajectory.

6. The time transformation, referred to as “time dilation” in this paper and other LSS literature, is required to keep a phase space trajectory and its shadow trajectory close (in phase space) for infinite time.

If we use two solutions that shadow one another in equation (5), 4 can be used to compute accurate sensitivities [13]. $v \equiv \frac{\partial u}{\partial c}$, called the “shadowing direction” can be computed by solving the following optimization problem with the tangent equation as a constraint:

$$\begin{aligned}
 & \min_{v, \eta} \frac{1}{2T} \int_0^T v^2 + \alpha^2 \eta^2 dt, \\
 \text{s.t.} \quad & \frac{dv}{dt} = \frac{\partial \mathcal{R}}{\partial u} v + \frac{\partial \mathcal{R}}{\partial c} + \eta \mathcal{R}(u; c) \\
 & 0 < t < T
 \end{aligned} \tag{6}$$

where $\eta = d\tau/dt - 1$ is the time dilation term, which corresponds to the time transformation from the shadowing lemma and α^2 is a weighting parameter for the optimization problem. More details on the implementation of LSS can be found in [8].

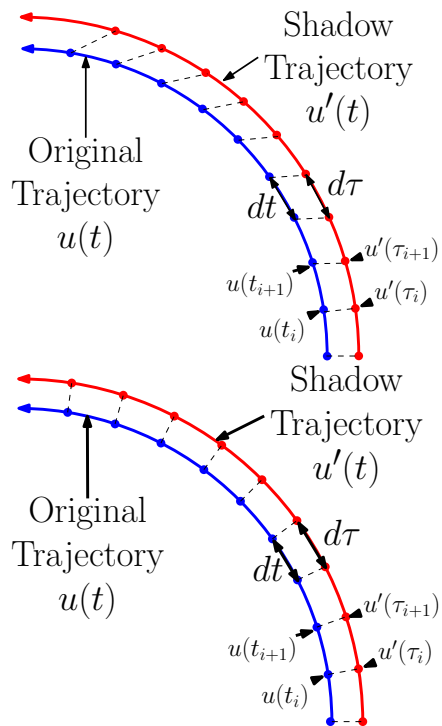


Figure 6: TOP: Original and shadow phase space trajectories without any time transformation ($d\tau/dt = 1$). BOTTOM: Original and shadow phase space trajectories with a time transformation $d\tau/dt = 1 + \eta$ that minimizes the distance between the two trajectories in phase space for all time.

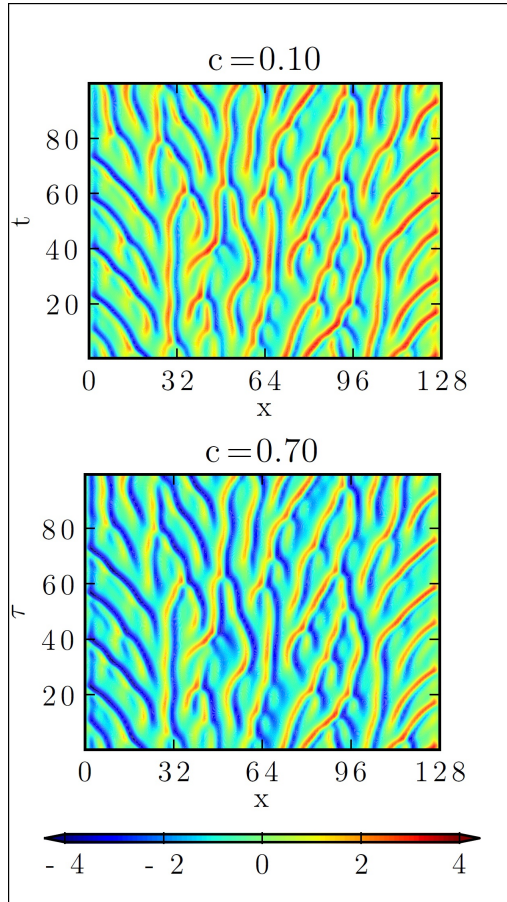


Figure 7: TOP: x vs. t plot for $u(x, t)$ for $c = 0.1$. BOTTOM: x vs. τ plot for the corresponding shadow trajectory $u_s(x, \tau)$ with $c = 0.7$. This shadow trajectory was computed using the algorithm outlined in [14].

5. Results

5.1. Shadowing for the Modified Kuramoto-Sivashinsky Equation

The concept of a shadow trajectory in phase space is easy to visualize for low dimensional dynamical systems such as the three degree of freedom Lorenz system. This is not the case for high dimensional systems, like our simulation of the modified K-S equation, which has between 127 and 511 dimensions. To understand what shadowing entails for a high dimensional ODE or a PDE, consider figure 7. We see that the shadow trajectory $u_s(x, \tau)$ has a lower average value than $u(x, t)$ from the coloring of the contours. Additionally, we see that $u_s(x, \tau)$ has spatio-temporal structures that has very similar to $u(x, t)$. This is because the spatio-temporal structures of $u_s(x, \tau)$ shadow those in $u(x, t)$. Also, note that the time scale of $u_s(x, \tau)$ is indistinguishable from $u(x, t)$, indicating that there is very little time dilation.

5.2. Least Squares Shadowing Sensitivity Analysis

As discussed in section 4, we can use shadow trajectories to compute sensitivities of long-time averaged quantities to system parameters. Figure 8 shows this for $\langle \bar{u} \rangle$ versus c . The gradients computed using 100 and 1000 time unit intervals in the light turbulence dominated region ($0 \leq c \leq 1.2$) are a good match for $\langle J \rangle = \langle \bar{u}^2 \rangle$, but slightly over-predict the magnitude of the gradient for $\langle J \rangle = \langle \bar{u} \rangle$. The insensitivity of $\langle \bar{u} \rangle$ and $\langle \bar{u}^2 \rangle$ to c in the steady region $c \geq 1.8$ is also computed by the LSS method.

However, LSS computes inaccurate gradients in the convection dominated regime ($1.2 < c \leq 1.8$) for both objective functions considered. Possible reasons for this, and slight error in the sensitivity of $\langle \bar{u} \rangle$ are discussed in the next section.

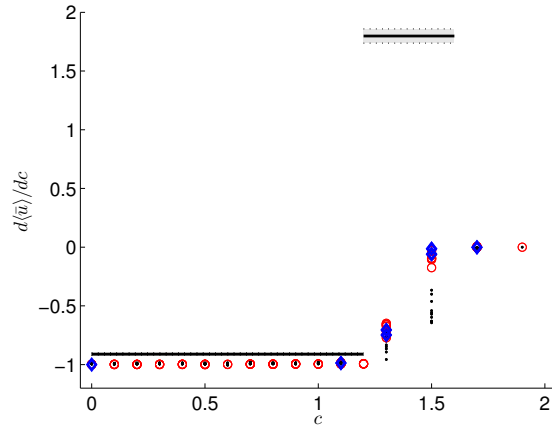
5.3. Sources of error

One source of error is the difference between the infinite time shadowing direction and its approximation by LSS. Our numerical shadowing direction $v_{num}(x, t)$ is just a finite time approximation of the infinitely long shadowing direction $v_{\infty}(x, t)$. It has been proven that the errors of this approximation are the greatest at the beginning ($t = 0$) and end ($t = T$) of the numerical shadowing direction [13]. Additionally, these errors decay exponentially in time. The $t = 0$ error decays forward in time at the rate of the smallest negative Lyapunov exponent. The $t = T$ error decays backward in time at the rate of the smallest positive Lyapunov exponent. Therefore, $v_{num}(x, t)$ most accurately approximates $v_{\infty}(x, t)$ in the middle of the time interval it is computed on [13].

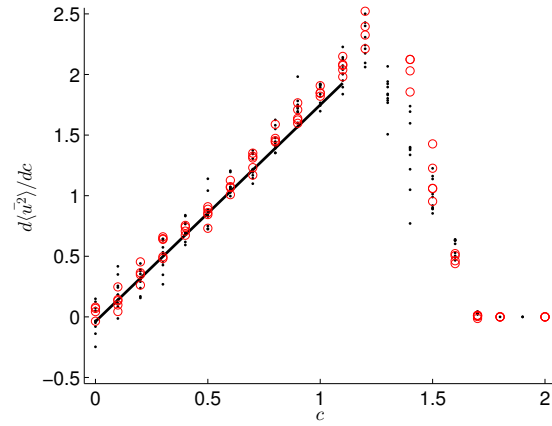
This property of LSS suggests that the attractor associated with the convection dominated region has some Lyapunov exponents with magnitudes smaller than those in the light turbulence dominated regime. This is shown to be the case in figure 9. The low magnitude Lyapunov exponents cause the error in $v(x, t)$ for c values in the convection dominated regime to decay very slowly as t increases from $t = 0$. This slower convergence can be seen in figure 10.

Additionally, figure 10 shows that the slower time scale oscillations associated with the shadowing direction $v(x, t)$ are greater in magnitude in the convection dominated region than those present in the light turbulence region. These longer time-scales require the use of a longer time interval to compute accurate gradients, as a low frequency periodic function would require a longer time interval to compute an accurate average. The longer time scales observed in $v(x, t)$ in the convection dominated regime arise because of the smaller magnitude Lyapunov exponents, whose reciprocals are related to the time scales of $v(x, t)$.

Overall, the convection dominated regime has a smaller magnitude non-zero Lyapunov exponent than the smallest non-zero Lyapunov exponent in the light turbulence dominated regime. This can contribute to the



(a) $J = \bar{u}$



(b) $J = \bar{u}^2$

Figure 8: (a) $d\langle \bar{u} \rangle / dc$ versus c and (b) $d\langle \bar{u}^2 \rangle / dc$. Each simulation was run for $T_0 = 1000$ time units before the LSS method was applied. Black dots indicate gradients computed with a 100 time unit intervals, the red circles indicate a 1000 time unit interval and the blue diamonds indicate a 4000 unit time interval respectively. The black lines are slopes of linear or quadratic regressions from the data in figure 1. The three σ error intervals, where σ is the standard error of the slope, are indicated by black dotted lines. $u(x, t)$ was solved for with $n = 511$ nodes, but LSS was conducted using $n = 127$ nodes.

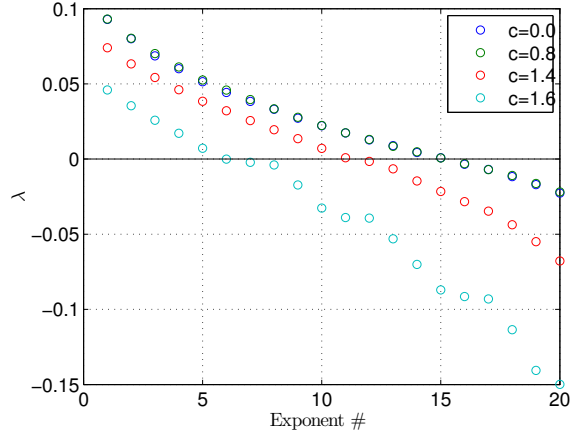


Figure 9: Spectrum of Lyapunov exponents λ for different values of c . Note that the spectra for c values in the convection dominated regime have two values very close to zero, as opposed to just one. One of these corresponds to the neutrally stable exponent $\lambda = 0$, the other is just a very small magnitude exponent. The exponents were computed using the method described by Benettin et al. [15], with $s = 100$ and $k = 20000$. Note that a time step size of $\Delta t = 0.05$ was used.

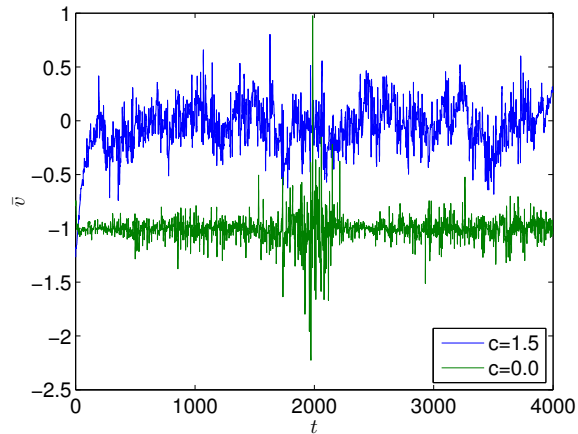


Figure 10: Spatially averaged shadowing direction \bar{v} versus t for $c = 0$ and $c = 1.5$. Each simulation was run for $T_0 = 1000$ time units before the LSS method was applied. While the mean shadowing direction for $c = 0$ converges to a quasi-steady state almost instantly, it takes until around $t = 300$ for the $c = 1.5$ shadowing direction to do so. Finally, note the much larger magnitude oscillations and time scales associated with $c = 1.5$.

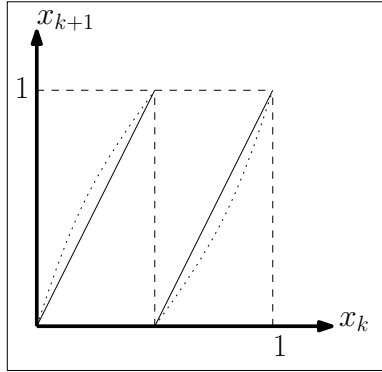


Figure 11: Sawtooth map. The map for $s = 0$ is indicated by the solid lines, the map for $s = 0.1$ is indicated by the dotted lines. Dashed lines indicated the limits of phase space on the x_k and x_{k+1} axes.

fact that the sensitivity computed over the same time interval is less accurate in the convection dominated regime.

However, even when a larger time interval is used, the gradients in the convection dominated regime appear to converge to the wrong value, as shown in figure 8. Therefore, small magnitude Lyapunov exponents are not the only source of error. To explore the other source of error present, we consider a different dynamical system, the sawtooth map:

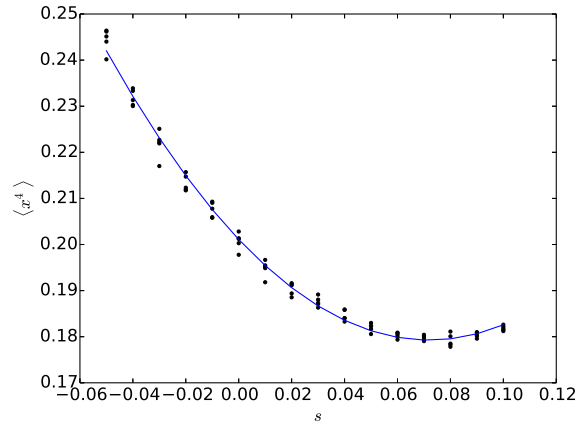
$$x_{k+1} = F(x_k; s) = \begin{cases} 2x_k + s \sin 2\pi x_k, & x_k \in [0, 0.5] \\ 2x_k - 1 + s \sin 2\pi x_k, & x_k \in (0.5, 1.0] \end{cases} \quad (7)$$

For a visual representation of equation (7), refer to figure 11. The quantity of interest we consider for the sawtooth map is long time averaged x^4 :

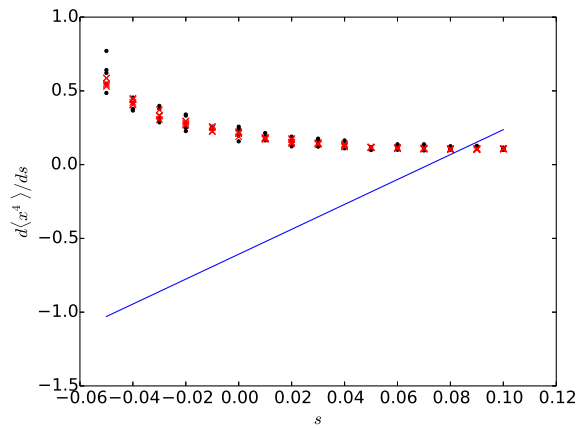
$$\langle x^4 \rangle = \frac{1}{n} \sum_{k=1}^n x_k^4 \quad (8)$$

In figure 12, we see that the gradients computed via LSS converge as n is increased from 1000 to 10000, but they do not converge to the proper value, indicated by the slope of a 2nd order polynomial curve fit through the data in figure 12 (a). This error is reminiscent of that observed for the K-S equation in the convection dominated regime.

To find the source of this error, we consider the stationary distribution of the sawtooth map. The stationary distribution, $\rho(x; s)$, is a distribution in phase space, or $x \in [0, 1]$ for the sawtooth map. The value of $\rho(x; s)$ corresponds to how often a trajectory $\{x_1, x_2, \dots, x_n\}$ passes through x , for a given value of s . It can be shown that time averaged quantities can be written as averages in phase space using the



(a) Quantity of Interest



(b) Gradients

Figure 12: (a) $\langle x^4 \rangle$ realizations for difference $n = 1^5$ length trajectories. The solid line is a 2nd order polynomial curve fit is. (b) $d\langle x^4 \rangle/ds$ realizations for $n = 1^3$ and $n = 1^4$ length trajectories in black and red, respectively. The solid line is the first derivatives of the curve fit in (a).

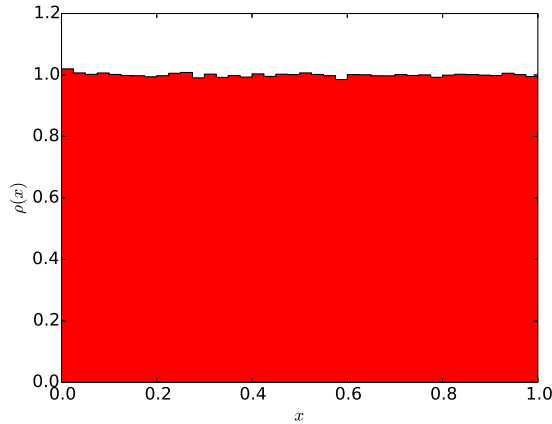


Figure 13: Typical empirical stationary distribution for the sawtooth map with $s = 0$. The distribution was formed using a $n = 1 \times 10^6$ length sequence of x_k 's.

stationary distribution [16], for instance:

$$\langle x^4 \rangle = \frac{1}{n} \sum_{k=1}^n x_k^4 = \int_0^1 x^4 \rho(x; s) dx \quad (9)$$

From equation (9), we can see that variations in $\langle x^4 \rangle$ with s can be expressed as variations of $\rho(x)$ with s . Therefore, sensitivities of $\langle x^4 \rangle$ can be expressed in terms of sensitivities of the distribution [16]:

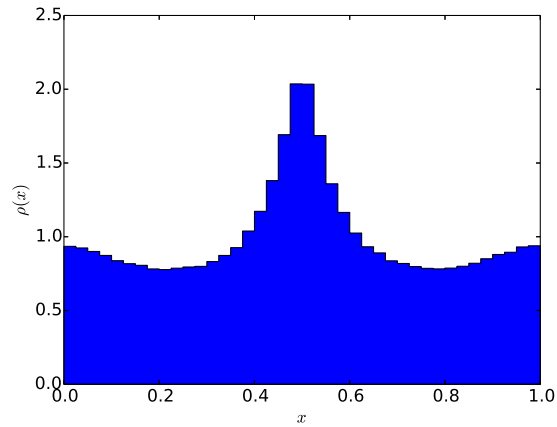
$$\frac{d \langle x^4 \rangle}{ds} = \lim_{\Delta s \rightarrow 0} \left(\int_0^1 x^4 \frac{(\rho(x; s + \Delta s) - \rho(x; s))}{\Delta s} dx \right) = \int_0^1 x^4 \frac{\partial \rho}{\partial s} dx \quad (10)$$

We compute $\rho_s(x)$ empirically with the following algorithm:

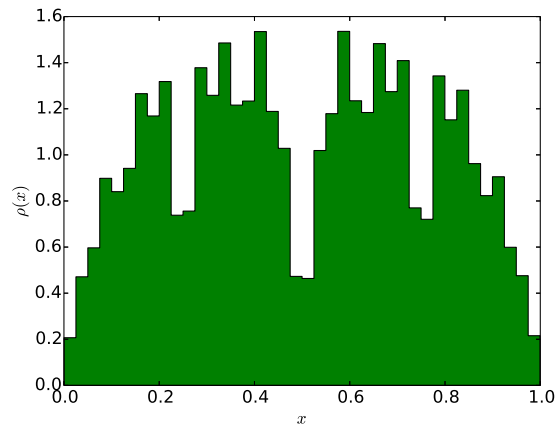
1. Compute a very long trajectory $\{x_1, x_2, \dots, x_n\}$. $n = 1 \times 10^6$ is sufficient for the sawtooth map.
2. Divide phase space into equally sized intervals. 40 were used for the results presented in this paper.
3. Compute the frequency that $\{x_1, x_2, \dots, x_n\}$ falls in each interval and normalize it with n to form a histogram.

For the case $s = 0$, figure 13 shows that the stationary distribution is approximately uniform ($\rho(x) = 1$). From figure 14 a), we see that the distribution changes as s is increased. Now, when we compute the $s = 0$ shadow trajectory of the $s = 0.1$ trajectory shown in figure 14 a), we expect to see a uniform distribution, as in figure 13. However, this is not the case.

From figure 14 b), we see the stationary distribution $\rho(x)$ of the shadow trajectory is not uniform at all. Therefore, the expression for the derivative in equation (10) is incorrect, as observed in figure 12. In other



(a) Reference Trajectory, $s = 0.1$



(b) Shadow Trajectory, $s = 0.0$

Figure 14: Typical empirical stationary distribution for the sawtooth map with $s = 0.1$ and the stationary distribution for the corresponding shadow trajectory with $s = 0.0$. Both distributions were formed using a $n = 1 \times 10^6$ length sequence of x_k 's.

words, a shadow trajectory with $s = s_1$ of some reference trajectory with $s = s_0$ does not necessarily have the same long time averaged quantities as a typical trajectory with $s = s_1$.

Obtaining different long time averaged quantities from solutions with the same values of s (or c for the K-S equation) contradicts the assumption of ergodicity. It can be shown that LSS will compute correct gradients if the long time averaged quantities are the same for all initial conditions [13]. However, in practice long time averaged quantities are only the same for *almost all* initial conditions. For example, the long time averaged quantities of the K-S equation are all zero for the initial condition $u(x, 0) = 0$ when $c = 0.5$, which is different from the values in figure 1. Also, if we consider a trajectory of the sawtooth map starting at $x = 1.0$ will stay at $x = 1.0$ for all iterations, resulting in $\langle x^4 \rangle = 1$, regardless of the value of s , in contrast to the trend shown in figure 12. Both the K-S equation and the sawtooth map are not strongly ergodic. That is, the same long time averaged quantities are obtained for *almost all*, not *all* initial conditions, so computation of the correct gradient is not guaranteed.

It is important to emphasize that LSS has been used to compute very accurate gradients for a number of smaller systems, including the Lorenz system [8], and that it computes fairly accurate gradients for the K-S equation when $|c| \leq 1.2$ and $|c| \geq 1.7$. Therefore, it seems that this breakdown in LSS occurs in largely varying degrees. Further work needs to be done to determine what specific properties of the sawtooth map cause this issue to occur.

6. Conclusion

In conclusion, the LSS method computes accurate gradients for a wide range of system parameter values for a modified K-S equation. In particular, it was found that LSS worked very well in the light turbulence dominated regime, especially for the quantity of interest $\langle \bar{u}^2 \rangle$. However, the method slightly over-predicted the magnitude of gradients of the quantity $\langle \bar{u} \rangle$ in the light turbulence regime and falsely predicted all sensitivities in the convection dominated regime. This breakdown in the method is also observed for smaller systems, such as the sawtooth map. It appears that while LSS works very well for some systems, including the Lorenz system and the K-S equation for certain values of c , it does not work very well or at all for other systems.

Future work needs to be done to determine what properties of the sawtooth map and others like it cause LSS to break down. The lessons learned from working with the sawtooth map can potentially be applied to making LSS more robust and learning what classes of PDEs and larger systems LSS is best suited for.

Acknowledgements

The authors would like to acknowledge AFSOR Award F11B-T06-0007 under Dr. Fariba Fahroo, NASA Award NNH11ZEA001N under Dr. Harold Atkins as well as financial support from the NDSEG fellowship.

References

References

- [1] Y. Kuramoto, T. Tsuzuki, Persistent propagation of concentration waves in dissipative media far from thermal equilibrium, *Prog. Theor. Phys.* 55 (1976) 356–369.
- [2] Y. Kuramoto, Diffusion-induced chaos in reactions systems, *Suppl. Prog. Theor. Phys.* 64 (1978) 346–367.
- [3] G. Sivashinsky, Nonlinear analysis of hydrodynamic instability in laminar flames, part i. derivation of basic equations, *Acta Astronautica* 4 (1977) 1177–1206.
- [4] G. Sivashinsky, Nonlinear analysis of hydrodynamic instability in laminar flames, part ii. numerical experiments, *Acta Astronautica* 4 (1977) 1207–1221.
- [5] G. Sivashinsky, D. Michelson, On irregular wavy flow of a liquid film down a vertical plane, *Progr. Theoret. Phys.* 63 (1980) 2112–2114.
- [6] D. Lea, M. Allen, T. Haine, Sensitivity analysis of the climate of a chaotic system, *Tellus* 52A (2000) 523–532.
- [7] Q. Wang, Forward and adjoint sensitivity computation of chaotic dynamical systems, *Journal of Computational Physics* 235 (2013) 1–13.
- [8] Q. Wang, R. Hui, P. Blonigan, Least Squares Shadowing sensitivity analysis of chaotic limit cycle oscillations. Accepted for publication in the *Journal of Computational Physics*. Preprint available at arXiv:1204.0159 (2013).
- [9] S. Y. Pilyugin, Shadowing in dynamical systems. *Lecture Notes in Mathematics* 1706.
- [10] J. M. Hyman, B. Nicolaenko. The kuramoto-sivashinsky equation: A bridge between pde’s and dynamical systems, *Physica D: Nonlinear Phenomena* 18:1-3 (1986) 113–126.
- [11] C. Brummitt, J. Sprott, A search for the simplest chaotic partial differential equation, *Physics Letters A* 373 (2009) 2717–2721.
- [12] D. Cavaglieri, P. Beyhaghi, T. Bewley. Low-storage IMEX Runge-Kutta schemes for the simulations of Navier-Stokes systems, 21st AIAA Computational Fluid Dynamics Conference, June 2013, San Diego, CA.
- [13] Q. Wang. Convergence of the least squares shadowing method for computing derivative of ergodic averages. *SIAM Journal on Numerical Analysis* 52 (2014) 156170. arXiv:1304.3635
- [14] Q. Wang, S. Gomez, P. Blonigan, A. Gregory, E. Qian. Towards scalable parallel-in-time turbulent flow simulations, *Physics of Fluids* 25 (11) (2013) 110818, arXiv:1211.2437.
- [15] G. Benettin, L. Galgani, A. Giorgilli, J.-M. Strelcyn. Lyapunov characteristic exponents for smooth dynamical systems and for hamiltonian systems; a method for computing all of them. part 2: Numerical application, *Meccanica* 15 (1) (1980) 21–30.
- [16] P. Blonigan, Q. Wang. Probability density adjoint for sensitivity analysis of the mean of chaos, submitted to the *Journal of Computation Physics*. In revision. Preprint available at arXiv:1306.3800 (2013).