The Most Minimal Seed for Transition to Turbulence in Shear Flow

Geoff Stanley

September 30, 2014

Abstract

A key question to develop our understanding of turbulence in shear flows is: what is the smallest perturbation to the laminar flow that causes a transition to turbulence, and how does this change with the Reynolds number R? Finding this so-called "minimal seed" is as yet unachievable in direct numerical simulations of the Navier-Stokes equations, but there exist low-dimensional dynamical systems that model those aspects of the full flow which are considered essential to turbulence. We search for the minimal seed in one such model, owing to Waleffe (1997). We employ an optimization technique, reviewed by Kerswell *et al.* (2014), to find non-linear optimal perturbations, with some modifications. In particular we apply the technique to a new regime in which the edge that is sought is an internal boundary to the basin of attraction of the laminar flow state. We find such a boundary in Waleffe's model and calculate the minimal seed on this edge.

1 Introduction

1.1 Shear Turbulence

It is well known that the Navier-Stokes equations for fluid flow present one of the greatest challenges of classical physics. The equations in general have not been solved analytically, and even numerical solutions have been unobtainable until recent decades owing to the vast range of scales that must be captured in a turbulent flow. One aim of study has been to consider the simplified case of a constant density, incompressible fluid with simple boundary conditions such as for flow through a pipe first studied by Reynolds, or between two concentric cylinders as first studied by Couette. Yet fluid flow even in these idealized settings remains far from understood. One major problem is how a laminar flow transitions to a turbulent flow.

For instance, plane Poiseuille flow[1] (PPF) admits a laminar flow that is linearly stable for all $R < R_c = 5772.22$, where $R = LU/\nu$ is the Reynolds number, R_c is the critical Reynolds number, L is the half-width of the channel, U is the maximum flow speed, and ν is

[1] This is flow between two infinite parallel plane sheets with a no-slip boundary condition and forced by a pressure gradient also parallel to the sheets.

the kinematic viscosity (Orszag, 1971). Yet laboratory experiments of such flows typically observe the laminar flow transitioning to turbulence at much lower Reynolds numbers, around $R \approx 1000$. A similar scenario exists for plane Couette flow[2] (PCF): the laminar flow is linearly stable for all R out to ∞ but turbulence is typically observed with Reynolds numbers as low as $R \approx 350$. This disagreement between experiment and linear stability theory ultimately owes to noise in the experimental setup: finite perturbations to the linear flow can, in dynamical systems parlance, move the system out of the basin of attraction of the stable laminar flow and into the basin of attraction of some other, turbulent state. Since the basin of attraction of the laminar flow diminishes in size with increasing R, a given experimental setup subject to a certain amount of noise will observe transition to turbulence at some sub-critical R.

Reversing this problem presents a key question, as posed in an illuminating paper by Trefethen *et al.* (1993): for a given R, what is the minimal perturbation to the laminar flow that causes the flow to transition to turbulence? This perturbation is called the "*minimal seed*", although sometimes this term simply refers to the magnitude of the actual perturbation. These authors conjectured that the magnitude of the minimal seed should be of order R^γ for some value of γ which will depend on the problem under consideration. For shear flows with stable eigenvalues for the laminar flow, they state $\gamma \leq -1$ citing a technical report that was later published (Kreiss *et al.*, 1994).

Following this conjecture there has been significant work to determine γ for different flows, using theory, laboratory experiments, direct numerical simulation, and low-dimensional model analogues. Many studies focus on a particular form of perturbation that is thought to be especially effective at generating turbulence—effective in the sense that a small energy injection to this mode then excites other modes in such a way as to amplify the total energy. By careful study of the asymptotics of flows with initial conditions in these particular modes,Chapman (2002) derived a R^{-1} scaling for PCF and $R^{-3/2}$ for PPF.

In the laboratory, Hof *et al.* (2003) searched for the minimal seed in pipe Poiseuille flow by injecting fluid through six small drill holes equally spaced around the circumference of the pipe downstream of where the flow was deemed to be laminar. The magnitude and total time of the injections was controlled and downstream observations made to determine whether the flow became turbulent. They found extremely good agreement with a R^{-1} scaling for the minimal seed, in this case the energy of their injections.

However, there is a common problem with the above studies: they begin with a specified form of initial perturbation, then determine the magnitude of that perturbation which is necessary for transition to turbulence. This is even true in the laboratory where the drill holes are specified and therefore allow only certain forms of perturbations. A complete solution to the minimal seed problem must determine the *form* of the perturbation as well as its magnitude. This is certainly a very difficult problem, and the above studies are important steps towards a full understanding.

Another approach, and the one taken here, is to study low-dimensional analogues of the full Navier-Stokes equations for shear turbulence. In these simpler systems it is possible (though still non-trivial) to find the exact form of the minimal seed. Many such models for

[2]This flow is the same as PPF but is forced instead by a constant velocity difference between the sheets.

various flow geometries were developed in the 1990's; Baggett & Trefethen (1997) compared them and found the particular γ for each. We focus in particular on the model by Waleffe (1997, hereafter W97). Waleffe formed by Galerkin truncation a four-dimensional model of sinusoidal flow[3] and speculated an R^{-1} scaling for the minimal seed of this flow, based on the fact that one component of the lower branch fixed point in the 4D model scales as R^{-1}. However, the minimal seed will involve non-zero values for the other three components; the possibility exists that the minimal seed could scale as R^{γ} with $\gamma < -1$. Waleffe's model also elucidated a "Self-Sustaining Process" that is supposedly a crucial building block of shear turbulence. This will be further discussed in Section 3.1; for now note that streamwise vortices are a critical flow structure in initiating turbulence: a small amount of energy in a streamwise vortex mode can, by advecting the mean shear, lead to a large transient growth of energy. Indeed these or their close cousins constitute the particular form studied intensively, as in the aforementioned studies.

The W97 model is a development of an earlier model (Waleffe, 1995,a) model that was included in the study by Baggett & Trefethen (1997). However, the 1995 and 1997 models are significantly different, and only much later was the first attempt made to find the minimal seed and the particular γ for the W97 model (Cossu, 2005). Cossu employed a three-dimensional search algorithm (having eliminated one dimension as described in Section 3.4). The particular "minimal seeds" found by Cossu are those minimal perturbations (for each R) which lead to *permanent* turbulence. However, initial conditions for the W97 model can be quickly found which demonstrate *transient* turbulence—they behave turbulently for a time but ultimately decay to the laminar state. It is these perturbations which, we argue, constitute candidates for the minimal seed.

To find these minimal seeds (over a range of Reynolds numbers), we will use a new optimization technique, recently reviewed by Kerswell *et al.* (2014) and discussed in Section 2. With some modifications, we apply it to a new class of problems wherein the turbulence is transient, versus those problems illustrated by Kerswell *et al.* or even the region of the W97 model that was studied by Cossu (2005) in which the turbulence is persistent. The W97 model itself and results from the search for its minimal seed are given in Section 3, before concluding in Section 4.

1.2 Definitions and Examples: Optimal Perturbations, The Edge, and the Minimal Seed

We now formulate the problem in dynamical systems terms and give a few brief definitions.

The Navier-Stokes equations are shifted so that the laminar state is at the origin. They may then be written in the following generic form:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = F(\mathbf{x}) = L\mathbf{x} + N(\mathbf{x}), \tag{1}$$

where L is a linear operator and N a non-linear function. This dynamical system defines a new function $\mathbf{X}(t, \mathbf{x}_0)$ that is the solution to (1) which also satisfies $\mathbf{X}(0, \mathbf{x}_0) = \mathbf{x}_0$.

The standard dot product is used together with the L_2 norm $\|\mathbf{x}\|_2$, written simply as $|\mathbf{x}|$. For Navier-Stokes, the state $\mathbf{x}$ represents the discretized velocity field $\mathbf{u}$ and hence the

[3]This is PCF but with free-slip boundary conditions, leading to a sinusoidal crosswise profile of the laminar streamwise flow.

kinetic energy $E = \frac{1}{2}\int \mathbf{u}\cdot\mathbf{u}\ \mathrm{d}V$ is technically given by the quadratic form $\frac{1}{2}\mathbf{x}\cdot\mathbf{x}$. Since this maps one-to-one onto $|\mathbf{x}| = (2\cdot E)^{1/2}$, we often refer to $|\mathbf{x}|$ as simply the energy and sometimes write it as $\mathcal{E}(\mathbf{x})$.

The *orbit* starting from an initial condition $\mathbf{x}_0$ is the set of points $\{\mathbf{X}(t,\mathbf{x}_0)\ :\ t\in[0,\infty)\}$.

A *fixed point* of a dynamical system is an $\mathbf{x}_{FP}$ for which $F(\mathbf{x}_{FP})=0$. The stability of the fixed point is found by computing the eigenvalues and eigenvectors of F linearized about $\mathbf{x}_{FP}$. If all eigenvalues have negative real part, $\mathbf{x}_{FP}$ is stable. If at least one eigenvalue has a positive real part, $\mathbf{x}_{FP}$ is unstable.

The *stable manifold* of the fixed point $\mathbf{x}_{FP}$, written $SM(\mathbf{x}_{FP})$, is an invariant set $\{\mathbf{x}_0\ :\ \lim_{t\to\infty}\mathbf{X}(t,\mathbf{x}_0)=\mathbf{x}_{FP}\}$.

The *basin of attraction* of a set A is $\{\mathbf{x}_0\ :\ \lim_{t\to\infty}\mathbf{X}(t,\mathbf{x}_0)\in A\}$.

For the Navier-Stokes equations shifted about the laminar shear flow solution, as in (1), the linear operator L is stable, having eigenvalues all with negative real part, but is non-normal, i.e. L does not commute with its adjoint. These properties of L imply that any solution of the linearized problem $\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}=L\mathbf{x}$ has $\lim_{t\to\infty}|\mathbf{x}(t)|=0$, but some solutions can exhibit a large transient growth before decay. The non-linear function N is quadratic in $\mathbf{x}$ and conserves the energy $\mathcal{E}$.

The *Linear Optimal Perturbation* (LOP) $\mathbf{x}_{LOP}$ satisfies

$$\max_t\{|\mathbf{X}^{(lin)}(t,\mathbf{x}_{LOP})|\}\ \geq\ \max_t\{|\mathbf{X}^{(lin)}(t,\mathbf{x})|\}\qquad\forall\ |\mathbf{x}|=1 \tag{2}$$

where $\mathbf{X}^{(lin)}(t,\mathbf{x}_0)$ is the solution to the linearized problem $\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}=L\mathbf{x}$, and having $\mathbf{X}^{(lin)}(0,\mathbf{x}_0)=\mathbf{x}_0$. In fact there are two LOPs, with $-\mathbf{x}_{LOP}$ also satisfying the above. In words, $\mathbf{x}_{LOP}$ maximizes the furthest distance from the origin obtained by the orbit of the linearized problem. When L is non-normal, transient growth of $|\mathbf{X}^{(lin)}(t,\mathbf{x}_0)|$ is possible even when all eigenvalues have negative real part (e.g. Farrell, 1988).

We define the *Non-Linear Optimal Perturbation* (NLOP) $\mathbf{x}_{NLOP}$, for a given d and T, to satisfy $|\mathbf{x}_{NLOP}|=d$ and

$$|\mathbf{X}(T,\mathbf{x}_{NLOP})|\ >\ |\mathbf{X}(T,\mathbf{x})|\qquad\forall\ |\mathbf{x}|=d. \tag{3}$$

That is, the NLOP maximizes, over all initial conditions with magnitude d, the distance from the origin of the orbit at time T.

The *Edge* E is the boundary of Ω, the basin of attraction of the origin. Recall that the boundary of a set Ω is the set defined as $\{\mathbf{x}\ :\ \forall\ \epsilon>0,\ B_\epsilon(\mathbf{x})\cap\Omega\neq\varnothing$ and $B_\epsilon(\mathbf{x})\cap\overline{\Omega}\neq\varnothing\}$, where $B_\epsilon(\mathbf{x}_0)=\{\mathbf{x}\ :\ |\mathbf{x}-\mathbf{x}_0|<\epsilon\}$, and $\overline{\Omega}$ is the compliment of Ω. That is, for a point $\mathbf{x}\in E$ and any $\epsilon>0$, the ball of radius ϵ centred at $\mathbf{x}$ contains a point in Ω and another point not in Ω. For all problems considered here, it will be a co-dimension one invariant manifold.

A *Strong Edge* is an edge, or subset of an edge, which separates orbits on one side which return to the origin from orbits on the other side which do not.

A *Weak Edge* is an edge, or subset of an edge, for which the orbits of initial conditions on either side both return to the origin but do so in a qualitatively different way: orbits started on one side return to the origin directly while orbits started on the other side take a more circuitous route and/or require more time to return to the origin.

The Edge is, in general, the union of the Strong Edge and the Weak Edge.

The *Minimal Seed* $\mathbf{x}_{MS}$ is the initial condition on the Edge that is closest to the laminar fixed point (the origin): $\mathbf{x}_{MS} \equiv \text{argmin}_{\mathbf{x}}\{|\mathbf{x}| \,:\, \mathbf{x} \in E\}$.

1.3 A 2D Example

An illustrative two-dimensional dynamical system was explored by Lebovitz (2012). The equations are:

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = -\delta x_1 + x_2 + x_1 x_2 - 3x_2^2, \tag{4}$$

$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = -\delta x_2 - x_1^2 + 3x_1 x_2 \tag{5}$$

where $\delta = 1/R$. The origin is a fixed point for all R. A saddle-node bifurcation occurs at $R = 2$: hence for $R > 2$ an additional pair of fixed points exists, the lower branch point $\mathbf{x}_{LB}$ and the upper branch point $\mathbf{x}_{UB}$. For a full bifurcation analysis, see Lebovitz (2012).

Let us consider this example system with $R = 2.45$, as shown in Figure 1. The LOP shows transient growth: its orbit at a particular time is actually further from the origin than where it began at a radius d from the origin. Most other orbits started at radius d do not have this property. Note that because $L(a\mathbf{x}) = aL\mathbf{x}$ for any scalar a, trajectories of the linearized system have a scale invariance, and hence so too does the LOP.

Four NLOPs are shown. In each case it is seen that its orbit (magenta curve) finishes (at time T) further from the origin than other orbits (black curves) started the same distance from the origin.

Here, the stable manifold of the lower branch point $SM(\mathbf{x}_{LB})$ extends from $\mathbf{x}_{LB}$ in two directions: positive in x_1 and negative in x_1, roughly speaking. The first extends to $+\infty$ in x_1, while the second winds around $\mathbf{x}_{UB}$ and hence does not extend to $-\infty$ in x_1. Thus, $SM(\mathbf{x}_{LB})$ does not divide phase space into two separate regions. Orbits started on either side of $SM(\mathbf{x}_{LB})$ eventually return to the origin, but they do so by qualitatively and significantly different paths. Thus, $SM(\mathbf{x}_{LB})$ forms the weak edge. Orbits started below the weak edge return to the origin with at most one instance of transient growth, which can be thought of either as due to the linearized dynamics or due to the orbit, started near $SM(\mathbf{x}_{LB})$, must move to near to $\mathbf{x}_{LB}$. Orbits started above $SM(\mathbf{x}_{LB})$ must wind their way around $\mathbf{x}_{UB}$ before returning to the origin.

The strong edge is, clearly, the periodic orbit P surrounding $\mathbf{x}_{UB}$. Orbits started outside P eventually return to the origin, while those started inside P eventually reach $\mathbf{x}_{UB}$.

The edge, being the union of its strong and weak components, is therefore the union of $SM(\mathbf{x}_{LB})$ and P.

The Minimal Seed $\mathbf{x}_{MS}$ is that point on the edge with least distance from the origin. This distance $|\mathbf{x}_{MS}|$ happens to be approximately 0.2384. Furthermore, $\mathbf{x}_{MS}$ is the NLOP for $d = |\mathbf{x}_{MS}|$, since $\mathbf{X}(T, \mathbf{x}_{MS})$ will be near $\mathbf{x}_{LB}$ while any other $\mathbf{x}$ with $|\mathbf{x}| = |\mathbf{x}_{MS}|$ has $\mathbf{X}(T, \mathbf{x})$ near the origin, for sufficiently large T. In fact, the $d = 0.2384$ ball is slightly below the edge everywhere, and hence the orbit of its NLOP (third magenta curve moving radially outward from the origin) goes towards $\mathbf{x}_{LB}$ then decreases towards the origin. Increasing this NLOP by a multiplicative factor of 1.0005 yields a new initial condition that is slightly above the edge: its orbit (cyan curve) goes to $\mathbf{x}_{LB}$ then increases (eventually returning to

the origin given enough time, but here it is only shown up to $T = 16$). Thus the exact minimal seed lies somewhere near these points and has magnitude between 0.2384 and 1.0005×0.2384. These bounds can be tightened considerably, but for illustrative purposes we shall leave it at that.

For small d where the linearized dynamics well approximate the full dynamics, the NLOP tends to be near to the LOP (scaled by d). This is not in general true for larger d. Note that it is merely coincidental that here the LOP scaled by $|\mathbf{x}_{UB}|$ is in fact $\mathbf{x}_{UB}$. It is also not in general true that the LOP scaled by 0.2384 is the NLOP with $d = 0.2384$.

2 Optimization Algorithm

In this section we describe the optimization technique, which is the subject of a recent review paper (Kerswell *et al.*, 2014), that we employ to find the NLOP. The basic idea is described in the next section, followed by a more careful discussion of the convergence criteria. We then discuss how this algorithm fits into a bisection of energy levels to find the minimal seed. The reader may find it useful to keep Figure 1 in mind throughout this section.

2.1 Basic Algorithm

The technique described here is an algorithm to determine the NLOP, i.e. the initial condition $\mathbf{x}_0 \equiv \mathbf{x}(0)$ having $|\mathbf{x}_0| = d$ which maximizes $|\mathbf{x}(T)|$. The leftmost branch of the flowchart in Figure 2, beginning with "K" and ending with "A", "U", or "MS", shows the algorithm. The underlying idea is to maximize a Lagrangian

$$\mathcal{L} = \mathcal{L}(\mathbf{x}, \boldsymbol{\nu}, \lambda; d, T) = |\mathbf{x}(T)|^2 + \int_0^T \boldsymbol{\nu} \cdot \left(\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} - L\mathbf{x} - N(\mathbf{x}) \right) \mathrm{d}t + \lambda \left(|\mathbf{x}(0)| - d \right) \tag{6}$$

where $\boldsymbol{\nu}$ and λ are the Lagrange multipliers which respectively ensure that $\mathbf{x}$ is a solution of the dynamical system (1), and that the initial condition lies on a sphere of radius d from the origin, i.e. has a specified energy.

Taking the variation of $\mathcal{L}$ with respect to $\mathbf{x}$ gives

$$\delta\mathcal{L} = \underbrace{[\mathbf{x}(T) + \boldsymbol{\nu}(T)]}_{A} \cdot \delta\mathbf{x}(T) - \int_0^T \underbrace{\left[\frac{\mathrm{d}\boldsymbol{\nu}}{\mathrm{d}t} + \boldsymbol{\nu} \cdot \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \right]}_{B} \cdot \delta\mathbf{x} \, \mathrm{d}t + \underbrace{[\lambda\mathbf{x}(0) - \boldsymbol{\nu}(0)]}_{C} \cdot \delta\mathbf{x}(0) \tag{7}$$

When $\mathcal{L}$ has been maximized, the perturbation $\delta\mathcal{L} = 0$ for an arbitrary perturbation $\delta\mathbf{x}(t)$, $t \in [0, T]$, and thus the quantities labelled A, B, and C must all vanish. Thus, the optimization procedure is as follows:

Step 0: Guess an initial condition $\mathbf{x}^{(n)}(0)$ with $n = 1$.

Step 1: Integrate the dynamical system (1) forwards from $\mathbf{x}^{(n)}(0)$ to obtain $\mathbf{x}^{(n)}(T)$.

Step 2 (Apply $A = 0$): Set $\boldsymbol{\nu}^{(n)}(T) = -\mathbf{x}^{(n)}(T)$.

Step 3 (Apply $B = 0$): Integrate the dual dynamical system,

$$\frac{\mathrm{d}\boldsymbol{\nu}^{(n)}}{\mathrm{d}t} = -\boldsymbol{\nu}^{(n)} \cdot \frac{\partial \mathbf{F}}{\partial \mathbf{x}}, \tag{8}$$

backwards in time[4] from $\boldsymbol{\nu}^{(n)}(T)$ to obtain $\boldsymbol{\nu}^{(n)}(0)$.

Step 4 (Apply C to maximize $\mathcal{L}$): Having now ensured $A = 0$ and $B = 0$, $\mathcal{L}$ varies with $\mathbf{x}(t)$ only through $\mathbf{x}(0)$. Therefore, use

$$\frac{\delta \mathcal{L}}{\delta \mathbf{x}^{(n)}(0)} = \lambda \mathbf{x}^{(n)}(0) - \boldsymbol{\nu}^{(n)}(0) \tag{9}$$

to maximize $\mathcal{L}$ by a maximum ascent method: for some small step-size $\epsilon^{(n)}$, the next iteration's guess for the NLOP is

$$\mathbf{x}^{(n+1)}(0) = \mathbf{x}^{(n)}(0) + \epsilon^{(n)} \left[\lambda \mathbf{x}^{(n)}(0) - \boldsymbol{\nu}^{(n)}(0) \right]. \tag{10}$$

Note that, although $\mathcal{L}$ is being maximized for arbitrary values of the Lagrange multipliers $\boldsymbol{\nu}$ and λ, their specific values are important, and here we must simultaneously choose λ to satisfy

$$d = \left| \mathbf{x}^{(n+1)}(0) \right| = \left| \mathbf{x}^{(n)}(0) + \epsilon^{(n)} \left[\lambda \mathbf{x}^{(n)}(0) - \boldsymbol{\nu}^{(n)}(0) \right] \right| \tag{11}$$

Using the squared L_2 norm $|\mathbf{x}| = \|\mathbf{x}\|_2^2 = \mathbf{x} \cdot \mathbf{x}$, this condition on λ yields an easily solved quadratic equation (dropping (n) superscripts for clarity):

$$(\epsilon^2 \ \mathbf{x}_0 \cdot \mathbf{x}_0)\lambda^2 + (2\epsilon \ \mathbf{x}_0 \cdot \mathbf{x}_0 - 2\epsilon^2 \ \mathbf{x}_0 \cdot \boldsymbol{\nu}_0)\lambda + (\epsilon^2 \ \boldsymbol{\nu}_0 \cdot \boldsymbol{\nu}_0 - 2\epsilon \ \mathbf{x}_0 \cdot \boldsymbol{\nu}_0) = 0. \tag{12}$$

The choice of the step-size ϵ is critically important for convergence and is discussed next. The choice of T is also important and may need to be adjusted as d changes, as discussed in the subsequent section.

2.2 Convergence

How one chooses the step size ϵ is of great practical importance. With ϵ too small, the method will take too long to feasibly complete, while too large an ϵ can lead to $\Delta \mathbf{x}_0 = \mathbf{x}(0)^{(n+1)} - \mathbf{x}(0)^{(n)}$ also being so large that $\Delta \mathcal{L} = \mathcal{L}^{(n+1)} - \mathcal{L}^{(n)} \approx (\delta \mathcal{L}/\delta \mathbf{x}_0) \cdot \Delta \mathbf{x}_0$ is a poor approximation and $\mathcal{L}$ actually decreases. We therefore employ some additional measures, applied in each iteration of the algorithm, to determine a good step-size ϵ, so as to prevent both the above problems.

First, we require $\epsilon \leq \epsilon_{\max,\lambda}$, where the latter is defined as the maximum ϵ for which λ can be real. Requiring $\lambda \in \mathbb{R}$ in turn requires the discriminant of (12) to be non-negative, giving another quadratic equation, this time for ϵ, the solution of which is

$$\epsilon_{\max,\lambda} = \left(\frac{(\mathbf{x}_0 \cdot \mathbf{x}_0)^2}{(\mathbf{x}_0 \cdot \mathbf{x}_0)(\boldsymbol{\nu}_0 \cdot \boldsymbol{\nu}_0) - (\mathbf{x}_0 \cdot \boldsymbol{\nu}_0)^2} \right)^{1/2}, \tag{13}$$

having chosen the positive root so as to maximize, not minimize, $\mathcal{L}$. This criterion is applied at each iteration (n).

[4]This requires knowing $\mathbf{x}(t)$ at each time step (and sub-steps) of the ODE integration method. For high dimensional dynamical systems such as a direct numerical simulation of the Navier-Stokes equations, the full forward integration of $\mathbf{x}(t)$ cannot be stored in memory, and therefore a check-pointing procedure is required, as discussed by Kerswell *et al.* (2014), wherein $\mathbf{x}(t)$ is saved only at some times t_j and these $\mathbf{x}(t_j)$'s are used, when needed in solving for $\boldsymbol{\nu}(t)$, to initiate another forward integration of $\mathbf{x}$ from t_j to t_{j+1}.

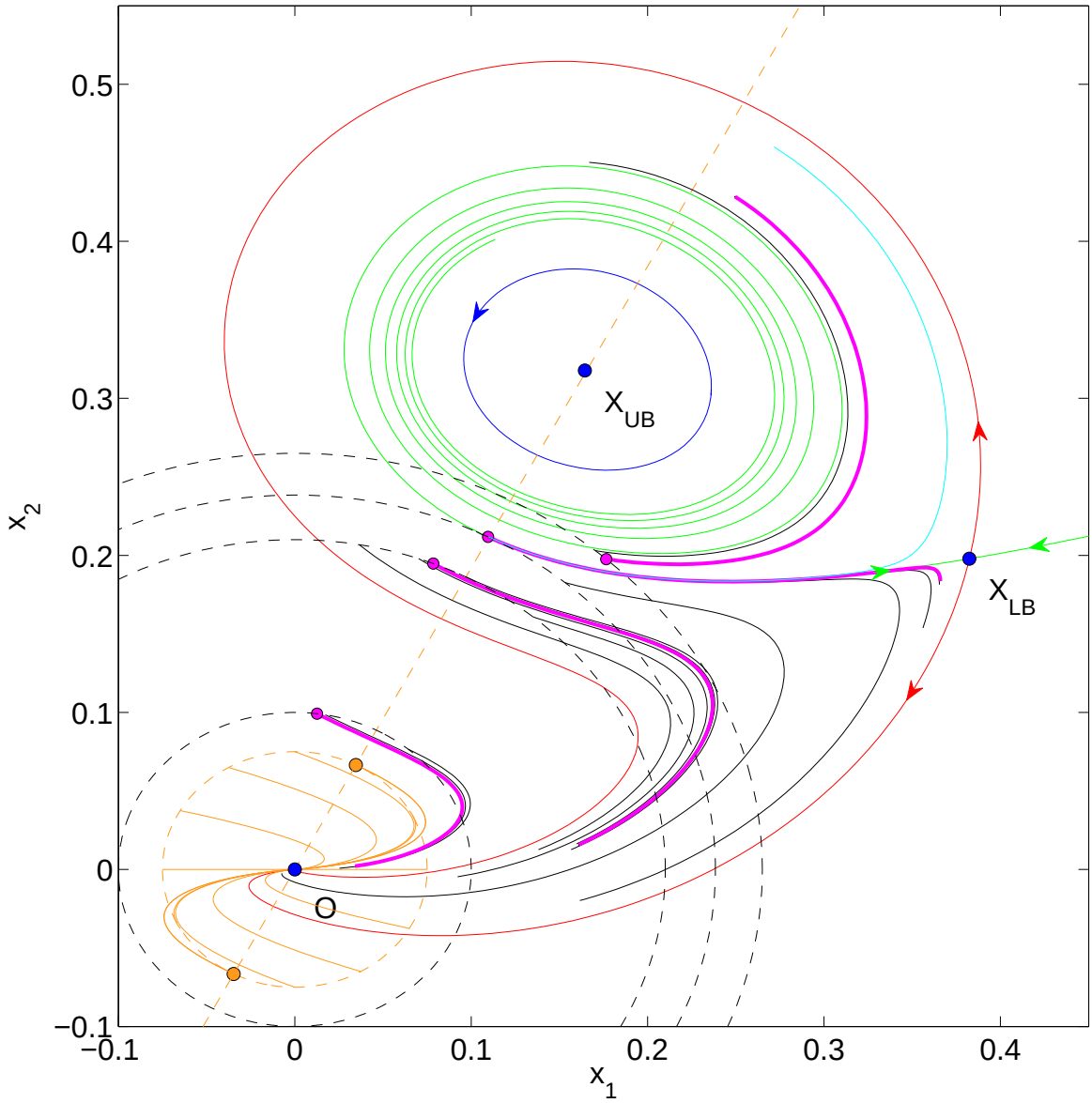

Figure 1: The system (4) with $R = 2.45$. Blue circles indicate the three fixed points: the origin O, the lower branch point $\mathbf{x}_{LB}$, and the upper branch point $\mathbf{x}_{UB}$. The unstable manifold of $\mathbf{x}_{LB}$ (red curves) consists of two orbits, beginning in opposite directions but both converging to O. The stable manifold of $\mathbf{x}_{LB}$ (green curves) winds around an unstable periodic orbit P (blue) which encloses the basin of attraction of the stable fixed point $\mathbf{x}_{UB}$. The LOP and its negative are indicated by the the orange circles, together with their orbits in the linearized system (orange curves). Other orange lines show orbits of the linearized system started at 30° intervals around the origin and at a radius of 0.075 (dashed orange circle). Any point on the dashed orange line is an LOP. The NLOP (magenta circles) is shown for each of $d = 0.1, 0.21, 0.2384, 0.265$ (black dashed circles), with $T = 8, 8, 16, 8$ respectively. The NLOP is calculated by the iterative procedure described in Section 2: successive orbits (black curves) begin with the orbit started at an angle from the $+x_1$ axis of 50° (for this illustration) and converge to the NLOPs orbit (magenta curves). The orbit started from the NLOP with $d = 0.2384$ scaled by a factor 1.0005 is shown (cyan curve).

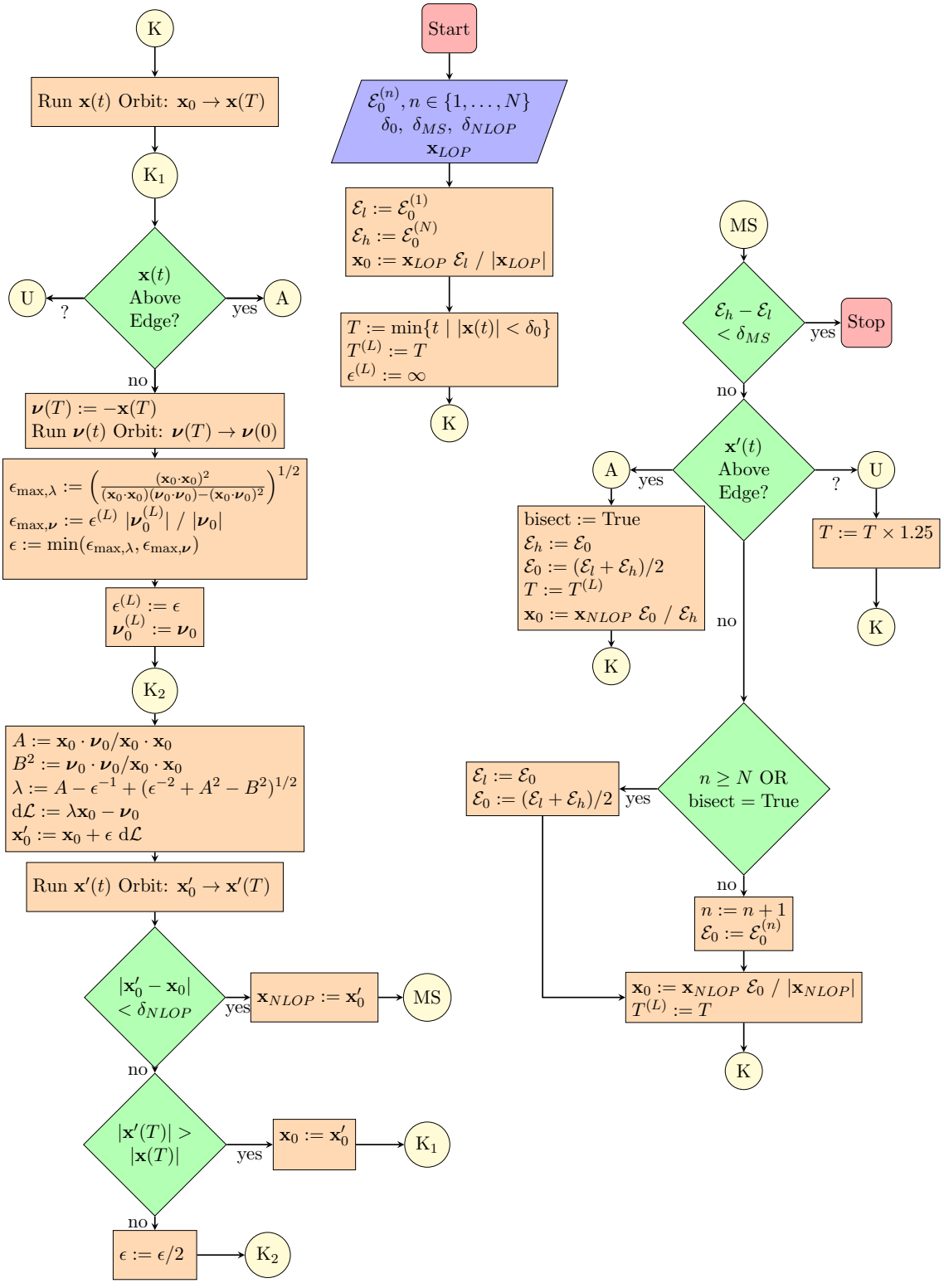

Figure 2: Flowchart showing the optimization procedure to calculate the NLOP (left branch, K to A, U, or MS), the energy incrementation and bisection procedure (right branch, MS to K or Stop), and the initialization for the whole algorithm (centre branch, Start to K).

The above criterion is useful when beginning the iterations, telling us how large a step size we can feasibly take. But it will not bring $\epsilon \to 0$, so we add another criterion which limits the contribution $\boldsymbol{\nu}$ can make to $\Delta\mathbf{x_0}$ in (10):

$$\epsilon^{(n)}_{\max,\boldsymbol{\nu}} = \epsilon^{(n-1)} \frac{|\boldsymbol{\nu}_0^{(n-1)}|}{|\boldsymbol{\nu}_0^{(n)}|}. \tag{14}$$

and require $\epsilon^{(n)} \leq \epsilon^{(n)}_{\max,\boldsymbol{\nu}}$.

Another criteria that was considered was to require that the jump in $\mathbf{x}_0$ between successive iterations does not increase. That is, $\left|\mathbf{x}_0^{(n+1)} - \mathbf{x}_0^{(n)}\right|^2 \leq \left|\mathbf{x}_0^{(n)} - \mathbf{x}_0^{(n-1)}\right|^2 \equiv D_0^{(n)}$. Using (10) this yields another upper bound for $\epsilon^{(n)}$, which is

$$\epsilon^{(n)}_{\max,\mathbf{x}_0} = \left(\frac{D_0(\mathbf{x}_0 \cdot \mathbf{x}_0) - D_0^2/4}{(\mathbf{x}_0 \cdot \mathbf{x}_0)(\boldsymbol{\nu}_0 \cdot \boldsymbol{\nu}_0) - (\mathbf{x}_0 \cdot \boldsymbol{\nu}_0)^2} \right)^{1/2}, \tag{15}$$

all evaluated at iterate (n). However, this was found to be too restrictive. Consider when the initial guess $\mathbf{x}_0$ begins in a fairly "flat" region of phase space (i.e. where $|\mathbf{x}(T)|$ changes little with changes in $\mathbf{x}(0)$) and is also not near the local maxima of $|\mathbf{x}(T)|$. Including this restriction forces all subsequent steps after the initial step in $\mathbf{x}(0)$, which is small owing to flatness, to be at least as small; hence convergence to the NLOP requires an enormous number of iterations. Without this condition, small initial steps are taken, followed by larger steps when $\mathbf{x}(0)$ has left the flat region.

The two criteria used above attempt to limit ϵ but do not themselves ensure that $\mathcal{L}$, and hence $|\mathbf{x}(T)|$, increases with each iteration. Since this is essential, one final condition is added: the dynamical system is integrated forward to obtain $\mathbf{x}^{(n+1)}(T)$ and $\left|\mathbf{x}^{(n+1)}(T)\right| > \left|\mathbf{x}^{(n)}(T)\right|$ is explicitly checked. This requirement becomes more stringent as the iterations proceed. If this check fails, $\epsilon^{(n)}$ is halved, $\mathbf{x}_0^{(n+1)}$ re-calculated from (10), and this criteria checked again. This halving process is repeated until the condition passes, or some maximum number of halvings is reached, currently set to 20.[5]

Not only does this ϵ-halving technique give us confidence that we are finding a local maximum of $\mathcal{L}$, numerical experiments suggest that it may even hasten the whole optimization algorithm by preventing overshooting of the NLOP and thereby reduce the number of iterations (n) required.

The final statement regarding convergence regards when the NLOP is deemed to be found and the full optimization algorithm finishes. During the forward integrations of the ϵ-halving procedure (which also serve as performing Step 1 in Section 2.1 above; see also Figure 2) we determine whether $\left|\mathbf{x}^{(n+1)}(0) - \mathbf{x}^{(n)}(0)\right| < \mathcal{E}_0^{\text{tol}}$, for some small number $\mathcal{E}_0^{\text{tol}}$, for which we tend to use a default value of 10^{-8}. When this is true we declare the NLOP found and finish the algorithm.

[5]For additional numerical safety, subsequent applications of the ODE solver for forward integration of $\mathbf{x}(t)$ at this iteration level (n) will do the following: sub-sample the output $\mathbf{x}(t)$ by a further factor of two, to help ensure the backwards integration of ν is accurate, and raise by tenfold the accuracy of the ODE solver (which reduces the size of the time steps).

2.3 Energy Incrementation and the Two Sides of the Edge

Having a procedure to find the NLOP on a given energy ball $\{\mathbf{x} \ : \ |\mathbf{x}| = \mathcal{E}_0\}$ (and for a given integration time T), we now desire an algorithm that applies the above procedure on successively larger energy spheres, in order to find the minimal seed. The basic idea here is to increase the size of the energy spheres until the edge is passed, and then to refine the energy by bisection down to a desired precision.

2.3.1 Which Side of the Edge?

This procedure will require a function that determines whether a particular orbit is above or below the edge, or whether this is unclear from the information available. The information available is the full orbit $\mathbf{x}(t)$, $t \in [0, T]$, as well as, in principle, all previously calculated orbits, although in practice only the orbits of the previously calculated NLOPs are needed and generally just some metric (such as the gain) of them.

This function will vary between problems. Clearly, if the problem has a strong edge (the basin of attraction of $\mathbf{x}_{UB}$) but not a weak edge, a criteria could be the following: if $|\mathbf{x}(T)| < \delta_O$ then $\mathbf{x}_0$ is below the edge; if $|\mathbf{x}(T) - \mathbf{x}_{UB}| < \delta_{UB}$ then $\mathbf{x}_0$ is above the edge; otherwise, it is unclear whether $\mathbf{x}_0$ is above or below the edge.

However, if the problem has a weak edge which is sought then a different function must be used. One approach is use the ratio of the gain $G \equiv \max_t\{|\mathbf{x}(t)|\} \ / \ |\mathbf{x}(0)|$ between the current orbit and the orbit of the last NLOP determined to be below the edge. If this ratio exceeds a certain value[6] then $\mathbf{x}_0$ is above the edge; if $|\mathbf{x}(T)| < \delta_O$ then $\mathbf{x}_0$ is below the edge; otherwise, it is unclear whether $\mathbf{x}_0$ is above or below the edge.

This approach has been successful in simple problems such as the 2D system studied by Lebovitz (2012) (Section 1.3), and also in the 2D example problem used by Kerswell *et al.* (2014) which only has a strong edge.

However, the geometry of the problem may prevent this gain-based approach from working. This is easily seen if we had $|\mathbf{x}_{LB}| > |\mathbf{x}_{UB}|$. Then two orbits started nearby but on either side of the edge would, as per the usual understanding, go near $\mathbf{x}_{LB}$; then one goes directly back to the origin O while the other goes around $\mathbf{x}_{UB}$ before returning to O. But the gain of both orbits would be approximately $|\mathbf{x}_{LB}|/|\mathbf{x}(0)|$, with no obvious difference between them.

For the W97 4D model, we do have the desirable $|\mathbf{x}_{LB}| < |\mathbf{x}_{UB}|$ but the separation is not so large and a formulation in terms of the gain, as above, was not robustly successful. The following, more reliable approach, was found instead. Find all local maxima of the function $|\mathbf{x}(t)|$, then for numerical reasons discard all those having $|\mathbf{x}(t)| < \delta_0$. If there are one or fewer local maxima—the orbit either directly relaminarizes or does so after a transient growth as per a non-normal linear system—the orbit is deemed below the edge. If there are two or more local maxima—the orbit first undergoes transient growth, then undergoes secondary growth when repulsed on the far side of the saddle that is $\mathbf{x}_{LB}$—and their magnitudes exceed some specified thresholds (for each maxima), the orbit is deemed

[6] This value need not be much great than 1. For the 2D problem of Section 1.3, we choose a value of 1.2. If the incremental increase of the energy spheres is small enough then, for two spheres below the edge, the ratio of the gains of their NLOPs will be nearly 1, while for two spheres straddling the edge this ratio grows without bound as the distance between the energy spheres shrinks.

above the edge. The thresholds chosen were 0 for the first peak (any size peak will do), and $|\mathbf{x}_{LB}| + \max(\delta_{LB}, 0.2 \cdot |\mathbf{x}_{UB} - \mathbf{x}_{LB}|)$, where δ_{LB} is defined below. This function, together with another improvement detailed below, is illustrated as a flowchart in Figure 3.

When, in all examples above, it is deemed unclear whether the orbit is above or below the edge, the correct thing to do is to choose a larger T. Clearly this works as desired in the first example above when all of phase space (except for a measure zero set, the stable manifold of the lower branch point) is in the basin of attraction of either O or $\mathbf{x}_{UB}$. In the second example, the added time may allow further gain to be realized (supporting an above-edge result), or it may simply bring the orbit closer to the origin (supporting a below-edge result). Thus, increasing T is still the correct response, but one must have carefully chosen the threshold for the ratio of the gains, based on previous experience with the particular problem.

When, by bisection of energy levels, the orbits under question are very near the minimal seed, they will track the stable manifold of $\mathbf{x}_{LB}$ until near $\mathbf{x}_{LB}$ itself; this point being a fixed point, velocities near $\mathbf{x}_{LB}$ are small so the orbit lingers for much time near $\mathbf{x}_{LB}$. If increasing T sufficiently to escape $\mathbf{x}_{LB}$ is infeasible, a complimentary condition may be added that applies when the orbit enters a small ball $\delta_{LB} \ll 1$ of $\mathbf{x}_{LB}$ wherein the linearized dynamics dominate. First, linearize F about $\mathbf{x}_{LB}$: $F_{LB} = \partial F/\partial \mathbf{x}\Big|_{\mathbf{x}_{LB}}$. Second, find all stable eigenvectors of F_{LB} and (as by Gram-Schmidt) orthonormalize them to get $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$. Third, calculate the vector rejection of the unstable eigenvector $\mathbf{v}_n$ of F_{LB}:

$$\mathbf{v}_O = \mathbf{v}_n - \sum_{i=1}^{n-1} (\mathbf{v}_n \cdot \mathbf{v}_i)\, \mathbf{v}_i. \tag{16}$$

This $\mathbf{v}_O$ is orthogonal to the hyperplane formed by the stable eigenvectors. Fourth, calculate the orbits starting from $\mathbf{x}_{LB} + \delta_{LB}\mathbf{v}_O$ and $\mathbf{x}_{LB} - \delta_{LB}\mathbf{v}_O$ and decide which is below the edge and which is above. The orbit below the edge may be said to be the one that requires less time to reach a δ_0-ball around the origin[7], or the one that has a lesser gain. If $\mathbf{x}_{LB} - \delta_{LB}\mathbf{v}_O$ is the one below the edge, then re-assign $\mathbf{v}_O := -\mathbf{v}_O$. Hence $\mathbf{v}_O$ points, following the dynamical system's trajectories, from $\mathbf{x}_{LB}$ towards the origin. Now, for any orbit $\mathbf{x}(t)$, find the time t_c which minimizes $D_{LB} = |\mathbf{x}(t_c) - \mathbf{x}_{LB}|$. If $D_{LB} > \delta_{LB}$ then it is unclear from this analysis whether the orbit is above or below the edge. If $D_{LB} < \delta_{LB}$ then the orbit $\mathbf{x}(t)$ is deemed below the edge if

$$S = (\mathbf{x}(t_c) - \mathbf{x}_{LB}) \cdot \mathbf{v}_O \tag{17}$$

is positive, and above the edge if S is negative.

2.3.2 Energy Incrementation

Now the task is to increase the size of the energy spheres upon which the NLOP is calculated, until a sphere intersects the edge. The fundamental outline of the algorithm is, in words, as follows.

[7]There are systems, such as the 2D system of Lebovitz (2012), where the orbit from $\mathbf{x}_{LB}$ that winds around $\mathbf{x}_{UB}$, thereby obtaining a much larger gain, actually relaminarizes (reaches the δ_0-ball centred at the origin for $\delta_0 \ll 1$) in less time.

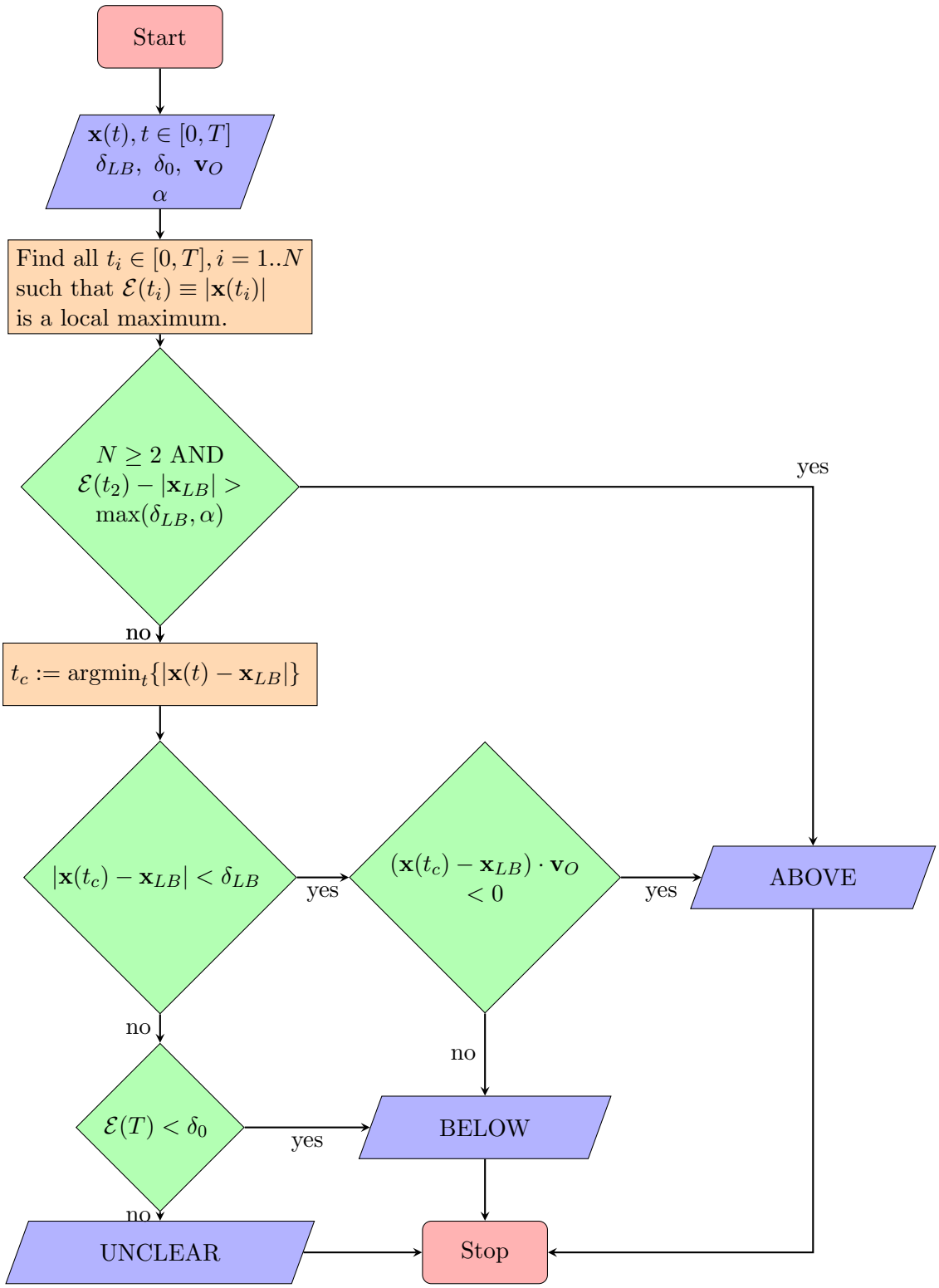

Figure 3: Flowchart to determine whether an orbit $\mathbf{x}(t)$ is above or below the edge, or whether this is unclear. This formulation was used to work on the W97 problem, with $\delta_{LB} = 10^{-2}$, and $\alpha = (|\mathbf{x}_{UB}| - |\mathbf{x}_{LB}|)/5$. δ_0 was initially chosen as 10^{-2} but can be made larger (as discussed in Section 2.3.2) once the minimal seeds for some values of R are available and an approximate scaling law for the minimal seed can be made.

Step 0: Choose a series of energy spheres $\mathcal{E}_l \equiv \mathcal{E}_1, \ldots \mathcal{E}_N \equiv \mathcal{E}_h$. Choose δ_0: an orbit entering the δ_0-ball from outside is deemed going to the origin and below the edge. Set $\mathbf{x}_0 := \mathcal{E}_l\ \mathbf{x}_{LOP}$, where $\mathbf{x}_{LOP}$ is the Linear Optimal Perturbation with unit magnitude. Run the orbit starting at $\mathbf{x}_0$ until the smallest time T such that $|\mathbf{x}(T)| < \delta_0$.

Step 1: Calculate, by the optimization algorithm above, $\mathbf{x}_{NLOP}$.

Step 2: Check whether the orbit from $\mathbf{x}_{NLOP}$ is above or below the edge, or if this is unclear.

If unclear: increase T.

If above: set $\mathcal{E}_h := \mathcal{E}_0$; reduce $\mathcal{E}_0$ by bisection between $\mathcal{E}_h$ and $\mathcal{E}_l$; scale $\mathbf{x}_0$ to $\mathbf{x}_0\ \mathcal{E}_0\ /\ |\mathbf{x}_0|$.

If below: set $\mathcal{E}_l := \mathcal{E}_0$; increase $\mathcal{E}_0$ by bisection between $\mathcal{E}_l$ and $\mathcal{E}_h$ if previously an orbit was found above the edge, or if the $\mathcal{E}_i$ set is exhausted; otherwise set $\mathcal{E}_0$ to the next element in $\mathcal{E}_i$; set $\mathbf{x}_0$ to $\mathbf{x}_{NLOP}\ \mathcal{E}_0\ /\ |\mathbf{x}_{NLOP}|$.

Step 3: Check convergence. If $\mathcal{E}_h - \mathcal{E}_l < \delta_{MS}$ for some specified δ_{MS}, exit the algorithm; otherwise, go to Step 1.

There are several additional points which can improve the basic algorithm, which we now discuss. The flowchart in Figure 2 includes many of these improvements.

In all problems considered so far, the condition for an orbit to be below the edge rests on whether $|\mathbf{x}(T)| < \delta_0$. When beginning this energy incrementing algorithm, this δ_0 must be chosen very small so that there is complete confidence that any orbit which crosses from outside to inside of the δ_0-ball centred at the origin will continue directly to the origin, i.e. with no later growth of $|\mathbf{x}(t)|$. However, this δ_0 can be increased to $\mathcal{E}_l$, the size of the energy sphere upon which the NLOP was last found to be below the edge: if the NLOP on the $\mathcal{E}_l$-ball goes to the origin and an orbit started on the $\mathcal{E}_0$-ball passes from outside to inside of the $\mathcal{E}_l$-ball, then this orbit must also go to the origin.

When an orbit is above the edge, we don't need to find the NLOP for that energy sphere. We simply decrease the energy and try again. Thus, at each iteration in the optimization algorithm, the orbit is passed to the function that determines whether it's above or below the edge. If it's above the edge, then the energy sphere is immediately shrunk ("A" in Figure 2).

Relatedly, if it is unclear whether the orbit is above or below the edge, but the orbit is in fact below the edge, then since further iterations of the optimization procedure will only increase $|\mathbf{x}(T)|$, the orbits of all further iterations will also be deemed unclear. Hence, T should be immediately increased and the optimization procedure restarted. Similarly, if it is unclear whether the orbit is above or below the edge, but the orbit is in fact above the edge, then increasing T will likely accelerate the whole algorithm and at the very least will not lead to an incorrect answer. Hence, shortcut "U" is taken in Figure 2.

When the NLOP is found to be below the edge, it is scaled radially outward to the next energy sphere. But note, when an $\mathbf{x}_0$ is detected above the edge (shortcut "A"), the NLOP is not found and so it is the last NLOP that is scaled to the next energy sphere. $\mathbf{x}_0$ should similarly not be updated when it is unclear whether the orbit of the optimization procedure is above or below the edge (but this case can only occur on the first iteration of the optimization procedure, before $\mathbf{x}_0$ could be updated.)

In the W97 model, the fourth and final component of an orbit that relaminarizes decays to 0 much faster than the other components. This component can quickly become less than

the absolute error tolerance of the ODE solver used to integrate the dynamical system—a situation that can cause numerical issues. Thus, it was found extremely helpful to nudge the initial condition $\mathbf{x}_0$ given to the optimization algorithm. In this case we chose to increase the fourth component to a minimum of 10^{-4}, while scaling down the other components of $\mathbf{x}_0$ such that the total energy $|\mathbf{x}(0)|$ is unchanged.

3 In Search of the Minimal Seed

In this section we give the results obtained by applying the optimization algorithm of Section 2 to the W97 four-dimensional model. We begin by describing this model, then discuss how the optimization technique operates for this model, then show our results and briefly compare them to the previous results of Cossu (2005).

3.1 The Waleffe (1997) 4D Model

While significant progress has been made recently, a complete understanding of shear turbulence arising from the full Navier-Stokes equations, even in simple geometries such as plane Couette flow or plane Poiseuille flow, remains out of reach. When viewed as a dynamical system, such a system would have roughly 10^5 degrees of freedom, obfuscating visualization as well as the implementation of many numerical techniques. A lower dimensional system that still captures the important features of a potentially turbulent shear flow is therefore desired.

One such low dimensional model for plane Couette flow was developed in the 1990's by Waleffe and collaborators, culminating in a 1997 paper (Waleffe, 1997). One feature of this work was to recognize that organized, coherent structures may exist within a turbulent shear flow. These structures, uncovered in their analysis, are streamwise roll vortices, streaks (spanwise perturbations of the streamwise flow), and a streamwise wave dominated by spanwise velocity perturbations. The streamwise rolls advect the mean shear, pulling positive streamwise flow down and lifting negative streamwise flow up, thereby creating streaks. Even for roll vortices that are weak (relative to the mean shear), the streaks represent significant spanwise variation of the streamwise flow and can be linearly unstable, leading to the growth of the streamwise wave. Through careful analysis, Waleffe showed that these waves non-linearly force a streamwise roll vortex very near to the original rolls, thereby amplifying the original perturbation. Thus, this is an energetically "self-sustaining process", thereby avoiding the decay to the laminar state.

In addition to the physical understanding of these structures and their interaction, this work led to a four-dimensional dynamical system. It has been much studied and seems to capture many important features of the full Navier-Stokes equations for shear turbulence.

Writing $\mathbf{x} = (m, u, v, w)^T$, the Waleffe's model in the general form (1) has

$$L = \begin{pmatrix} -k_m^2 R^{-1} & 0 & 0 & 0 \\ 0 & -k_u^2 R^{-1} & \sigma_u & 0 \\ 0 & 0 & -k_v^2 R^{-1} & 0 \\ 0 & 0 & 0 & -k_w^2 R^{-1} \end{pmatrix} \tag{18}$$

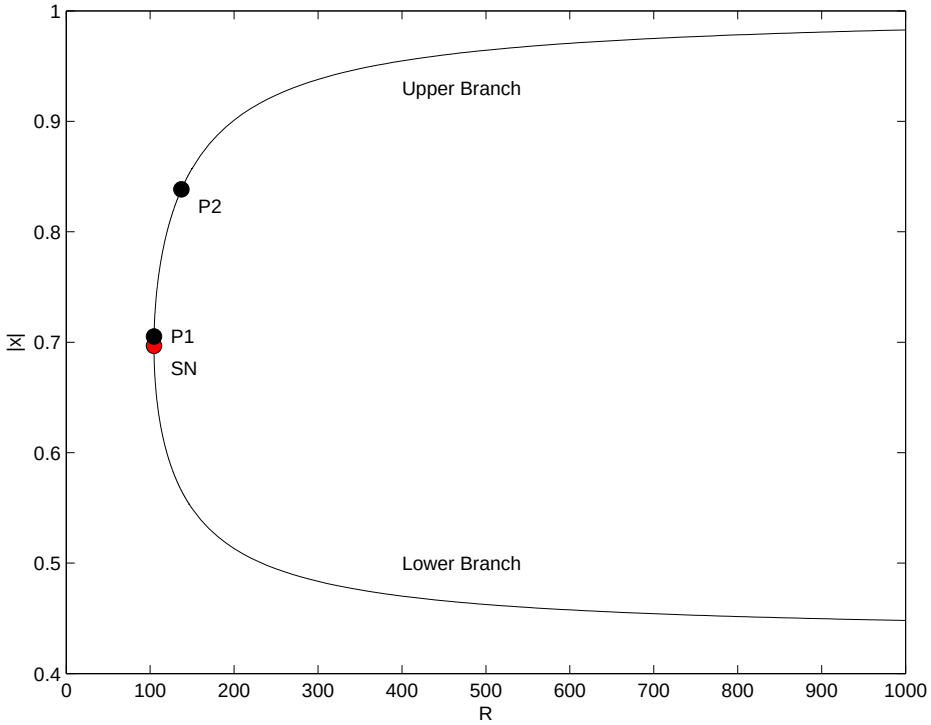

Figure 4: Magnitude of fixed points of the W97 model over a range of Reynolds numbers R. See text for details.

and

$$N(\mathbf{x}) = \begin{pmatrix} \sigma_m w^2 - \sigma_u uv \\ -\sigma_w w^2 + \sigma_u mv \\ \sigma_v w^2 \\ (\sigma_w u - \sigma_m m - \sigma_v v)w \end{pmatrix}. \tag{19}$$

The non-linear terms conserve the energy, recycling it between the different modes. The linearized problem admits a globally attracting fixed point at the origin—the laminar state. The same is true for the full non-linear problem when R is small.

For values of $R > R_{SN} = 104.5467$ where a saddle-node bifurcation occurs, two additional fixed points are born: the lower branch fixed point, $\mathbf{x}_{LB}$, and the upper branch fixed point $\mathbf{x}_{UB}$. Their magnitudes are shown in Figure 4. A sub-critical Hopf-type bifurcation occurs at $R_{P2} = 137.2569$: an unstable periodic orbit is born and for $R > R_{P2}$, $\mathbf{x}_{UB}$ becomes a stable attractor—it has two complex pairs of eigenvalues, both pairs having negative real parts. For $104.6434 = R_{P1} < R < R_{P2}$, $\mathbf{x}_{UB}$ one pair of complex eigenvalues has positive real part, so $\mathbf{x}_{UB}$ is unstable. For $R_{SN} < R < R_{P1}$, these eigenvalues become real but remain positive and hence $\mathbf{x}_{UB}$ is unstable here as well.

This 4D model has been a popular choice for study, but so far only one attempt has been made to find the minimal seed systematically over a range of R values. That study, by Cossu (2005), found that $|\mathbf{x}_{MS}|$ scaled closely with R^{-1}. However, the orbits started

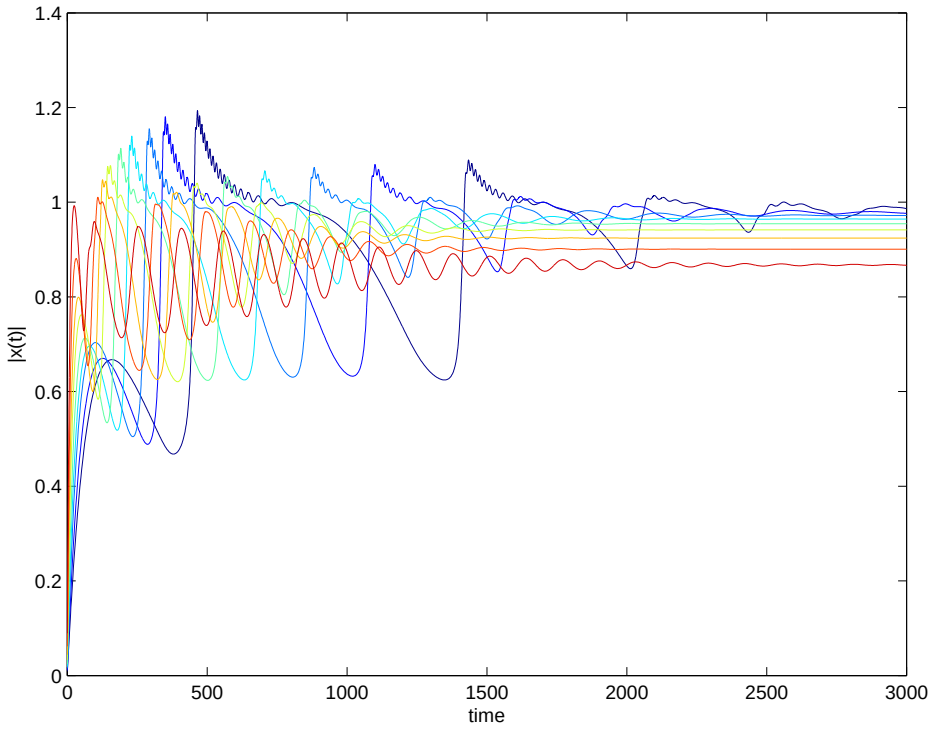

Figure 5: Magnitude of orbits beginning from the minimal seeds found by Cossu (2005). Colours are for different R values. All minimal seeds have orbits that are attracted to the upper branch point. Hence, they are within the basin of attraction of the upper branch point, not the origin, and are on the far side of the strong edge.

from these minimal seeds are seen (Figure 5) to go to the upper branch point[8] rather than to return to the origin. Thus, Cossu's results are for the minimal distance from the origin to the strong edge. If there exists a weak edge, it is entirely possible and perhaps likely that it contains a point closer to the origin than any point on the strong edge, in which case the minimal seed will be smaller than stated by these previous results. We will give results from applying the search method of Section 2 to the W97 model, which first requires a discussion of local maxima.

3.2 Optimizing to the Global Maximum

The optimization procedure of Section 2.1 finds a local maximum of $\mathcal{L}$, i.e. it maximizes $|\mathbf{x}(T)|$. A common difficulty with such optimization algorithms lies in whether the result is the global maximum or merely a lesser, local maximum.

We therefore calculate $|\mathbf{x}(T)|$ for two degrees of freedom of initial conditions $\mathbf{x}(0)$, constrained to have a specified w and a specified $|(m,u,v)| = \mathcal{E}_0$ (Figure 6). When these constraints place all $\mathbf{x}(0)$ well below the edge (top panel), $|\mathbf{x}(T)|$ appears a smooth function of initial conditions with only two local maxima, and for very small $\mathcal{E}_0$ these are nearly equal, as expected from the linearized problem which has a -1 symmetry (i.e. $\mathbf{X}(t,\mathbf{x}_0) = -\mathbf{X}(t,-\mathbf{x}_0)$;

[8] Cossu (2005) only studied values of R for which the upper branch point is a stable attractor.

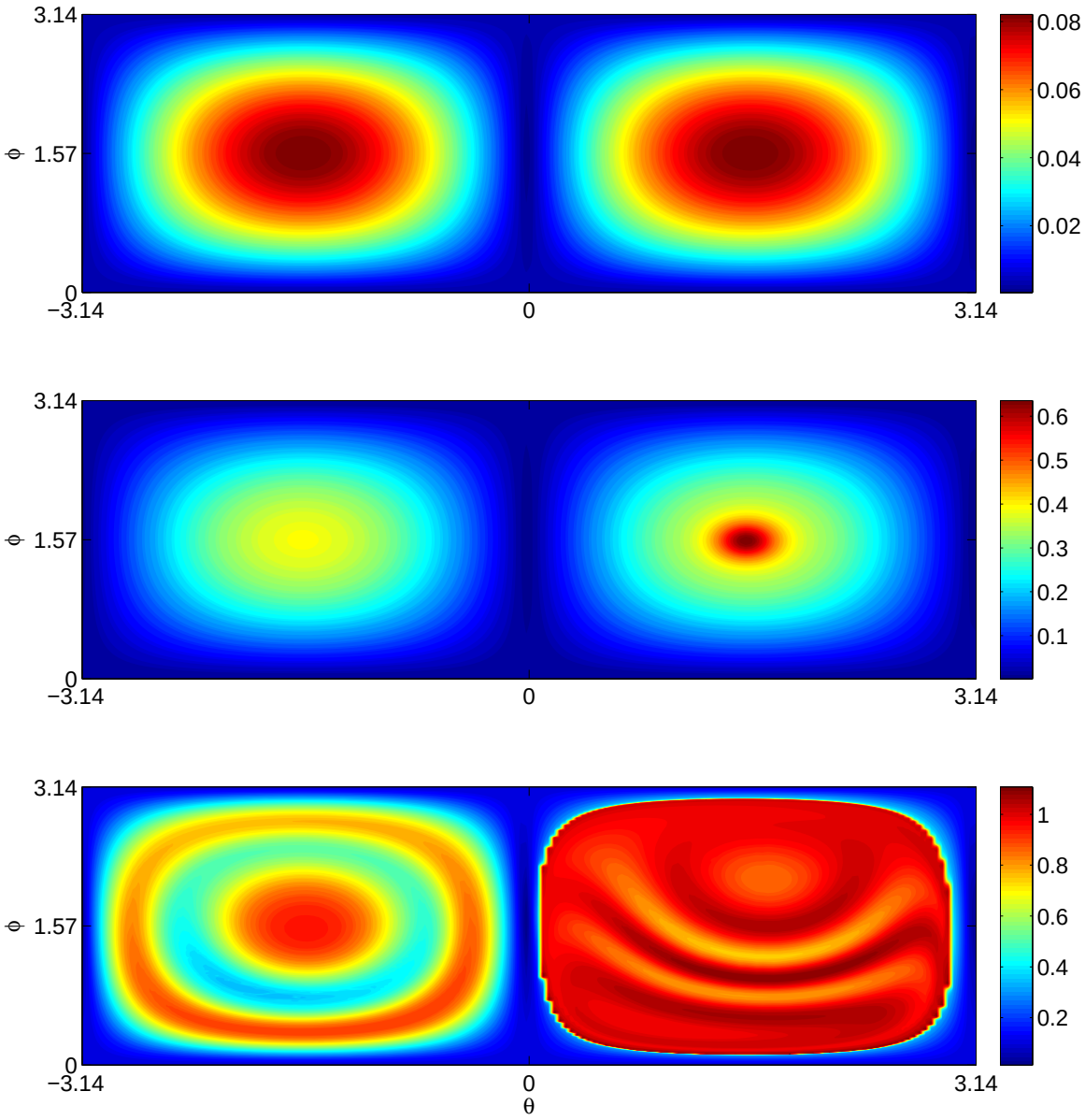

Figure 6: The final energy $|\mathbf{x}(T)|$ (colours), as a function of $\phi \in [0, \pi]$ and $\theta \in [-\pi, \pi]$ which define an initial condition $\mathbf{x}(0) = (\mathcal{E}_0 \cos\phi, \mathcal{E}_0 \cos\theta \sin\phi, \mathcal{E}_0 \sin\theta \sin\phi, W_0)$. The W97 system with $R = 200$ and $T = 80$ then gives $\mathbf{x}(T)$. The values $\mathcal{E}_0$, W_0 are respectively (top) 10^{-2}, $2.1894 \cdot 10^{-3}$, (middle) $4.3 \cdot 10^{-2}$, $9.4142 \cdot 10^{-3}$, and (bottom) 0.3, $6.5681 \cdot 10^{-2}$. The chosen $\mathcal{E}_0$ place the top, middle, and bottom panels well below, just below, and well above the edge, respectively. To complement this W_0 was chosen as $w_{MS} \cdot (\mathcal{E}_0/|\mathbf{x}_{MS}|)$, but the qualitative results are not sensitive to small changes of W_0. Well below the edge (top) there are two nearly equal local maxima of $|\mathbf{x}(T)|$, one the negative of the other, as expected from the non-normal linearized problem. Just below the edge (middle), there is one clearly global maxima and one lesser, local maxima. Well above the edge (bottom) there are many local maxima.

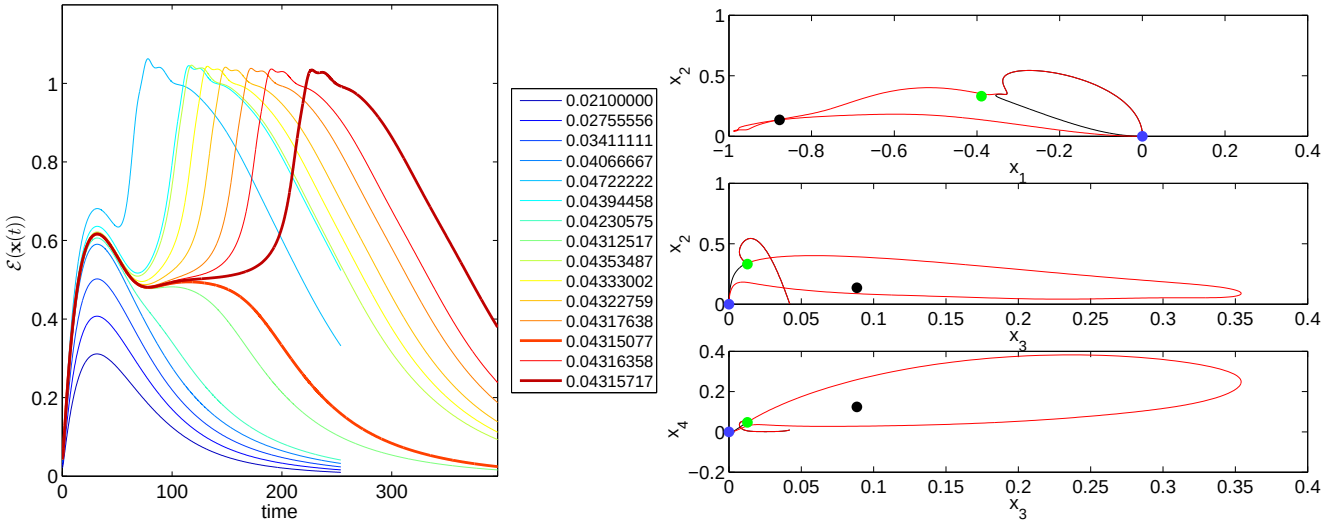

Figure 7: Left: The energy, $\mathcal{E}(\mathbf{x}(t)) \equiv |\mathbf{x}(t)|$, of successive orbits $\mathbf{x}(t)$ in the iterative procedure to find the minimal seed, for the W97 model with $R = 200$. Each curve corresponds to the orbit started from an initial condition with energy indicated by colour. For orbits below the edge, the initial condition is the NLOP. Thick lines show the orbits of those initial conditions straddling the edge. Right: The two thick-lined orbits from the left panel, shown in phase space. Both orbits begin near the origin (blue circle) and track each other closely until they near the lower branch point (green circle), at which point the black orbit (below the edge) returns directly to the origin while the red orbit (above the edge) circumnavigates the upper branch point (black circle) before returning to the origin. The top, middle, and bottom panels show x_1 vs. x_2, x_3 vs. x_2, and x_3 vs. x_4, respectively.

also note that the W97 model is symmetric about the $w = 0$ hyperplane). For larger $\mathcal{E}_0$ up to and below the minimal seed, one local maxima dominates and becomes the global maxima (middle panel). For $\mathcal{E}_0$ above the minimal seed (bottom panel), $SM(\mathbf{x}_{LB})$ folds (back and forth, or perhaps around $\mathbf{x}_{UB}$) across this domain of $\mathbf{x}(0)$. In the limit of $T \to \infty$, all orbits started off of $SM(\mathbf{x}_{LB})$ will have $|\mathbf{x}(T)| \approx 0$ or $|\mathbf{x}_{UB}|$, while those started on $SM(\mathbf{x}_{LB})$ will have $|\mathbf{x}(T)| \approx |\mathbf{x}_{LB}|$, suggesting the folds of $SM(\mathbf{x}_{LB})$ represent different local maxima of $|\mathbf{x}(T)|$ with finite but large T. Thus, for such high $\mathcal{E}_0$, many local maxima of $|\mathbf{x}(T)|$ emerge.

These calculations provide evidence that the optimization procedure of Section 2, applied to successively larger energy spheres and initiated in the correct hemisphere (the right hemisphere in Figure 6) will find the global maximum when below the minimal seed. Above the minimal seed, there is no guarantee that the optimization procedure will find the global maximum, which is why the algorithm was chosen so that finding the NLOP for energy spheres above the minimal seed is not necessary.

3.3 Results

We have found the minimal seed $\mathbf{x}_{MS}$ for the W97 model for values of R ranging from 105 to 2000. Let us first illustrate the search: the left panel of Figure 7 shows the progression of orbits in the search for the minimal seed when $R = 200$. The orbit of an NLOP below the edge shows transient growth with one well-defined maximum in energy. The growth is transient as would be had in the linearized system, with the orbit returning to the origin, hence producing a single maximum.

An orbit started above the edge shows similar initial transient growth followed by at least one further growth episode. When the initial condition is near the edge, the initial transient growth sends the orbit into a region of phase space where it is attracted towards the lower branch point. This much of the orbit, so far, has therefore been on or near the non-linear stable manifold of the lower branch point. The lower branch point being a saddle point, it has one unstable eigenvector $\mathbf{v}_u$. An orbit that began just above the edge will, when it gets near to $\mathbf{x}_{LB}$, be pushed off the stable manifold of $\mathbf{x}_{LB}$ by its unstable dynamics $\mathbf{v}_u$ and *away* from the origin, generating a second local maxima in the orbit's magnitude. An orbit started just below the edge, however, would be pushed *towards* the origin by the unstable dynamics of $-\mathbf{v}_u$. This weak edge, which is the stable manifold of the lower branch point, separates initial conditions whose orbits follow qualitatively different routes back to the origin (right panel of Figure 7). These orbits that straddle the edge also linger for a long time near $\mathbf{x}_{LB}$ itself ($t \sim 80$—125 in Figure 6): since $\mathbf{x}_{LB}$ is a fixed point, the velocity through phase space in nearby regions is small.

Note that there are often additional small maxima of energy at large time (not visible), owing to numerical inaccuracies when the orbit becomes extremely close to the origin. This necessitates the condition involving α in Figure 3.

Before showing the main result of this work—how the minimal seed scales with R—we ask how the minimal seed scales relative to $|\mathbf{x}_{LB}|$. In this model we know that $|\mathbf{x}_{LB}|$ asymptotes to a non-zero value for large R, and that the minimal seed $|\mathbf{x}_{MS}|$ asymptotes to zero for large R. Hence the ratio $|\mathbf{x}_{LB}|/|\mathbf{x}_{MS}|$ should go to ∞ as $R \to \infty$. But how does it behave for finite R, in particular for $R \to R_{SN}$. This is shown in Figure 8 over the range of R tested. The relationship is remarkably linear over almost all R, shallowing only for R very close to R_{SN} where the lower branch point is born and its magnitude changes rapidly with R.

Now in Figure 9 we show the scaling of the minimal seed against R. First, note that for each R the size of the minimal seeds found here is less than that found by Cossu (2005). This is indeed because we have found the minimal distance to the weak edge, which lies closer to the origin than does the strong edge. Second, we are now able to add to the discussion regarding how the minimal seed should scale with R. Cossu found reasonably good agreement with an asymptotic R^{-1} scaling, and we do too. For comparison, we found that the asymptotic (i.e. requiring the $R = 2000$ data point to be matched perfectly) scaling with a minimal least squares fit is $R^{-1.0361}$—reasonably close to R^{-1}. However, the exceptionally linear relationship shown in Figure 8 suggests we can do even better if we include not just R but also the (easily attainable) information of $|\mathbf{x}_{LB}|$: we find the scaling $|\mathbf{x}_{LB}|R^{-1}$ fits especially well. Indeed, $|\mathbf{x}_{LB}|$ is asymptotically flat at large R (recall Figure 4), so this scaling is equivalent to the R^{-1} scaling in the large R limit. However, the increase of

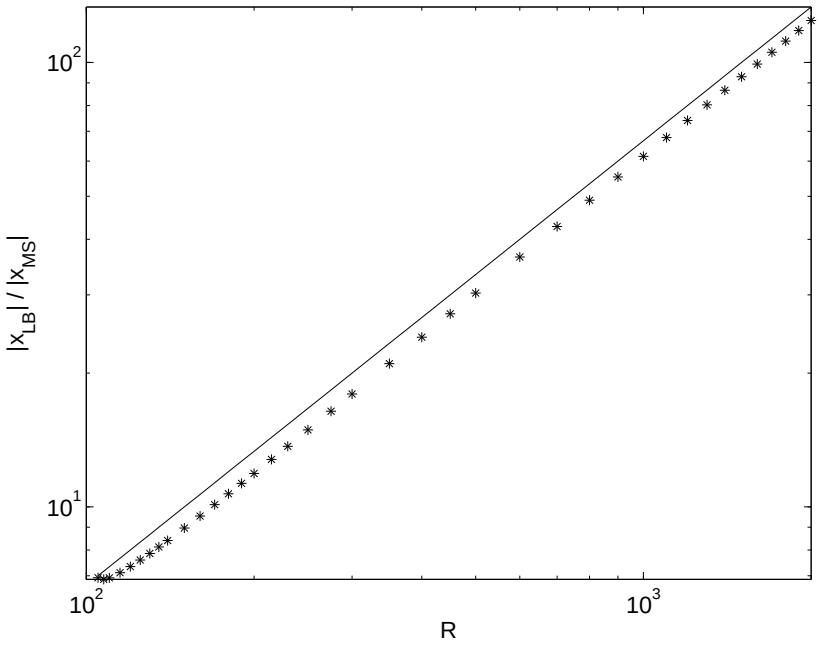

Figure 8: The ratio between the magnitude of the lower branch point $|\mathbf{x}_{LB}|$ and the minimal seed $|\mathbf{x}_{MS}|$ shown in black stars over a range of R, for Waleffe's 4D model. The upper solid line with a slope of 1 has been added to guide the eye, indicating a linear relationship between R and $|\mathbf{x}_{LB}|/|\mathbf{x}_{MS}|$.

$|\mathbf{x}_{LB}|$ going to smaller R allows this scaling to capture a great deal of how the minimal seed grows (as R decreases) faster than the R^{-1} scaling in the finite, non-asymptotic R regime.

Finally, we show in Figure 10 how the first, second, and fourth components of the minimal seed compare against its third component v, over a range of R. Recall that in the W97 model, v represents the amplitude of streamwise rolls. That v is always the largest component confirms that, in this model, perturbing the laminar flow with streamwise rolls is, broadly speaking, the most effective way to transition to turbulence. However, the other components being non-zero, the minimal seed—the most effective perturbation—is somewhat different from a pure streamwise roll perturbation, having finite but smaller perturbations from the laminar state in all modes (components of $\mathbf{x}$).

3.4 Cossu's Necessary Condition

Cossu (2005) gives a necessary condition for an initial condition $\mathbf{x}_0$ to be that point with minimal magnitude $\mathcal{E}(\mathbf{x}_0)$ which satisfies

$$\lim_{t\to\infty} \mathcal{E}(\mathbf{X}(t, \mathbf{x}_0)) \neq 0, \tag{20}$$

i.e. for $\mathbf{x}_0$ to be what Cossu had called the minimal seed. Following his argument, consider a candidate $\mathbf{x}_0$ for the minimal seed, and perturb it by following its orbit forward for a

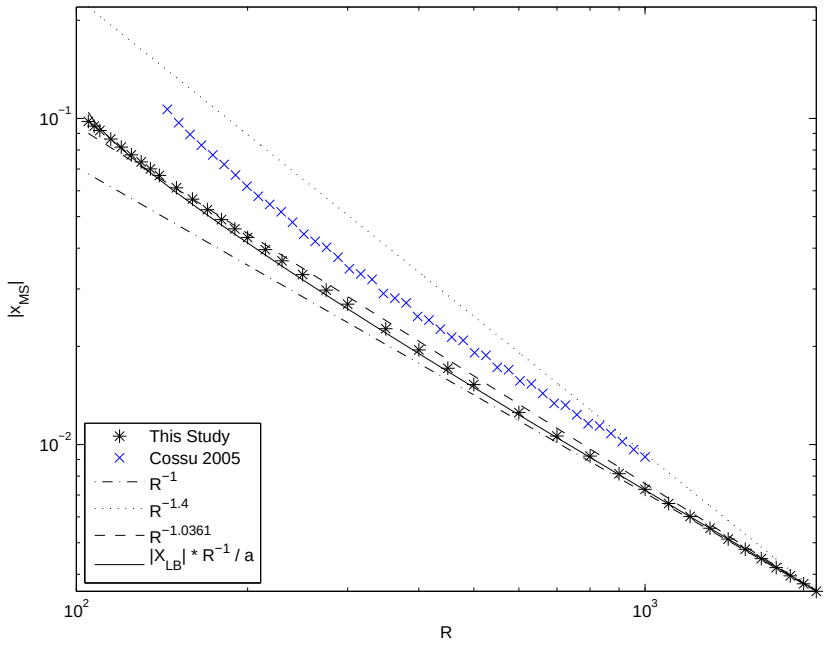

Figure 9: The Minimal Seed as a function of R for Waleffe's 4D model. Previous calculations by Cossu (2005) are shown in blue crosses, while our data is shown in black stars. The dash-dot and dotted lines show the R^{-1} and $R^{-1.4}$ scalings, respectively, while the dashed line shows the R power law with a least squares fit; since these are asymptotic ($R \to \infty$) scalings they are all fit to match the largest R data point. The solid line is the curve $|\mathbf{x}_{LB}|R^{-1}/a$ where a is the least-squares fit to $|\mathbf{x}_{LB}|/|\mathbf{x}_{MS}|$ vs. R from Figure 8.

short time Δt. The perturbation energy, in a Taylor series expansion, is

$$\mathcal{E}_0(\Delta t) \equiv \mathcal{E}(\mathbf{X}(\Delta t, \mathbf{x}_0)) = \mathcal{E}_0(0) + \Delta t \left(\frac{\mathrm{d}\mathcal{E}_0}{\mathrm{d}t}\bigg|_{\Delta t=0} \right) + \frac{\Delta t^2}{2} \left(\frac{\mathrm{d}^2\mathcal{E}_0}{\mathrm{d}t^2}\bigg|_{\Delta t=0} \right) + \mathcal{O}(\Delta t^3). \quad (21)$$

If $\frac{\mathrm{d}\mathcal{E}_0}{\mathrm{d}t}\Big|_{\Delta t=0} < 0$, then $\mathcal{E}(\mathbf{X}(\Delta t, \mathbf{x}_0)) < \mathcal{E}(\mathbf{x}_0)$. So $\mathbf{X}(\Delta t, \mathbf{x}_0)$, since it shares an orbit with $\mathbf{x}_0$, must also satisfy (20), and hence $\mathbf{x}_0$ is not the minimal seed. Similarly if $\frac{\mathrm{d}\mathcal{E}_0}{\mathrm{d}t}\Big|_{\Delta t=0} > 0$, then $\mathbf{X}(-\Delta t, \mathbf{x}_0)$ shows that $\mathbf{x}_0$ is not the minimal seed. Therefore, it is necessary that the minimal seed $\mathbf{x}_{MS}$ must satisfy

$$\frac{\mathrm{d}\mathcal{E}(\mathbf{X}(t, \mathbf{x}_{MS}))}{\mathrm{d}t}\bigg|_{t=0} = 0 \quad \text{and} \quad \frac{\mathrm{d}^2\mathcal{E}(\mathbf{X}(t, \mathbf{x}_{MS}))}{\mathrm{d}t^2}\bigg|_{t=0} \geq 0 \quad (22)$$

This argument also applies when finding the minimal seed when the edge contains a weak component, i.e. when the minimal seed is on the stable manifold of, in this case, the lower branch point: $\mathbf{x}_{MS} \in SM(\mathbf{x}_{LB})$. But note, however, that this argument does not apply so easily when searching for the NLOP subject to $\mathcal{E}(\mathbf{x}_0)$ equal to some specified d.

These ideas afford us a way to check our results. First, for the W97 model and any

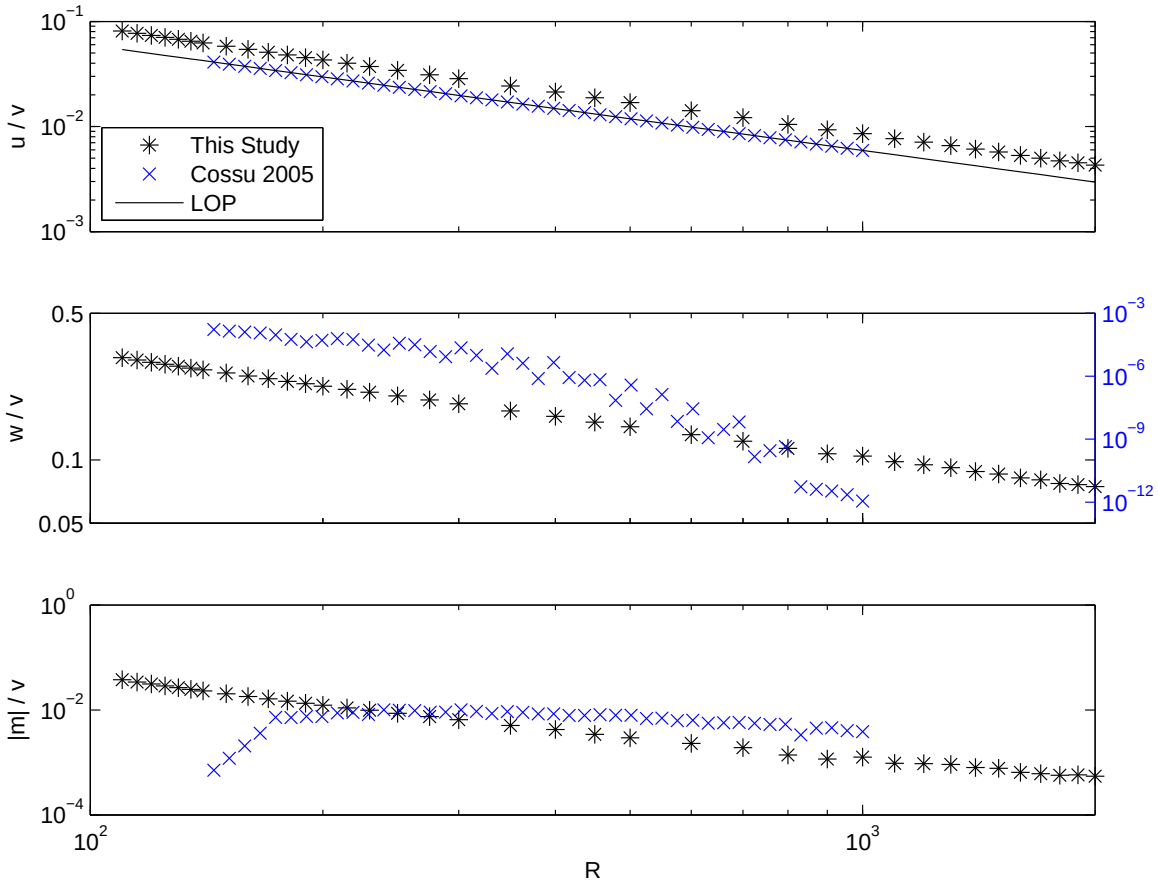

Figure 10: Ratio of the first (m), second (u), and fourth (w) components of the minimal seed to the third component v. We found $m > 0$ for all minimal seeds, but Cossu (2005) had $m < 0$ for most R, hence $|m|$ is shown here.

other in which the non-linear terms are energy preserving,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\mathcal{E}(\mathbf{x})^2\right) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}\cdot\mathbf{x} = \mathbf{x}\cdot(L\mathbf{x}) + \mathbf{x}\cdot(N(\mathbf{x})) = \mathbf{x}\cdot(L\mathbf{x}). \tag{23}$$

This measure of the error is immediately calculable and shown in Figure 11, scaled by $\mathbf{x}_{MS}\cdot\mathbf{x}_{MS}$, over the range of R tested. That there is no particular pattern suggests the error is due to numerical noise.

However, what is an acceptable magnitude for the above error? If we shift $\mathbf{x}_{MS}$ along its orbit by a small time Δt, then $\mathcal{E}^2/2$ will change by approximately $\Delta t\ \mathbf{x}_{MS}\cdot(L\mathbf{x}_{MS})$. But the size of Δt is presently unknown: too large a Δt and higher order terms in (21) will become significant. Thus, for a given R and estimate of $\mathbf{x}_{MS}$, we follow the orbit from $\mathbf{x}_{MS}$ backwards, or forwards, in time until $\mathrm{d}\mathcal{E}/\mathrm{d}t = 0$ is had, and therefore a more minimal $\mathbf{x}_{MS}$ is found. Carrying this out on the $\mathbf{x}_{MS}$ for $R = 120$ manages to reduce $|\mathbf{x}_{MS}|$ by only $5\cdot10^{-12}$, or $2\cdot10^{-11}$ for $R = 800$, or $2\cdot10^{-13}$ for $R = 2000$, suggesting our estimate of $\mathbf{x}_{MS}$ is quite accurate.

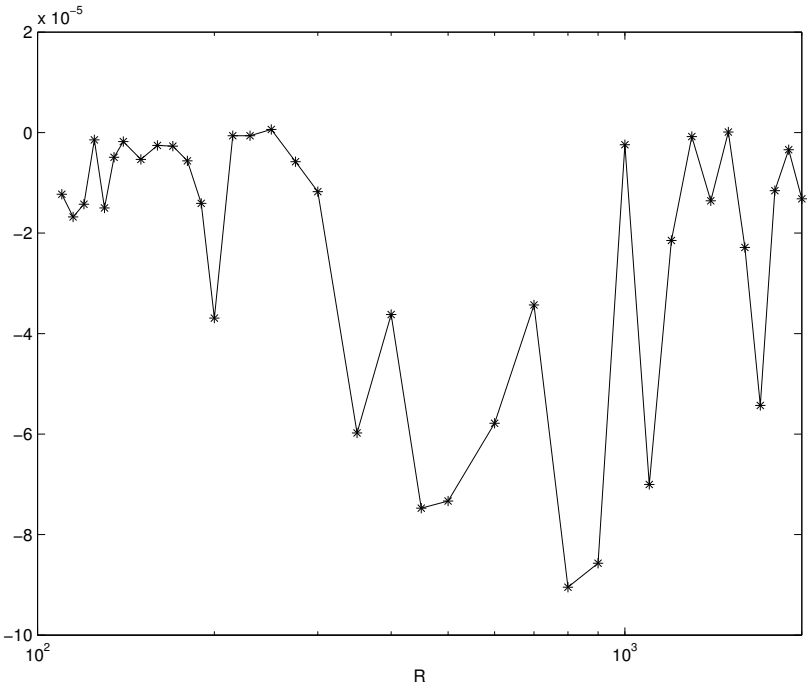

Figure 11: $(\mathbf{x}_{MS} \cdot (L\mathbf{x}_{MS}))$ / $(\mathbf{x}_{MS} \cdot \mathbf{x}_{MS})$ as a function of R. Deviations from zero indicate error in the estimation of $\mathbf{x}_{MS}$.

4 Conclusion

We have undertaken a search, across a range of Reynolds numbers R, for the minimal seed in the low-dimensional model of W97. One previous attempt (Cossu, 2005) at this was made. Cossu kindly shared his data with us in private communication, from which we discovered that his results for the minimal seed are in fact the least distance from the origin (the laminar flow state) to the boundary of the basin of attraction of the upper branch fixed point (a flow state that exists within turbulence yet itself is stable). The latter marks the ("strong") edge of the basin of attraction of the laminar fixed point, separating orbits that eventually return to the laminar flow state from those that never return. However, we have shown that in the W97 model there exists a boundary that is internal to the basin of attraction of the laminar fixed point: this so-called "weak" edge separates orbits that both return to the laminar flow state given sufficient time but do so by qualitatively different routes: orbits below this "weak" edge return fairly directly to the origin, while orbits above undergo a circuitous route that must wind around the upper branch fixed point before returning to the origin. The latter orbit, it would be said, did indeed transition to turbulence, despite that it ultimately relaminarizes. Thus the true minimal seed lies on this "weak" edge, and is in fact smaller than any point on the "strong" edge. We have found a more minimal seed than was previously known.

We have applied the relatively new non-linear optimization technique, reviewed by Kerswell *et al.* (2014). In this problem, the demarcation between states that are above and those that are below the weak edge is more subtle than between states that are above or below

the strong edge in which the demarcation is simply whether the orbit eventually relaminarizes. This marks new territory for this optimization technique and some modifications were required.

Considering the magnitudes of the minimal seeds we found over a range of R, our results support the hypothesis for a R^{-1} scaling law of the minimal seed in the W97 model. However, we discovered an even better scaling can be obtained if an additional piece of information is used: the magnitude of the lower branch point. Now, the minimal seed is the point on the stable manifold of the lower branch point closest to the origin. At large R the lower branch point is relatively insensitive to increases in R, and the decreasing minimal seed with increasing R is due to this stable manifold shifting closer to the origin. At finite R however, the lower branch point moves towards the origin with increasing R, translating its stable manifold closer to the origin along with it. This shifting of the lower branch point, it seems, causes the minimal seed to decrease somewhat faster than R^{-1}. In fact, we found that the ratio of the lower branch point to the minimal seed was exceptionally linear with R, leading to a new scaling for the minimal seed that goes as the product of the magnitude of the lower branch point with R^{-1}. The agreement of this scaling law with our results is quite encouraging, but whether such a scaling law is unique to the particular model studied here is not presently known.

5 Acknowledgements

I am deeply grateful for the opportunity to participate in the Geophysical Fluid Dynamics Program. Norman Lebovitz deserves my deepest thanks, for developing a summer research project that was both intriguing and accessible to me, and for skillfully guiding me through it. I also thank Colm-cille Caulfield for many useful discussions related to this research project. I would like to thank Glenn Flierl, Raf Ferrari, and Antonello Provenzale for organizing and directing the 2014 school, Geoff Vallis and Kerry Emanuel for their lectures and mentorship, and all the seminar speakers for sharing their knowledge particularly with the fellows. Lastly, thanks to all the fellows for your camaraderie and joviality.

References

J. S. Baggett and L. N. Trefethen, *Low-dimensional models of subcritical transition to turbulence*, Physics of Fluids (1994-present), 9 (1997), pp. 1043–1053.

S. J. Chapman, *Subcritical transition in channel flows*, Journal of Fluid Mechanics, 451 (2002), pp. 35–97.

C. Cossu, *An optimality condition on the minimum energy threshold in subcritical instabilities*, Comptes Rendus Mécanique, 333 (2005), pp. 331 – 336.

B. F. Farrell, *Optimal excitation of perturbations in viscous shear flow*, Physics of Fluids (1958-1988), 31 (1988), pp. 2093–2102.

B. Hof, a. Juel, and T. Mullin, *Scaling of the turbulence transition threshold in a pipe*, Physical Review Letters, 91 (2003), p. 244502.

R. R. Kerswell, C. C. T. Pringle, and A. P. Willis, *An optimization approach for analysing nonlinear stability with transition to turbulence in fluids as an exemplar*, Reports on Progress in Physics, 77 (2014), p. 085901.

G. Kreiss, A. Lundbladh, and D. S. Henningson, *Bounds for threshold amplitudes in subcritical shear flows*, Journal of Fluid Mechanics, 270 (1994), pp. 175–198.

N. R. Lebovitz, *Boundary collapse in models of shear-flow transition*, Communications in Nonlinear Science and Numerical Simulation, 17 (2012), pp. 2095 – 2100. Special Issue: Mathematical Structure of Fluids and Plasmas Dedicated to the 60th birthday of Phil Morrison.

S. A. Orszag, *Accurate solution of the orr-sommerfeld stability equation*, Journal of Fluid Mechanics, 50 (1971), pp. 689–703.

L. N. Trefethen, A. E. Trefethen, S. C. Reddy, and T. A. Driscoll, *Hydrodynamic stability without eigenvalues*, Science, 261 (1993), pp. 578–584.

F. Waleffe, *Hydrodynamic stability and turbulence - beyond transients to a self-sustaining process*, Studies In Applied Mathematics, 95 (1995), pp. 319–343.

——, *Transition in shear flows. nonlinear normality versus non-normal linearity*, Physics of Fluids (1994-present), 7 (1995), pp. 3060–3066.

——, *On a self-sustaining process in shear flows*, Physics of Fluids (1994-present), 9 (1997), pp. 883–900.