

## Identifying spatially localized instability mechanisms using sparse optimization

Talha Mushtaq\* and Maziar S. Hemati<sup>†</sup>

*Department of Aerospace Engineering and Mechanics, University of Minnesota, Minneapolis, Minnesota 55455, USA*



(Received 14 October 2024; accepted 23 January 2026; published 6 April 2026)

Recent investigations have established the physical relevance of spatially localized instability mechanisms in fluid dynamics and their potential for technological innovations in flow control. In this paper, we show that the mathematical problem of identifying spatially localized optimal perturbations that maximize perturbation-energy amplification can be cast as a sparse (cardinality-constrained) optimization problem. Unfortunately, cardinality-constrained optimization problems are nonconvex and combinatorially hard to solve in general. To make the analysis viable within the context of fluid dynamics problems, we propose an efficient iterative method for computing suboptimal spatially localized perturbations. Our approach is based on a generalized Rayleigh quotient iteration algorithm followed by a variational renormalization procedure that reduces the optimality gap in the resulting solution. The approach is demonstrated on a subcritical plane Poiseuille flow at  $Re = 4000$ , which has been a benchmark problem studied in prior investigations on identifying spatially localized flow structures. Notably, we find that a subset of the perturbations identified by our method yield a comparable degree of energy amplification as their global counterparts. We anticipate our proposed analysis tools will facilitate further investigations into spatially localized flow instabilities, including within the resolvent and input-output analysis frameworks.

DOI: [10.1103/llrp-vkln](https://doi.org/10.1103/llrp-vkln)

### I. INTRODUCTION

Nonmodal stability analysis is conducted by determining the maximum amplification a flow perturbation can exhibit over a finite time horizon [1]. The associated optimal flow perturbations that result in the maximum transient growth (i.e., energy amplification) are global in nature, tending to involve many flow quantities distributed across the full spatial domain. However, it is often of physical interest to consider optimal perturbations that are sparse and spatially localized: identification of spatially localized optimal perturbations would reveal specific regions of the spatial domain and the associated flow quantities that are most pertinent to triggering instabilities. Recent investigations have shown that such information can guide the placement of actuators and inform the development of effective flow control strategies [2–4].

Spatially localized stability analysis is naturally posed as a cardinality-constrained optimization problem, as will be shown in this paper. However, cardinality-constrained optimization problems are notoriously difficult to solve due to the inherent need for a combinatorial search. As such, prevailing methods relax the cardinality constraint and instead work with a modified but tractable optimization problem that promotes sparsity in the resulting solution through some heuristic

---

\*Present address: Gulfstream Aerospace Corporation, Savannah GA 31408, USA.

†Contact author: [mhemati@umn.edu](mailto:mhemati@umn.edu)

[2,5–8]. For example, sparsity-promoting  $\ell_1$  norms have been used to sparsify optimal forcing modes in resolvent analysis [2,6] and to sparsify optimal perturbations in nonlinear nonmodal stability analysis [7,8]. Despite their successes, the drawback of employing sparsity-promoting heuristics is that such analyses do not guarantee a given degree of sparsity *a priori*, and so will often require computationally expensive parameter sweeps to realize sparse and physically meaningful solutions.

In this paper, we revisit cardinality-constrained optimization as the natural setting for conducting spatially localized instability analysis. For brevity, this paper focuses on linear nonmodal stability analysis (see Sec. II A), but the ideas developed are equally valid within the context of any analysis method based on generalized Rayleigh quotient maximization (e.g., resolvent and input-output analysis). As noted earlier, cardinality-constrained optimization problems are nonconvex and combinatorially hard to solve in general. To make the analysis tractable, we propose an efficient iterative method for computing sub optimal spatially localized perturbations (see Sec. II B). Our approach employs two key steps: (i) find a “good” sparsity pattern for a given cardinality  $k$ , followed by (ii) determine an optimal solution for that sparsity pattern. In particular, we formulate a generalized Rayleigh quotient iteration algorithm to achieve (i), then apply a variational renormalization procedure to reduce the optimality gap in the solution to achieve (ii). The approach is demonstrated on a subcritical plane Poiseuille flow at  $\text{Re} = 4000$ , which has been a benchmark problem studied in prior investigations on identifying spatially localized flow structures (see Sec. III). Notably, we find that a subset of the perturbations identified by our method yield a comparable degree of energy amplification as their global counterparts. Further, our proposed approach decisively reveals the well-known lift-up mechanism as a driver of transient growth: sparse and spatially localized optimal perturbations are comprised solely of the wall-normal component of velocity, which results in transient amplification of streamwise streaks. Concluding remarks are given in Sec. IV.

## II. PROBLEM FORMULATION AND SOLUTION APPROACH

### A. Problem formulation

Consider the linear dynamics of perturbations about a steady base flow in the spatially discretized form,

$$\frac{d}{dt}q(t) = Lq(t), \quad (1)$$

where  $q \in \mathbb{C}^n$  denotes the state vector [9]. The state response due to an initial perturbation  $q(t_0) = q_0$  at any instant  $t \geq t_0$  can be determined as  $q(t) = \Phi(t, t_0)q_0$ , where  $\Phi(t, t_0)$  is the state transition matrix. Thus, given the knowledge of the state at any time  $t$ , the energy can be computed as  $E(t) = q(t)^H Q q(t)$ , where  $(\cdot)^H$  denotes Hermitian transpose and  $Q = Q^H > 0$ . For example, when  $E$  is the kinetic energy, then  $Q$  would be an appropriate matrix of energy and quadrature weights. The amplification of energy over a prescribed time interval can then be expressed in terms of the generalized Rayleigh quotient

$$R(q_0, P, Q) = \frac{q_0^H P q_0}{q_0^H Q q_0}, \quad (2)$$

where  $P = P(t, t_0) := \Phi^H(t, t_0) Q \Phi(t, t_0)$ . If  $L^H Q + Q L \not\leq 0$ , then there will exist an initial condition  $q_0$  that triggers a nonunity transient growth, i.e., a nontrivial amplification of energy  $R(q_0, P, Q) = E(t)/E(t_0) > 1$  for some  $t > t_0$ . In nonmodal stability analysis, the maximum amplification is of interest. The *optimal perturbation*  $q_0^*$  that maximizes  $R(q_0, P, Q)$  can be found by solving the generalized eigenvalue problem

$$P q_0 = \lambda Q q_0. \quad (3)$$

The corresponding maximum energy amplification  $G(t)$  is simply

$$G(t) = R(q_0^*, P, Q) = \lambda_{\max}(P, Q), \quad (4)$$

where  $\lambda_{\max}(P, Q)$  denotes the maximum generalized eigenvalue of the pair  $(P, Q)$ . We note that this generalized Rayleigh quotient maximization problem is equivalent to the following quadratically constrained quadratic program (QCQP),

$$\begin{aligned} & \underset{q_0}{\text{maximize}} && q_0^H P q_0 \\ & \text{subject to} && q_0^H Q q_0 = 1. \end{aligned} \quad (5)$$

Further, the generalized eigenvalue problem results because we are aiming to maximize the generalized Rayleigh quotient  $R(q_0, P, Q)$ . Using a simple change of variables  $y(t) = Q^{1/2}q(t)$  and defining  $S := Q^{-1/2}PQ^{-1/2} > 0$ , we could instead find  $G(t)$  by solving a standard eigenvalue problem. In particular,  $G(t) = R(y_0^*, S)$ , where  $y_0^*$  is the optimal perturbation represented in the transformed coordinates and  $R(y_0, S) := R(y_0, S, I)$  is the standard Rayleigh quotient with  $I$  denoting the  $n \times n$  identity matrix.

In general, optimal perturbations  $q_0^*$  are global in nature, i.e., the unit energy associated with the perturbation is distributed among all state variables and therefore across all of the spatial domain. In this work, we seek “sparse” solutions  $q_{0,\text{sp}}^*$  that correspond to a subset of the state variables and tend toward spatially localized perturbations, i.e., the perturbation energy is concentrated locally, rather than distributed globally. To this end, we propose solving the cardinality-constrained QCQP,

$$\begin{aligned} & \underset{q_0 \in \mathbb{C}^n}{\text{maximize}} && q_0^H P q_0 \\ & \text{subject to} && q_0^H Q q_0 = 1 \\ & && \text{card}(q_0) \leq k, \end{aligned} \quad (6)$$

which yields solutions that are at most  $k$  sparse (i.e., the number of nonzero entries in  $q_{0,\text{sp}}^*$  will be at most  $k$ ). For the case where  $k = n$ , the cardinality constraint becomes redundant and the optimization in (6) simplifies to (5). However, when  $k < n$ , the optimization in (6) is nonconvex and the optimal solution requires a combinatorial search over all possible sparsity patterns.

The combinatorial nature of the solution to (6) makes optimal solutions intractable for most fluid dynamics applications due to the high dimension of the state space. However, for relatively small problem instances (e.g.,  $n \lesssim 30$ ) optimal solutions can be computed using combinatorial-optimization methods, e.g., dedicated branch-and-bound algorithms [10]. As such, various heuristics and convex relaxations are often used to obtain suboptimal solutions. For instance, it is common to incorporate sparsity-promoting  $\ell_1$  norms to regularize the optimization problem and obtain sparse solutions to a different—but now tractable—optimization problem. However, methods based on sparsity-promoting norms generally do not guarantee a particular degree of sparsity *a priori*; rather, the sparse and “physically meaningful” solutions are identified only after sweeping over a large grid of regularization parameters and solving each associated optimization problem.

To circumvent the issues inherent to the relaxed solution methods, we present a computationally efficient numerical algorithm that is simple to implement and converges to suboptimal solutions of the cardinality-constrained QCQP in (6). Any solution to (6) is guaranteed *a priori* to be  $k$  sparse [i.e.,  $\text{card}(q_0) = k$ ]. As we will see, the resulting sparse solutions tend toward spatially localized perturbations that reveal pertinent flow physics and instability mechanisms.

Before we discuss our proposed solution approach, it is instructive to consider again the change of variables  $y(t) = Q^{1/2}q(t)$  discussed above. In this case, the equivalent cardinality-constrained QCQP will be

$$\begin{aligned} & \underset{y_0}{\text{maximize}} && y_0^H S y_0 \\ & \text{subject to} && y_0^H y_0 = 1 \\ & && \text{card}(Q^{-1/2}y_0) \leq k. \end{aligned} \quad (7)$$

The representation of the QCQP in (7) explicitly highlights the fact that the cardinality constraint is imposed on a different set of variables than the set used in defining the energy norms, i.e., cardinality is imposed on  $q_0 = Q^{-1/2}y_0$ , and not on  $y_0$ . This is an important point to note because it explicitly shows that the underlying problem structure is *not* equivalent to a sparse PCA (principle components analysis) problem for which  $\text{card}(y_0) \leq k$ , but rather a sparse LDA (linear discriminant analysis) problem for which  $\text{card}(Q^{-1/2}y_0) \leq k$  [10]. In the next section, we will describe how we can extend existing algorithms for sparse PCA to accommodate the special sparse LDA problem structure inherent to (6).

### B. Solution approach

We begin with the following observation: For any  $k$ -sparse vector  $q_{0,\text{sp}} \in \mathbb{C}^n$  and a given matrix pair  $(P, Q)$ ,

$$\lambda = R(q_{0,\text{sp}}, P, Q) = R(z, P_k, Q_k), \quad (8)$$

where  $z \in \mathbb{C}^k$  is the subvector containing only the  $k$  nonzero indices of  $q_{0,\text{sp}}$  and  $P_k \in \mathbb{C}^{k \times k}$  and  $Q_k \in \mathbb{C}^{k \times k}$  are, respectively, the principal minors obtained from deleting the rows and columns of  $P$  and  $Q$  associated with zero indices of  $q_{0,\text{sp}}$ . Thus, the  $z$  that maximizes  $\lambda$  in (8) will be the principal generalized eigenvector of the pair  $(P_k, Q_k)$ . It follows that an optimal subvector  $z^*$  obtained from *any* method solving (6) must satisfy  $z^* = v_{\max}(P_k, Q_k)$  and consequently  $\lambda^* = \lambda_{\max}(P_k, Q_k)$  [10]. Therefore, computing an optimal sparse solution from (6) inherently requires a combinatorial search over all possible sparsity patterns for  $k$ -sparse vectors  $q_{0,\text{sp}}$ .

Here we propose a two-step approach to solving (6) that avoids a combinatorial search:

- (i) find a “good” sparsity pattern for a given cardinality  $k$ , then
- (ii) determine an optimal solution for that sparsity pattern.

In principle, step (i) involves using any systematic (or *ad hoc*) approach to find a sparsity pattern. A simple approach can be devised by noting that any (nonsparse) candidate solution  $q_0$  can be “sparsified” to a prescribed cardinality  $k$  via Euclidean projection  $\mathbb{P}_k$  onto the set  $\{\|q_0\|_0 = k\} \cap \{\|Q^{1/2}q_0\|_2 = 1\}$ . This approach is simple and efficient because it amounts to setting the  $(n - k)$  smallest entries in  $q_0$  to zero, then renormalizing the resulting vector to have unit norm. A straightforward way to do this is to use the “simple thresholding” technique, which involves zeroing out  $(n - k)$  components of  $q_0$  with the smallest absolute values. This technique is popular for sparse PCA problems and avoids any combinatorial searches (see [10] for details). Of course, there is no guarantee that the resulting solution will possess a “good” sparsity pattern for optimizing the underlying problem; however, this would be a simple way of sparsifying existing solutions for (nonsparse) optimal perturbations  $q_0^*$  determined from the generalized eigenvalue problem in (5). Here, we propose an efficient numerical method inspired by the generalized Rayleigh quotient iteration (GRQI) algorithm originally presented in [11] to solve sparse PCA problems. We will introduce our extension of the GRQI method to the more general setting of sparse LDA momentarily.

Step (ii) is central to our approach as it ensures that a suboptimal solution to the original problem posed in (6) is determined. The key observation we exploit in conducting step (ii) is as follows: Fixing the sparsity pattern on the decision variable  $q_0$  in (6) fixes the  $k$  working indices for the problem and the original cardinality constraint becomes redundant. In this case, we can use knowledge of the working indices to define  $q_{0,\text{sp}}$  and its associated subvector  $z_0$ . It then follows that the optimal sparse perturbation  $q_{0,\text{sp}}^*$  for  $k$  working indices can be determined by maximizing the generalized Rayleigh quotient  $R(z, P_k, Q_k)$ , which is simply the largest generalized eigenvalue  $\lambda_{\max}(P_k, Q_k)$ . This is sometimes referred to as *variational renormalization* in the machine learning literature [10]. An important feature of variational renormalization is that it can be used to reduce the optimality gap in candidate solutions to (6) derived from *any* method, including those based on sparsity-promoting norms and problems for which the cardinality constraint is imposed on a different variable. In addition, maximizing  $R(z, P_k, Q_k)$  in this step is computationally inexpensive because we work with the reduced matrices  $P_k$  and  $Q_k$ . We emphasize here that this variational

renormalization approach works just as well in reducing the optimality gap in problems of sparse resolvent and input-output analysis [2,12].

It is important to note that the variational renormalization in step (ii) yields an optimal solution for a given sparsity pattern identified in step (i). The *generalized inclusion principle* [10] can be used to derive upper and lower bounds on the maximum amplification for  $k$ -sparse perturbations, regardless of the sparsity pattern. The bounds are given in terms of the generalized eigenvalues as

$$\lambda_k(P, Q) \leq \lambda_{\max}(P_k, Q_k) \leq \lambda_n(P, Q), \quad (9)$$

where  $\lambda_k$  is the  $k$ th smallest generalized eigenvalue of  $(P, Q)$  and  $\lambda_n(P, Q)$  is the largest eigenvalue associated with the maximum amplification in the nonsparse setting. The upper bound establishes that  $k$ -sparse perturbations cannot induce a larger amplification than their nonsparse counterparts over a given time horizon  $t = T$ . The lower bound can guide the range of  $k$  to grid over when pursuing sparse optimal perturbations that are able to impart a transient growth greater than a given threshold.

Returning to step (i), we will now seek “better” solutions than might be obtained from the simple Euclidean projection  $\mathbb{P}_k$  of the nonsparse  $q_0$  followed by a variational renormalization. To this end, we propose a further generalization of the GRQI of [11] in order to solve sparse LDA problems given by the QCQP in (6). We will refer to this generalization as a modified generalized Rayleigh quotient (mGRQI) algorithm so as to distinguish it from the original GRQI algorithm from [11]. As the name suggests, the original GRQI method generalizes the standard Rayleigh quotient iteration (RQI) algorithm [13] to find sparse maximizers of the Rayleigh quotient  $R(y_0, S) = R(y_0, S, I)$ . The GRQI method is attractive for fluids applications because it inherits the rapid (cubic) convergence of RQI and requires only modest computational effort when  $k \ll n$ , which tends to be the case when seeking sparse and spatially localized perturbations: The number of floating point operations per iteration scales as  $\sim \mathcal{O}(k^3 + nk)$  [11]. This is in contrast to generalized power methods [14]—sparsity-promoting variants of the standard power iteration—for which convergence is linear and the number of floating point operations per iteration scales as  $\sim \mathcal{O}(n^2 + nk)$ .

The original GRQI algorithm solves the sparse PCA problem, but not the more general sparse LDA problem in (6). Recall that the distinction is most apparent in the equivalent QCQP in (7), which is obtained through the change of variables  $y_0 = Q^{1/2}q_0$ : in particular, the cardinality constraint is imposed on the inverse bijection of  $y_0$ , i.e.,  $\text{card}(Q^{-1/2}y_0) = k$ . This constraint is difficult to satisfy short of a combinatorial search since it requires that  $z_0 = Q_k^{-1/2}v_0$ . However, we can use this relation between  $z_0$  and  $v_0$  to estimate a “good” working set for  $k$  indices of  $q_0$  within our mGRQI algorithm, as described below. During the course of our investigation, we found that our proposed mGRQI algorithm yielded better sparsity patterns than either a simple Euclidean projection  $\mathbb{P}_k$  of the nonsparse optimal  $q_0^*$  or application of the standard GRQI—in the sense that variational renormalization tended to result in larger transient growth in an overwhelming majority of cases investigated.

The implementation of mGRQI uses the matrix pair  $(P, Q)$  to find a  $k$ -sparse  $q_0$ , denoted as  $q_{0,\text{sp}}$ . Specifically, mGRQI performs a Rayleigh update on the set of  $k$  nonzero indices of  $q_0$  for a fixed sparsity pattern followed by a power update to obtain improvement in the working set for  $q_0$  at each iterate. The power update is similar in nature to the one used in the original GRQI; however, mGRQI performs the update on the full indices of  $y_0$  instead of  $q_0$  to maintain numerical stability and improve solution convergence. Ideally, using  $Q^{-1}P$  for the power update makes sense since it avoids transforming the system into  $y_0$  coordinates and  $q_0$  is directly related to the principal generalized eigenvector of  $Q^{-1}P$ . However,  $Q^{-1}P$  is not generally a self-adjoint operator, which can cause solutions to oscillate and introduce numerical errors. The resulting  $y_0$  from the power update is projected using  $\mathbb{P}_k$  to get  $y_{0,\text{sp}} = \mathbb{P}_k(y_0)$  and consequently, its associated subvector  $v_0$  containing  $k$  nonzero indices. These indices can then be mapped back to the  $k$  indices of  $q_0$  using the relation  $z_0 = Q_k^{-1/2}v_0$ . The resulting  $z_0$  vector can be used to perform a variational renormalization to satisfy

ALGORITHM 1. mGRQI ( $P, Q, q, J, k, \epsilon$ ).

---



---

```

1: Initialize  $j \leftarrow 0, \lambda^{(j)} \leftarrow \lambda_{\max}(P, Q), q_0^{(j)} \leftarrow \mathbb{P}_k(q)$ 
2: repeat
3:    $\mathcal{W} \leftarrow \{i | q_{0_i}^{(j)} \neq 0\}$ 
4:    $q_{0_{\mathcal{W}}}^{(j)} \leftarrow (P_{\mathcal{W}} - \lambda^{(j)} Q_{\mathcal{W}})^{-1} q_{0_{\mathcal{W}}}^{(j)}$  (Rayleigh update)
5:    $q_{0_{\text{new}}} \leftarrow \frac{q_0^{(j)}}{\|Q^{1/2} q_0^{(j)}\|_2}$ 
6:    $y_{0_{\text{new}}} \leftarrow Q^{1/2} q_{0_{\text{new}}}$ 
7:   if  $j < J$  then
8:      $y_{0_{\text{new}}} \leftarrow (Q^{-1/2} P Q^{-1/2}) y_{0_{\text{new}}}$  (power update)
9:   end if
10:   $y_0^{(j)} \leftarrow \mathbb{P}_k(y_{0_{\text{new}}})$ 
11:   $\mathcal{W}_y \leftarrow \{i | y_{0_i}^{(j)} \neq 0\}$ 
12:   $q_{0_{\mathcal{W}_y}}^{(j+1)} = Q_{\mathcal{W}_y}^{-1/2} y_{0_{\mathcal{W}_y}}^{(j)}$ 
13:   $q_0^{(j+1)} \leftarrow \frac{q_0^{(j+1)}}{\|Q^{1/2} q_0^{(j+1)}\|_2}$ 
14:   $\lambda^{(j+1)} \leftarrow (q_0^{(j+1)})^H P q_0^{(j+1)}$ 
15:   $j \leftarrow j + 1$ 
16: until  $\|q_0^{(j)} - q_0^{(j-1)}\|_2 < \epsilon$ 
17: return  $q_0^{(j)}$ 

```

---



---

(8) and obtain  $q_{0_{\text{sp}}}$ . The detailed mGRQI algorithm is reported as Algorithm 1, and the combined two-step algorithm for mGRQI with variational renormalization is reported as Algorithm 2.

### III. RESULTS

#### A. Model problem: Plane Poiseuille flow

We will use the incompressible plane Poiseuille flow model as an example to demonstrate our proposed analysis framework. This is a canonical flow model that has been studied in prior works on stability analysis, including recent techniques investigating methods that promote spatial localization of instability mechanisms [2,5,12]. Plane Poiseuille flow is a pressure-driven flow between two infinite parallel plates separated a distance  $2h$  apart. The steady baseflow is a parabolic profile in the streamwise component of velocity  $U(y) = 1 - y^2$ , where  $y \in [-1, 1]$  is the nondimensional coordinate in the wall-normal direction and  $U(y)$  has been normalized by the centerline velocity. The Reynolds number  $\text{Re} = Uh/\nu$  is based on the channel half-height  $h$ , centerline velocity  $U = U(y = 0)$ , and kinematic viscosity  $\nu$ . We next consider the dynamics of perturbations  $(u, v, w, p)$  about this baseflow. No-slip boundary conditions are imposed at both

ALGORITHM 2. mGRQI-variational renormalization ( $P, Q, J, k, \epsilon$ ).

---



---

```

1:  $q \leftarrow$  largest column of  $Q^{-1}P$ 
2:  $q_0 \leftarrow$  mGRQI( $P, Q, q, J, k, \epsilon$ )
3:  $\mathcal{W} \leftarrow \{i | q_{0_i} \neq 0\}$ 
4: if  $q_{0_{\mathcal{W}}} \neq v_{\max}(P_{\mathcal{W}}, Q_{\mathcal{W}})$  then
5:    $q_{0_{\mathcal{W}}} \leftarrow v_{\max}(P_{\mathcal{W}}, Q_{\mathcal{W}})$ 
6: end if

```

---



---

walls. Since the flow is incompressible, we can equivalently express the dynamics in terms of wall-normal velocity perturbations  $v(x, y, z, t)$  and wall-normal vorticity perturbations  $\eta(x, y, z, t)$ . Homogeneous Dirichlet boundary conditions are imposed on  $\eta$  at the walls, while homogeneous Dirichlet and Neumann boundary conditions are imposed on  $v$  at the walls. Spatial Fourier expansions are applied in the streamwise ( $x$ ) and spanwise ( $z$ ) directions with associated wave numbers  $\alpha$  and  $\beta$ , respectively, such that  $v(x, y, z, t) = \hat{v}_{\alpha, \beta}(y, t)e^{(i\alpha x + i\beta z)}$  and  $\eta(x, y, z, t) = \hat{\eta}_{\alpha, \beta}(y, t)e^{(i\alpha x + i\beta z)}$ . Retaining only linear terms (i.e., considering the dynamics of infinitesimal perturbations about the baseflow) yields the familiar Orr-Sommerfeld and Squire equations. Discretizing in the wall-normal direction using Chebyshev polynomials yields the spatially discretized form,

$$\underbrace{\frac{d}{dt} \begin{bmatrix} \hat{v} \\ \hat{\eta} \end{bmatrix}}_{\hat{q}} = \underbrace{\begin{bmatrix} L_{OS} & 0 \\ L_C & L_{SQ} \end{bmatrix}}_L \underbrace{\begin{bmatrix} \hat{v} \\ \hat{\eta} \end{bmatrix}}_q, \quad (10)$$

where  $L_{OS} = M^{-1}(-i\alpha UM - i\alpha U'' - M^2/\text{Re})$ ,  $L_{SQ} = -i\alpha U - M/\text{Re}$ ,  $L_C = -i\beta U'$ ,  $M = (\alpha^2 + \beta^2 - D^2)$ , and  $D$  is the Chebyshev differentiation matrix. The associated perturbation kinetic energy is computed as  $E(t) = q^H(t)Qq(t)$ , where  $Q$  is determined from the kinetic energy integral in  $(\hat{v}, \hat{\eta})$  variables via Clenshaw-Curtis quadrature.

The flow state  $q(t) = (\hat{v}(t), \hat{\eta}(t)) \in \mathbb{C}^{2N}$  is discretized in the wall-normal direction  $y \in [-1, 1]$  using Chebyshev polynomials [1], where  $N$  is the number of collocation points. In this study, we select  $N = 100$  as it provides sufficient accuracy in capturing the transient energy growth: specifically, we observe  $< 0.1\%$  change in (nonsparse) transient growth when  $N$  is doubled beyond this value). Since the objective function in (6) is dependent on  $t$ , we solve the optimization problem by sweeping over  $N_T$  values of time horizons  $t = T$ . We use a linearly spaced grid of  $N_T = 101$  values for the time horizons in the interval  $[10, 30]$ , i.e., the values increment by 0.2.

The analysis results to be presented momentarily are all based on the specific modeling and discretization choices that we have outlined here. The optimization framework operates on the discretized model operators to yield a  $k$  sparse state vector. For a given cardinality  $k$ , different spatial discretization schemes can yield different sparsity patterns in the resulting solution of the optimization. Since the optimization can only reveal critical spatial locations that are included in the model operators provided to it, a sufficiently dense spatial grid should be used to ensure that physically critical spatial locations are included within the discretization grid. In the model problem of plane Poiseuille flow, Fourier modes are used to model the spatially homogeneous streamwise and spanwise directions; thus, sparse optimal perturbations in this context will correspond to perturbations that are spatially localized in the wall-normal direction for a given wave-number pair  $(\alpha, \beta)$ . This restriction is imposed by the modeling choice, not by the optimization framework. Spatial localization within the volume can be achieved by incorporating alternative models within the same optimization framework.

## B. Spatially localized optimal perturbations and transient growth analysis

Sparse optimal perturbations  $q_{0, \text{sp}}$  are computed using Algorithm 2 for the linearized plane Poiseuille flow (10). In this paper, we only report results for the case  $\text{Re} = 4000$ ,  $\alpha = 1$ , and  $\beta = 2$ , which are representative of our findings for other wave-number pairs and Reynolds numbers. The algorithm parameters are set as  $J = \infty$  and  $\epsilon = 10^{-6}$  for the power method and the stopping criterion, respectively. Sparse optimal perturbations were computed for  $k = \{1, 2, \dots, 100\}$ . Here, we report and closely analyze three sets of results that highlight interesting properties of the algorithm and the flow physics, namely,  $k = 10, 20$ , and 50.

The transient growth  $G(t = T)$  is reported in Fig. 1. The solid black curve is associated with  $G(T)$  for (nonsparse) optimal perturbations for which  $k = 2N = 200$ , while each red curve corresponds to the  $G(T)$  associated with a sparse optimal perturbation with  $k = \{10, 20, 50\}$ . The relative magnitude of transient growth for different cardinalities  $k$  follows from the generalized inclusion principle (9) as expected, with the nonsparse  $G(T)$  serving as an upper bound to transient growth

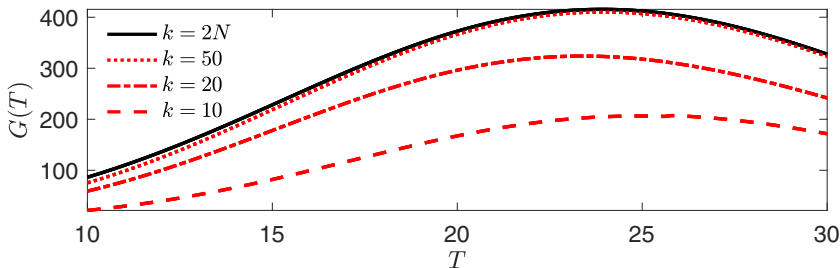


FIG. 1. Transient growth in kinetic energy  $G(t = T)$  for a specified time horizon  $T$  for plane Poiseuille flow at  $\text{Re} = 4000$  with  $(\alpha, \beta) = (1, 2)$  due to sparse optimal perturbations (red) approach that of the (non-sparse,  $k = 2N = 200$ ) optimal perturbation as  $k$  is increased. At  $k = 50$ , sparse optimal perturbations yield comparable transient growth for all time horizons, including the maximum transient growth over all time horizons  $G = G_{\max}$ . Associated optimal perturbations and responses are reported in Fig. 2.

due to any sparse optimal perturbation. As  $k \rightarrow 2N$ ,  $G(T)$  for the sparse and non-sparse cases will be equivalent. Strikingly, sparse optimal perturbations with  $k = 50 < 2N = 200$  yield  $G(T)$  within  $<13\%$  of the non-sparse transient growth response for all  $T$  considered. This suggests that only a subset of flow quantities and/or spatial locations are essential to initiating the transient growth that is characteristic to this flow. We will discuss the perturbations and the associated response in more detail momentarily. When  $k = 50$ , the maximum transient growth  $G_{\max} = G(T^*) = \max_T G(T)$ , i.e., the peak in the transient growth curve, also occurs at the same time  $T^* = 24.0$  as in the non-sparse case. The relative error between  $G_{\max}$  for  $k = 50$  and  $k = 2N = 200$  is  $<1.3\%$ . Interestingly, for  $k = \{10, 20\}$ , the maximum transient growth  $G_{\max}$  occurs at a different time  $T^*$  than in the non-sparse case. For  $k = 20$ , the maximum transient growth leads the non-sparse case ( $T^* = 23.4 < 24.0$ ); whereas, for  $k = 10$ , the maximum transient growth lags the non-sparse case  $T^* = 25.8 > 24.0$ .

Figure 2 reports optimal perturbations (left) and associated responses (right) corresponding with  $G_{\max} = G(T^*)$ . The black solid curves correspond to the non-sparse analysis ( $k = 2N = 200$ ), whereas the red curves correspond to the cardinality-constrained transient growth analysis ( $k = \{10, 20, 50\}$ ). Notably, the non-sparse optimal perturbations include contributions from both wall-normal velocity ( $\hat{v}$ ) and wall-normal vorticity  $\hat{\eta}$ , and both quantities are “active” across the entirety of the wall-normal direction ( $y$ ). In contrast, all three of the sparse optimal perturbations reported only include contributions from wall-normal velocity, with absolutely no contribution from the wall-normal vorticity. These sparse solutions reveal a lift-up mechanism that commonly arises in subcritical transition in shear flows: perturbations are comprised solely of the wall-normal component of velocity, which results in transient amplification of streamwise streaks in the response. We note that since an optimal perturbation is defined to have unit energy, this fact results in the sparse optimal perturbations having an apparently larger magnitude of  $\hat{v}$  than the non-sparse perturbation, which consequently distributes some of the unit energy into  $\hat{\eta}$ . The sparsest solution reported ( $k = 10$ ) highlights the capability of the proposed approach for achieving spatially localized optimal perturbations. For the cases of  $k = 10$  and  $k = 20$ , the localized  $\hat{v}(y)$  structure reported favors proximity to the upper wall; however, a symmetric and equally optimal sparse perturbation also exists near the lower wall and was identified by our method by perturbing the initial iterate, but is not reported here. As  $k$  is increased from  $k = 10$  to  $k = 50$ , the localized structure distributes further across the domain. In the case of  $k = 50$ , the sparse optimal perturbation profile for  $\hat{v}$  closely resembles that of the non-sparse optimal perturbation. Notably, the flow response at time  $t = T^*$  shown in the figure is visually indistinguishable between  $k = 50$  and  $k = 2N = 200$ . This striking similarity is consistent with the observation that  $G_{\max}$  for  $k = 50$  has a relative error of  $<1.3\%$  with respect to the non-sparse case.

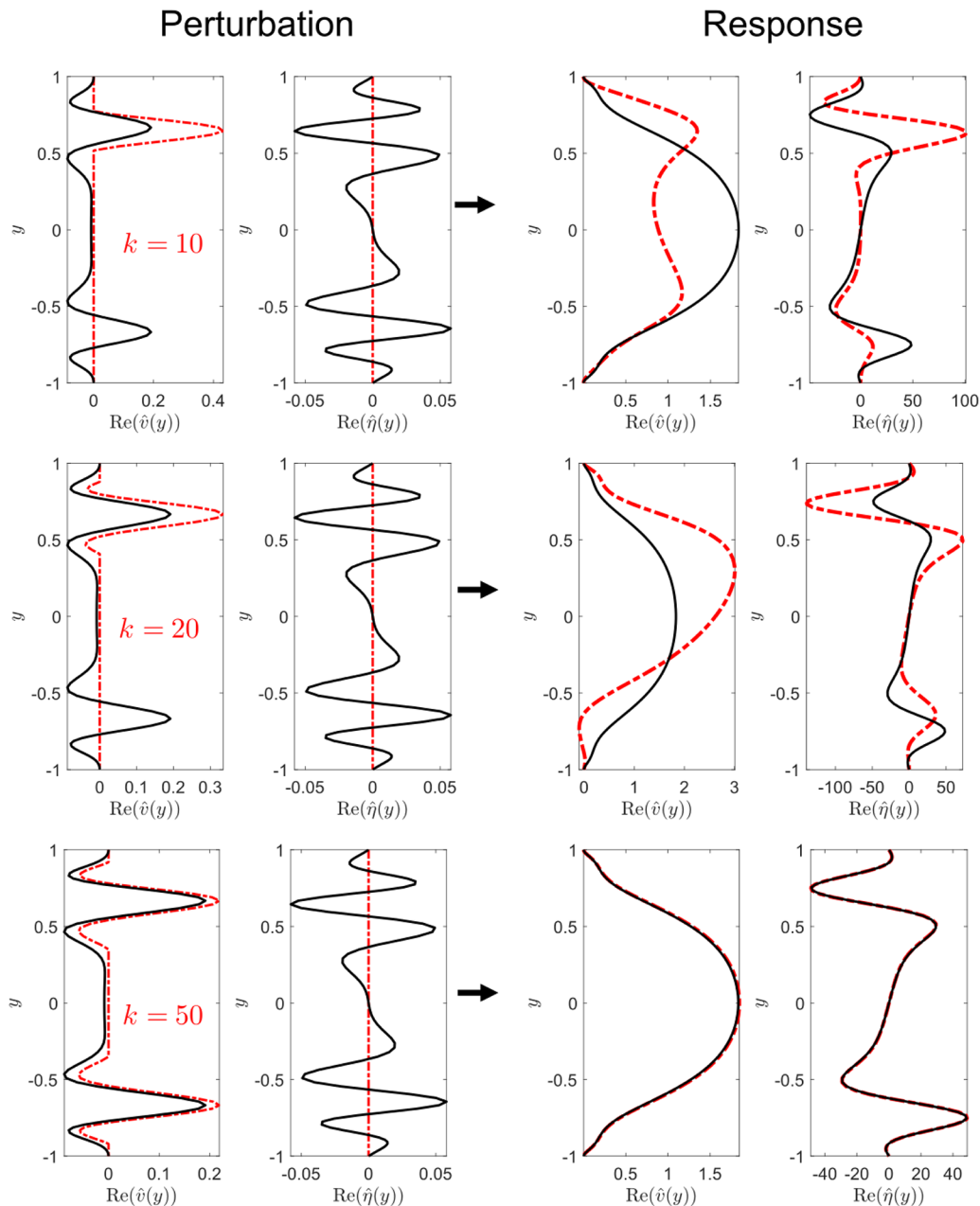


FIG. 2. Optimal perturbations (left) and associated responses (right) resulting in  $G = G_{\max}$  for  $\text{Re} = 4000$  and  $(\alpha, \beta) = (1, 2)$ . Nonsparse optimal perturbations (black) are composed of wall-normal velocity  $v$  and wall-normal vorticity  $\eta$  distributed across the wall-normal direction ( $y$ ). Sparse (sub)optimal perturbations (red) with  $k \leq 50$  are composed only of wall-normal velocity.

### C. Computational scaling

Here, we investigate the computational complexity of the mGRQI approach, i.e., Algorithms 1 and 2. Note that the computational complexity of the original GRQI algorithm was extensively studied in [11], where GRQI was reported to have a cubic rate of convergence and an operation

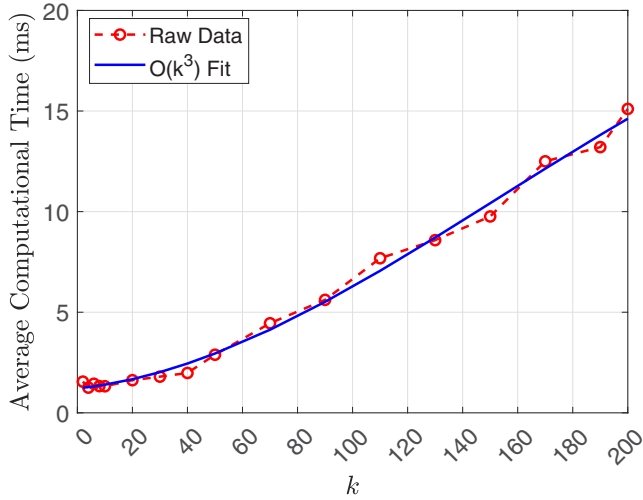


FIG. 3. Computational time per iteration shown for  $N = 100$  resulting in  $\max G(T)$  for  $\text{Re} = 4000$  and  $(\alpha, \beta) = (1, 2)$  at each  $k$ .

count per iteration scaling roughly as  $\mathcal{O}(k^3 + nk)$  flops. Here, for the given operator of size  $n = 2N$  and for each  $k$  value, we set the time horizon  $T = T^*$  based on the time at which transient growth is maximized, i.e.,  $G_{\max} = G(T = T^*)$ . The computational scaling per iteration for each  $k$  with  $N = 100$  is provided in Fig. 3, where a cubic polynomial fit to the data establishes consistency with the cubic scaling with  $k$  of the original GRQI approach for fixed  $n = 2N$ . Note that reported computational times per iteration are taken as the average over all iterations required to convergence, discussed in more detail below. To evaluate the scaling with operator size, we fix  $k = 50$  and grid over  $N$ . We find that computational time scales quadratically with  $N$  (see Fig. 4), which is higher than the linear scaling with  $N$  reported for the GRQI method in [11]. Note, however, that the scaling with  $k$  is the more central parameter than  $N$  in conducting spatially localized stability analysis:  $N$

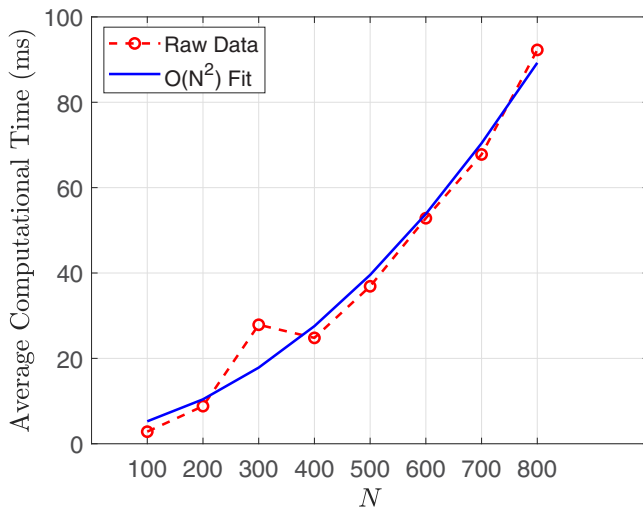


FIG. 4. Computational time per iteration shown for  $k = 50$  resulting in  $\max G(T)$  for  $\text{Re} = 4000$  and  $(\alpha, \beta) = (1, 2)$  at each  $N$ .

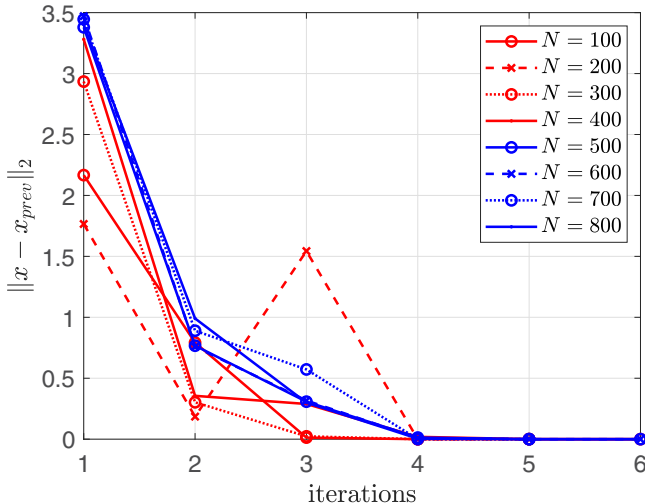


FIG. 5. Convergence shown for  $k = 50$  resulting in  $\max G(T)$  for  $\text{Re} = 4000$  and  $(\alpha, \beta) = (1, 2)$ .

is dictated by initial grid convergence studies for resolving the flow and is a fixed parameter at the point that the mGRQI method would be employed. Finally, the rate of convergence is evaluated empirically for different operator sizes with  $k = 50$  (see Fig. 5). We find that the number of iterations required for convergence remains constant despite the increasing size of the operator, and cubic convergence is observed in all cases considered. The insensitivity of convergence rate with respect to operator size is attributed to the fact that Algorithms 1 and 2 predominantly utilize only  $k$  nonzero elements of the initial state vector, which not only significantly reduces the convergence time but also improves the error convergence since the Rayleigh update only utilizes the  $k$  elements and is agnostic to the size of the initial matrix operator.

#### IV. CONCLUSIONS

In this paper, we proposed an iterative method for identifying sparse and spatially localized instability mechanisms as suboptimal solutions to a cardinality-constrained QCQP. In contrast to prevailing sparsity-promoting methods, the proposed approach determines suboptimal perturbations with a prescribed degree of sparsity, thereby removing the need for computationally demanding parameter sweeps. In our study, the approach successfully identified spatially localized perturbations in a plane Poiseuille flow at  $\text{Re} = 4000$  that achieved comparable transient growth as their global counterparts. By increasing the cardinality of the solution, the method systematically revealed a lift-up instability mechanism. While this paper focused on linear nonmodal stability analysis, the proposed methods are equally applicable to related methods for stability analysis based on generalized Rayleigh quotient maximization (e.g., resolvent and input-output analysis). We anticipate that our proposed algorithms will facilitate further research into sparse and spatially localized instabilities in the future.

#### ACKNOWLEDGMENTS

This material is based upon work supported by the Air Force Office of Scientific Research under Award No. FA9550-21-1-0434, the Office of Naval Research under Award No. N00014-22-1-2029, and the National Science Foundation under Award No. CBET-1943988.

## DATA AVAILABILITY

The data supporting the findings of this study are available from the authors upon reasonable request.

- 
- [1] P. J. Schmid and D. S. Henningson, *Stability and Transition in Shear Flows* (Springer-Verlag, New York, 2001).
  - [2] C. Skene, C. Yeh, P. Schmid, and K. Taira, Sparsifying the resolvent forcing mode via gradient-based optimisation, *J. Fluid Mech.* **944**, A52 (2022).
  - [3] P. Tamilselvam, K. Asztalos, and S. Dawson, Transient growth analysis of flow over an airfoil for identifying high-amplification, spatially-localized inputs, *AIAA Paper* (2022).
  - [4] D. Bhattacharjee, B. Klose, G. Jacobs, and M. Hemati, Data-driven selection of actuators for optimal control of airfoil separation, *Theor. Comput. Fluid Dyn.* **34**, 557 (2020).
  - [5] D. Foures, C. Caulfield, and P. J. Schmid, Localization of flow structures using  $\infty$ -norm optimization, *J. Fluid Mech.* **729**, 672 (2013).
  - [6] B. Lopez-Doriga, E. Ballouz, H. J. Bae, and S. T. Dawson, A sparsity-promoting resolvent analysis for the identification of spatiotemporally-localized amplification mechanisms, *AIAA Paper* (2023).
  - [7] A. L. Heide and M. Hemati, An optimization framework for analyzing nonlinear stability due to sparse finite-amplitude perturbations, in *AIAA AVIATION 2023 Forum* (2023), p. 4258.
  - [8] A. L. Heide and M. S. Hemati, An optimization framework for analyzing nonlinear stability due to sparse finite-amplitude perturbations, *Phys. Rev. Fluids* **10**, 014401 (2025).
  - [9] Here we work with complex variables because we will work with operators formulated using spectral methods in later sections. For real variables, the methods will be analogous to the ones derived here.
  - [10] B. Moghaddam, Y. Weiss, and S. Avidan, Generalized spectral bounds for sparse LDA, in *Proceedings of the 23rd International Conference on Machine Learning* (Association for Computing Machinery, New York, NY, 2006).
  - [11] V. Kuleshov, Fast algorithms for sparse principal component analysis based on Rayleigh quotient iteration, in *Proceedings of the 30th International Conference on Machine Learning, Proceedings of Machine Learning Research* (Proceedings of Machine Learning Research, 2013), Vol. 28, p. 1418 .
  - [12] B. Lopez-Doriga, E. Ballouz, H. J. Bae, and S. T. M. Dawson, Sparse space–time resolvent analysis for statistically stationary and time-varying flows, *J. Fluid Mech.* **999**, A87 (2024)
  - [13] B. N. Parlett, *The Symmetric Eigenvalue Problem* (Society for Industrial and Applied Mathematics, Philadelphia, 1998).
  - [14] M. Journée, Y. Nesterov, P. Richtárik, and R. Sepulchre, Generalized power method for sparse principal component analysis, *J. Mach. Learn. Res.* **11**, 517 (2010).