# An optimization framework for analyzing nonlinear stability due to sparse finite-amplitude perturbations

A. Leonid Heide* and Maziar S. Hemati†
*University of Minnesota, Minneapolis, MN 55455, USA*

**In this paper, we present an optimization framework for computing sparse and spatially-localized finite-amplitude perturbations that maximize transient growth in nonlinear systems. A variational approach is used to derive the first-order necessary conditions for optimality, which form the basis of our iterative direct-adjoint looping numerical solution algorithm. The method is demonstrated on an illustrative 2-state dynamical system that possesses key features of the incompressible Navier-Stokes equations. We then apply the method to analyze a reduced-order model of a sinusoidal shear flow at $Re = 20$. Our results establish the power of the proposed optimization framework for revealing dominant modal interactions and sparse perturbation mechanisms for transient growth and instability in fluid flows.**

## I. Introduction

This paper presents a framework for computing sparse and spatially-localized optimal perturbations that maximize transient growth of perturbation energy in dynamic systems. Transient growth refers to the amplification of perturbations over a finite-time horizon, and is an important mechanism for instability in fluids systems [1]. Perturbation energy can grow over finite transient time-horizons, even when the dynamics are linear and stable. This is the basis of non-modal stability theory [2]. The transient growth phenomenon arises due to a large-degree of non-normality in the linear operator that governs the dynamics. Within the context of incompressible flows, transient growth of the linearized dynamics is often a necessary condition for instability of perturbations governed by the associated nonlinear dynamics [1].

Worst-case analysis is a common feature of transient growth studies. Among all admissible finite-amplitude perturbations, we are most interested in the one that maximizes some norm (e.g., kinetic energy) over a given time horizon. Such a perturbation is referred to as an *optimal perturbation*. For linear perturbation dynamics, the optimization problem for a *linear optimal perturbation (LOP)* reduces to maximizing a Rayleigh quotient, which corresponds to solving an eigenvalue problem [1–3]. For nonlinear dynamics, a *nonlinear optimal perturbation (NLOP)* of a prescribed amplitude can be computed by solving the associated optimization problem using the calculus of variations. This is the basis of the so-called nonlinear non-modal stability theory [4, 5]. An optimization over the perturbation amplitude and time-horizon can be used to determine the *minimal seed* for instability in a nonlinear system. An overview of the details of these approaches will be given in Section II. For additional details, the reader is pointed to several excellent review articles [2–5].

In general, optimal perturbations that maximize transient growth in fluids systems are spatially extended and lack sparsity—i.e., many flow variables contribute to an optimal perturbation, rather than a smaller (sparse) subset of flow variables. Yet, it is often of interest to identify localized and sparse optimal perturbations. Sparse and spatially-localized optimal perturbations would reveal specific spatial locations where perturbations of a specific quantity would dominate in driving baseflow instability. Similarly, obtaining sparse optimal perturbations in a suitable modal basis would reveal dominant modes and associated physical processes. Such solutions could also provide guidance on actuator placement and inform flow control strategies by revealing specific locations and quantities where the effect of control can be most pronounced [6–8].

In this paper, we propose an optimization framework for computing sparse NLOP solutions. This is done by augmenting the standard transient growth objective function in NLOP analysis with a sparsity promoting norm. We then formulate a variational method to solve the associated sparse NLOP problem.

The framework is demonstrated on two finite-dimensional nonlinear systems based on the incompressible Navier-Stokes equations: A simple 2-state model that exhibits important features of the incompressible Navier-Stokes equations, and a reduced-order model of a sinusoidal shear flow. The 2-state model is amenable to graphical demonstrations and is

*Graduate Student, Aerospace Engineering and Mechanics, AIAA Student Member.
†Associate Professor, Aerospace Engineering & Mechanics, AIAA Associate Fellow.

used to highlight important features of the sparse NLOP problem and associated solutions. The reduced-order model of a sinusoidal shear flow provides a non-trivial benchmark and highlights the utility of the proposed method in extracting important instability mechanisms.

The paper is organized as follows: In Section II, we review the standard NLOP optimization problem and associated direct-adjoint looping (DAL) solution algorithm. In Section III, we introduce the proposed sparse NLOP optimization problem and a corresponding DAL solution algorithm. The approach is demonstrated on two finite-dimensional systems with results presented in Sections IV.A and IV.B. Section V concludes the paper.

## II. Optimal Perturbations

Consider the finite-dimensional dynamical system

$$\dot{x} = f(x; X_0) \tag{1}$$

where $x = x(t) \in \mathbb{R}^n$ is a perturbation away from a steady-state attractor $X_0$ of the nonlinear governing equations. The nonlinear optimal perturbation (NLOP) is defined as the solution to [4, 5]

$$\underset{x(0)}{\text{maximize}} \quad \|x(T)\|_2^2 \tag{2a}$$

$$\text{subject to} \quad \dot{x} - f(x; X_0) = 0, \tag{2b}$$

$$\|x(0)\|_2^2 - d^2 = 0, \tag{2c}$$

where the initial perturbation energy $d^2$ and the length of the time-horizon $T$ are given. Note that defining the energy as $\|x(t)\|_2^2$ is without loss of generality. Bisection over $d$ and $T$ can be used to determine the so-called "minimal seed" and the associated upper-bound for the perturbation energy $\|x(0)\|_2^2$ required to trigger instability away from the steady attractor $X_0$.

The equality constrained optimization problem can be converted to an unconstrained optimization problem by introducing the Lagrangian

$$\mathcal{L} = \|x(T)\|_2^2 + \int_0^T p^\mathsf{T}(t)\,(\dot{x}(t) - f(x(t)))\,\mathrm{d}t + \lambda(\|x(0)\|_2^2 - d^2) \tag{3}$$

where $\lambda \in \mathbb{R}$ and $p(t) \in \mathbb{R}^n$ are Lagrange multipliers. The $p(t)$ is referred to as the *co-state* or *adjoint* in the optimal control literature. Considering the first variation of the Lagrangian with respect to each of the variables yields the first-order necessary conditions for optimality:

$$\dot{x} = f(x) \tag{4a}$$

$$\dot{p} = -\left(\frac{\partial f}{\partial x}\right)^\mathsf{T} p(t) \tag{4b}$$

$$0 = 2\lambda x(0) - p(0) \tag{4c}$$

$$0 = 2x(T) + p(T) \tag{4d}$$

$$0 = \|x(0)\|_2^2 - d^2. \tag{4e}$$

This system of two differential and three algebraic equations can be solved iteratively, for example, using gradient methods. A basic implementation follows roughly as:

1) Initialize $x^{(0)}(0)$ to satisfy $\|x^{(0)}(0)\|_2^2 = d^2$.
2) Given $x^{(i)}(0)$, integrate the primal system (4a) forward in time from $t = 0$ to $t = T$ and store the solution $x^{(i)}(t)$.
3) Given $x^{(i)}(t)$ and $p(T) = -2x(T)$, solve the co-state equation (4b) backward in time from t=T to t=0. Store $p^i(0)$.
4) Evaluate the stopping criterion and terminate if $\left\| \frac{x^{(i)}(0)^\mathsf{T} p^{(i)}(0)}{d\|p^{(i)}(0)\|} \right| - 1 \Big| < \epsilon$. Otherwise, solve for $\lambda$ such that $\|x^{(i)}(0) + \Delta(2\lambda x^{(i)}(0) - p^{(i)}(0))\|_2^2 = d^2$ and repeat from step 2 using $x^{(i+1)}(0) \leftarrow x^{(i)}(0) + \Delta(2\lambda x^{(i)}(0) - p^{(i)}(0))$, where $\Delta > 0$ is a parameter that defines the gradient step size.

There are several practical issues that must be considered when implementing NLOP. The first is that the co-state $p(t)$ depends on the state $x(t)$; however, storing $x(t)$ over the full time horizon can be expensive. Thus, a 'check-pointing' procedure (see [9, 10]) can be implemented to reduce the storage requirements. This is done by recalculating the (primal) state over short intervals during the backward integration of the co-state. This approach requires that $x(t)$ be stored at regular intervals $t = \tau_k$ during the forward time integration. The second practical consideration for gradient-ascent methods is selecting a suitable step size $\Delta$. To find a step size that yields sufficient accuracy for each iteration—while also balancing the rate of convergence—we perform an inexact line search using Armijo's rule [11]. We thereby ensure that the change in the optimization variable is proportional to the step length and gradient (the search direction).

**A. Optimal Perturbations: The Linear Case**

Now consider the special case of linear time-invariant perturbation dynamics $\dot{x}(t) = Ax(0)$. The state at time $T$ is related to the initial state as $x(T) = \Phi(T, 0)x(0)$, where $\Phi(T, 0)$ is the state transition matrix. The transient amplification $G(T) := \|x(T)\|_2^2/\|x(0)\|_2^2$ will be invariant to the initial perturbation magnitude $d$ by virtue of linearity, and so we set it to unity. Thus, for linear dynamics, the optimization in (2a) reduces to

$$\begin{array}{ll} \text{maximize} & \|\Phi(T, 0)x(0)\|_2^2. \\ \|x(0)\|_2^2 = 1 & \end{array} \tag{5a}$$

It is straightforward to show that the Lagrangian in this case is

$$\mathcal{L} = x^{\mathsf{T}}(0)P(T)x(0) - \lambda(x^{\mathsf{T}}(0)x(0) - 1) \tag{6}$$

where $P(T) := \Phi^{\mathsf{T}}(T, 0)\Phi(T, 0) > 0$. Setting the first variation of $\mathcal{L}$ with respect to $x(0)$ to zero yields

$$(P(T) - \lambda I)\, x(0) = 0. \tag{7}$$

This is a standard eigenvalue problem, with the optimal $x(0)$ corresponding to the eigenvector associated with the maximum eigenvalue $\lambda$ of $P(T)$.

## III. Sparse and Spatially-Localized Optimal Perturbations

The standard NLOP problem can further be extended to investigate sparse optimal perturbations. Sparse NLOP can be desirable when seeking the single element in the admissible set of perturbations to concentrate the available perturbation energy $d^2$. In principle, the sparse (localized) NLOP problem is combinatorially difficult. Here, we exploit $\ell_1$-regularization to promote sparsity in the localized NLOP problem. When $\|x(0)\|_1$ is relatively small, the solution tends to be sparser than when $\|x(0)\|_1$ is relatively large. Although $\ell_1$-regularization is a heuristic, it allows for relatively efficient numerical procedures for solving the associated sparse NLOP problem:

$$\begin{array}{lll} \underset{x(0)}{\text{maximize}} & \|x(T)\|_2^2 - \sigma\|x(0)\|_1 & \text{(8a)} \\ \text{subject to} & \dot{x} - f(x; X_0) = 0, & \text{(8b)} \\ & \|x(0)\|_2^2 - d^2 = 0 & \text{(8c)} \end{array}$$

where again the initial perturbation energy $d^2$ and the length of the time-horizon $T$ are given. Here, we have introduced a sparsity promoting parameter $\sigma \geq 0$ into the objective function. Solving the optimization problem for different values of $\sigma$ results in a family of optimal perturbations that define a Pareto front. When $\sigma = 0$, we recover the standard NLOP problem/solution. As $\sigma$ is increased, the solution will tend to become sparser. In the limit $\sigma \to \infty$, $x^*(0)$ will be empty.

The Lagrangian for this problem follows as

$$\mathcal{L} = \|x(T)\|_2^2 + \int_0^T p^{\mathsf{T}}(t)\,(\dot{x}(t) - f(x(t)))\,\mathrm{d}t + \lambda(\|x(0)\|^2 - d^2) - \sigma\|x(0)\|_1. \tag{9}$$

Note that this Lagrangian is simply the sum of the original Lagrangian (3) from the standard NLOP problem, and a sparsity promoting term $-\sigma\|x(0)\|_1$. This problem structure enables the use of an important class of solution algorithms known as *proximal gradient methods* [12]. We will derive a convenient and easy-to-implement method based on the *iterative soft thresholding algorithm (ISTA)*, which is commonly used to solve optimization problems with a composite

objective function that includes an $\ell_1$-regularization term [12]. ISTA makes use of the soft threshold function, which is defined as

$$S_\sigma(y) := \begin{cases} y+\sigma & \text{if } y < -\sigma \\ 0 & \text{if } -\sigma \le y \le \sigma \\ y-\sigma & \text{if } y > \sigma. \end{cases} \tag{10}$$

Thus, $S_\sigma$ serves to reduce the absolute value of the argument by no more than a given threshold on $\sigma \ge 0$. It follows that the sparse NLOP problem can be solved simply by solving a modified version of the standard NLOP problem—i.e., using the gradient ascent algorithm presented in Section II with a modified gradient calculation—with an application of $S_\sigma$ to each gradient step prior to advancing to the next iteratate. Care must be taken with the gradient calculation, since the gradient of $\|\cdot\|_1$ consists of signum functions and is not continuously differentiable. The sub-gradient at sign(0) is defined as the interval $[-1, 1]$, and algorithms can be developed accordingly. Here, we simply choose remove the singularity at zero in the signum function by defining a regularized gradient

$$\nabla_\epsilon \|y\|_1 = W_\epsilon y, \qquad W_\epsilon := \mathrm{diag}(w_1, \ldots, w_n), \quad w_i := \begin{cases} 1/|y_i| & \text{if } |y_i| \ge \epsilon \\ 0 & \text{if } |y_i| < \epsilon. \end{cases} \tag{11}$$

A sparsity promoting DAL algorithm based on ISTA and a regularized gradient (11) roughly follows as:

1) Initialize $x^{(0)}(0)$ to satisfy $\|x^{(0)}(0)\|^2 = d^2$.
2) Given $x^{(i)}(0)$, integrate the primal system (4a) forward in time from t=0 to t=T, storing $x^{(i)}(t)$ at $k$ predetermined checkpoints $t = \tau_k$. Store the solution $x^{(i)}(t)$.
3) Given $x^{(i)}(t)$ and $p(T) = -2x(T)$, solve the co-state equation (4a) backward in time from t=T to t=0 utilizing the checkpoints $x^{(i)}(\tau_k)$. Store $p^i(0)$.
4) Evaluate the stopping criterion and terminate if $\left\| \frac{x^{(i)}(0)^\mathsf{T} p^{(i)}(0)}{d\|p^{(i)}(0)\|} \right| - 1 \Big| < \epsilon$. Otherwise, solve for $\lambda$ such that $\|x^{(i)}(0) + \Delta(2\lambda x^{(i)}(0) - p^{(i)}(0))\|^2 = d^2$ and repeat from step 2 using $x^{(i+1)}(0) \leftarrow S_\sigma\left(x^{(i)}(0) + \Delta(2\lambda x^{(i)}(0) - p^{(i)}(0) - \sigma W_\epsilon x^{(i)}(0))\right)$.

Just as for standard NLOP discussed in section II, we implement check-pointing for backward time integration of the co-state. Furthermore, an appropriate step size $\Delta$ is computed using an inexact line-search at each iterate.

### A. Sparse and Spatially-Localized Optimal Perturbations: The Linear Case

For linear systems, several simplified solution procedures can be derived. Specifically, consider the linear system $\dot{x}(t) = Ax(0)$. The state at time $T$ is related to the initial state as $x(T) = \Phi(T, 0)x(0)$. It is straightforward to show that the Lagrangian in this case reduces to

$$\mathcal{L} = x^\mathsf{T}(0)P(T)x(0) - \lambda(x^\mathsf{T}(0)x(0) - 1) - \sigma\|x(0)\|_1 \tag{12}$$

where $P(T) := \Phi^\mathsf{T}(T, 0)\Phi(T, 0) > 0$ and we have set $d = 1$ as before. Considering the variation with respect to $x(0)$ and regularizing the signum function as in (11), the first-order necessary condition for optimality follows as

$$(P(T) + \sigma W_\epsilon)x(0) = \lambda x(0). \tag{13}$$

For $\sigma = 0$, there will be no sparsity promotion, and we recover the standard eigenvalue problem associated with the optimal perturbation for maximizing the linear transient growth at time $t = T$: i.e., the optimal $x(0)$ will be the eigenvector associated with the maximum eigenvalue $\lambda$ determined from (13). For $\sigma \neq 0$, this is no longer a standard eigenvalue problem due to the fact that $W_\epsilon = W_\epsilon(x(0))$. A very simple solution method is to iteratively solve a sequence of eigenvalue problems to obtain a sparse optimal perturbation $x(0)$. That is, first solve (13) for the standard optimal perturbation $x(0)$, then use this $x(0)$ to determine $W_\epsilon(x(0))$. This fixes $W_\epsilon$, and so now (13) can be solved as a standard (but modified) eigenvalue problem. This process can be repeated until convergence. Alternatively, this problem can be solved using related methods used for solving sparse principal component analysis (PCA) problems (see, e.g., [13]).

# IV. Results

## A. An Illustrative Example

Consider the nonlinear dynamics of perturbations for a two-dimensional system given by

$$f(x;R) = A(R)x + Q(x)x \tag{14}$$

where

$$A(R) = \begin{bmatrix} -1/R & 1 \\ 0 & -1/R \end{bmatrix}, \qquad Q(x) = -Q^{\mathsf{T}}(x) = \begin{bmatrix} 0 & -x_1 \\ x_1 & 0 \end{bmatrix}. \tag{15}$$

This simple model possesses many features of the incompressible Navier-Stokes equations: the scalar parameter $R > 0$ acts like the Reynolds number; the linear term is non-normal; and the nonlinear term $Q(x)x$ is quadratic and energy conserving. For a simple two-dimensional system, sparse optimal perturbations can be determined with relative ease using a brute force search. Nonetheless, this simple example is instructive with regards to the features of sparse NLOP solutions in comparison with standard NLOP solutions. The example also serves to validate the algorithm, which is facilitated by visual inspection of phase portraits.

The energy threshold to instability for a sparse NLOP is found to be greater than that for a standard NLOP. This is to be expected; in general, the energy threshold to instability for a sparse minimal seed will be no less than the energy threshold for a non-sparse minimal seed. For this two dimensional example, a sparse non-zero solution will have a single non-zero element. Here we will investigate the case with $R = 2\sqrt{2}$, which is greater than the critical $R$ for transient growth in this system. Applying bisection on $d$ with a tolerance of 0.01 and fixing $T = 3.6$ and $\sigma = 0.2d$, we find upper bounds on the disturbance threshold for instability: $d^* \leq 0.12$ for the NLOP and $d^*_{sp} \leq 0.13$ for the sparse NLOP. We note that the NLOP and sparse NLOP problems can only provide upper bounds to instability thresholds. Any additional bisection over the other optimization parameters could only serve to provide lower values for these upper bounds. Recently proposed convex-optimization-based methods can be used to obtain lower bounds on these instability thresholds for the non-sparse NLOP problem in quadratic systems [14–17], which would also provide lower-bounds for the sparse NLOP problem.

The system response for three disturbance levels are reported in Figure 1. For $d = 0.10$, the NLOP $x^*$ and the sparse NLOP $x^*_{sp}$ solutions computed are

$$x^* = \begin{bmatrix} 0.0384 & 0.0923 \end{bmatrix}^{\mathsf{T}}, \qquad x^*_{sp} = \begin{bmatrix} 0 & 0.1 \end{bmatrix}^{\mathsf{T}}.$$

For $d = 0.12$, we find

$$x^* = \begin{bmatrix} 0.0485 & 0.1098 \end{bmatrix}^{\mathsf{T}}, \qquad x^*_{sp} = \begin{bmatrix} 0 & 0.12 \end{bmatrix}^{\mathsf{T}}$$

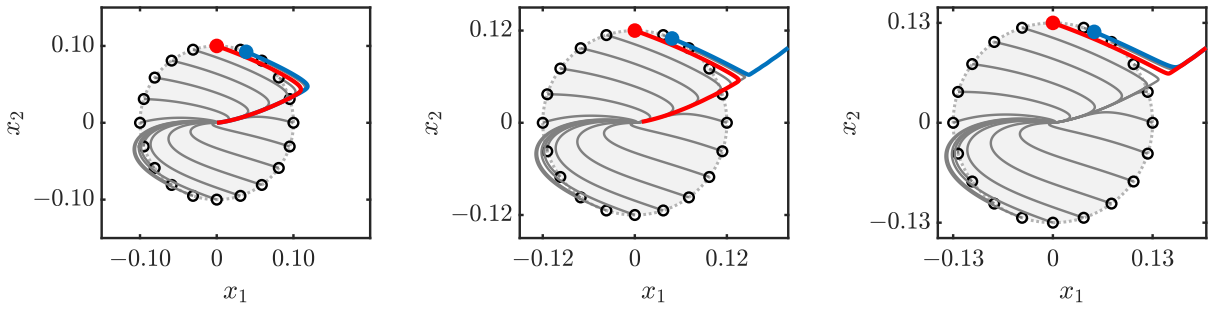
and for $d = 0.13$ we find

$$x^* = \begin{bmatrix} 0.0538 & 0.1183 \end{bmatrix}^{\mathsf{T}}, \qquad x^*_{sp} = \begin{bmatrix} 0 & 0.13 \end{bmatrix}^{\mathsf{T}}.$$

Next, we consider the associated linear dynamics about the stable attractor. Figure 2 highlights the responses due to a linear optimal perturbation and a sparse linear optimal perturbation found by applying the method presented in Section III.A:

$$x^* = \begin{bmatrix} 0.4472 & 0.8944 \end{bmatrix}^{\mathsf{T}}, \qquad x^*_{sp} = \begin{bmatrix} 0 & 1 \end{bmatrix}^{\mathsf{T}}.$$

Of note in this example is that the sparse linear optimal perturbation identified coincides with the sparse NLOP identified previously. Of course, for a linear system, the (sparse) optimal solution will never be unique—even when a unique optimal eigendirection is identified: if $x(0) = x^*$ is optimal, then so is $x(0) = -x^*$. Nonetheless, this provides a heuristic for efficiently finding sparse NLOP solutions: First, solve for all sparse linear optimal perturbations $\{x^*_i\}$, then seed the nonlinear simulation with $\{d \cdot x^*_i\}$, and find the trajectory with $\|x_i(T)\|$ being the maximum. In most cases, this will require two simulations for a given perturbation amplitude $d$ and time-horizon $T$.




**(a)** $d = 0.1$ **(b)** $d = 0.12$ **(c)** $d = 0.13$

**Fig. 1 The nonlinear optimal perturbation (blue) drives the state further from the steady attractor than the sparse nonlinear optimal perturbation (red) as shown in (a). The NLOP triggers instability at a lower threshold disturbance than the sparse NLOP as shown in (b) and (c). Nonlinear system responses for a set of sub-optimal initial perturbations are also plotted in gray. The light gray circle indicates the region for which energy is less than $d^2$.**

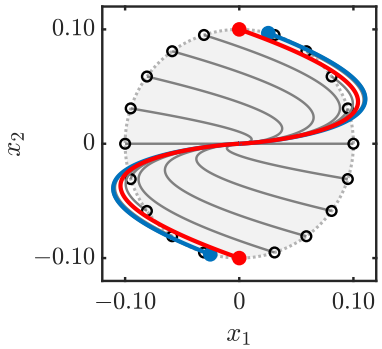



**Fig. 2 Linear optimal perturbations (blue) drive the state further from the steady attractor than sparse linear optimal perturbations (red). Linear system responses for a set of sub-optimal initial perturbations are also plotted in gray.**

### B. Reduced-Order Model of a Sinusoidal Shear Flow

Next, we demonstrate the sparse NLOP framework on a reduced-order model of a shear flow wherein the fluid experiences a a sinusoidal body force between two free-slip moving walls [18].This model has the form

$$f(x; Re) = A(Re)x + Q(x)x, \tag{16}$$

where the state vector $x = x(t) \in \mathbb{R}^9$. The linear term $A(Re) \in \mathbb{R}^{9\times 9}$ is Hurwitz, and is parameterized by Reynolds number $Re$. The nonlinear term $Q(x)x$ in equation (16) is a quadratic function

$$Q(x)x = \begin{bmatrix} x^T Q^{(1)} x \\ \vdots \\ x^T Q^{(9)} x \end{bmatrix}, \tag{17}$$

where $Q^{(1)}, ..., Q^{(9)} \in \mathbb{R}^{9\times 9}$ are symmetric matrices defined such that the quadratic nonlinearity is lossless.

The 9-state sinusoidal shear flow model presented here is a generalization of the well-established 8-state Waleffe model [19] and has been used in a number of previous works to demonstrate novel stability analysis methods (see, e.g., [20–23]). The model also benefits from being physically interpretable as each of the nine Fourier modes describes

a well-known feature that contributes to the dynamics of turbulence in many shear flows. The introduction of the ninth mode—which describes the turbulence-induced modification of the basic flow profile—is the distinguishing feature of this 9-state model over the 8-state Waleffe model [18]. Note that the dynamics in (16) possess a fixed-point at the origin. This was achieved by shifting the state vector $\tilde{x} \in \mathbb{R}^9$ from the original dynamical equations in [18] for which the fixed-point is located at $\tilde{x}_1 = 1, \tilde{x}_2 = ... = \tilde{x}_9 = 0$ for all $Re$. The coordinate shift is accounted for in the linear term as $A(Re) = \tilde{A}(Re) + W$, where $\tilde{A}(Re)$ is the linear term from [18], and $W \in \mathbb{R}^{9\times9}$ is defined such that for $c = [1, 0, ..., 0]^\mathsf{T} \in \mathbb{R}^9$:

$$W\tilde{x} = Q(\tilde{x})c + Q(c)\tilde{x}. \tag{18}$$

The state vector $\tilde{x}$ contains modal coefficients that provide physical interpretability of the dynamics and allow for resonctruction of the velocity field $\boldsymbol{u}(z, t)$ as,

$$\boldsymbol{u}(z, t) = \sum_{m=1}^{9} \tilde{x}_m(t)\boldsymbol{\phi}_m(z), \tag{19}$$

where $\boldsymbol{\phi}_m(z)$ are the velocity modes and $\tilde{x}_m(t)$ are the associated modal coefficients. An overview of each mode and its physical relevance is provided in Table 1.

| Mode | Name | Description |
|---|---|---|
| $\tilde{x}_1$ | basic profile | Describes the mean velocity profile of the flow. |
| $\tilde{x}_2$ | streak | Captures spanwise variation of the streamwise velocity. |
| $\tilde{x}_3$ | downstream vortex | Describes streamwise vorticies spanning the entire gap. |
| $\tilde{x}_4, \tilde{x}_5$ | spanwise flow modes | Sinusoidal streak-flow instabilities causing velocity perturbations of the streaky flow. |
| $\tilde{x}_6, \tilde{x}_7$ | normal vortex modes | Generated by the advection of $\tilde{x}_4, \tilde{x}_5$ by the streak ($\tilde{x}_2$) and the vortex ($\tilde{x}_3$) modes. |
| $\tilde{x}_8$ | 3D mode | Fully three-dimensional mode created by the interaction of modes $\tilde{x}_2$–$\tilde{x}_7$. |
| $\tilde{x}_9$ | base flow modifier | Modification to the structure of the basic velocity profile ($\tilde{x}_1$) by the turbulence. |

**Table 1 The modes described by the 9-state model [18] are physically interpretable and carefully constructed to model observed dynamics in turbulent flow simulations.**

We now demonstrate the sparse NLOP method for $Re = 20$. To begin our analysis, we must set a perturbation magnitude $d$. To obtain an initial approximation for $d$, we employ the method described in [22], which estimates bounds on permissible perturbations using a quadratic constraint framework. This quadratic constraint approach indicates that a perturbation amplitude of $d = d_0 := 0.676$ is certifiably below the threshold for baseflow instability. Thus, beginning our investigation with $d = d_0$ will allow us to study the transient amplification phenomenon within the region of attraction of the steady baseflow. Using $d_0$, we can perform sparse NLOP to find an optimal perturbation that maximizes energy at time $T$, where energy is defined as:

$$E(T) = ||x(T)||_2^2. \tag{20}$$

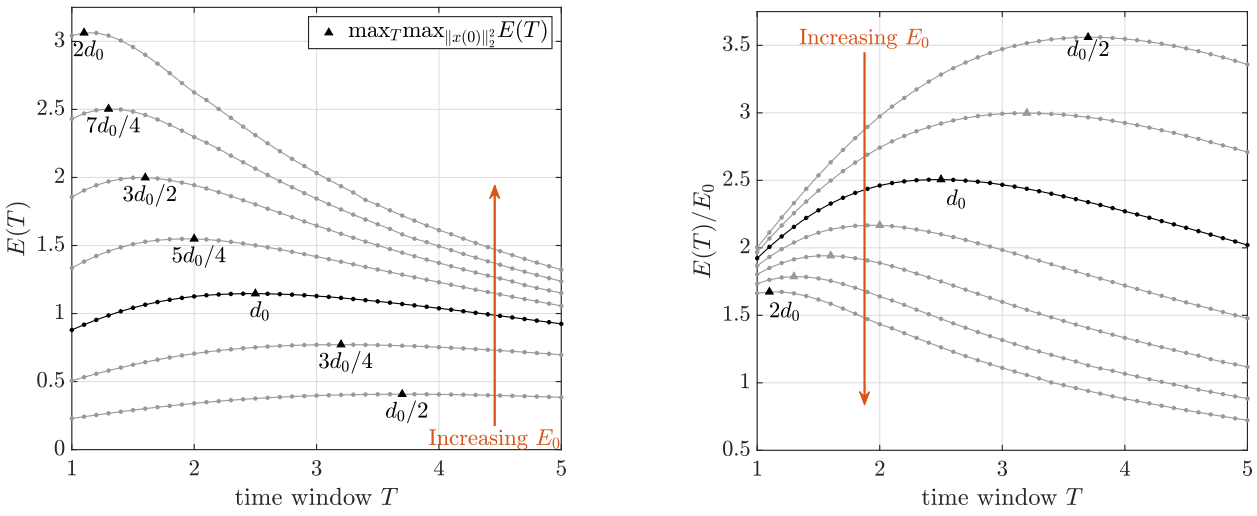
In addition to the initial perturbation energy $d_0^2$, we also need to define the length of the time-horizon $T$. The black curve in figure 3a shows the maximum $E(T)$ obtained by performing sparse NLOP over a grid of $T$ values. For all the results in figure 3a, the sparsity promoting parameter $\sigma$ was chosen so that the cardinality of the optimal perturbation would be $k = 4$ in all cases. For $d = d_0$, the maximum $E(T)$ over all $T$ is found to correspond to a time horizon of $T = 2.5$. For this case, the maximized energy $E(T = 2.5) = 1.28$ was the largest out of forty tested $T$ values. To examine the amplification of the initial perturbation, we consider the transient growth

$$\frac{E(T)}{E_0} = \frac{||x(T)||_2^2}{d^2}. \tag{21}$$

The black curve in figure 3b shows that for $d = d_0$ and $T = 2.5$, the initial perturbation energy, $E_0$, is amplified by a factor of 2.5.

As is demonstrated by the gray lines in figure 3, increasing the perturbation magnitude $d$ results in an earlier and higher spike in the maximum energy $E(T)$. In other words, the identified initial perturbations $x(0)$ tend to drive the flow away from its steady attractor over a shorter time as $E_0$ is increased. However, because the amplification is maximized

at an earlier time, increasing the initial energy $E_0$ causes a decrease in the maximum transient energy growth. In other words, the identified initial perturbations become less efficient at driving the flow away from the steady attractor, due to a higher input energy being required.




**(a)** $E$ **(b)** $E/E_0$

**Fig. 3 Increasing the initial perturbation size $E_0$ pushes the the energy-maximizing flow state closer to $t = 0$, resulting in a quicker return to energy decay $E(T)/E_0 < 1$. The sparsity parameter $\sigma$ was tuned so that the cardinality of the initial perturbations would be $k = 4$. As shown in figure 5, the energies are approximately the same for cardinalities between $k = 9$–$k = 4$, so we chose to show $k = 4$ as the sparsest of these cases.**

The sparsity promoting framework described in III has no guarantee of producing a specific cardinality $k$. As such, it is necessary to search over a range of $\sigma$ values to obtain optimal perturbations with different cardinalities. To do this, $d$ and $T$ must first be fixed as constants. We report results for the perturbation size $d = d_0 = 0.676$, which was shown in figure 3 to correspond to a time horizon of $T = 2.5$. As discussed in section IV.A, a linear estimate is used to seed the non-sparse NLOP, which gives the optimal perturbation vector $x(0)$ corresponding to the column labeled ($\sigma_9 = 0$) in figure 4 (the subscript on $\sigma$ indicates the cardinality $k$). The resultant solution is used as a seed to perform sparse NLOP for a range of $\sigma$ values, revealing optimal perturbations as shown in figure 4. Note that we do not report a result for $k = 2$, as we were unable to find a corresponding $\sigma$ value using our search grid. The energy amplification associated with each sparse optimal perturbation presented in figure 4 is shown in figure 5. The energies at $T$ associated with the optimal perturbations for $\sigma_9$ through $\sigma_4$ do not differ significantly, while the two sparsest models correspond to significantly smaller maximum energies at time $T$. The comparable transient energy growth between the non-sparse optimal perturbation ($k = 9$) and the sparse optimal perturbation with $k = 4$ suggests a strong sensitivity to the initial seed and algorithmic constants such as the convergence tolerance $\epsilon$. By seeding the non-sparse NLOP with a linear heuristic, more gradient ascent steps were needed to converge to tolerance $\epsilon$ when compared to sparse NLOP. We note that the method is also highly sensitive to the gradient step size $\Delta$, which is why we used an inexact line search to find $\Delta$ in-the-loop. The convergence properties of gradient-based optimization are a well known weakness of this class of methods [12], and are likely the cause of the 0.1% increase in transient energy between $k = 9$ and $k = 4$. We also note that although the energies shown in figure 4 are maximized, the sparsification values may differ depending on the initial guess for $x(0)$.

The trajectories of the individual modes and the associated transient energies are shown in figure 6 for cardinalities $k = 9$, $k = 5$, and $k = 1$. We note that the spanwise flow modes and normal vortex modes are the first to be pruned, as in the case of $Re = 20$ their growth is an order of magnitude lower than the other modes. This results in $x_4$–$x_8$ experiencing no amplification when $k = 1$. The sparsity patterns shown in figure 4 reveal that actuation of the vortex mode $x_3$ is central to maximizing transient energy growth. Furthermore, the sparsity pattern highlights the dominance of modes $x_1$ and $x_9$. This result is particularly noteworthy because the authors of the 9-state model emphasize the importance of the basic profile modification mode $x_9$ in the formulation of the model [18]; the sparse NLOP identifies $x_9$ as a dynamically significant mode in the model without any prior knowledge of this fact.

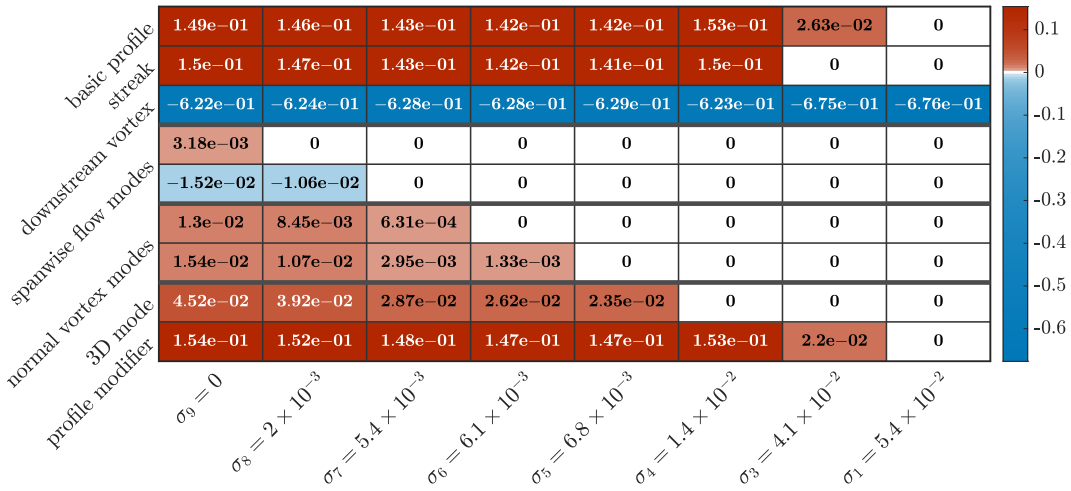



**Fig. 4 Sparsity patterns in the optimal perturbations found for the case where $Re = 20$, $T = 2.5$, and $d_0 \approx 0.676$. The columns each correspond to a sparsity parameter $\sigma$ that maximized energy for the particular sparsity value, and the rows correspond to the Fourier modes. Note that every optimal perturbation satisfies $||x(0)||^2 = d^2$.**

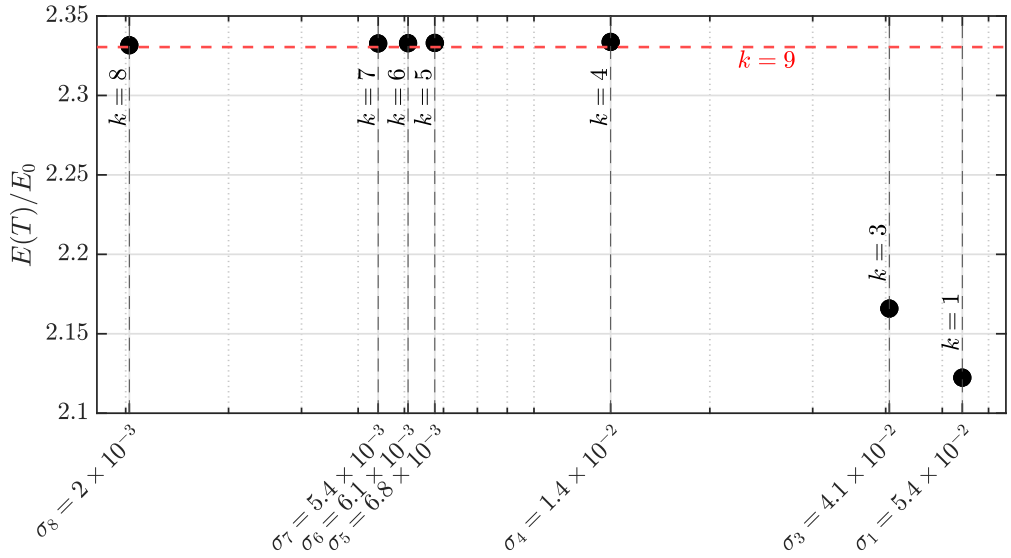



**Fig. 5 The transient energy growths corresponding to each sparsified models shown in figure 4 vs. the sparsification variable $\sigma$.**

We will now show that the sparse optimal perturbations identified lead to mechanisms that are similar to those observed during a transition to turbulence. Using equation 19 and the optimal initial perturbations, we first obtain velocity fields for the sinusoidal shear flow. We observe that for $Re = 20$, the flow exhibits similar features for all initial perturbations shown in figure 3, regardless of cardinality. Furthermore, the downstream vortex mode $x_3$ is maximized for all aforementioned optimal perturbations. A comparison of flow states corresponding to two disturbance sizes and cardinalites is provided in figure 8.

As shown in figures 8 and 9, maximizing mode $x_3$ in the optimal perturbation leads to the formation of strong stream-wise vortices. The stream-wise vortices in turn form streaks, which are clearly visible in the $x - z$ plots for $t > 0$. This behavior is similar to the first step of the 'self-sustaining process' [19, 24], which consists of three phases: (1) formation of streaks by stream-wise vortices, (2) the breakdown of streaks, and (3) the regeneration of the stream-wise vortices. When a sufficiently high Reynolds number is considered, each stage preempts the next, resulting in a self-regenerating turbulent cycle. Although the process is not regenerative for the configuration examined here,

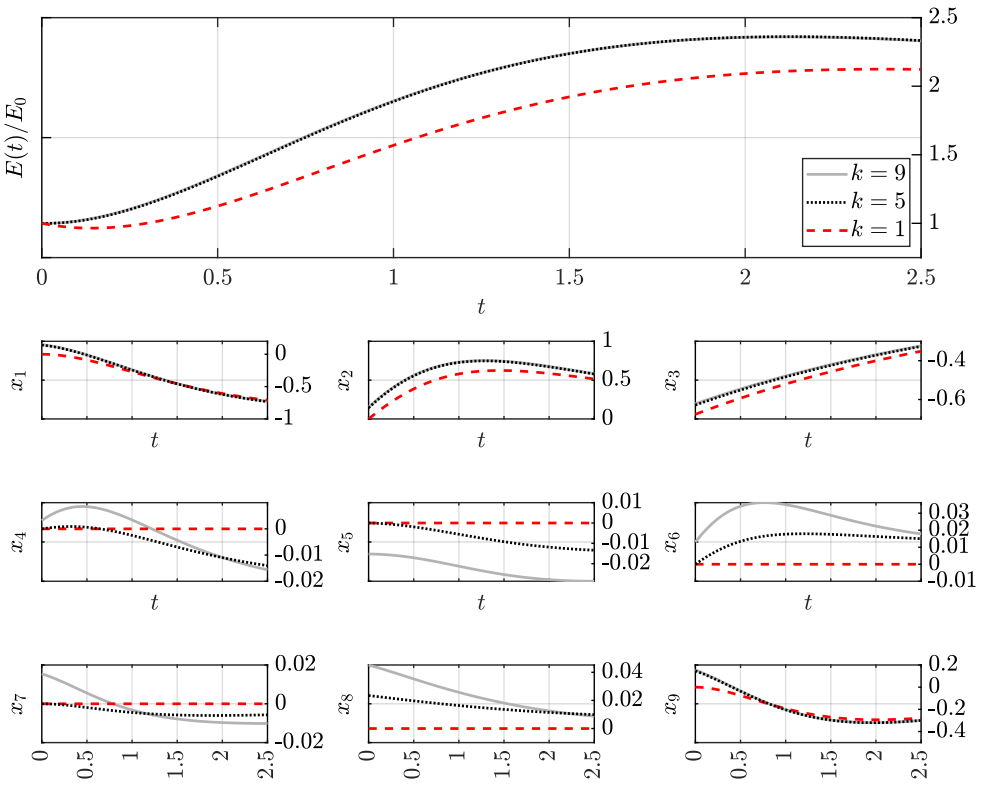
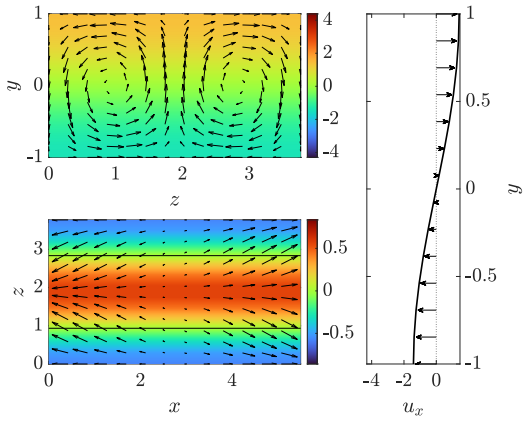



**Fig. 6 Energy and mode-trajectories for the models obtained from $\sigma_9$, $\sigma_5$, and $\sigma_1$ in figure 4. The energy $E(T = 2.5)/E_0$ for $k = 1$ is lower than for the non-sparse case, as modes $x_4$ through $x_8$ remain at zero for all $t$.**

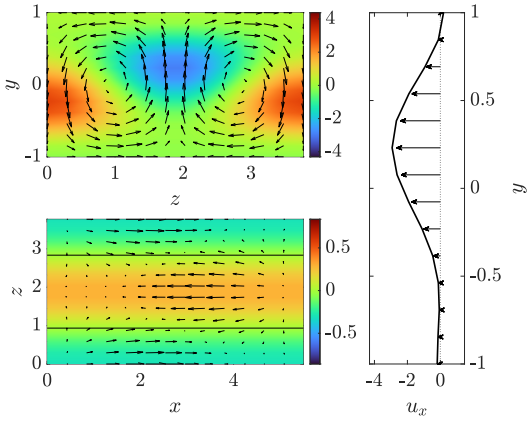
features that resemble the transition are present. To show this, an overview of a $k = 1$ flow with $d = d_0$ is given for several time steps in figure 9. Here, the stream-wise velocity is amplified over the time interval $[0, T]$ which in the case of $k = 1$ is solely the result of vortex and streak induced cross-stream disturbances. This behavior, known as the 'lift-up' mechanism [25], drives the flow to its maximum transient energy at time $T$. The lifting of the streak to form a locally high inflectional velocity profile at time $t = T$ resembles the disturbance growth mechanism that is key in the transition to turbulence [26, 27].

The result reported in figure 4 is not the only one; multiple solution branches emerged for this system. By changing the seed from which sparse NLOP is performed, the optimal perturbations and associated $\sigma$ values change, as shown figure 10. The simultaneous existence of multiple solution branches as a result of differing initial guesses is a feature of these types of optimization frameworks, as shown in [28]. Although the signs and exact values of the optimal perturbations in figure 10 differ from those in figure 4, the sparsity pattern is similar. Furthermore, the maximum amplifications corresponding to each cardinality are the same as those presented in figure 5. The flows generated using the optimal perturbations from figure 10 exhibit the same mechanisms discussed previously, though flipped about the coordinate axes. This flip makes sense as a property of the sinusoidal shear flow is its top-down symmetry. We note that the mean flow profile shape at $t = T$ is also the same which together with the previous results suggests that the flow is most sensitive around $y = \pm 0.2$ for $Re = 20$.

Likely due the low Reynolds number, after achieving its maximum amplification at time $T$, the energy $E(t > T)$ decays monotonically. Increasing the initial perturbation energy $E_0$ effectively 'shifts' the maximizing $T$ closer to $t = 0$, as is shown in figure 3, resulting in a quicker decay in $E(t)$. This is a consequence of the higher $E_0$ resulting in stronger initial vortices, which in turn cause stronger streaks. These streaks are then lifted up over a shorter time window $T$, causing the velocity profile to reach its locally maximized inflection sooner for higher $E_0$. Although we do not observe the onset self-sustaining process, these fundamental behaviors bear a strong resemblance to well-established mechanisms that lead to sustained turbulence in shear flows[18, 24, 26].

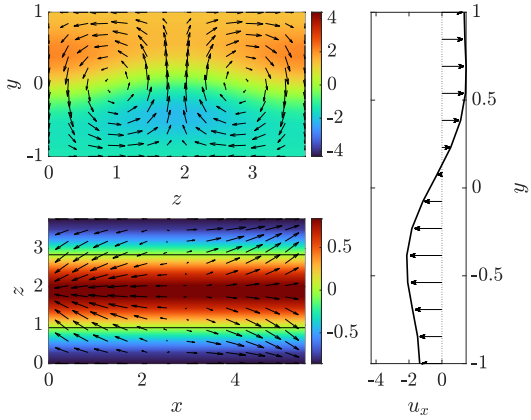



**(a)** $t = 0$ **with** $k = 1$ **and** $d = d_0$

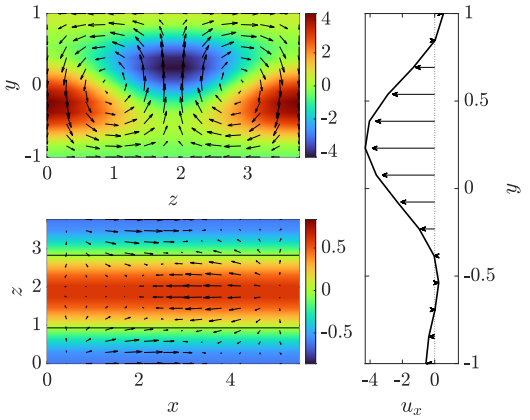



**(b)** $T = 2.5$ **with** $k = 1$ **and** $d = d_0$

**Fig. 7 Velocity fields corresponding to the sparsest ($\sigma_1$) optimal perturbation from figure 4 with perturbation size $d_0$. Within each block, the top left plots show the flow averaged in the streamwise direction, with the arrows indicating the $(y, z)$ plane flow and the color indicating the averaged x-velocity. The bottom left plots show the flow in the midplane between plates (at $y = 0$). Here the arrows show the $(x, z)$ plane velocity and the colors indicate the $y$ velocity. Finally, the right plots shows the mean velocity profile.**

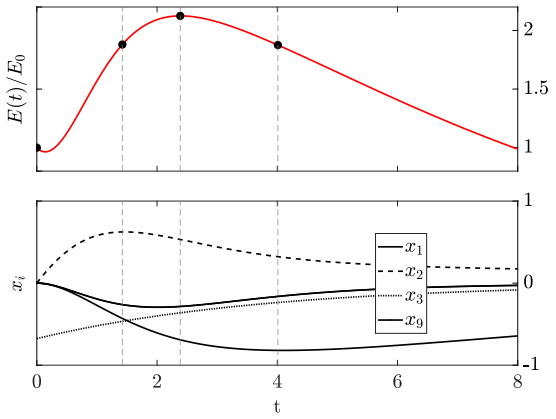



**(a)** $t = 0$ **with** $k = 4$ **and** $d = 7d_0/4$

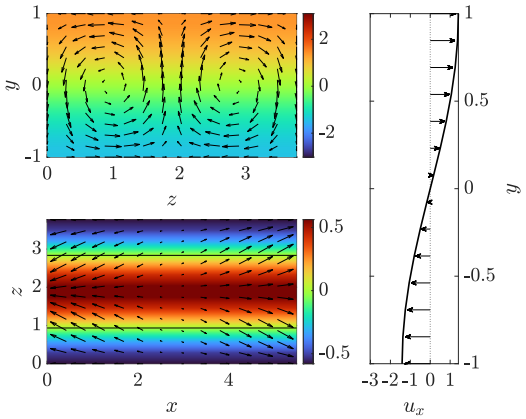



**(b)** $T = 0.4$ **with** $k = 4$ **and** $d = 7d_0/4$

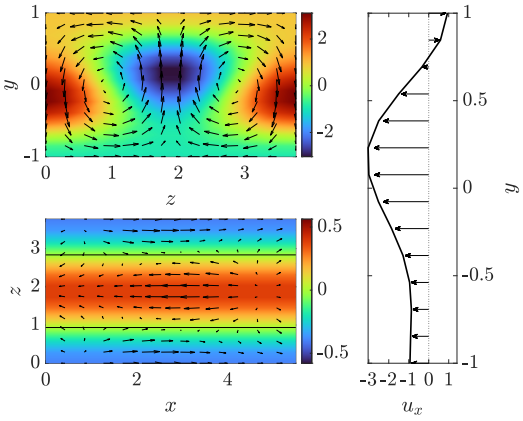
**Fig. 8 The flows at $T$ exhibit the same fundamental behavior for all $k$ and perturbation sizes that were shown in figure 3. Here, $k = 4$ and the perturbation is $7d_0/4$. The lift-up mechanism is the same as seem in figure 7, though due to the higher $E_0$ the magnitude of the disturbances are greater.**



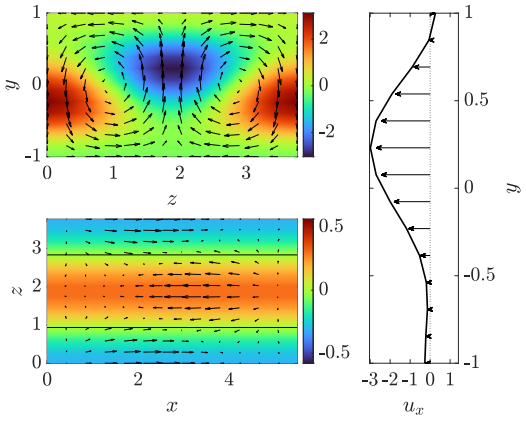

**(a) For $k = 1$, modes $x_4$–$x_8$ are zero for all $t$.**



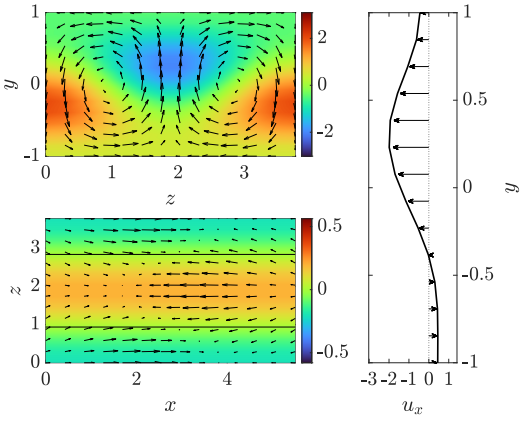

**(b) $t = 0$: Streak formation through initial vortices.**



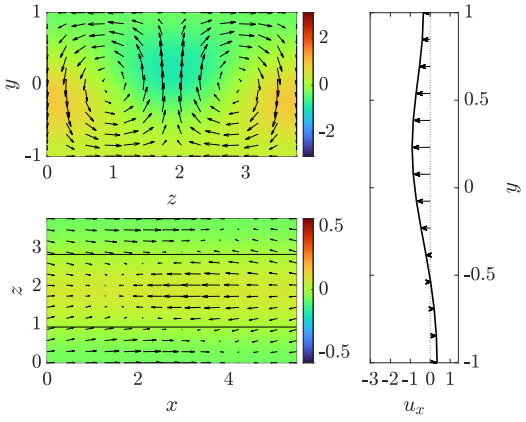

**(c) $t = 1.4$: The streak mode ($x_2$) is maximized.**




**(d) $t = 2.5$: Transient energy $E(t)/E_0$ is maximized.**

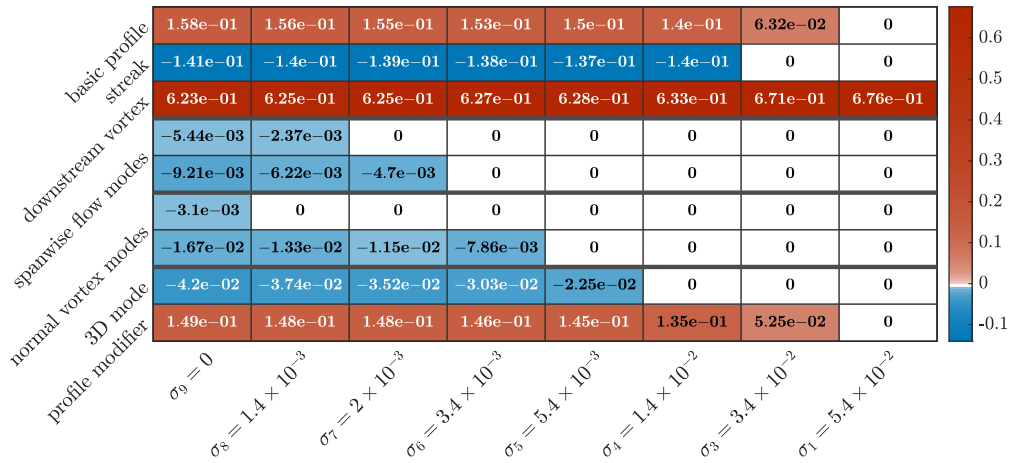



**(e) $t = 4$: Magnitude of $x_1$ maximized, the streaks begin to break down.**

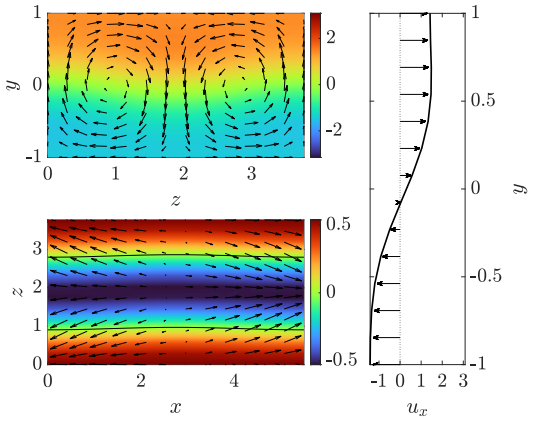



**(f) $T = 8$: Monotonic energy decay from this point $E(t) = E_0$.**

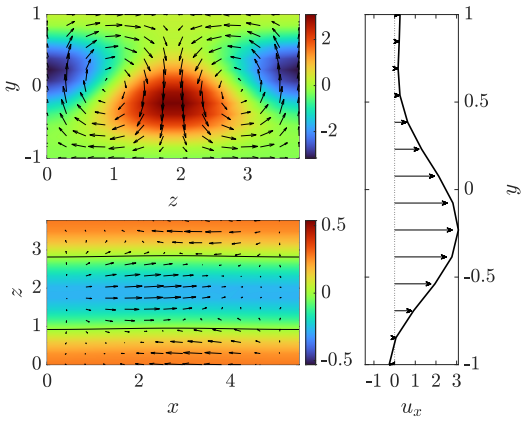
**Fig. 9 By maximizing the streak-formation at $t = 0$, energy is maximized for time $T$, revealing a mechanism similar to the beginning of a turbulent 'self sustaining' process.**




**(a) Just as for figure 4,the optimal perturbations are found for the case where $Re = 20$, $T = 2.5$, and $d_0 \approx 0.676$. However, the initial seed is different from that used to find figure 4.**




**(b) $t = 0$ with $k = 5$: Streak formation.**




**(c) $T = 2.5$ for $k = 5$: $E(T)$ is maximized.**

**Fig. 10 Even though the solutions to the spare NLOP problem are non-unique, the energy-maximizing mechanism is the same. The streak formation and base profile deflection is the same as in figures 9 and 8, though it is flipped across $y = 0$. The energies $E(T)$ associated with this result are the same as for the optimal perturbations in figure 4.**

## V. Conclusions

In this work, we presented a framework for computing sparse and localized optimal perturbations that maximize the transient growth of perturbation energy in nonlinear dynamic systems. The proposed variational solution method enables the analysis of stability and transient growth by revealing a set of sparse optimal initial perturbations. We first demonstrated the sparse NLOP framework on a simple 2-state model to develop an intuition for the approach and the underlying optimization problem. The graphical analysis afforded by this simple example also facilitated the development of a simple heuristic for seeding the sparse NLOP problem using a linear estimate. Subsequently, we applied the sparse NLOP framework to analyze a reduced-order (9-state) model of a sinusoidal shear flow at a $Re = 20$. Notably, the method identified the basic profile modifier mode, a core feature of the 9-state model, as an important mode for driving instabilities. In addition, the sparse NLOP framework identified perturbations giving rise to the formation of streaks through strong downstream vortices and as the driver of instability. Our observation is consistent with the literature and resembles the initial steps of the 'self sustaining process' that drives recurring turbulence. In future work we plan to implement sparse NLOP directly on the full Navier-Stokes equations. Sparse NLOP will also be employed to address questions pertaining to optimal actuator placement for active flow control applications. The framework outlined here opens new possibilities for the analysis and control of nonlinear flows.

## Acknowledgments

This material is based upon work supported by the Air Force Office of Scientific Research under award number FA9550-21-1-0106 and FA9550-21-1-0434, and the National Science Foundation under grant number CBET-1943988.

## References

[1] Schmid, P. J., and Henningson, D. S., *Stability and Transition in Shear Flows*, Springer–Verlag, New York, 2001.

[2] Schmid, P. J., "Nonmodal Stability Theory," *Annual Review of Fluid Mechanics*, Vol. 39, No. 1, 2007, pp. 129–162.

[3] Schmid, P. J., and Brandt, L., "Analysis of Fluid Systems: Stability, Receptivity, Sensitivity," *Applied Mechanics Review*, 2014.

[4] Kerswell, R., Pringle, C., and Willis, A., "An optimization approach for analysing nonlinear stability with transition to turbulence in fluids as an exemplar," *Reports of Progress in Physics*, Vol. 77, No. 085901, 2014.

[5] Kerswell, R., "Nonlinear Nonmodal Stability Theory," *Annual Review of Fluid Mechanics*, 2018.

[6] Bhattacharjee, D., Klose, B., Jacobs, G., and Hemati, M., "Data-driven selection of actuators for optimal control of airfoil separation," *Theoretical and Computational Fluid Dynamics*, Vol. 34, 2020.

[7] Skene, C., Yeh, C., Schmid, P., and Taira, K., "Sparsifying the resolvent forcing mode via gradient-based optimisation," *Journal of Fluid Mechanics*, Vol. 944, No. A52, 2022.

[8] Tamilselvam, P., Asztalos, K., and Dawson, S., "Transient growth analysis of flow over an airfoil for identifying high-amplification, spatially-localized inputs," *AIAA Paper*, 2022.

[9] Berggren, M., "Numerical Solution of a Flow-Control Problem: Vorticity Reduction by Dynamic Boundary Action," *SIAM Journal on Scientific Computing*, Vol. 19, No. 3, 1998, pp. 829–860.

[10] Hinze, M., Walther, A., and Sternberg, J., "An optimal memory-reduced procedure for calculating adjoints of the instationary Navier-Stokes equations," *Optimal Control Applications and Methods*, Vol. 27, 2006.

[11] Nocedal, J., and Wright, S. J., "Numerical Optimization," 1999.

[12] Beck, A., *First-Order Methods in Optimization*, Society for Industrial and Applied Mathematics, Philadelphia, 2017.

[13] M. Journée, P. R., Y. Nesterov, and Sepulchre, R., "Generalized Power Method for Sparse Principal Component Analysis," *Journal of Machine Learning Research*, Vol. 11, 2010, p. 517–553.

[14] Kalur, A., Seiler, P., and Hemati, M., "Nonlinear stability analysis of transitional flows using quadratic constraints," *Physical Review Fluids*, Vol. 6, No. 044401, 2021.

[15] Kalur, A., Mushtaq, T., Seiler, P., and Hemati, M., "Estimating regions of attraction for transitional flows using quadratic constraints," *IEEE Control Systems Letters*, Vol. 6, 2022.

[16] L. F. Toso, R. D., and Duncan, S. R., "Regional stability analysis of transitional fluid flows," *IEEE Control Systems Letters*, 2022.

[17] Liao, S., Hemati, M., and Seiler, P., "Quadratic Constraints for Local Stability Analysis of Quadratic Systems," *IEEE Conference on Decision and Control*, 2022.

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

[19] Waleffe, F., "On a self-sustaining process in shear flows," *Physics of Fluids*, Vol. 9, No. 4, 1997, pp. 883–900.

[20] Goulart, P. J., and Chernyshenko, S., "Global stability analysis of fluid flows using sum-of-squares," *Physica D: Nonlinear Phenomena*, Vol. 241, No. 6, 2012, pp. 692–704.

[21] Liu, C., and Gayme, D. F., "Input-output inspired method for permissible perturbation amplitude of transitional wall-bounded shear flows," *Physical Review E*, Vol. 102, No. 6, 2020.

[22] Kalur, A., Mushtaq, T., Seiler, P., and Hemati, M. S., "Flows Using Quadratic Constraints," Vol. 6, 2022, pp. 482–487.

[23] Mushtaq, T., Seiler, P., and Hemati, M. S., "A framework for feedback stabilization of incompressible flows using quadratic constraints," *AIAA AVIATION 2022 Forum*, 2022, pp. 1–10.

[24] Hamilton, J. M., Kim, J., and Waleffe, F., "Regeneration mechanisms of near-wall turbulence structures," *Journal of Fluid Mechanics*, Vol. 287, 1995, p. 317–348.

[25] Ellingsen, T., and Palm, E., "Stability of linear flow," *The Physics of Fluids*, Vol. 18, No. 4, 1975, pp. 487–488.

[26] Landahl, M. T., "Dynamics of boundary layer turbulence and the mechanism of drag reduction," *The Physics of Fluids*, Vol. 20, No. 10, 1977, pp. S55–S63.

[27] Brandt, L., "The lift-up effect: The linear mechanism behind transition and turbulence in shear flows," *European Journal of Mechanics, B/Fluids*, Vol. 47, No. March 2014, 2014, pp. 80–96.

[28] Foures, D. P., Caulfield, C. P., and Schmid, P. J., "Localization of flow structures using1-norm optimization," *Journal of Fluid Mechanics*, Vol. 729, 2013, pp. 672–701.