Variational framework for flow optimization using seminorm constraints

D. P. G. Foures, C. P. Caulfield, Peter J. Schmid

Variational framework for flow optimization using seminorm constraints

D. P. G. Foures,[1] C. P. Caulfield,[2,1] and P. J. Schmid[3]

[1]*DAMTP, University of Cambridge, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, United Kindgom*
[2]*BP Institute, University of Cambridge, Madingley Rise, Madingley Road, Cambridge CB3 0EZ, United Kingdom*
[3]*LadHyX, Ecole Polytechnique, 91128, Palaiseau cedex, France*

We present a general variational framework designed to consider constrained optimization and sensitivity analysis of spatially and temporally evolving flows defined as solutions of partial differential equations. We particularly focus on seminorm constraints which naturally arise for instance when the quantity which we wish to optimize can have contributions from several terms in the PDE through different physical mechanisms in a specific physical system. We show that this case implicitly requires that constraints be placed on the magnitude of complementary (with respect to the first constraining seminorm) seminorms of initial perturbations such that the sum of these complementary seminorms defines a total "true" norm of the state vector. A simple (true) norm constraint naturally satisfies this property. Therefore, the use of this framework requires the introduction of new parameters which describe the relative magnitude of the initial perturbation state vector calculated using the various constrained complementary seminorms to the magnitude calculated using the true norm, even for linear problems. We demonstrate that any required optimization has to be carried out by prescribing these new parameters as initial conditions on the admissible perturbations; the influence and significance of each seminorm component, partitioning the initial total norm of the perturbation, can then be considered quantitatively. To demonstrate the utility of this framework, we consider an idealized problem, the (linear) nonmodal stability analysis of a mean flow given by a "Reynolds averaging" of the one-dimensional stochastically forced Burgers equation. We close the mean flow equation by introducing a turbulent viscosity to model the turbulent mixing, which we allow to evolve subject to a new transport equation. Since we are interested in optimizing the relative amplification of the perturbation kinetic energy (i.e., the perturbation's "gain") this problem naturally requires the use of our new framework, as the kinetic energy is a seminorm of the full state velocity-viscosity vector, with a new adjustable parameter, describing the ratio of an appropriate viscosity seminorm to the sum of this viscosity seminorm and the kinetic energy seminorm. Using this framework, we demonstrate that the dynamics of the full system, allowing the turbulent viscosity to evolve subject to its transport equation, is qualitatively different from the behavior when the turbulent viscosity is "frozen" at a fixed, mean value, since a new mechanism of perturbation energy production appears, through the coupling of the evolving turbulent viscosity perturbation and the mean velocity field.

I. INTRODUCTION

Much fluid dynamical research, dating from the pioneering work of Osborne Reynolds [1], has been focused on the identification of a critical value of the "Reynolds number" of a flow for the onset of unsteadiness, significant perturbation growth, or indeed the transition to turbulence of initially laminar flows. The classical approach involves linearizing the governing equations around a steady state (also referred to as a "base flow"), and then investigating the properties of the eigenvalues of the corresponding operator. This modal analysis approach yields good agreement with experiments for a variety of flows (a prime example being Rayleigh-Bénard convection), but typically fails for shear flows. Such a modal stability analysis predicts a critical Reynolds number of 5772 for plane Poiseuille flow [2] and predicts no (infinitesimal) instability at all for the case of Couette flow [3], although experiments show that the transition to turbulence actually occurs for Reynolds number around 1000 for plane Poiseuille [4] flow and around 360 for Couette flow [5].

The concept of nonmodal stability analysis emerged more recently and allowed for a description of the perturbation for intermediate times, instead of focusing on the infinite time interval implicitly considered in a standard modal analysis. Indeed, because of the nonnormality of the Navier-Stokes operator [6], transient growth of the energy is possible for short times, even though all the (normal) modes are exponentially decaying. This phenomenon has been widely studied in shear flows [7–10], and it is now well known that an optimal perturbation can experience transient energy gain (i.e., the ratio of the kinetic energy at the end "target time" of a finite time interval to the kinetic energy at the start time) of several orders of magnitude. Exploiting the properties of the underlying linear operator, the gain can be calculated through a singular value decomposition of the evolution operator [10]. This very rich linear process could perhaps explain how a linearly stable flow experiences a sufficiently large energy increase for nonlinear effects to become significant, and thus possibly trigger a transition to turbulence [6].

In order to find this optimal gain, an alternative Lagrangian variational formulation was proposed [11,12], allowing more flexibility in the way we describe and consider "optimal" perturbations. Indeed, this formulation can take into account nonautonomous operators (e.g., a time-dependent base flow), nonlinear operators [13], and nonquadratic measures of the perturbation, to name just a few examples. More generally, this

variational approach can be used to optimize any functional of the state vector (not necessarily based on a norm). Moreover, the adjoint variables, which are dual variables of state variables (Lagrange multipliers in the variational formulation imposing the requirement that the state variables satisfy the underlying partial differential equation), can yield valuable information concerning the sensitivity of the optimized quantity to any input of the problem, including, e.g., the sensitivity to the chosen initial conditions, boundary conditions, physical or modeling coefficients, or even the flow geometry and base flow. Therefore, the particular objective functional can be chosen specifically for a sensitivity analysis [14], rather than for optimizing an initial perturbation over a finite time interval. Using a variational formulation, it is possible to derive the sensitivity of an eigenvalue (or a singular value) of the system with respect to any variable of the problem, which is a far more efficient way to gain insight into the impact of a parameter on the dynamics of a flow, rather than performing a more time-consuming finite-difference analysis.

Formally, the conventional problem of investigating perturbation kinetic energy gain in simple cases (for example, incompressible constant-density fluid flow) is essentially a problem of determining the initial conditions which maximize the 2-norm of the state vector at the end of the prescribed time interval. This has undoubted technical attractions, as the amplitudes of all the state vector components are simultaneously constrained in magnitude within the objective functional of the optimization problem when the objective functional is the kinetic energy gain. However, it is often of interest to consider problems where the objective functional is distinct from the 2-norm of the state vector. In this paper, we will investigate the implications of relaxing this requirement, and present an algorithmic framework to consider a general class of different problems, specifically where the objective functional is a *seminorm*. It is important to appreciate that this restriction to seminorms is not essential, as it is certainly conceivable, as noted above, that it may be of interest to consider problems where the objective functional bears no relationship to a norm. However, we believe that the key concepts which we wish to explore when the objective functional is no longer actually a norm can be most simply considered when the objective functional takes the form of a seminorm, and so that is the situation on which we focus in this paper.

At this stage, it is important to define clearly the difference between a norm and a seminorm. The null space (or kernel) of a norm on a vector space has, by definition, a unique element, the zero vector of that vector space. Although perhaps tautological, for clarity we will refer to such a norm as a "true" norm. A "nontrivial" seminorm, however, is a functional that satisfies the triangle inequality, and has the positive scalability property; i.e., for a scalar c (in the cases of interest c is a member of the real number field)

$$\|c\mathbf{q}\| = |c|\|\mathbf{q}\|, \tag{1}$$

where $\|\mathbf{q}\|$ denotes the seminorm acting on a state vector $\mathbf{q}$, but also has a larger null space, which means that nonzero vectors can have a null measure with respect to this seminorm. Within this nomenclature, a "true" norm is thus a "trivial" seminorm. We draw this extra distinction since there are many physical circumstances of undoubted fluid-dynamical interest

where the natural objective functional (or a natural constraint) is not a true norm but is rather defined in terms of a seminorm on the state vector space.

Two simple examples of optimization problems, where the nature of the objective functional requires the use of seminorm constraints, are related to a partitioning in space, where we are interested in maximizing the perturbation energy growth strictly in a subregion of the flow domain, and to a partitioning of the state vector, where we are interested in maximizing the gain of some (but strictly not all) components of the state vector. The former example may arise in an industrial context, where one is interested in maximizing perturbation growth, e.g., in the immediate vicinity of an injector, while the latter example may arise in situations where the fluid density is not constant (due to compositional, thermal, or compressible effects) and thus the state vector does not only include the flow velocity components but also the density field. We could be interested in maximizing the gain in the kinetic energy or the potential energy of a perturbation in a stratified yet incompressible flow, or alternatively maximizing the acoustic energy in a compressible flow; each of these optimizations would yield an objective functional which is most naturally defined in terms of a seminorm of the state vector.

For such classes of problems, we are then faced with the challenge of identifying techniques to constrain the elements of the state vector which are in the kernel (i.e., the null space) of this seminorm. A central aim of this paper is to present an algorithmic framework to address this challenge. The key idea is to impose "complementary seminorm" constraints on the allowable initial conditions for the state vector.

The use of seminorm constraints seems to be the appropriate approach which will separate the different effects contributing to the objective functional, while still allowing us to consider optimal seminorm gains. However, we here stress that seminorm constraints can be useful even if the objective functional is a (true) norm. In fact, the present paper presents a method that allows the consistent use of seminorm constraints, no matter what the nature of the objective functional.

As explained more precisely below in Sec. II, the "complementary seminorms" are defined so that they have two useful properties. First, some set of them must appropriately constrain the amplitudes of state vectors in the kernel of the objective functional. Second, the kernels of all the seminorms are distinct (except for the zero state vector) such that their direct sum must completely span the state vector space. This latter property effectively means that the initial constraints imposed by the complementary seminorms can be imposed independently. Therefore, the relative importance of the dynamics associated with the initial values of these complementary seminorms and the initial value of the objective functional itself can be investigated in a self-consistent and clear manner by considering parameters quantifying the relative size of these initial values.

A particular attraction of the proposed framework is its flexibility and generality, allowing the treatment of problems that go beyond the obvious (at best, weighted) 2-norm of the state vector [15]. We will focus on seminorm constraints and objective functionals, as we believe they are the most natural generalization of the conventional, "true" norm-based functionals to demonstrate our framework and algorithmic

approach, but physically motivated functionals (completely unrelated to norms) could also be used with only minor changes to the proposed framework. After discussing our framework in a quite general fashion, we demonstrate its utility by considering a simple idealized fluid dynamical problem, namely, parameterized turbulence flow modeling where use of such a framework is essential to yield the correct results for the natural perturbation kinetic energy gain optimization problem.

A classical approach to parameterized turbulence flow modeling has been to use the averaging method first proposed by Reynolds, which leads to the set of equations now commonly referred to as the Reynolds-averaged Navier-Stokes (RANS) equations. Naturally, due to the quadratic nonlinearity of the advection term in the underlying Navier-Stokes equations, a turbulence "closure" is required to close the system of equations, and one of the simplest (and most commonly used) closures is to assume that the (second-order in velocity) Reynolds stress tensor can be related to the (first-order) mean stress tensor through an (in general) temporally and spatially varying coefficient, the "eddy" or "turbulent viscosity." Of course, such a closure naturally leads to the further question of how the turbulent viscosity should be modeled, and in particular if it is allowed to vary in space and time, one or indeed several extra empirical equations may be required to describe the physical processes acting on this new quantity.

Recently, turbulence modeling techniques have been applied in various stability problems, and it appears that stability analysis of a Reynolds-averaged set of equations, coupled with an appropriate simple turbulence model, can be successful in predicting the onset of instabilities affecting mean flows [16], allowing the appropriate description of large-scale (compared to the turbulence length scale) instability processes in such turbulent flows. In the more general case of transient temporal perturbation growth, due to the nonnormality of the underlying linearized Navier-Stokes operator, some research has been conducted on the turbulent boundary layer, assuming a RANS base flow, yet, critically, perturbations in the velocity and pressure variable only, while fixing the turbulent viscosity at a constant value throughout the flow evolution (the so-called frozen turbulent viscosity approach; see [17] for more details). Crucially, however, the influence of the closure on the actual flow evolution is still largely unknown and in particular the robustness of the results to relaxing the frozen turbulent viscosity assumption is an open question.

If the turbulent viscosity is rather allowed to vary spatially and temporally (subject to an appropriately constructed evolution equation) then the state vector of the system formally involves not only the velocity components, but also the turbulent viscosity. Therefore, even if the problem of interest is the conventional problem of maximization of the perturbation kinetic energy gain over some finite time, the objective functional for the optimization problem naturally becomes a seminorm of the state vector, and so we obtain a relatively simple example of the type of optimization problem for which we have developed our generalized framework. In this problem, we have to impose a constraint (using a complementary seminorm) on the initial magnitude of the turbulent viscosity, and so this problem has (in a very simple way) the central characteristics of interest illustrating the utility of our framework.

Indeed, we wish to consider an extremely simple one-dimensional problem which nevertheless contains the salient features of turbulence: time dependence, nonlinearity, enhanced diffusivity, and stochastic forcing. An appropriate choice is the stochastically forced Burgers equation. This equation is a good one-dimensional analog of the Navier-Stokes equations, where the analog of "turbulence" is artificially introduced by a (stochastic) forcing term. Furthermore, it has been shown that there exists an equivalence between the Kuramoto-Sivashinsky equation (which is one of the more famous one-dimensional turbulence model equations [18]) and such a stochastically forced Burgers equation [19], suggesting that this is an appropriate model system to consider.

Therefore, the rest of this paper is organized as follows. In Sec. II, we describe our variational framework involving the required use of complementary seminorm constraints in some generality. In Sec. III, we then demonstrate the application of this framework to the model problem described above. Specifically, we derive the Reynolds-averaged Burgers (RAB) equations and apply a turbulent viscosity closure with an evolution equation for the turbulent viscosity, based on the well-known Spalart-Allmaras turbulence model [20]. We will then consider the problem of the identification of "optimal" perturbations (where optimality is defined in various ways) as an example to show the potential usefulness of our variational framework not only for identification of optimal initial conditions but also for sensitivity analysis [14,21]. In Sec. IV we present our results, focusing in particular on demonstrating the flexibility (and superiority when compared to other methods) of this framework for considering different objective functionals to optimize when there is no "natural" choice of an objective functional corresponding to a "true" norm of the state vector space. In Sec. V, we briefly discuss other potential fluid-dynamical applications of this framework, and finally, in Sec. VI, we draw our conclusions.

II. VARIATIONAL FRAMEWORK

A. Governing equations

We consider an arbitrary state vector $\mathbf{q}$ from a vector space Ω, defined on the time interval $[0,T]$. We choose $\mathbf{q} \in H^1(\Omega)$, a Sobolev space of order 2. We choose this space so that $\mathbf{q}$ and its gradient on Ω are both in $L^2(\Omega)$ (space of square integrable functions on Ω), which means that the state vectors are appropriately well behaved for the types of differential operations we wish to consider. We now consider a hierarchy of constraints which we wish to impose upon $\mathbf{q}$. The first constraint is that we wish $\mathbf{q}$ to satisfy a partial differential equation, the most general form of which is

$$\partial_t \mathbf{q} - \mathcal{N}(\mathbf{q}) = \mathbf{q}_f, \tag{2}$$

where $\mathcal{N}$ is a nonlinear operator acting on the variable $\mathbf{q}$, and $\mathbf{q}_f$ a forcing term. In this section for simplicity and clarity, we will however focus on linear homogeneous equations, although it is important to stress that this framework can be applied straightforwardly to the case of forced and/or nonlinear equations. In this simpler case, (2) reduces to

$$\partial_t \mathbf{q} - \mathbf{L}\mathbf{q} = 0, \tag{3}$$

where $\mathbf{L}$ is a linear operator. For a well-posed problem, we must of course impose (as constraints) initial conditions

$$\mathbf{q}(\mathbf{x},0) - \mathbf{q}_0 = 0, \quad \forall \mathbf{x} \in \Omega, \tag{4}$$

and boundary conditions defined on $\partial\Omega$, for all t:

$$\mathbf{q}(\mathbf{x},t) - \mathbf{q}_{\partial\Omega} = 0, \quad \forall (\mathbf{x},t) \in \partial\Omega \times [0,T]. \tag{5}$$

Once again, for clarity, we restrict ourselves to Dirichlet boundary conditions, although Neumann boundary conditions can be treated in the same way in the following framework, subject to conventional consistency conditions (for example associated with the divergence theorem) being satisfied.

B. Objective functional and seminorm considerations

Depending on the particular problem studied, we will define "the objective functional" $\mathcal{J}(\mathbf{q})$ which takes as its input the state vector $\mathbf{q}$. The particular functional form of $\mathcal{J}(\mathbf{q})$ is unique to any problem and can be of many different forms corresponding to some physical quantity of interest. Obvious examples include (some measure of) the flow's energy, enstrophy, drag, or mixing efficiency. In all cases, the functional $\mathcal{J}$ outputs a real number, which we may want to optimize, or alternatively we may wish to investigate the sensitivity of that output real number to small variations of some parameters of the problem. Variational frameworks of this form have conventionally been used to optimize the energy growth over some finite time interval by identifying an optimal initial condition for the state vector, which can also be identified (for linear operators) by considering a singular value decomposition (see [10] for further details).

However, a variational framework is much more flexible, and there is no formal requirement to restrict attention to optimization of perturbation energy gain. Indeed, the objective functional can describe the receptivity of a system to an external forcing, the sensitivity of the least stable eigenvalue (in the case of a linearized equation only) with respect to parameters or to a base flow modification [14,21], and more generally, any (real) quantity derived from the state vector $\mathbf{q}$. Another specific interesting application of a variational framework is data assimilation and consists of minimizing (ideally of course reducing to zero) the difference between a calculated state vector solution of the underlying partial differential equation and a (measured) target vector, and thus to identify "optimal" choices for coefficients or parameter functions within the governing equation [22].

Although the goal of this particular section is to present a variational framework in as general a fashion as possible, actual calculations cannot be carried out without specifying the kind of problem we are considering, because the objective functional as well as the various constraints we will consider naturally change depending on the chosen problem. In order to demonstrate the framework, we will therefore focus on the case of the identification of optimal perturbations, i.e., finding the optimal initial condition $\mathbf{q}_0$ which maximizes (the output of) an objective functional. Such an optimal perturbation is sometimes referred to as the most dangerous perturbation (in the sense of what is optimized).

It is important to note that we will use true norms or seminorms for the objective functional, but in general the

positive definiteness is not required, and any functional can be used. We will in this paper consider the following generic objective functional:

$$\mathcal{J}(\mathbf{q}) = \|\mathbf{q}(T)\|_O^2, \tag{6}$$

which defines a quantity of interest given by an objective (in general semi-) norm at the target time T (without loss of generality we will always assume that the time interval for optimization starts at $t = 0$ and so the target time is T and the time interval for optimization is also T). We stress again that this objective functional is not uniquely defined, and the norm (or seminorm) can be changed depending on the specific problem being considered. For example, the objective functional often describes the kinetic energy of a perturbation evolving around a base flow state. However, it can also describe the total energy, summing the kinetic energy and some form of potential energy. For example the internal energy in a gas or a fluid can be quantified as a function of the temperature of the system [23], the potential energy density in a stratified fluid can be straightforwardly calculated from the density distribution, and the electrostatic energy due to the presence of an electrostatic field [24,25] or magnetic energy associated with magnetic field [26] can similarly be evaluated in space and time.

These are only a few examples of the other types of "energies" which can be defined, and indeed the objective functional does not have to be a conventional energy of the physical system under consideration. For example, to find the energy threshold leading to a turbulent state in a Couette flow configuration, an objective functional defined as the time and space average of the viscous dissipation has been used successfully in [27].

Of particular interest are problems where the objective functional is actually defined in terms of a "seminorm," as discussed in detail in the introduction. Such objective functionals naturally arise when we are interested only in some partitioning of the state vector, either in space where we are interested in optimizing the energy in some compact set of the domain, or in terms of components of the state vector where (for example) we are interested in only some part of the total energy of the system. As noted in the introduction, a (nontrivial) seminorm has a nontrivial null space or kernel, defined for the particular vector space which we are considering as the set of state vectors $\mathbf{q}$ such that the seminorm $\|\cdot\|$ returns a zero value; i.e.,

$$\ker(\|\cdot\|) = \{\mathbf{q} \in H^1(\Omega); \quad \|\mathbf{q}\| = 0\}. \tag{7}$$

For a "true" norm the kernel is trivial, containing only the zero state vector. We then define the complementary space to this kernel implicitly as

$$\ker(\|\cdot\|) \oplus \ker^*(\|\cdot\|) = H^1(\Omega), \tag{8}$$

for any (semi)norm, where $\oplus$ stands for the space direct sum which has for definition for three arbitrary ensembles A, B, and C:

$$A \oplus B = C \Leftrightarrow \begin{cases} A + B = C, \\ A \cap B = \{\mathbf{0}\}, \end{cases} \tag{9}$$

with $\{0\}$ the appropriate zero state vector. The kernel complementary space is thus in fact the restriction of the space $H^1(\Omega)$ for which the seminorm $\|\cdot\|$ [on $H^1(\Omega)$] becomes a norm.

As we discuss in the following subsection, optimization of gain defined by such a nontrivial seminorm requires a special treatment of further constraints. In order to address the development of a variational framework where the objective functional may potentially use a seminorm, we define a particularly simple expression capable of describing both norms and seminorms. Our objective (in general semi-) norm is then defined as

$$\|\mathbf{q}(T)\|_O^2 = \frac{1}{2} \int_\Omega \mathbf{q}(T)^H \mathbf{W}_O \mathbf{q}(T)\, d\Omega, \qquad (10)$$

where the superscript H denotes the transpose conjugate and the matrix $\mathbf{W}_O$ is a weight matrix. If $\mathbf{W}_O$ is singular (noninvertible) then $\|\mathbf{q}(T)\|_O$ is a (nontrivial) seminorm, while if $\mathbf{W}_O$ is invertible then this expression defines a (true) norm. This weight matrix can in general be a function of position, and so one obvious way in which it can be noninvertible is if it is nonzero in only a compact subregion of the flow domain (i.e., it is space partitioning). Another obvious way in which it may be noninvertible is if $\mathbf{W}_O$ is nonzero only in a block, so that certain components of the state vector (for example the fluid's density, or as we shall see below, a spatially and temporally varying eddy or turbulent viscosity) do not have any effect on the value of the "energy" norm (i.e., it is state partitioning). This class of parameterized (through the weight matrix $\mathbf{W}_O$) quadratic norms will be the only one considered in this paper. However, a more general seminorm could take into account the total time-evolving flow [27], i.e., the full space-time evolution of the state vector, and be for example the evaluation of (at least some component of) the energy integrated over space and time. Moreover, we are not in general constrained to choose a quadratic norm.

C. Lagrangian framework using seminorm constraints

A sensitivity analysis identifies the impact of a small variation of an input of the optimization problem on the value of the objective functional, and so in a particularly natural way, a Lagrangian variational framework enables the performance of a sensitivity analysis subject to constraints. Indeed, the Lagrangian framework allows us to add as many dimensions to the problem as we have constraints. By adding these extra degrees of freedom, we are then able to investigate the impact of variations of the constraints on the returned value of the objective functional. As a consequence, if the variables whose magnitude we wish to optimize (or to consider within a sensitivity analysis) are part of the formulation of the constraints acting on the system, we then have to embed them in an augmented functional which takes into account the objective functional and the constraints at the same time. In other words, when we allow the constraints to vary, we have to include them in the augmented functional (i.e., the Lagrangian) of the problem, in order to retrieve the sensitivity information.

In many situations, we are interested in optimizing a given quantity (for example the initial condition, the external forcing or the boundary conditions) which will have an impact on the space-time evolution of the state vector $\mathbf{q}$, and as a

consequence the objective functional $\mathcal{J}$ not only depends implicitly on the optimized quantity, but also inevitably on the full space-time evolution of the state vector $\mathbf{q}$. Therefore, as already noted, the first constraint which we must impose is that the state vector must satisfy the evolution equation (3). Then, depending on the problem we are solving, different constraints must be imposed. In general, a correct, well-posed problem statement involves appropriate boundary conditions; although it is of course possible to optimize with respect to such boundary conditions, in an entirely equivalent way to optimizing with respect to initial conditions (see [28] for a fuller discussion), for clarity in this paper we opt to restrict our attention to problems where the boundary conditions are chosen conveniently and appropriately to not enter explicitly into the variational problem of interest. Rather we wish to focus on identifying optimal perturbations, and so initial conditions play a central role, so that we add an appropriate (and essentially self-evident) initial condition constraint (4). Furthermore, in order to avoid the final state vector amplitude becoming arbitrarily large during the optimization process, we have to impose a normalization (and hence scalar) constraint on the initial condition; i.e.,

$$\|\mathbf{q}_0\|_N^2 - N_0 = 0,$$
$$\|\mathbf{q}_0\|_N^2 = \frac{1}{2} \int_\Omega \mathbf{q}_0^H \mathbf{W}_N \mathbf{q}_0\, d\Omega, \qquad (11)$$

where the subscript N stands for normalization and emphasizes the fact that this (true) norm is used for a normalization purpose. This (true) norm is defined in an analogous way to $\|\cdot\|_O$ defined in (10) but is in general defined by a weight matrix $\mathbf{W}_N$ different from the energy weight matrix $\mathbf{W}_O$. In particular, since we wish all possible state vectors to be constrained, we require $\mathbf{W}_N$ to be nonsingular, and so the (true) norm used for the normalization of the initial condition can be different from the objective norm (or seminorm) used to define the optimized quantity $\mathcal{J}$. Indeed, in general, we can optimize the value of a certain seminorm, given that the initial perturbation is normalized with respect to a different (true) norm.

In the specific case where we wish to optimize a "gain" (i.e., the ratio of final to initial objective value), we need to optimize (the value at the final time) and constrain (the value at the initial time) the same quantity. It is therefore natural to choose the same norm for optimization and normalization and so $\mathbf{W}_O = \mathbf{W}_N$, such that the normalization constraint is simply

$$\|\mathbf{q}_0\|_O^2 - O_0 = 0, \qquad (12)$$

with O_0 describing an initial value of the objective functional. The gain in the objective functional is then straightforwardly defined as

$$G_O = \frac{\|\mathbf{q}(T)\|_O}{O_0}. \qquad (13)$$

Since $\|\mathbf{q}(T)\|_O$ is the optimized quantity and O_0 is fixed, at the end of the optimization process, the gain found will be optimal.

In general, the normalization constraint has to act on the totality of the state vector $\mathbf{q}$ in order to have a well-posed

optimization problem. In particular, imposing constraint (12) with a singular matrix $\mathbf{W}_O$ is not an appropriate constraint, as this constraint will not affect any vector which is part of the kernel of the seminorm involved in the definition of the objective functional, and will as a consequence remain unbounded. Although an optimization investigating the optimal final state objective value $\|\mathbf{q}(T)\|_O$ under a seminorm constraint defining the initial seminorm of the state vector O_0 can still be conducted, and an optimal $\mathbf{q}_0$ can in principle be identified, it is very likely that the objective functional will diverge with the magnitude of the nonconstrained part of the state vector, a typically undesired and unrealistic behavior.

As a consequence, in the more general situation (for which we wish to construct a framework) where we want to constrain the state vector with the help of a seminorm, to define a well-posed problem we have to add (at least) a further constraint on the part of the state vector which is in the kernel of the seminorm. From now on, the seminorm which we wish to impose as a constraint will be denoted with a subscript K_0. A natural way to do this is to appreciate that there is some flexibility in the construction of the constraint, and especially that we can have several normalization constraints. Therefore, to be able to constrain the seminorm $\|\cdot\|_{K_0}$ of interest and the magnitude of all possible state vectors at the same time, we are then led to the necessity of (at least) a second initial condition constraint beyond the normalization through $\|\cdot\|_{K_0}$ in order to impose an appropriate constraint equivalent to (11). A very convenient way to do this is through defining a set of "complementary seminorms" $\|\cdot\|_{K_i}$:

$$\|\mathbf{q}_0\|_{K_0}^2 - K_{00} = 0,$$
$$\|\mathbf{q}_0\|_{K_1}^2 - K_{10} = 0, \tag{14}$$

where $\|\cdot\|_{K_1}$ is constructed such that the norms $\|\cdot\|_{K_0}$ and $\|\cdot\|_{K_1}$ are "complementary." In this context, we wish to refer to a set of seminorms as being complementary if the direct sum of the kernel complementary spaces of these two seminorms is the entire state vector space (9); i.e.,

$$\ker^*(\|\cdot\|_{K_0}) \oplus \ker^*(\|\cdot\|_{K_1}) = H^1(\Omega). \tag{15}$$

By construction, this complementary seminorm $\|\cdot\|_{K_1}$ constrains the initial magnitude of the state vectors in the kernel of the seminorm $\|\cdot\|_{K_0}$ and vice versa, such that the full space $H^1(\Omega)$ is constrained without any interference between the two normalizations. Therefore, for a general state vector $\mathbf{q}$, we define the total normalization norm through $\|\cdot\|_N$

$$\|\mathbf{q}\|_N^2 = \|\mathbf{q}\|_{K_0}^2 + \|\mathbf{q}\|_{K_1}^2. \tag{16}$$

This is clearly a straightforward construction when the first seminorm constraint considers only a compact subregion of the flow domain (i.e., when the problem is partitioned in space) or partitions the state vector by its components (e.g., when $\|\cdot\|_{K_0}$ only considers the kinetic energy of a stratified flow). Therefore, we can define the initial (true) norm value as

$$N_0 = K_{00} + K_{10}. \tag{17}$$

Consequently, a new (adjustable) parameter arises which quantifies the relative size of the initial magnitude given by

the energy seminorm to the total normalization norm; i.e.,

$$R_0 = \frac{K_{00}}{N_0} = \frac{K_{00}}{K_{00} + K_{10}} = 1 - \frac{K_{10}}{N_0}. \tag{18}$$

In order actually to find the optimal perturbation, we also have to optimize with respect to this parameter (and not with respect to the total norm N_0 since the problem is linear). Indeed, this framework offers the possibility to perform a multiscale stability analysis where the initial amplitude of the perturbation is different in each component of the state vector. Optimizing on the parameter R_0 would then maximize the corresponding objective functional. However, in some cases, the ratio R_0 will be fixed physically or be an input if one wants to investigate a certain case. For example, in the case where we want to constrain the initial condition to lie only within a compact subregion of the domain, we would enforce the initial condition on this subregion and on the complementary subregion independently with seminorms weighted by spatial (mask) functions, and would set the ratio R_0 to be zero which forces the initial condition to be completely free of any component of the kernel of the seminorm $\|\cdot\|_{K_0}$, and ensure the initial localization of the perturbation in the desired subregion.

The situation is somewhat more straightforward if the problem of interest is one of optimization of a gain defined by a seminorm. In this particular case, the optimization seminorm has to coincide with one of the constraint seminorms, and so $\|\cdot\|_{K_0} \equiv \|\cdot\|_O$ and so we may write the complementary seminorm as $\|\cdot\|_K$. The associated seminorm initial values are denoted O_0 and K_0, and the corresponding full-norm initial value N_0 is still the sum of these two values. The new single parameter arising is defined in the same way as in (18):

$$R_0 = \frac{K_0}{N_0} = \frac{K_0}{O_0 + K_0} = 1 - \frac{O_0}{N_0}. \tag{19}$$

The gain is then defined in the exact same way as in (13) as the ratio of the final value of the objective functional norm $\|\mathbf{q}(T)\|_O$ to its initial value O_0.

More generally, if there are other (multiple) physically motivated constraints we wish to impose upon the problem (for example by requiring the initial conditions to have specific magnitudes in different subregions of the flow domain) we can impose a larger complete yet complementary set of initial constraints:

$$\|\mathbf{q}_0\|_{K_i}^2 - K_{0_i} = 0, \quad \text{for } i \text{ from } 0 \text{ to } n_c,$$
$$\bigoplus_{i=0}^{n_c} \ker^* \left(\|\cdot\|_{K_i} \right) = H^1(\Omega), \tag{20}$$

where the symbol $\oplus$ denotes the direct sum [explicitly written in (15)] and where the number of complementary constraints is $n_c + 1$, with implicitly $n_c + 1$ different complementary seminorms which satisfy

$$\sum_{i=0}^{n_c} \|\mathbf{q}_0\|_{K_i}^2 = \sum_{i=0}^{n_c} K_{0_i} = \|\mathbf{q}_0\|_N^2 = N_0. \tag{21}$$

The number of new parameters to optimize over is n_c (because the system is linear) and can be defined as [generalizing (18)]

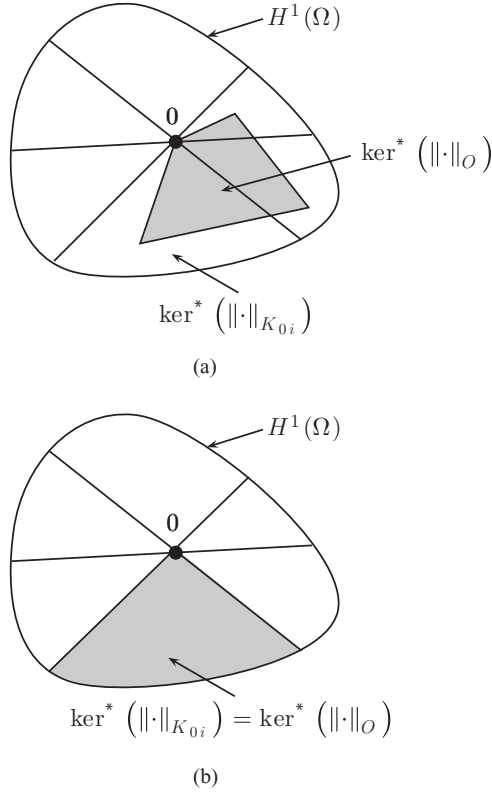$$R_{0_i} = \frac{K_{0_i}}{N_0}, \quad i = 1 \ldots n_c, \tag{22}$$

(a)

(b)

FIG. 1. Schematic representation of the partition of the space $H^1(\Omega)$ through the choice of $n_c + 1$ (here $n_c = 5$) seminorm constraints. (a) Schematic representation of the generic case of final value optimization, where the objective seminorm is different from all the (initial) constraint seminorms. (b) Schematic representation of the special case of gain optimization, where the objective seminorm coincides with one of the constraint seminorms (and as a consequence, so do their kernel complementary spaces).

and we will retain this general form for the constraints to construct our general framework. This general situation is shown in Fig. 1(a), making it explicit that the seminorm used to define the objective functional does not need to correspond to any of these constraint seminorms.

As before, the situation is simpler if the problem of interest corresponds to a problem where we wish to optimize a gain, because then one of the constraint seminorms has to coincide with the objective seminorm, and so without loss of generality, we define $\|\cdot\|_{K_0} = \|\cdot\|_O$ [see Fig. 1(b)]. We have decided to express the objective functional with a norm denoted with a subscript O for "objective" to highlight that final energy or energy gain optimization is only a single possibility allowed by this framework. In an energy gain optimization case, we choose to write $\|\cdot\|_O \equiv \|\cdot\|_E$ and $O_0 \equiv E_0$. If optimized, the gain will then be an energy gain and denoted

$$G_E(T) = \frac{\|\mathbf{q}(T)\|_E}{E_0}. \tag{23}$$

We are now able to express the appropriate Lagrangian functional for our optimization problem embedding the constraints, provided we define the different scalar products we will need to use. We will use three different scalar products in this study: one related to space, one to time, and one to both

space and time. Respectively, these scalar products are

$$\langle f,g \rangle(t) = \int_\Omega f^H g \, d\Omega,$$

$$[f,g](x) = \int_0^T f^H g \, dt, \tag{24}$$

$$(f,g) = \int_\Omega \int_0^T f^H g \, d\Omega dt.$$

Using these definitions, the augmented Lagrangian functional for our optimal perturbation problem can now be written in a rather general way:

$$\mathcal{L}(\mathbf{q},\mathbf{q}_0,\mathbf{q}^\dagger,\mathbf{q}_0^\dagger,\lambda_i) = \|\mathbf{q}(T)\|_O^2 - (\mathbf{q}^\dagger, \partial_t \mathbf{q} - \mathbf{L}\mathbf{q})$$
$$- \langle \mathbf{q}_0^\dagger, \mathbf{q}(0) - \mathbf{q}_0 \rangle - \sum_{i=0}^{n_c} \lambda_i \left(\|\mathbf{q}_0\|_{K_i}^2 - K_{0_i} \right), \tag{25}$$

where the objective functional $\mathcal{J}(\mathbf{q}) = \|\mathbf{q}(T)\|_O^2$ consistently with (6).

D. Optimality conditions

We wish to find an extremum of the augmented Lagrangian functional $\mathcal{L}$ by ensuring that the variations with respect to all the considered variables vanish. The total variation of the (augmented) Lagrangian is

$$\delta \mathcal{L} = \frac{\delta \mathcal{L}}{\delta \mathbf{q}} \delta \mathbf{q} + \frac{\delta \mathcal{L}}{\delta \mathbf{q}_0} \delta \mathbf{q}_0 + \frac{\delta \mathcal{L}}{\delta \mathbf{q}^\dagger} \delta \mathbf{q}^\dagger + \frac{\delta \mathcal{L}}{\delta \mathbf{q}_0^\dagger} \delta \mathbf{q}_0^\dagger$$
$$+ \sum_{i=0}^{n_c} \frac{\partial \mathcal{L}}{\partial \lambda_i} \delta \lambda_i = 0. \tag{26}$$

Since all the variables of the problem are independent, all the terms in the previous equation have to vanish. Variations with respect to $\mathbf{q}^\dagger$ and $\mathbf{q}_0^\dagger$ yield the "direct" or "forward" partial differential equation (3) and the initial conditions for $\mathbf{q}$, while the first variation with respect to the λ_i will simply yield the constraints on the normalization of the initial perturbation.

Requiring variations with respect to the direct variable $\mathbf{q}$ to be zero leads (typically after some integration by parts, and application of appropriate boundary conditions) to the adjoint evolution equation, defined as

$$-\partial_t \mathbf{q}^\dagger - \mathbf{L}^\dagger \mathbf{q}^\dagger = 0. \tag{27}$$

The integration by parts of the time derivative yields the final condition

$$\mathbf{q}^\dagger(T) = \frac{\delta \mathcal{J}}{\delta \mathbf{q}} = \mathbf{W}_O \mathbf{q}(T). \tag{28}$$

Because of the Laplacian structure of the diffusive term in equations of interest, the adjoint equation turns out to be an antidiffusive equation which, for well-posedness reasons, has to be integrated backward in time from $t = T$ to $t = 0$ to calculate $\mathbf{q}^\dagger(0)$ which can then be used to find the sensitivity of the Lagrangian to the chosen initial condition of the state vector. By requiring that the boundary terms play no role (and hence are homogeneous), the natural boundary conditions for the adjoint are found straightforwardly to be $\mathbf{q}^\dagger(\partial\Omega) = 0$.
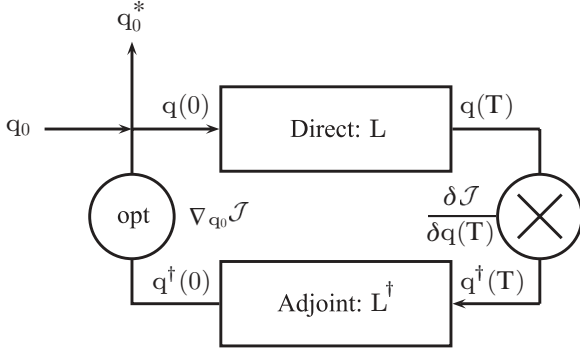
FIG. 2. A schematic representation of the "Direct-Adjoint" loop process in order to find the optimal perturbation. We start with a guess $\mathbf{q}_0$, apply the initial condition constraint (4), then integrate the direct equation forward in time. This gives the direct state vector at time T which allows us to define the "final" condition for the adjoint state vector using (28). We then integrate the adjoint equation backward in time from this final condition to obtain the "initial" adjoint state vector which allows us to compute the sensitivity with respect to the chosen initial condition on the (forward) state vector $\mathbf{q}_0$ using (29). We then use an appropriate optimization method in order to find the "best" initial condition achieving the maximum value of the objective functional defined in (6).

Taking variations with respect to the initial condition $\mathbf{q}_0$ leads to the following expression for the gradient of the objective functional with respect to the initial condition:

$$\nabla_{\mathbf{q}_0}\mathcal{J} = \mathbf{q}_0^\dagger - \sum_{i=0}^{n_c} \lambda_i \mathbf{W}_{K_i}\mathbf{q}_0, \qquad (29)$$

where $\mathbf{q}_0^\dagger = \mathbf{q}^\dagger(0)$. Ideally, at the stationary point of the Lagrangian (when the solution to the underlying variational problem has then been identified), this gradient vanishes. However, this is not true for generic initial conditions, and we have to employ an optimization technique in order to reach the (solution) condition. The λ_i, the Lagrange multipliers imposing the various amplitude complementary seminorm constraints on the initial state vector, will be determined at each iteration of the optimization algorithm by satisfying the initial normalization conditions. This determination will however depend on the particular iterative optimization algorithm used. The whole loop process is represented schematically in Fig. 2.

Eventually, at the end of the optimization, we have the optimal value of the objective functional $\mathcal{J}^* = \mathcal{J}(\mathbf{q}^*(T))$, associated with the optimal set of direct and adjoint state vectors $(\mathbf{q}^*, \mathbf{q}^{*\dagger})$ at all times, and in particular at the initial time $t = 0$, thus identifying the optimal initial condition $\mathbf{q}_0^*$, for which the gradient given by (29) vanishes by definition. We will now see that the adjoint state vector can also yield information on the sensitivity of the objective functional with respect to every varying field or coefficient taken into account in our Lagrangian framework.

E. Sensitivity analysis

In this section, we will describe the sensitivity analysis possibilities that our variational framework allows. In some sense, the optimal perturbation framework presented above is

already a sensitivity analysis, with the appropriate sensitivity information [the gradient in (29)] with respect to the choice of initial conditions being used to find the optimal perturbation. A general sensitivity analysis will allow us to find what is the impact of a small variation of a parameter p on the value of a functional at the optimal state vector point $\mathbf{q}^*$. As the sensitivity analysis can be performed on a functional which is totally different from the (optimized) objective functional, we will define a general functional $\mathcal{I}$ which is *a priori* different from the original optimized objective functional $\mathcal{J}$.

The sensitivity may then be defined as

$$\nabla_p \mathcal{I}(\mathbf{q}^*(p),p) = \frac{\partial \mathcal{I}}{\partial p}(\mathbf{q}^*) + \left(\frac{\delta \mathcal{I}}{\delta \mathbf{q}}, \frac{\partial \mathbf{q}^*}{\partial p}\right), \qquad (30)$$

where $\nabla_p \mathcal{I}$ is just a condensed way to write the total derivative of $\mathcal{I}$ with respect to p, and where the chain rule appears under the form of a scalar product on the state vector space. The first term on the right-hand side of the equation represents the explicit contribution of p to the functional while the second term is the implicit contribution of p to $\mathcal{I}$ through the (optimal) state vector $\mathbf{q}^*$.

We consider two qualitatively different cases, depending on the particular properties of the parameter p. We can define two broad classes of parameters: constraint parameters p_c which will modify the constraints while keeping the functional $\mathcal{I}$ unchanged, and external parameters p_e which will change the functional $\mathcal{I}$ without changing the constraints. An example of a constraint parameter is a coefficient of the underlying partial differential equation satisfied by the state vector, such as a viscosity coefficient or a modeling parameter, while an example of an external parameter is a parameter directly involved in the definition of the energy seminorm.

Focusing first on sensitivity with respect to constraint parameters (p_c), the first term on the right-hand side of Eq. (30) is zero by definition of a constraint parameter, as it does not appear directly in the functional $\mathcal{I}$. Therefore,

$$\nabla_{p_c} \mathcal{I}(\mathbf{q}^*(p_c),p_c) = \left(\frac{\delta \mathcal{I}}{\delta \mathbf{q}}, \frac{\partial \mathbf{q}^*}{\partial p_c}\right). \qquad (31)$$

This implicit contribution can be expressed analogously to before using a Lagrangian framework. We can add the constraint into a yet further new augmented functional $\mathcal{K}$ combining the functional $\mathcal{I}$ as well as the dynamical PDE constraint on the (optimal) state vector $\mathbf{q}^*$:

$$\mathcal{K}(\mathbf{q}^*, \mathbf{q}^{\dagger*}) = \mathcal{I}(\mathbf{q}^*) - (\mathbf{q}^{*\dagger_\mathcal{I}}, \partial_t \mathbf{q}^* - \mathbf{L}\mathbf{q}^*), \qquad (32)$$

where we have added a subscript $\mathcal{I}$ since the adjoint will depend on the functional $\mathcal{I}$ and is in general different from the adjoint state vector $\mathbf{q}^{*\dagger}$ associated with the optimization of the underlying objective functional $\mathcal{J}$.

The required implicit derivative can be obtained by calculating the partial derivative of the augmented Lagrangian functional $\mathcal{K}$ with respect to p_c since the constraints have been embedded in this augmented functional. The direct state vector $\mathbf{q}^*$ is defined by its initial condition $\mathbf{q}_0^*$ (the optimal for maximizing the original, underlying objective functional $\mathcal{J}$) and the adjoint state vector $\mathbf{q}^{*\dagger_\mathcal{I}}$ (which carries the sensitivity information) will be retrieved through the backward integration of the adjoint equations, the structure of which is not changed by this algorithm. However, the chosen
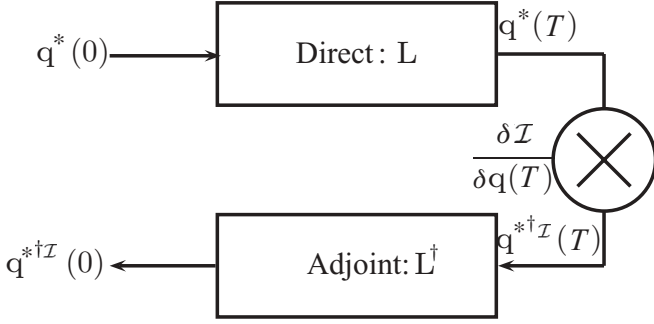
FIG. 3. A schematic representation of the algorithm used to calculate the sensitivity of a general functional $\mathcal{I}$ of the optimal initial condition state vector $\mathbf{q}_0^*$ to constraint parameters p_c. We start from the optimal initial condition state vector $\mathbf{q}_0^*$ we obtained using the optimization framework and integrate the direct equation to obtain the "final" state vector $\mathbf{q}^*(T)$. (This step may not actually be required if the final state of the optimal direct state vector $\mathbf{q}^*$ has been saved in the last iteration of the optimization framework.) We then construct a new final adjoint state vector $\mathbf{q}^{*\dagger\mathcal{I}}(T)$, which construction depends on the particular choice of the functional $\mathcal{I}$. Finally a backward integration of the adjoint state vector leads to a new "initial" adjoint state vector $\mathbf{q}^{*\dagger\mathcal{I}}(0)$ which is needed in order to determine the required sensitivity.

starting form of the "final" adjoint state vector $\mathbf{q}^{*\dagger\mathcal{I}}(T)$ is now determined by the gradient of the new functional $\mathcal{I}$, and so in general is different from the final adjoint state vector $\mathbf{q}^{*\dagger}$ associated with the optimization of the original underlying objective functional $\mathcal{J}$. Using this new final adjoint state vector $\mathbf{q}^{*\dagger\mathcal{I}}(T)$, a single backward-in-time evolution using the adjoint equations yields the sensitivity information. This means that the sensitivity to a constraint parameter of a functional (potentially different from the original optimized objective functional) satisfies
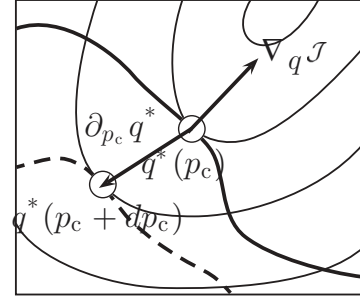
$$\frac{\partial \mathcal{K}}{\partial p_c} = \left(\frac{\delta \mathcal{I}}{\delta \mathbf{q}}, \frac{\partial \mathbf{q}^*}{\partial p_c} \right). \qquad (33)$$
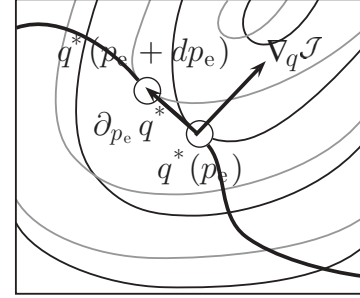
As a direct consequence,

$$\nabla_{p_c} \mathcal{I}(\mathbf{q}^*(p_c), p_c) = \frac{\partial \mathcal{K}}{\partial p_c}, \qquad (34)$$

where $\mathcal{K}$ is the secondary augmented Lagrangian functional. This expression is a scalar product between a function of the (optimal) direct state vector $\mathbf{q}^*$ and the adjoint state vector $\mathbf{q}^{*\dagger\mathcal{I}}$ corresponding to the functional $\mathcal{I}$. A schematic representation of this particular algorithm is shown in Fig. 3. For the particular special case where $\mathcal{I}$ actually is the original optimized functional $\mathcal{J}$, then the gradient is given by the same equation, where the adjoint vector $\mathbf{q}^{*\dagger\mathcal{J}} = \mathbf{q}^{*\dagger}$ was already evaluated during the optimization problem.

In the other case of an external parameter, the objective functional $\mathcal{I}$ depends explicitly on the parameter p_e, so the first term on the right-hand side of Eq. (30) will be different from zero. The gradient of the functional with respect to an explicit parameter p_e can be found in many cases analytically (for example for functionals defined in terms of integrals), and so the principal issue remains to evaluate the second term on the right-hand side of Eq. (30). We believe that the calculation of the second term of the product (i.e., the gradient of the optimal state vector $\mathbf{q}^*$ with respect to p_e), requires the use of

FIG. 4. (a) A schematic representation illustrating the sensitivity of the optimized functional $\mathcal{J}$ to a constraint parameter. (b) A schematic representation illustrating the sensitivity of the optimized functional $\mathcal{J}$ to an external parameter. Black lines are the level lines of the objective functional $\mathcal{J}$ [gray lines of part (b) of the figure correspond to the level lines of the functional for $p_e = p_e + \delta p_e$]. Thick black lines are the constraints (thick dashed line is the constraint for $p_c = p_c + \delta p_c$). Black circles represent the optimal locations in solution space for the state vector. In the case of the sensitivity with respect to an external parameter we can see that the terms $\frac{\delta \mathcal{J}}{\delta \mathbf{q}}$ and $\frac{\partial \mathbf{q}^*}{\partial p}$ are orthogonal whereas they are not in the case of a constraint parameter.

a simple, yet computationally costly, finite-difference method. Indeed, we have to utilize our variational framework to identify the optimal state vector $\mathbf{q}^*$ to the problem for a particular value of p_e, then for $p_e + \delta p_e$, and then evaluate:

$$\frac{\partial \mathbf{q}^*}{\partial p} \simeq \frac{\mathbf{q}^*(p_e + \delta p_e) - \mathbf{q}^*(p_e)}{\delta p_e}. \qquad (35)$$

Once there is a need to use finite differences however, in general there is no need to evaluate the terms in (30) independently, because sensitivity can of course also be directly estimated using finite difference:

$$\begin{aligned} \nabla_{p_e} &\mathcal{I}(\mathbf{q}^*(p_e), p_e) \\ &\simeq \frac{\mathcal{I}(\mathbf{q}^*(p_e + \delta p_e), p_e + \delta p_e) - \mathcal{I}(\mathbf{q}^*(p_e), p_e)}{\delta p_e}. \end{aligned} \qquad (36)$$

The situation is substantially more straightforward in the special case when the functional $\mathcal{I}$ whose sensitivity is being investigated is actually the same as the underlying optimized functional $\mathcal{J}$. In this specific case, we observe that the second term on the right-hand side of (30) actually vanishes. Since the objective functional $\mathcal{J}$ is (by definition) optimized, variations of the state vector while still satisfying all the

imposed constraints cannot improve the value of the objective functional $\mathcal{J}$ (see Fig. 4).

Formally, the gradient of the objective functional with respect to the state vector is perpendicular to the subspace defined by all the imposed constraints. Equivalently, the level lines of $\mathcal{J}$ are parallel to the constraint subspace at the optimal point $\mathbf{q}^*$. On the other hand, $\partial_{p_e}\mathbf{q}^*$ is tangent to the subspace defined by the constraints (since the optimized state vector must always satisfy all the constraints by definition) which subspace does not change as p_e varies, by the definition of an external parameter. Therefore, combining these two observations, the gradient of $\mathcal{J}$ with respect to the state vector is normal to the variation of $\mathbf{q}^*$ (confined to the subspace defined by the constraints) with respect to the external parameter, and so the second term on the right-hand side of (30) (which is simply the scalar product of these two quantities) is exactly zero. We can then simply express the sensitivity of the optimized objective functional $\mathcal{J}$ to variations in an external parameter as

$$\nabla_{p_e}\mathcal{J}(\mathbf{q}^*(p_e), p_e) = \frac{\partial \mathcal{J}}{\partial p_e}. \tag{37}$$

Here, we have only discussed variations with respect to a parameter. However, it could also be of interest to consider the sensitivity of a functional to a function, either associated with the definition of the objective functional or the constraints. For example if the operator $\mathbf{L}$ describes the linear evolution of a small perturbation evolving on a base flow defined by a base state vector $\mathbf{q}_B$ (which is a function governing the dynamics of the perturbation), it is possible to derive the sensitivity of a functional to this base flow in an analogous fashion to the algorithm described above to investigate sensitivity to parameters.

In the particular example of considering the sensitivity to the base flow state vector, the base flow must satisfy base flow equations which can be expressed in the same form as Eq. (3), where the implicit coefficients in the operators (such as the flow's Reynolds number) are "constraint parameters" p_c and are shared by the base flow and perturbation equations (since the perturbation equation is derived from the base flow equation). In general, small variations in these coefficients will affect both the perturbation state vector $\mathbf{q}$ and the base state vector $\mathbf{q}_B$. As a consequence, the requirement (effectively another constraint) that the base state vector satisfies the base flow equation must be embedded within the Lagrangian functional, with the constraint imposed by a new Lagrange multiplier $\mathbf{q}_B^\dagger$.

III. REYNOLDS-AVERAGED BURGERS EQUATION (RAB) OPTIMAL PERTURBATION PROBLEM FORMULATION

A. Derivation of the Reynolds-averaged Burgers equations

As a relatively simple demonstration example of our variational framework, we will in this section construct a model problem of interest, where optimization of the perturbation (kinetic) energy gain inevitably leads to an objective functional which is defined in terms of a (nontrivial) seminorm of the state vector. We study the evolution of a velocity-like variable

defined on $[0,1] \times [0,T]$ and governed by the stochastically forced Burgers equation, entirely defined by the viscosity coefficient ν, with Dirichlet boundary conditions and a well-posed (in particular appropriately smooth) initial condition. This can be formulated as

$$\partial_t u + u \partial_x u - \nu \partial_{xx} u = s, \tag{38}$$

with $u(0,t) = u_l$, $u(1,t) = u_r$, and $u(x,0) = u_0$, and s a stochastic forcing of zero ensemble average [which is needed in order to later be consistent with (3) where no forcing term is present], and vanishing at the boundary. To obtain nontrivial energy production dynamics, we consider a symmetric focusing base flow, and so we choose the boundary conditions to be $u_l = -u_r = 1$. The solution $u(x,t)$ is stochastic because of the nature of the forcing, but can be expressed as the superposition of a coherent field $\langle u \rangle$ and a stochastic field u_s; i.e.,

$$u(x,t) = \langle u \rangle (x,t) + u_s(x,t),$$
$$\text{such that} \quad \langle u_s \rangle = 0, \tag{39}$$

where $\langle \cdot \rangle$ denotes ensemble averaging. We interpret u_s as the "turbulent" component of the flow, and so this decomposition of the flow into two variables with different spatial and temporal scales of variation constitutes a so-called "Reynolds decomposition."

We introduce this decomposition into the governing equation, and then ensemble-average to obtain the mean flow equation for $\langle u \rangle$, which is

$$\partial_t \langle u \rangle + \langle u \rangle \partial_x \langle u \rangle + \langle u_s \partial_x u_s \rangle - \nu \partial_{xx} \langle u \rangle = 0. \tag{40}$$

In this equation all the terms except the third one are expressed in terms of the ensemble-average velocity of the flow. Indeed, when Reynolds-averaging a nonlinear state equation, higher order terms inevitably appear which cannot directly be expressed as a function of the first-order "mean" quantities, leading to a classic "closure" problem. Here, this term is the equivalent of the gradient of the Reynolds stress tensor in the Reynolds-averaged Navier-Stokes equations, a second-order quantity in a first-order equation. In this particularly simple one-dimensional context, we can rewrite this term as the spatial derivative of the turbulent kinetic energy, defined as

$$\langle u_s \partial_x u_s \rangle = \partial_x e_t, \quad \text{with} \quad e_t = \frac{1}{2}\langle u_s^2 \rangle. \tag{41}$$

Therefore, to close the evolution equation for the mean velocity (40), we need to add a model in order to express the turbulent kinetic energy density [defined in (41)] as a function of mean quantities.

B. Turbulent viscosity closure

We here follow the classical Boussinesq [29] turbulence hypothesis, by assuming that e_t is proportional to the gradient of the mean velocity field with a viscosity-like coefficient of proportionality ν_t, which in general is itself a function of space and time:

$$\frac{1}{2}\langle u_s^2 \rangle = -\nu_t \partial_x \langle u \rangle. \tag{42}$$

Simple assumptions of this kind are widely used as closures for RANS equations. In the highly idealized model situation we are considering, it is thoroughly plausible that the stochastic
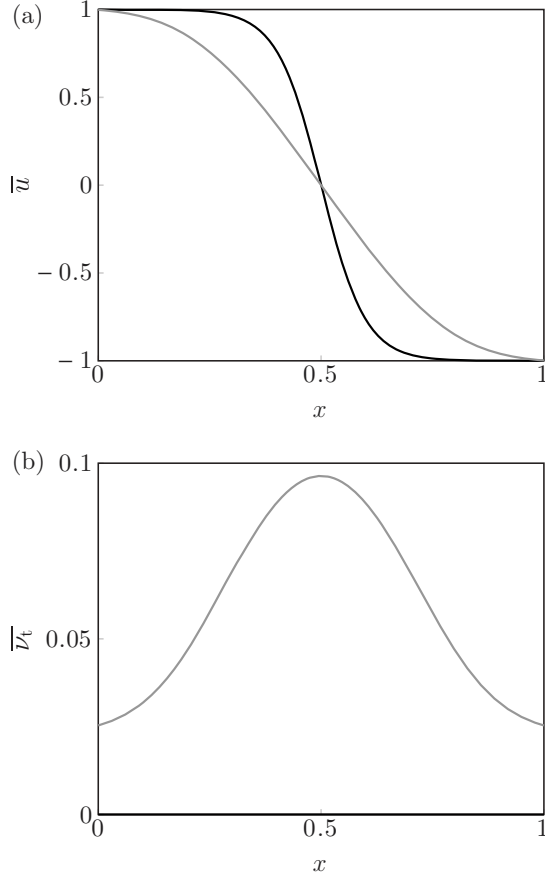
FIG. 5. (a) Velocity $\overline{u}$ and (b) turbulent viscosity $\overline{\nu_t}$ solutions of the steady RAB equations (47) with boundary conditions $\overline{u}(0) = 1$, $\overline{u}(1) = -1$ and $\overline{\nu_t}(0) = r\nu$, $\overline{\nu_t}(1) = r\nu$. The value of the laminar viscosity is set to $\nu = 0.05$. Production and destruction coefficients are chosen to have the illustrative values $c_1 = 0.75$ and $c_2 = 2$. For $r = 0$, no turbulent viscosity is created, and the flow is laminar. When $r \neq 0$, turbulent viscosity is generated in the zones of high gradient, which are, as a consequence, smoothed for $r = 0.5$.

field will have the effect of increasing the total viscosity of the flow by a certain amount ν_t. Furthermore, in this special case of Burgers equation, the left-hand side of (42) is always positive. As a consequence, the product $\nu_t \partial_x \langle u \rangle$ has to be negative, which means that the turbulent (eddy) viscosity must have the same sign as $-\partial_x \langle u \rangle$. Therefore, the slope of the ensemble average of u cannot be positive (flow going toward the edges of the domain), as that would require the turbulent viscosity to be negative, which is physically inconsistent. We will always respect the positivity constraint of the modeled turbulent kinetic energy e_t since we restrict ourselves to consideration of a focusing flow [with negative slope, as shown in Fig. 5(a)].

Combining (40) and (42) we obtain the following equation for the mean field $\langle u \rangle$:

$$\partial_t \langle u \rangle + \langle u \rangle \, \partial_x \langle u \rangle - \partial_x \left[(\nu + \nu_t) \, \partial_x \langle u \rangle \right] = 0. \quad (43)$$

At this point, we could simply close the equation by choosing a fixed value of this new viscosity, based, for example, on the evaluation of a mixing length scale. However,

this naive technique is typically not appropriate since the properties of a turbulent flow naturally vary in space and time. As a consequence, it is appropriate to develop a model equation allowing us to describe the spatial and temporal evolution of the new turbulent (or eddy) viscosity coefficient.

A widely used approach is to produce a transport equation for the turbulent viscosity. We follow this approach and assume that ν_t is a solution of an advection-diffusion equation with production and destruction terms. We assume that the production of turbulent viscosity is driven by the magnitude of the first derivative of the mean field $\langle u \rangle$ (equivalent to the shear in a real flow) while we assume the destruction term is quadratic in the turbulent viscosity (see [20]). This leads to the following equation for ν_t:

$$\partial_t \nu_t + \langle u \rangle \, \partial_x \nu_t - \partial_x \left[(\nu + \nu_t) \, \partial_x \nu_t \right]$$
$$- c_1 \left| \partial_x \langle u \rangle \right| \nu_t + c_2 \nu_t^2 = 0, \quad (44)$$

where c_1 and c_2 are two real coefficients defining the strength of production and destruction mechanisms. We also add a Dirichlet boundary condition for the viscosity:

$$\nu_t(0,t) = \nu_t(1,t) = r\nu, \quad (45)$$

where r is the ratio between the turbulent viscosity at the boundaries of the flow domain and the laminar viscosity. This parameter just controls the amount of "turbulence" we want to introduce at the boundaries. Equation (44) has also the property of preserving the positivity of the turbulent viscosity. Equations (43) and (44) then constitute a closed set for the Reynolds-averaged Burgers (RAB) equations, acting on the state vector $\langle \mathbf{q} \rangle (x,t) = (\langle u \rangle, \nu_t)^\top$:

$$\partial_t \langle u \rangle + \langle u \rangle \, \partial_x \langle u \rangle - \partial_x((\nu + \nu_t) \, \partial_x \langle u \rangle) = 0,$$
$$\partial_t \nu_t + \langle u \rangle \, \partial_x \nu_t - \partial_x((\nu + \nu_t) \, \partial_x \nu_t) \quad (46)$$
$$- c_1 \left| \partial_x \langle u \rangle \right| \nu_t + c_2 \nu_t^2 = 0.$$

By considering steady flows (i.e., $\partial_t = 0$), we can search for steady states as a solution of the coupled equations. We denote the steady flow solution of such an equation as $\overline{\mathbf{q}}(x) = (\overline{u}, \overline{\nu_t})^\top$. This steady state must satisfy

$$\overline{u} d_x \overline{u} - d_x((\nu + \overline{\nu_t}) \, d_x \overline{u}) = 0,$$
$$\overline{u} d_x \overline{\nu_t} - d_x((\nu + \overline{\nu_t}) \, d_x \overline{\nu_t}) \Leftrightarrow \overline{\mathbf{L}}(\overline{\mathbf{q}}) = 0, \quad (47)$$
$$- c_1 \left| d_x \overline{u} \right| \overline{\nu_t} + c_2 \overline{\nu_t}^2 = 0,$$

defining a set of ordinary differential equations.

We plot a solution of this steady set of equations in Fig. 5. We first notice that in the two cases $r = 0$ (laminar) and $r \neq 0$ (turbulent), the flow indeed corresponds to a focusing, with a positive velocity (from left to right) in the left part of the domain and a negative velocity (from right to left) in the right part of the domain. Then, in the turbulent case ($r \neq 0$), we can see that the turbulent viscosity equation has the effect of enhancing the turbulent viscosity in the middle of the domain (where the gradients of $\overline{u}$ are strong), which has the secondary effect of smoothing the gradient of velocity in the middle part of the domain.

C. Perturbation equations

Now that we have constructed a (steady) base flow, we can investigate its stability properties by introducing small

perturbations, which are in general functions of space and time. The ensemble average of the flow can now be decomposed as follows:

$$\langle \mathbf{q} \rangle (x,t) = \overline{\mathbf{q}}(x) + \tilde{\mathbf{q}}(x,t), \qquad (48)$$

with $\tilde{\mathbf{q}} = (\tilde{u}, \tilde{v})^\top$, so the perturbation state vector in general involves a perturbation to both the velocity and the turbulent viscosity. The perturbation velocity may be thought of as a "coherent" velocity perturbation, as it has a nonzero ensemble average. We assume that the magnitude and gradients associated with this perturbation state vector are sufficiently "small" relative to the base flow for a linearization to be well posed. Substituting this decomposition into the full equations (43) and (44), imposing the mean flow equations (47), and neglecting nonlinear terms, we obtain the (full) linearized perturbation equation:

$$\partial_t \tilde{u} + \overline{u}\partial_x \tilde{u} + \tilde{u}\partial_x \overline{u} - \partial_x((\nu + \overline{\nu_t})\,\partial_x \tilde{u}) = \partial_x\,(\tilde{v}\partial_x \overline{u})$$
$$\partial_t \tilde{v} + \overline{u}\partial_x \tilde{v} - \partial_x\,(\tilde{v}\partial_x \overline{\nu_t}) - \partial_x((\nu + \overline{\nu_t})\,\partial_x \tilde{v})$$
$$-c_1\,|\partial_x \overline{u}|\,\tilde{v} + 2c_2\tilde{v}\overline{\nu_t} = c_1\mathrm{sgn}(\partial_x \overline{u})\overline{\nu_t}\partial_x \tilde{u} - \tilde{u}\partial_x \overline{\nu_t}$$
$$\Leftrightarrow \partial_t \tilde{q} - \tilde{\mathbf{L}}\tilde{q} = 0, \qquad (49)$$

where $\tilde{\mathbf{L}}$ is the (full) linearized operator of the closed RAB equations. We write the differential operator in block matrix form in Appendix A.

In each evolution equation, transport terms of the relevant dependent variable are on the left-hand sides, while the right-hand sides may be interpreted as forcing terms since they are independent of the relevant dependent variable. A particular point to note is that this full linearized system of equations has a forcing term for the mean flow perturbation velocity $\tilde{u}$ equation involving the perturbation turbulent viscosity and the gradient of the base mean flow

$$F_\nu = \partial_x\,(\tilde{v}\partial_x \overline{u})\,, \qquad (50)$$

a term which plays a crucial role in the stability analysis of the total flow when the turbulent viscosity is allowed to vary in space and time, and so $\tilde{v}$ is nonzero.

It is also mathematically possible to consider a perturbation to the mean flow velocity only. Indeed, the simplest way to deal with stability analysis of mean flows (or more precisely stability analysis of the coherent flow) is to consider a steady mean flow solution of the previous system which constitutes the base flow $\overline{\mathbf{q}} = (\overline{u}, \overline{\nu_t})^\top$ and then apply a perturbation $\tilde{q}_F = (\tilde{u}, 0)^\top$ which has a perturbation component in the (mean flow as opposed to the stochastic field) velocity, but does not allow any variation in the turbulent viscosity, which is as a consequence "frozen" in its base state. This is equivalent to considering only the first equation of system (49), and imposing $\tilde{v} = 0$. It leads to the "frozen turbulent viscosity perturbation equation," defined as

$$\partial_t \tilde{u} + \overline{u}\partial_x \tilde{u} + \tilde{u}\partial_x \overline{u} - \partial_x((\nu + \overline{\nu_t})\,\partial_x \tilde{u}) = 0$$
$$\Leftrightarrow \partial_t \tilde{u} - \tilde{\mathbf{L}}_{11}\tilde{u} = 0, \qquad (51)$$

where the operator $\tilde{\mathbf{L}}_{11}$ is the operator describing the evolution of the perturbation in a frozen turbulent viscosity context. By comparison of (49) and (51), it is apparent that if represented

in matrix form $\tilde{\mathbf{L}}_{11}$ corresponds to the top left block matrix of the full linearized perturbation operator $\tilde{\mathbf{L}}$, as written in Appendix A.

D. Energy evolution equation

In order to understand the various growth mechanisms, it is useful to consider an evolution equation for an appropriately defined perturbation "energy."

A natural choice of course is to define the perturbation energy in terms of the coherent velocity perturbation; i.e.,

$$E = \frac{1}{2}\int_0^1 \tilde{u}^2 dx. \qquad (52)$$

The perturbation kinetic energy evolution equation can thus be derived by multiplying the first perturbation equation (for $\tilde{u}$) of system (49) by the perturbation velocity $\tilde{u}$, to obtain (after various integrations by parts)

$$\partial_t E = -\underbrace{\int_0^1 \frac{1}{2}\tilde{u}^2 \partial_x \overline{u}\,dx}_{P_{E_1}} - \underbrace{\int_0^1 \tilde{v}\partial_x \tilde{u}\partial_x \overline{u}\,dx}_{P_{E_2}}$$
$$- \underbrace{\int_0^1 (\nu + \overline{\nu_t})\,(\partial_x \tilde{u})^2\,dx}_{D_E}. \qquad (53)$$

The first term on the right-hand side (labeled P_{E_1}) is associated with the production or destruction of energy due to the interaction between the perturbation in the coherent velocity field and the base mean flow velocity. We notice that the sign of this quantity only depends on the sign of the base mean flow gradient: Since this gradient is negative by definition [required by (42)], this term is a (perturbation) energy production term. The second term (due to the presence of a forcing term in the coherent perturbation equation) also involves the perturbation in the turbulent viscosity, and is as a consequence denoted P_{E_2}. Assuming a negative base mean flow gradient, this term will be a source of perturbation kinetic energy when $\tilde{v}\partial_x \tilde{u}$ is positive, and will be a sink when $\tilde{v}\partial_x \tilde{u}$ is negative. It represents a (in general not sign-definite) catalytic term describing the amount of energy we are able to extract from the base mean flow due to variations in the turbulent viscosity.

Of course, in the case of the frozen turbulent viscosity analysis, this term is not present, which removes a possible mechanism for perturbation kinetic energy production. The last term quantifies the (appropriately linearized) dissipation of perturbation kinetic energy by both laminar and turbulent viscosity, and is always negative.

To complete the "energy" analysis, we can also consider the evolution of the squared norm of the second component $\tilde{v}$. We define this quantity as

$$K = \frac{1}{2}\int_0^1 \tilde{v}^2 dx. \qquad (54)$$

An evolution equation for this quantity can be derived following the same method described above for the perturbation energy. The evolution of the quantity K is governed by the

following equation:

$$
\begin{aligned}
\partial_t K = & \underbrace{\frac{1}{2} \int_0^1 \tilde{v}^2 \left(\partial_x \overline{u} + \partial_{xx} \overline{v_t} \right) dx}_{S_{K_1}} \\
& \underbrace{- \int_0^1 (\nu + \overline{v_t}) \left(\partial_x \tilde{v} \right)^2 dx}_{D_{K_1}} + \underbrace{\int_0^1 c_1 \left| \partial_x \overline{u} \right| \tilde{v}^2 \, dx}_{P_{K_1}} \\
& \underbrace{- \int_0^1 2 c_2 \overline{v_t} \nu^2 \, dx}_{D_{K_2}} + \underbrace{\int_0^1 c_1 \tilde{v} \partial_x \tilde{u} \, \mathrm{sgn} \left(\partial_x \overline{u} \right) \overline{v_t} \, dx}_{S_{K_2}} \\
& \underbrace{- \int_0^1 \tilde{u} \tilde{v} \partial_x \overline{v_t} \, dx}_{P_{K_2}} .
\end{aligned}
\tag{55}
$$

Since $\partial_x \overline{u} + \partial_{xx} \overline{v_t}$ is always negative, the first term S_{K_1} trivially acts like a sink for K. D_{K_1} is also always a diffusive term and so negative for all time while P_{K_1} and D_{K_2} are trivially associated with production and destruction of turbulent viscosity. At first sight, the term S_{K_2} has no obvious sign. However, the optimization of E suggests from term P_{E_2} that the product $\tilde{v} \partial_x \tilde{u}$ is positive. Furthermore, the term $\mathrm{sgn}(\partial_x \overline{u}) \overline{v_t}$ being negative, this term (in the optimization of E context) will act as a new sink of K. Finally, the last term P_{K_2} has no obvious sign and will depend on the perturbation symmetry.

Most importantly, there is no equivalent term to P_{E_2} in this equation, meaning that there is no transfer from one component of the state vector to the other. Therefore, the perturbation kinetic energy E can grow substantially by extracting energy from the mean flow with $\tilde{v}$ acting as a catalyst, rather than a direct source of energy, while K may well vary more slowly.

E. Optimal perturbation Lagrangian formulation

We are now in position to define an optimization problem using a Lagrangian approach based upon the general framework developed in Sec. II. We are particularly interested in the effect on our results of the application (or not) of a range of increasingly more restrictive assumptions. We will first formulate the "FULL" problem using the full linearized set of equations, allowing for coherent perturbations from the base mean flow velocity and the turbulent viscosity [i.e., using (49)]. We then formulate the "FROZ" problem by deriving the equations for the frozen turbulent viscosity analysis from the complete set of equations by not allowing any perturbations for the turbulent viscosity; i.e., $\tilde{v} = 0$. Finally, we consider a specific particularly simple example, the "LAM" problem of a completely laminar Burgers equation, removing all turbulent viscosity from the evolution equations ($\overline{v_t} = 0$). We present a summary of the key features of each of these three cases in Table I.

1. Seminorm gain

In all cases, we are interested in the gain of the perturbation kinetic energy over a finite time interval $[0, T]$, and so we

TABLE I. Summary of the different cases considered. LAM: laminar analysis; FROZ: frozen turbulent viscosity analysis; FULL: full linearized analysis.

	LAM	FROZ	FULL
Mean steady flow $\overline{\mathbf{q}}$	$(\overline{u}, 0)^\top$	$(\overline{u}, \overline{v_t})^\top$	$(\overline{u}, \overline{v_t})^\top$
Perturbation $\tilde{\mathbf{q}}$	$(\tilde{u}, 0)^\top$	$(\tilde{u}, 0)^\top$	$(\tilde{u}, \tilde{v})^\top$

define the objective functional which we optimize as

$$
\mathcal{J}(\tilde{\mathbf{q}}) = E(T) = \| \tilde{\mathbf{q}}(T) \|_E^2,
\tag{56}
$$

where the "energy" seminorm is defined as

$$
\begin{aligned}
\| \tilde{\mathbf{q}}(T) \|_E^2 &= \frac{1}{2} \int_0^1 \tilde{\mathbf{q}}(T)^\top \mathbf{W}_E \tilde{\mathbf{q}}(T) \, dx, \\
\mathbf{W}_E &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.
\end{aligned}
\tag{57}
$$

This is clearly a specific (and very simple) example of the energy seminorms described in Sec. II, and we can also define the appropriate complementary seminorm, acting on the kernel of $\| \cdot \|_E$ defined as

$$
\begin{aligned}
\| \tilde{\mathbf{q}}(T) \|_K^2 &= \frac{1}{2} \int_0^1 \tilde{\mathbf{q}}(T)^\top \mathbf{W}_K \tilde{\mathbf{q}}(T) \, dx, \\
\mathbf{W}_K &= \mathbf{I} - \mathbf{W}_E,
\end{aligned}
\tag{58}
$$

with associated normalization norm defined as

$$
\begin{aligned}
\| \tilde{\mathbf{q}}(T) \|_N^2 &= \frac{1}{2} \int_0^1 \tilde{\mathbf{q}}(T)^\top \mathbf{W}_N \tilde{\mathbf{q}}(T) \, dx, \\
\mathbf{W}_N &= \mathbf{I} = \mathbf{W}_E + \mathbf{W}_K.
\end{aligned}
\tag{59}
$$

Here it is clear that the appropriate "energy" is thus defined as a seminorm of the state vector $\tilde{\mathbf{q}}$. We use this expression, because we have seen in (52) that from a stability point of view, the most relevant quantity to look at is the kinetic energy of the perturbation velocity $\tilde{u}$ from the base mean flow. The kernel of this energy thus exclusively contains perturbations of the turbulent viscosity $\tilde{v}$. As we have shown in the method developed in Sec. II, the normalization constraints are thus more subtle, because the initial magnitude of these turbulent viscosity perturbations must also be constrained.

Constraints for the fully linearized system (FULL) are the dynamical constraint (49) requiring that the state vector satisfies the appropriate evolution equation and the initial condition

$$
\begin{aligned}
\tilde{\mathbf{q}}(x, 0) - \tilde{\mathbf{q}}_0 &= 0, \\
\tilde{\mathbf{q}}_0 &= (\tilde{u}_0, \tilde{v}_0)^\top.
\end{aligned}
\tag{60}
$$

The required constraints for normalization are

$$
\begin{aligned}
E_0 - \| \tilde{q}_0 \|_E^2 &= 0, \\
K_0 - \| \tilde{q}_0 \|_K^2 &= 0,
\end{aligned}
\tag{61}
$$

with E_0 the initial amount of perturbation energy from the (coherent) velocity and K_0 the initial amount of turbulent

viscosity perturbation in the system; i.e.,

$$E_0 - \frac{1}{2}\int_0^1 \tilde{u}_0^2\, dx = 0,$$

$$K_0 - \frac{1}{2}\int_0^1 \tilde{v}_0^2\, dx = 0. \tag{62}$$

As we explained in Sec. II, we introduce a new parameter R_0 describing the relative contribution of E_0 and K_0 to the initial normalization of the perturbation state vector N_0, where R_0 is defined as

$$R_0 = \frac{K_0}{N_0} = 1 - \frac{E_0}{N_0}. \tag{63}$$

Since our problem is linear, the total norm N_0 has no influence on the dynamics of the flow and as a consequence the ratio C_0 of K_0 and E_0 is sometimes a more relevant parameter. This ratio can be straightforwardly related to the parameter R_0:

$$C_0 = \frac{K_0}{E_0} = \frac{R_0}{1 - R_0}. \tag{64}$$

These quantities represent the initial structure of the state vector. These quantities are however of interest for any time, and we will as a consequence extend their definition for all t such that

$$R(t) = \frac{K(t)}{N(t)},\ R(0) = R_0,$$

$$C(t) = \frac{K(t)}{E(t)},\ C(0) = C_0. \tag{65}$$

We notice that we can also rewrite the coefficient $C(t)$ the following ways:

$$C(t) = \frac{R(t)}{1 - R(t)} = \frac{\int_0^1 \tilde{v}(t)^2 dx}{\int_0^1 \tilde{u}(t)^2 dx}. \tag{66}$$

Therefore, we can express the Lagrangian functional of our problem as

$$\mathcal{L}(\tilde{\mathbf{q}}, \tilde{\mathbf{q}}_0, \tilde{\mathbf{q}}^\dagger, \tilde{\mathbf{q}}_0^\dagger, \lambda_E, \lambda_K)$$
$$= \|\tilde{\mathbf{q}}(T)\|_E^2 - (\partial_t\tilde{\mathbf{q}} - \tilde{\mathbf{L}}\tilde{\mathbf{q}}, \tilde{\mathbf{q}}^\dagger) - \langle\tilde{\mathbf{q}}_0 - \tilde{\mathbf{q}}(x,0), \tilde{\mathbf{q}}_0^\dagger\rangle$$
$$- \lambda_E\left(E_0 - \|\tilde{\mathbf{q}}_0\|_E^2\right) - \lambda_K\left(K_0 - \|\tilde{\mathbf{q}}_0\|_K^2\right). \tag{67}$$

All the variations of the Lagrangian functional with respect to the parameters have to vanish; i.e., $\delta\mathcal{L} = 0$. Once again, we note that taking variations with respect to the adjoint variables will yield the constraints on the initial condition, and the underlying evolution equation. Conversely, taking variations with respect to the direct variables yields the adjoint set of equations

$$-\partial_t\tilde{\mathbf{q}}^\dagger + \tilde{\mathbf{L}}^\dagger\tilde{\mathbf{q}}^\dagger = 0. \tag{68}$$

The operator $\tilde{\mathbf{L}}^\dagger$ is the adjoint of the direct (full) perturbation RAB operator $\tilde{\mathbf{L}}$ defined in (49) and in Appendix A, where $\tilde{\mathbf{L}}^\dagger$ is also written out in full.

This adjoint equation also has the "final" condition:

$$\tilde{\mathbf{q}}^\dagger(T) = \mathbf{W}_E\tilde{\mathbf{q}}(T). \tag{69}$$

Since $\mathbf{W}_E\tilde{\mathbf{q}}(T) = \tilde{u}$ by definition, the "final" condition for the adjoint turbulent viscosity perturbation is zero. Nevertheless,

because of the coupling terms in the adjoint equation (68), the adjoint turbulent viscosity does not remain zero during its evolution. We also obtain homogeneous Dirichlet boundary conditions $\tilde{\mathbf{q}}^\dagger(\partial\Omega) = 0$, and a natural compatibility condition linking $\tilde{\mathbf{q}}_0^\dagger$ to the initial condition of the adjoint problem; i.e., $\tilde{\mathbf{q}}_0^\dagger = \tilde{\mathbf{q}}^\dagger(0)$.

Finally, we take variations of $\mathcal{L}$ with respect to the initial condition $\tilde{\mathbf{q}}_0$, which gives us the expression for $\frac{\delta\mathcal{L}}{\delta\tilde{\mathbf{q}}_0}\delta\tilde{\mathbf{q}}_0$ which immediately yields gradient information to optimize the objective functional $\mathcal{J}$:

$$\nabla_{\tilde{\mathbf{q}}_0}\mathcal{J} = \tilde{\mathbf{q}}_0^\dagger - (\lambda_E\mathbf{W}_E + \lambda_K(\mathbf{I} - \mathbf{W}_E))\tilde{\mathbf{q}}_0. \tag{70}$$

In order to find the maximum of our functional, we will use this gradient information to find the optimal initial condition realizing maximum energy at time $t = T$. Therefore in summary, this framework allows us to find the optimal initial perturbation associated with the maximum energy gain $G_E(T)$ over an optimization time interval, where $G_E(T)$ is defined as

$$G_E(T) = \frac{E(T)}{E_0}. \tag{71}$$

2. Full norm gain and frozen turbulent viscosity analysis

In order to compare our framework to already existing tools, we will also perform a classical SVD analysis of the system. In this case, the norm optimized is the total norm defined [as in (59)] by

$$\|\tilde{\mathbf{q}}\|_N^2 = \|\tilde{\mathbf{q}}\|_E^2 + \|\tilde{\mathbf{q}}\|_K^2$$
$$= \frac{1}{2}\int_0^1 (\tilde{u}^2 + \tilde{v}^2)dx. \tag{72}$$

As a consequence, the gain we identify is not an energy gain, but a "total" gain (which has no particular physical meaning). It is simply defined as

$$G_N(T) = \frac{N(T)}{N_0} = G_E(T)g(C_0), \tag{73}$$

where

$$g(C_0) = \frac{1 + C(T)}{1 + C_0} = \frac{1 + h(C_0)}{1 + C_0}, \tag{74}$$

where h is a function embedding all the dynamics relating the initial ratio C_0 [as defined in (65)] to the final ratio at the end of the optimization time interval $C(T)$. This expression shows that the SVD optimization by construction cannot correctly describe the physics of G_E since it implicitly optimizes the product of G_E and a nontrivial function of the initial ratio C_0, which product has no physical meaning.

In the frozen turbulent viscosity case (FROZ), the perturbation vector is $\tilde{\mathbf{q}} = (\tilde{u}, 0)^\top$, i.e., $\tilde{v} = 0$ in the system of direct and adjoint equations. The state vector has only one component and the energy norm is once again a true norm, $\|\tilde{\mathbf{q}}\|_E = \|\tilde{\mathbf{q}}\|_N$, allowing the use of the well-known singular value decomposition (SVD) analysis technique, which is explained briefly in Appendix C. The laminar case (LAM) is a further special case where the frozen turbulent viscosity $\overline{\nu_t}$ is set precisely to zero. In these two cases, $C_0 = 0$ (also true for all times) so that the two gains G_E and G_N defined in (71) and (73), respectively, are naturally equal.

F. Sensitivity analysis

We emphasized in Sec. II that our Lagrangian framework was not only a way to perform an optimization subject to constraints, but also a way to analyze the sensitivity of the objective functional to those constraints. In the particular class of problems under consideration, the state vector is constrained by a partial differential equation which is entirely defined by the base mean flow $\overline{\mathbf{q}}$ and the parameters ν, c_1, and c_2. The sensitivity of the optimized objective functional to a small change in any of these parameters can be retrieved thanks to the additional sensitivity information in the adjoint state vector. In terms of the nomenclature of Sec. II, all these parameters are constraint parameters, as they do not feature explicitly in the objective functional defined by (56).

1. Sensitivity with respect to the mean flow

We first consider the sensitivity of the objective functional with respect to the mean flow. This consists of computing the change of the Lagrangian functional $\mathcal{L}$ [defined in (67)] when we allow a small variation in the mean flow components. Since the mean flow is time-independent, so are the infinitesimal variations $\delta\overline{\mathbf{q}} = (\delta\overline{u}, \delta\overline{v_t})^\top$.

The sensitivities information is computed by taking the functional derivative of $\mathcal{L}$ with respect to the base flow. The full sensitivity vector ($\boldsymbol{\nabla}_{\overline{\mathbf{q}}}\mathcal{J}$) has two components and they can be expressed, after some integration by parts, as

$$\begin{aligned}\boldsymbol{\nabla}_{\overline{u}}\mathcal{J} &= \int_0^T \tilde{\mathbf{S}}_{\overline{u}}(\tilde{\mathbf{q}}^\dagger, \tilde{\mathbf{q}})\,dt,\\ \boldsymbol{\nabla}_{\overline{v_t}}\mathcal{J} &= \int_0^T \tilde{\mathbf{S}}_{\overline{v_t}}(\tilde{\mathbf{q}}^\dagger, \tilde{\mathbf{q}})\,dt,\end{aligned} \tag{75}$$

where the explicit expressions of the sensitivity vectors $\tilde{\mathbf{S}}_{\overline{u}}(\overline{\mathbf{q}})$ and $\tilde{\mathbf{S}}_{\overline{v_t}}$ are given in Appendix B. We notice that the sensitivity with respect to the base mean flow is a time scalar product (time integral). This means that this sensitivity is the cumulative contribution of the base mean flow variation at each time step. Therefore, in a practical situation, the longer the time interval for the optimization, the larger will be the error in the evaluation of the objective functional if there is any uncertainty in the mean flow state vector.

2. Sensitivity with respect to parameters

This model has three parameters: ν, c_1, and c_2. In order to derive the sensitivity of the optimal energy with respect to these parameters, we first need to notice that a change in their value will not only change the dynamics governing the perturbation equation but also the mean flow equation and, therefore, the mean flow itself. Thus, we need to define a new functional to account for the change in the base mean flow because of the small variation we allow in the parameters. We then incorporate the base mean flow equations (47) in the Lagrangian functional. This Lagrangian can be expressed as an extension of the one defined previously in (67):

$$\mathcal{L}'(\tilde{\mathbf{q}}, \tilde{\mathbf{q}}^\dagger, \overline{\mathbf{q}}, \overline{\mathbf{q}}^\dagger) = \mathcal{J}(\tilde{\mathbf{q}}) - (\tilde{\mathbf{L}}\tilde{\mathbf{q}}, \tilde{\mathbf{q}}^\dagger) - \langle\overline{\mathbf{L}}(\overline{\mathbf{q}}), \overline{\mathbf{q}}^\dagger\rangle. \tag{76}$$

Consequently, we have a new Lagrange multiplier which is the base flow adjoint state vector $\overline{\mathbf{q}}^\dagger = (\overline{u}^\dagger, \overline{v_t}^\dagger)^\top$. In order to fulfill the optimality condition of the problem, we have to satisfy the condition

$$\frac{\partial\mathcal{L}'}{\partial\overline{\mathbf{q}}}\delta\overline{\mathbf{q}} = 0. \tag{77}$$

This condition leads to the definition of the base flow adjoint variables. We notice that the adjoint base flow system is no longer homogeneous, but is additionally forced by the sensitivity with respect to the base flow:

$$\overline{\mathbf{L}}^\dagger\overline{\mathbf{q}}^\dagger = \boldsymbol{\nabla}_{\overline{\mathbf{q}}}\mathcal{J}. \tag{78}$$

Therefore,

$$\begin{aligned}\boldsymbol{\nabla}_\nu\mathcal{J} &= (\tilde{\mathbf{q}}^\dagger, \tilde{\mathbf{S}}_\nu\tilde{\mathbf{q}}) + \langle\overline{\mathbf{q}}^\dagger, \overline{\mathbf{S}}_\nu(\overline{\mathbf{q}})\rangle,\\ \boldsymbol{\nabla}_{c1}\mathcal{J} &= (\tilde{\mathbf{q}}^\dagger, \tilde{\mathbf{S}}_{c1}\tilde{\mathbf{q}}) + \langle\overline{\mathbf{q}}^\dagger, \overline{\mathbf{S}}_{c1}(\overline{\mathbf{q}})\rangle,\\ \boldsymbol{\nabla}_{c2}\mathcal{J} &= (\tilde{\mathbf{q}}^\dagger, \tilde{\mathbf{S}}_{c2}\tilde{\mathbf{q}}) + \langle\overline{\mathbf{q}}^\dagger, \overline{\mathbf{S}}_{c2}(\overline{\mathbf{q}})\rangle.\end{aligned} \tag{79}$$

The expression of the sensitivity matrices $\tilde{\mathbf{S}}$ and vectors $\overline{\mathbf{S}}(\overline{\mathbf{q}})$ are also given in Appendix B. The sensitivities have two contributions: a space-time scalar product accounting for the sensitivity due to the perturbation equation, and a space-only scalar product accounting for the sensitivity due to the base mean flow change induced by variation of the relevant parameter.

IV. RESULTS

A. Optimal perturbations

The results are presented in three parts, in order of increasing complexity. First of all, we will consider the laminar case "LAM" as summarized in Table I, i.e., the stability analysis of the Burgers equation, without any coupling with another partial differential equation, and in particular constant viscosity with no turbulent viscosity contribution. We then consider the "FROZ" case for a particular constant nonzero choice $\overline{v_t}$ of the turbulent viscosity, and considering the stability of the RAB equations with only a (coherent) perturbation velocity $\tilde{u}$, which allows us to understand the impact of a constant turbulent viscosity on the system. Finally, we consider the behavior of the full linearized model "FULL," using both our seminorm based framework and an SVD analysis based on optimizing the total true norm of the system. By considering the results of our framework and the unphysical SVD in tandem, we are able to identify the significance or otherwise of the output of the SVD analysis in a consistent manner. Fixing the turbulent viscosity to its mean value is a common simplifying assumption, and we are very interested in the robustness of our results to the application of this assumption.

1. Laminar analysis: The LAM case

Let us consider the Burgers equation, with a constant and uniform eddy viscosity $\nu = 0.05$. The equation governing the evolution of a perturbation of the form $\tilde{\mathbf{q}} = (\tilde{u}, 0)^\top$ is thus given by (51) with $\overline{v_t} = 0$. In this case, the perturbation kinetic energy is simply the 2-norm of the state vector, and so we can use an SVD analysis (as described in Appendix C). The optimal gain is then given by the largest singular value of the evolution operator. Here, the production of energy can only come from the coupling between the coherent perturbation $\tilde{u}$ and the base mean flow $\overline{u}$ [i.e., via term P_{E_1} of Eq. (53)], and since we are considering only focusing base mean flows (with
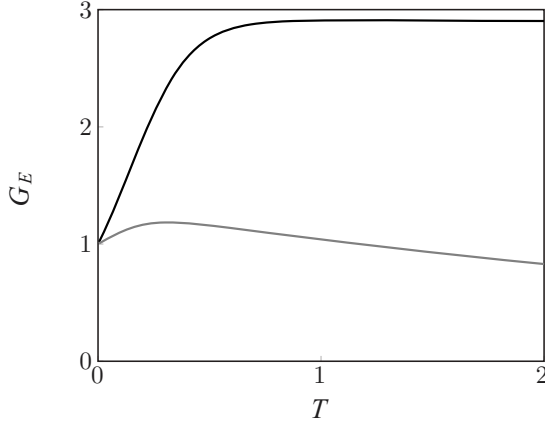
FIG. 6. Optimal gain against time for the "LAM" case (plotted with a black line) and the "FROZ" case (plotted with a gray line). The amplification in the LAM case is due to the positive energy production term P_{E_1} of Eq. (53), which can be seen as a focusing of the perturbation in the middle part of the domain. The gain then decreases very slowly due to the low value of the viscosity. In the FROZ case with $r = 0.5$ [as defined in (45)] the transient growth remains but is reduced dramatically because of the larger total viscosity. For sufficiently large values of r, any transient growth can be completely suppressed due to strong viscous damping.

negative slopes), we will have some energy production in the middle of the domain, due to the focusing of the perturbation.

In Fig. 6, we show the optimal coherent perturbation energy gain (71) against time for such a LAM case. We identify optimal transient growth which reaches its maximum gain for $T_{E_{opt}} = 1.15$, subsequent to which the gain decays slowly due to the relatively low value of the viscosity we have chosen. In Fig. 7, we show the structure of the optimal perturbation state vector both initially and at $T_{E_{opt}} = 1.15$, where the energy reaches its maximum value. We can see that the initial perturbation is not localized, but has a constant value over much of the domain, only decreasing at the edge of the domain to satisfy the boundary conditions, while the final perturbation has been strongly localized in the center of the domain and has a much larger amplitude than the initial state vector.

2. Frozen turbulent viscosity analysis: FROZ case

The laminar analysis we just performed is equivalent to a frozen turbulent viscosity stability analysis [perturbation of the form $\tilde{\mathbf{q}} = (\tilde{u}, 0)^\top$] for $r = 0$ (and so no turbulent viscosity, or indeed any turbulent property in the system). In the FROZ case, the base mean flow $\overline{\mathbf{q}}$ is a solution of the complete set of base flow equations (47) with a nonzero viscosity ratio r, defined by (45). The parameters in this set of equations are essentially arbitrary, and we choose to use $\nu = 0.05$, $r = 0.5$, $c_1 = 0.75$, and $c_2 = 2$, which are a good set of parameters in order to produce enough turbulent viscosity to have an effect on the base mean flow, but not to remove all the dynamics of the system. This appeared to be a balanced choice of parameters allowing us to examine all the interesting features of the system.

The observed gain for the FROZ case must be smaller than in the LAM case for two reasons. First of all, the total base mean flow viscosity will be larger than the laminar
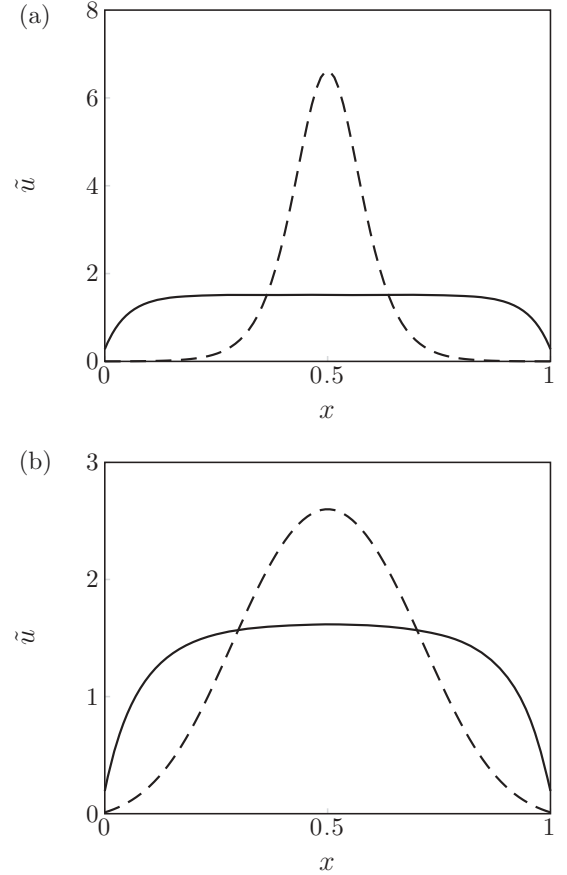

FIG. 7. (a) LAM case; (b) FROZ case. Optimal initial conditions (in solid lines) and final state (in dashed lines) at global maximum gain time $T_{E_{opt}} = 1.15$, for $r = 0.5$ as defined in (45). The perturbation is concentrated in the middle of the domain under the action of the base mean flow. For the FROZ case, the focusing is weaker (smaller base flow gradient), and the damping is larger.

value because it now includes the space-dependent turbulent viscosity. The damping term D_E in equation (53) will as a consequence be stronger. Moreover, a direct consequence of having more viscosity is a smaller slope for the base mean flow velocity which is directly involved in the production of energy P_{E_1} which will therefore be smaller than in the LAM case. The optimal gain curve for the FROZ case must then be beneath that of the LAM case. In Fig. 6, we plot the optimal curves corresponding to $r = 0$ (LAM case) and $r = 0.5$ (FROZ case). We note that even if the ratio of the turbulent viscosity to the laminar viscosity is small, the optimal gain curve is substantially affected. However, this depends in a nontrivial way on the modeling coefficients c_1 and c_2. We also notice that the optimal horizon time decreases as the amount of turbulence (modeled by the parameter r) increases.

3. Full linearized analysis: FULL case

a. Total norm optimization: SVD analysis. The analysis of the full linearized system of equations (49) requires the use of a norm for the two-component perturbation state vector $\tilde{\mathbf{q}} = (\tilde{u}, \tilde{v})^T$. We will present the results for the time dependence of the objective functional $\mathcal{J}$ of interest defined in (56), first for the total gain $G_N(T)$ [defined in (73)] optimized

in the total normalization norm $\| \cdot \|_N$ defined by (59) using SVD analysis, and then optimizing gain with respect to the energy seminorm $\| \cdot \|_E$ defined by (57). Let us first start with the case of the total norm ($\| \cdot \|_N$) gain optimization. For this total norm optimization, we will consider the gain in the energy seminorm for comparison with the other cases (in particular with the results obtained using our variational framework based on seminorm constraints) although the coherent perturbation kinetic energy is not actually the quantity being optimized. Other quantities which are also of interest to characterize the nature of the state vector are the time-dependent generalizations of the initial condition ratios R_0 and C_0 defined in (65).

These quantities measure the relative importance of the turbulent viscosity perturbation to the coherent velocity perturbation. A state vector having a high value of C (or equivalently $R \simeq 1$) will be identified with a "turbulent" state, while a state vector with a low value of C (or equivalently $R \simeq 0$) will be associated with a "laminar" state.

We plot $G_N(T)$ [as defined in (73)] against optimization time interval T for the optimal and first suboptimal state vectors (i.e., the two first singular values of the evolution operator, as discussed in Appendix C) in Fig. 8(a). In this case, two modes are competing in order to define the overall optimal perturbation: a transient mode (plotted with a black line) which has strong transient growth of the value of the total norm at short times, and the least stable mode (plotted with a gray line) responsible for the weakest possible long time decay. For the sake of simplicity, we will denote these two modes by STO (short time optimal) perturbation and LTO (long time optimal) perturbation. The STO perturbation reaches its maximum ($G_{N_{\text{opt}}} = 3.46$) for $T_{N_{\text{opt}}} = 0.38$ and the switching time for which the two modes have the same (total) gain is $T_s \simeq 1.25$. The main result of this SVD analysis is that there is a competition between two modes for which a clear transient growth is observed, meaning that two perturbation growth mechanisms are relevant. Even if this dynamics cannot be associated exclusively with coherent velocity perturbation energy production (as in the LAM and FROZ cases discussed above), we can now say that the presence of the second perturbation evolution equation [for $\tilde{\nu}$ in (49)] introduces new dynamics to the system's behavior. Indeed, the term P_{E_2} of Eq. (53) can now be a new source of energy. This term is directly proportional to both the magnitude of the slope of the coherent velocity perturbation and to the perturbation turbulent viscosity, and is thus responsible for much richer dynamics of the system.

However, it is legitimate to question whether this transient growth is associated with transient growth of the kinetic energy seminorm of the perturbation or merely with growth of the seminorm of the turbulent viscosity perturbation. We can determine the value of C_0 (and hence the relative magnitudes of the perturbation velocity and the turbulent viscosity) for each optimization interval for both the initial state vector $C_0 = C(0)$ and for the final state vector $C(T)$, as plotted in Fig. 8(b). We notice that the (short-time) STO perturbation is associated with "turbulent" ($C_0 \gg 1$) initial perturbations for sufficiently long optimization intervals, which evolves towards a more laminar [$C(T) \lesssim 1$] state, with a larger contribution from the coherent perturbation velocity $\tilde{u}$. This
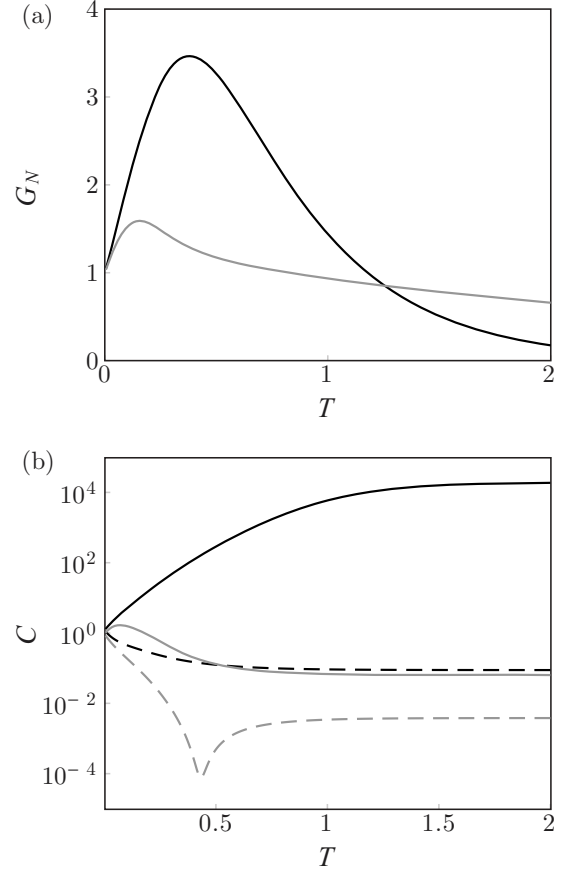

FIG. 8. (a) The variation of $G_N(T)$ [defined in (73)] with optimization time T of the short time optimal (STO) perturbation (plotted with a black line) and long time optimal (LTO) perturbation (gray line) total norm for the FULL perturbation equations case, when the total norm $\| \cdot \|_N$ defined in (59) is optimized using an SVD analysis. (b) The variation of the ratio C_0 as defined in (64) (plotted with solid lines) and $C(T)$ as defined in (65) (plotted with dashed lines) with optimization time T of the STO perturbation (black lines) and the LTO perturbation (gray lines) for both the optimal (black) and suboptimal (gray) mode. We notice that there is a competition between two modes: the short time optimal (STO) perturbation and the long time optimal (LTO) perturbation. The STO perturbation is typically associated with large values of C_0, meaning that the STO perturbation is initially "turbulent," although it evolves to have $C(T) \ll 1$, meaning that ultimately the perturbation is almost exclusively composed of coherent perturbation velocity. On the other hand, the LTO perturbation is dominated at all times by the coherent perturbation velocity [$C(T) \ll C_0 \ll 1$] and thus we refer to it as a "laminar" perturbation.

implies that this transient growth is due to perturbation kinetic energy production. The (late-time) LTO perturbation (which is actually optimal starting from $T = T_s \simeq 1.25$) is on the contrary initially mostly laminar (with low values of C_0) and evolves toward an almost completely laminar state, even more dominated by perturbation velocity [i.e., with $C(T) \ll 1$].

In order to identify the associated transient growth of the perturbation kinetic energy, we now consider the evolution of the perturbation velocity kinetic energy gain $G_E(T)$ as defined in (71). We plot this quantity (for both the STO perturbation and the LTO perturbation) in Fig. 9. We can see that this gain
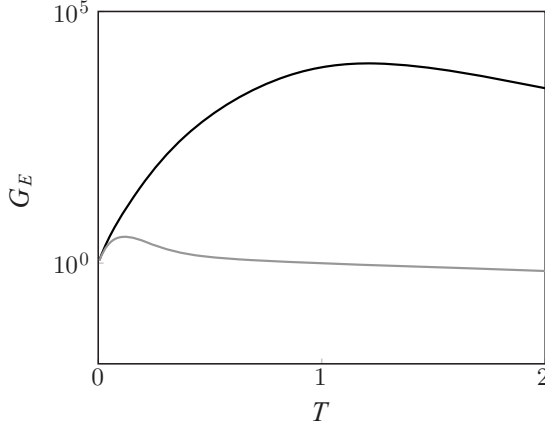
FIG. 9. Variation of energy gain $G_E(T)$ as defined in (71) with the optimization time interval T for the full linearized perturbation system of equations (49) for perturbations which optimize the total norm gain $G_N(T)$, as defined in (73). The STO perturbation (plotted with a black line) achieves very large values of $G_E(T)$, while the LTO perturbation (plotted with a gray line) is associated with substantially smaller values of $G_E(T)$, comparable to those obtained in the FROZ case. The STO perturbation has a larger energy gain than the LTO perturbation in the plotted time window but eventually has a larger decay rate and is thus less efficient in preserving a large energy gain.

(which is not the optimal energy gain) can have very large values compared to the previous studies. Indeed, the perturbation kinetic energy is amplified by approximately five orders of magnitude and reaches its maximum for $T_{E_{opt}} = 1.21$ with a gain of $G_{E_{opt}} = 9.3 \times 10^4$. We consider these two quantities as the "optimal" time and gain values extracted from a full-norm gain optimization. Since $G_E(T)$ is not optimized, but is just derived from the results of the full-norm optimal perturbation for the entire perturbation velocity-viscosity state vector, it is not surprising to see that the STO perturbation remains the mode with larger $G_E(T)$ after $T = T_s \simeq 1.25$. For sufficiently long times, the LTO perturbation will ultimately become dominant with respect to kinetic energy gain as well. Besides, this figure clearly underlines the difference in the order of magnitude between the "laminar" perturbations of the LTO perturbations and the "turbulent" perturbations of the STO perturbations. Indeed, the turbulent STO perturbation is associated with very large values of energy gain while the laminar LTO perturbation shows an energy gain which can be compared most closely to the LAM case optimal gain shown in Fig. 6.

The linear optimal gain analysis of this small problem brings to light the fact that there is a competition between essentially laminar and turbulent perturbations. The turbulent perturbation is strikingly associated with a very large transient growth of the perturbation kinetic energy [due to the term P_{E_2} of Eq. (53)], and is optimal for short times. On the other hand, the laminar perturbation is only optimal for longer times, and is in fact the least stable perturbation, responsible for the weakest possible decay. We show the initial and final time (at $T = T_{E_{opt}}$) structure of the STO perturbations and LTO perturbations in Figs. 10 and 11, respectively. The STO (turbulent) perturbations are antisymmetric in $\tilde{u}$ and symmetric in $\tilde{v}$ whereas the LTO (laminar) perturbations have the opposite symmetries.
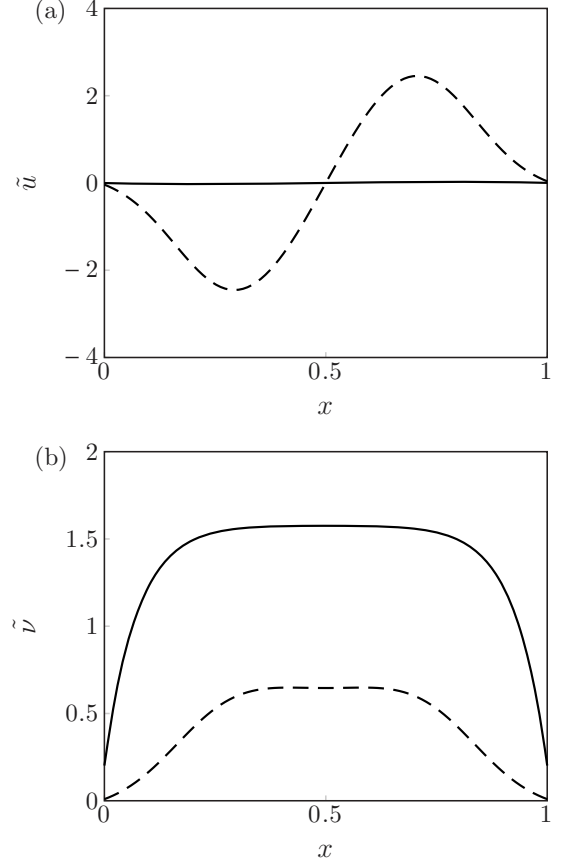

FIG. 10. Variation with x of the initial perturbation (plotted with solid lines) and final perturbation (at $T = T_{E_{opt}}$, plotted with a dashed line) of (a) coherent perturbation velocity $\tilde{u}$ and (b) perturbation turbulent viscosity $\tilde{v}$ for the STO perturbation. We notice a decay of the turbulent viscosity perturbation [through diffusion, destruction, and other mechanisms described by Eq. (55)] giving rise to large perturbation velocities in the final perturbation state vector. This perturbation is not present in either the LAM case or the FROZ case, and arises from the richer dynamics in the FULL case where the turbulent viscosity evolves in space and time.

This SVD analysis used to find the optimal perturbation associated with the largest achievable total norm gain given by (73) shows the major difference between the FROZ case and the FULL case (as defined in Table I). Indeed, we see that a new type of perturbation appears, for which the turbulent viscosity component $\tilde{v}$ of the state vector is much larger than the coherent perturbation velocity $\tilde{u}$, which leads to a substantially larger kinetic energy gain. The main drawback of this method is that we only access the energy gain information through the optimization of a nonphysical norm (the total norm). Therefore, the calculated energy gain is not optimal in any sense. Indeed, the optimized total gain defined in (73) is a product between the energy gain and a term which is a nontrivial function of the parameter C_0. As a consequence, in this analysis this contribution parameter C_0 is an output of the optimization [see Fig. 8(b)] instead of being an input parameter.

The seminorm framework introduced in Sec. II allows us to consider the optimization of the energy gain directly, and it is interesting to investigate what, if any, are the points of
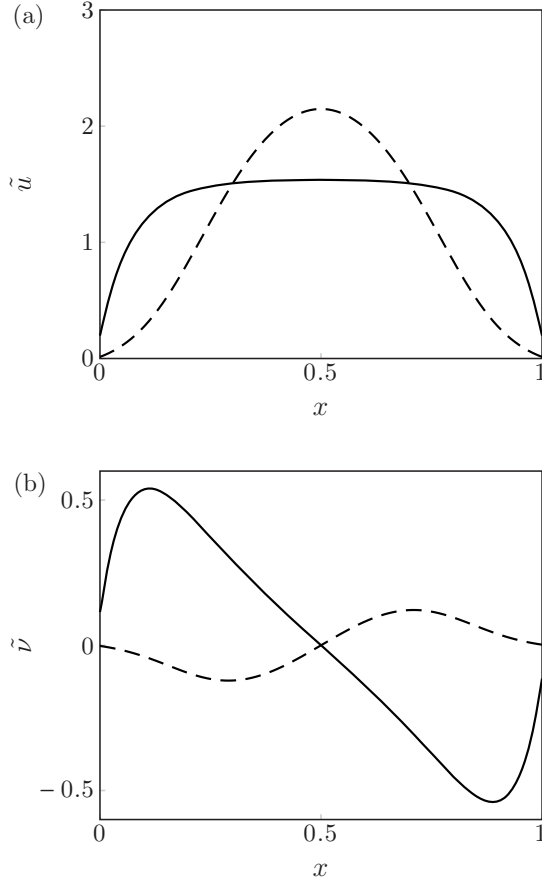
FIG. 11. Variation with x of the initial perturbation (plotted with solid lines) and final perturbation (at $T = T_{E_{\mathrm{opt}}}$, plotted with a dashed line) of (a) coherent perturbation velocity $\tilde{u}$ and (b) perturbation turbulent viscosity $\tilde{\nu}$ for the LTO perturbation. The turbulent viscosity perturbation has a minor role in the dynamics. Indeed, the behavior of the coherent perturbation velocity $\tilde{u}$ is very similar to the behavior observed in the FROZ case [see Fig. 7(b)].

connection between the results of calculations based on the seminorm framework, and the SVD analysis based around optimization of the gain expressed in terms of the total norm $\| \cdot \|_N$.

b. Energy optimization: Seminorm constraints. The perturbation kinetic energy gain G_E defined in (71) is a highly relevant quantity in the dynamics of the system. As a consequence we will now use the seminorm framework developed in Sec. II to optimize this quantity over finite time intervals. We present in Fig. 12 the variation of $G_E(T)$ with optimization time interval T for different values of C_0 [as defined in (64)]. Gray curves represent energy gains for $C_0 < 1$ (i.e., for "laminar" perturbations) whereas the black curves represent energy gains for $C_0 > 1$ (i.e., for "turbulent" perturbations) while the black curve corresponds to the balanced case $C_0 = 1$. We can clearly see that laminar perturbations corresponding to low values of C_0 are associated with very low gain values, while the turbulent perturbations (with $C_0 > 1$) can lead to much higher gains. We do not plot on the figure the optimal curves for large values of C_0 because we find that the gain increases linearly with C_0 for large C_0.
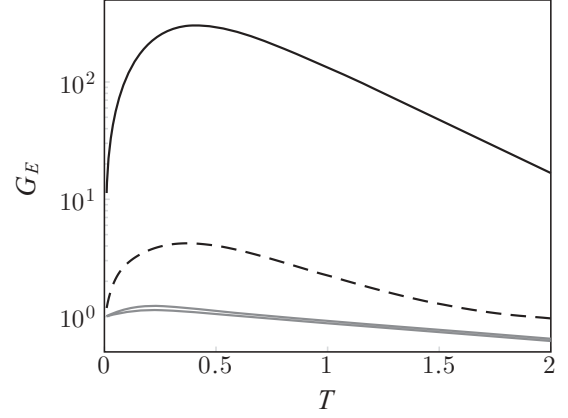

FIG. 12. Variation of the energy seminorm gain $G_E(T)$ as defined in (71) with optimization time interval T for values of $C_0 = K_0/E_0$ ranging from 10^{-2} to 10^1. Gray curves corresponds to $C_0 = 10^{-2}$ and $C_0 = 10^{-1}$, the black dashed curve corresponds to the balanced case $C_0 = 1$, and the black curve corresponds to $C_0 = 10$. We notice that when the contribution of the turbulent viscosity perturbation increases, higher gains are achieved. A whole family of black curves exists with gains evolving linearly with C_0 for larger values of C_0.

We plot in Fig. 13 the maximum gain $G_{E_{\mathrm{opt}}}$ as a function of the ratio C_0 as defined in (64). The perturbation velocity energy gain in the limit of low C_0 is constant ($G_{E_{\mathrm{opt}}} \simeq 1.13$). In order to compare the limit of the FULL case and the frozen turbulent viscosity FROZ case [using the same values for $\nu = 0.05$ and $r = 0.5$, where r is defined in (45)], we plot the two corresponding optimal gain curves in Fig. 14. The first observation is that the FULL case does not converge toward the FROZ case when the contribution from the turbulent viscosity perturbation is very small compared to the coherent velocity perturbation (i.e., in the limit $C_0 \to 0$). This means that no matter how small C_0 is, the contribution of the viscosity perturbation to the dynamics is never negligible. Indeed, the maximum gain we obtained for FROZ case was approximately $G_{E_{\mathrm{opt}}} = G_{N_{\mathrm{opt}}} \simeq 1.18$. This means that the stability analysis of mean flows is a singular problem, since in the limit of low values of the turbulent viscosity perturbation, we do not recover the results of the FROZ case.

Moreover, in the limit of large C_0 (turbulent perturbations), $G_{E_{\mathrm{opt}}}$ evolves linearly with C_0:

$$G_{E_{\mathrm{opt}}} \propto C_0, \tag{80}$$

with a proportionality coefficient of about 3 in our case. This is the signature of the linear relation involved in the energy production mechanism [term P_{E_2} of Eq. (53)]. As a consequence, we can linearly extract energy from the base flow via the perturbation turbulent viscosity perturbation as the ratio C_0 increases. Indeed, the growth of the turbulent perturbations is larger for initially large values of K because for large values of $\tilde{\nu}$, we consequently have a large production term P_{E_2} in (52). In fact, $\tilde{\nu}$ catalyses very strong growth of E for the turbulent modes. For the laminar modes however, this mechanism exists but is not proportionally significant. (It is totally absent for the FROZ case.)

Turning our attention to time dependence, Fig. 13(b) shows that turbulent perturbations are associated with larger optimal
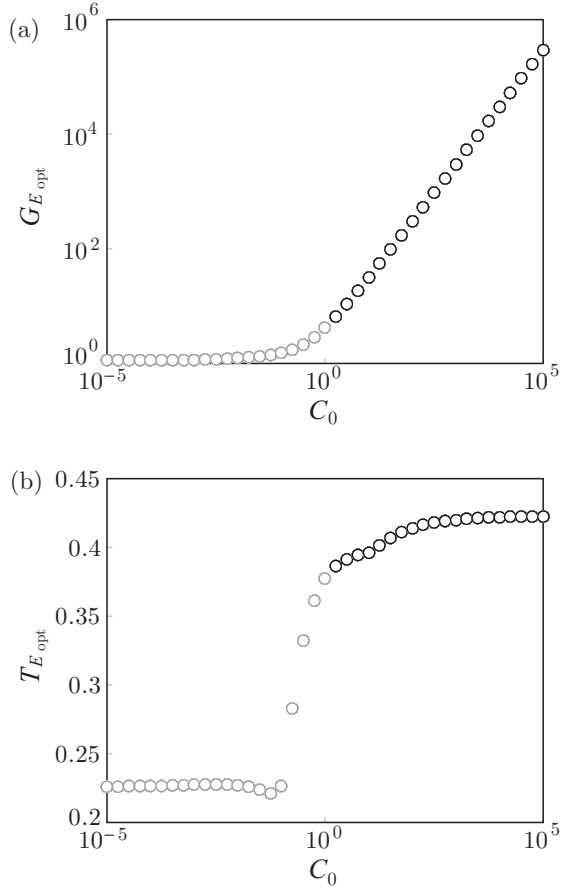
FIG. 13. Variation with $C_0 = K_0/E_0$ as defined in (64) of (a) optimal gain $G_{E_{\mathrm{opt}}}$ and (b) optimal times $T_{E_{\mathrm{opt}}}$ for the FULL case. We see that the optimal gain increases linearly with C_0, meaning that we can linearly extract as much energy as we want from the base flow and turbulent viscosity perturbation. The optimal time is also increasing with C_0, but not in such a neat fashion. The line types are the same as in Fig. 12.

times than laminar perturbations, though once again, the optimal time as $C_0 \to 0$ for the FULL case is still different from the optimal time for the FROZ case. In summary, we present the results of the three different analyses in Table II. We distinguish between laminar perturbations ($C_0 \ll 1$) and turbulent perturbations ($C_0 \gg 1$) for the FULL case, i.e., solutions of the full linearized equation system defined by (49).

Another way to understand the mechanism of energy production is to consider the plot of the kinetic energy of the coherent perturbation velocity as a function of the seminorm of the perturbation turbulent viscosity, parameterized by time t. We consider all the optimal perturbations for a given optimization time interval $T = 0.3$ for different values of the ratio $C_0 = K_0/E_0$ [defined in (64)] and evolve the corresponding optimal perturbations in time. In Fig. 15 we plot parametric curves defined by $X_{\mathrm{opt}}(t), Y_{\mathrm{opt}}(t)$ where $X_{\mathrm{opt}}(t)$ and $Y_{\mathrm{opt}}(t)$ are defined as

$$X_{\mathrm{opt}}(t) = \frac{K(t)}{E_0 + K_0} = G_E(t)C(t)(1 - R_0), \quad (81)$$

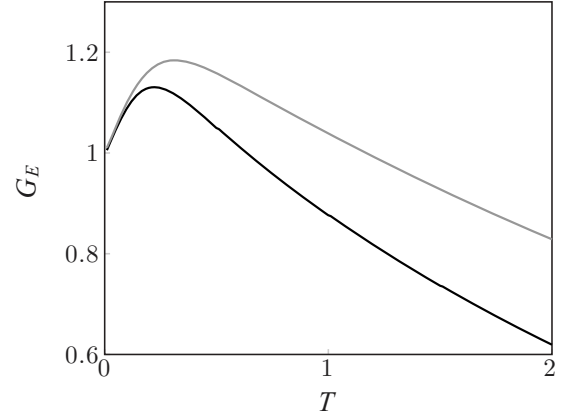$$Y_{\mathrm{opt}}(t) = \frac{E(t)}{E_0 + K_0} = G_E(t)(1 - R_0). \quad (82)$$

FIG. 14. Variation of coherent perturbation kinetic energy gain $G_E(T)$, as defined in (71) for the FROZ case (plotted with a gray line) and the FULL case in the limit $C_0 \to 0$ (plotted with a black line). These results have been obtained with the same set of parameter values, and especially with $r = 0.5$. Even in this limit, the results for the two cases differ quantitatively, both in optimal time and optimal gain predictions.

In the figure initial conditions are shown as black dots, which from the definitions (82) and (81) clearly must lie on the curve $x + y = 1$, which is plotted with a dashed line. Using the same line types as Fig. 12, gray curves denote "laminar" perturbations ($C_0 \ll 1$) and black curves denote "turbulent" perturbations ($C_0 \gg 1$) while a black dashed curve denotes the marginal perturbation with $C_0 = 1$.

The gain G_E can be retrieved from this figure by evaluating the difference in y coordinate from an initial condition to the maximum of the corresponding curve, due to the logarithmic scalings. The gray curves (for the laminar perturbations) are associated with very low gains since there is little vertical variation while the black curves (turbulent perturbations) show a very large energy gain since the distance from an initial condition to the maximum of the curves is getting larger and larger as the amount of perturbation turbulent viscosity in the initial perturbation increases. Moreover, we can see that the energy production for the turbulent perturbations does

TABLE II. Summary of the results section. LAM: laminar case using the frozen turbulent viscosity assumption $\tilde{\nu} = 0$ with $r = 0 = \overline{\nu_t}$; FROZ: frozen turbulent viscosity case with $\tilde{\nu} = 0$; FULL: full linearized case solving (49); SVD: singular value decomposition; SN: seminorm; STO: short time optimal; LTO: long time optimal; LAM: laminar ($C_0 \ll 1$); TURB: turbulent ($C_0 \gg 1$). The bold font means that this quantity is optimal.

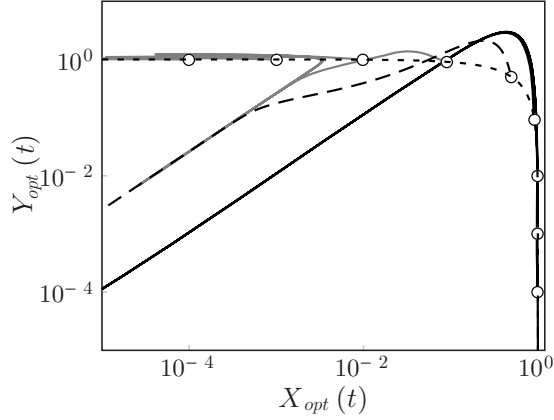| | | | FULL | | | |
| | | | SVD | | SN | |
	LAM	FROZ	LTO	STO	LAM	TURB
r	0	0.5	0.5	0.5	0.5	0.5
C_0	0	0	$\lesssim 1$	$\gg 1$	$\ll 1$	$\gg 1$
$T_{E_{\mathrm{opt}}}$	**1.15**	**0.31**	0.12	1.21	**0.22**	**0.42**
$T_{N_{\mathrm{opt}}}$	**1.15**	**0.31**	**0.16**	**0.38**	0.22	0.39
$G_{E_{\mathrm{opt}}}$	**2.91**	**1.18**	3.34	9.3×10^3	**1.13**	$\boldsymbol{\propto C_0}$
$G_{N_{\mathrm{opt}}}$	**2.91**	**1.18**	1.59	**3.46**	1.13	3.43

FIG. 15. Evolution diagram with time of scaled perturbation turbulent viscosity $X_{opt}(t)$ and perturbation velocity $Y_{opt}(t)$, defined by (81) and (82), respectively, parameterized by time t for an optimal perturbation with optimization interval $T = 0.3$, for different values of the initial ratio C_0. We can clearly see the two different types of perturbation: laminar perturbations ($C_0 \ll 1$) are plotted with gray lines, and turbulent perturbations ($C_0 \gg 1$) are plotted with black lines which are the two limit cases. Initial conditions are represented with black circles, and lie on the curve $x + y = 1$ (plotted with a short dashed line). The laminar perturbations do not have significant transient growth, while the turbulent perturbations in black have very large transient growth of the energy. An intermediate state (plotted with a dashed line) exists for $C_0 = 1$, corresponding to the point $(0.5, 0.5)$ on this figure.

FIG. 16. Evolution diagram with time of scaled perturbation turbulent viscosity $X_{opt}(t)$ and perturbation velocity $Y_{opt}(t)$, defined by (81) and (82), respectively, parameterized by time t for a full-norm optimal perturbation (for $T = T_{E_{opt}}$, $T = 1.21$ for the STO perturbation and $T = 0.12$ for the LTO perturbation) for the optimization of the total norm $\| \cdot \|_N$ defined in (59) using the SVD analysis as described in Sec. IV A 3 a. The gray curve plots the time evolution of the LTO perturbation and the black curve plots the time evolution of the STO perturbation. By comparison with Fig. 15, there is apparently a close relationship between the STO perturbation and the turbulent perturbation from the FULL case, and the LTO perturbation and laminar perturbation from the FULL case.

not immediately lead to a decrease of the turbulent viscosity perturbation norm, implying (of course) that the energy has been taken from the mean flow. Indeed, the optimal turbulent perturbation has the best symmetry and shape for both $\tilde{u}$ and $\tilde{v}$ in order to maximize this production, through the large size of the catalytic term P_{E_2} in (53).

In order to compare these results with the SVD analysis presented in Sec. IV A 3 a, we plot the equivalent parametric curves for an STO perturbation and an LTO perturbation (for the same values of the parameters as before) in Fig. 16. We can see that these two perturbations correspond closely to what we have identified using the optimization of the FULL case using seminorm constraints. The STO perturbations appear to be appropriately classified as "turbulent" perturbations, while the LTO perturbations appear to be appropriately classified as "laminar" perturbations.

However, even if the SVD analysis appears to give some relevant information, it is completely unable to predict the optimal gain and time. Indeed, the SVD also finds an (unphysical) optimal value for the ratio C_0 for all optimization time intervals (e.g., the initial condition used to produce Fig. 16), which prevents the calculated G_E gain from being "optimal" in any meaningful sense.

The addition of the dynamic constraint on the turbulent viscosity perturbation brings new dynamical properties to the system of equations, in terms of both optimal time and gain value, compared to the simpler FROZ case with a frozen turbulent viscosity. Such a method misses a fundamental physical process, the possibility of extracting energy from the mean flow due to spatial and temporal variations in the turbulent viscosity. Indeed, the frozen turbulent viscosity
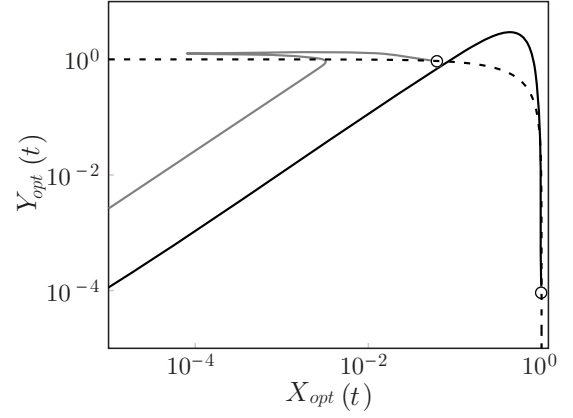
model yields a different result from the full linearized analysis even with an initial ratio $C_0 = K_0/E_0 \to 0$. In other words, in order to capture the true dynamics of the system, analysis of the full linearized system of equations seems to be indispensable, because of the new source of perturbation kinetic energy associated with fluctuating turbulent viscosity perturbation $\tilde{v}$.

Applying our seminorm framework allows us to identify two qualitatively different perturbations associated with two different dynamics. The first perturbation is driven principally by coherent perturbation velocity and hence energy extraction from the mean flow via the term P_{E_1} in (53). Such perturbations do not typically have a large perturbation kinetic energy gain, and we refer to them as "laminar" perturbations since they typically have small turbulent viscosity perturbation. The other perturbation is on the contrary essentially driven by a dominant turbulent viscosity perturbation, and is associated with very large energy production. Indeed, the energy gain increases linearly with the ratio $C_0 = K_0/E_0$ describing the relative contribution of the initial turbulent viscosity to the initial kinetic energy in the perturbation vector. In essence, we have performed a multiscale stability analysis where we are able to control the type of perturbation we imposed to the system, and we find that small-scale perturbations (parameterized by spatially varying turbulent viscosity) are much more efficient at driving the growth of velocity perturbations (through the term P_{E_2}). We can in fact extract as much energy as we want from the base mean flow, given that the magnitude of the term $\tilde{v}\partial_x\tilde{u}$ is large and has the appropriate symmetry.

However, nonlinear effects as well as feedback on the mean flow will certainly lead to saturation of this energy production and then lead to an identification of an optimal C_0 defining an energy gain (G_E) curve which will be different from that which was obtained using the singular value decomposition. We

indeed think nonlinear saturation is needed in this particular case to find the optimal initial ratio $C_0(T)$, but another set of linear equations describing a whole different problem might well be suitable for the identification of an optimal initial ratio $C_0(T)$ associated with an optimal energy gain $G_E(T)$ which will be different from the couple $[G_E(T), C_0(T)]$ obtained when optimizing $G_N(T)$ through a SVD analysis. Moreover, it is reasonable to suppose that in some particular problems, physical arguments may lead to estimates of the appropriate size of the initial turbulent viscosity (or more generally the respective size of the different components of the state vector), and hence give us a physically acceptable range for the parameter C_0.

The underlying physics described is of course consistent with SVD analysis optimizing the (nonphysical) gain of the total 2-norm of the perturbation state vector, since we showed that the dynamics of the calculated short time optimal (STO) perturbation and long time optimal (LTO) perturbation corresponds respectively to the turbulent perturbations and laminar perturbations dynamics of the full linearized system of equations. The new framework yields detailed information on the problem and, in fact, allowed us to have a full understanding thanks to the possibility of separation of scales in the perturbation vector. Furthermore, the framework allows us to approach the problem from a physical point of view by choosing the type of perturbation to impose and then identifying the associated dynamics.

B. Sensitivity analysis

Naturally, we are also able to conduct a sensitivity analysis, which will give us some information about how a change in the mean flow $\overline{\mathbf{q}}$ or any of the three parameters ν, c_1, or c_2 influences the optimal value of the objective functional $\mathcal{J}$. We focus on the most interesting case where $C_0 \gg 1$, meaning that we have a large transient growth due to a draining of energy from the velocity mean flow to the velocity perturbation through the catalysis allowed by the optimal symmetry chosen by both $\tilde{u}$ and $\tilde{v}$.

We plot in Fig. 17(a) the sensitivity of the final optimal energy $E_{opt}(T)$ to the mean flow velocity $\nabla_{\overline{u}}\mathcal{J}$ and in Fig. 17(b) the sensitivity of the final optimal energy to the mean flow turbulent viscosity $\nabla_{\overline{\nu}_t}\mathcal{J}$. The general structure of these space-time diagrams (where time is the optimizing time interval T) is useful to understand the role of the base mean flow structure in the production of energy. Indeed, we can see by looking at the general trend of these two figures that the zones of the domain which give the largest sensitivity magnitude are situated on both sides of the middle of the domain for the base mean flow velocity and centered in the middle of the domain for the mean turbulent viscosity. This can be simply explained by the fact that an increase of the slope of the mean flow velocity (increase of the mean velocity for $0 < x < 1/2$, or decrease for $1/2 < x < 1$) will give rise to a larger energy production (see for instance Fig. 12). Moreover, since the perturbation is mostly localized in the middle of the domain, an increase in the mean turbulent viscosity will be more dramatic in this zone compared to the edges of the domain. However, when looking more carefully at the sensitivity functions, we notice that a zone of high sensitivity
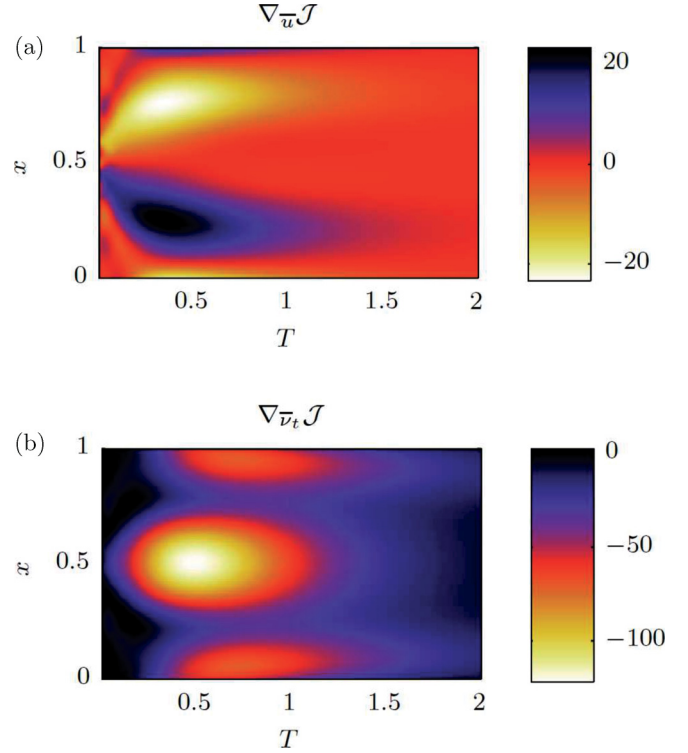

FIG. 17. (Color) Space-time plots of the sensitivity of the final optimized energy $E_{opt}(T)$ for varying optimization time intervals T with respect to (a) the mean flow velocity $\nabla_{\overline{u}}\mathcal{J}$ and (b) the mean flow turbulent viscosity $\nabla_{\overline{\nu}_t}\mathcal{J}$ for $C_0 \gg 1$.

appears on each side of the domain for $T \simeq 0.2$ for the mean flow velocity. These layers have an opposite sign to what we would simply expect to increase the negative slope amplitude $\partial_x \overline{u}$. In the same way, two other zones of sensitivity appear for times $T \simeq 0.4$ in the $\nabla_{\nu_t}\mathcal{J}$ plot and are situated close to the boundaries of the domain. These zones are also associated with negative sensitivity but a smaller absolute value compared to what is observed in the middle of the domain. This is consistent with the zones of high sensitivities at the boundaries in $\nabla_{\overline{u}}\mathcal{J}$ and suggests that the steeper the positive slope at the boundary, the larger the effect on the energy production.

In Fig. 18, we plot the sensitivities to the constraint parameters of the problem. For the mean flow, both $\nabla_{\nu}\mathcal{J}$ and $\nabla_{c_1}\mathcal{J}$ are negative since they are respectively associated with dissipation and production of viscosity (indirect dissipation), while for $\nabla_{c_2}\mathcal{J}$, this term is positive since the destruction of viscosity decreases the energy dissipation of the mean flow. Interestingly, there is qualitatively different behavior when we look at the perturbation contribution. Indeed, except for ν for which the sign of the sensitivity is the same for mean and perturbation sensitivities, we observe an opposite trend between these two contributions for c_1 and c_2 sensitivities. As we have seen, the base flow total viscosity (composed of both laminar and turbulent viscosity) is only responsible for decay while the perturbation turbulent viscosity actually triggers large energy growth. As a consequence, increasing base flow turbulent viscosity will lead to a decrease of $\mathcal{J}$ whereas adding some perturbation turbulent viscosity will boost the energy growth. The trends can now be explained
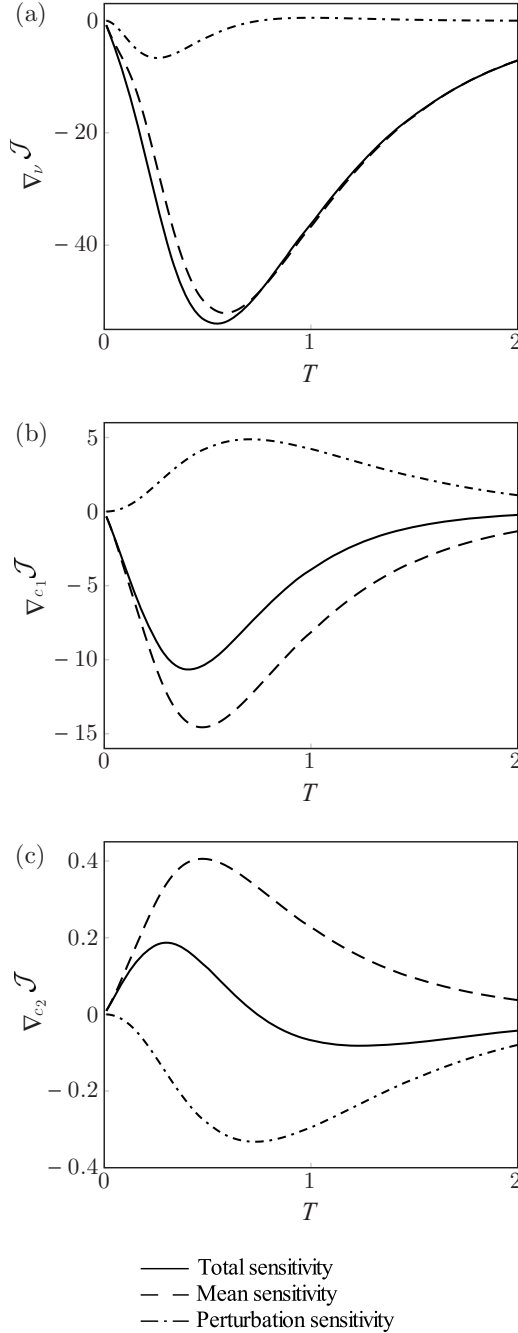
FIG. 18. Variation of the sensitivity of the final optimal energy $E_{opt}(T)$ with optimization time interval T to (a) the viscosity $\nabla_\nu \mathcal{J}$, (b) the turbulent viscosity production coefficient $\nabla_{c_1} \mathcal{J}$, and (c) the turbulent viscosity destruction coefficient $\nabla_{c_2} \mathcal{J}$. Solid lines show the variation with optimization time interval of the total sensitivities, dashed lines show the variation of the mean flow sensitivities, and dashed-dotted lines show the variation of the perturbation sensitivities. Increasing ν decreases the value of the objective functional $\mathcal{J}$, and mean and perturbation contributions act in the same way, with a larger sensitivity due to the perturbation. A positive variation of c_1 also leads to a decrease of the objective functional, although the perturbation and mean flow sensitivity act oppositely, with the perturbation sensitivity being positively correlated with the total sensitivity. The change in $\mathcal{J}$ due to a variation in c_2 is opposite (in both mean and perturbation sensitivity) to the one observed with c_1, but is an order of magnitude smaller.

by noticing that this process of production and destruction of either base or perturbation turbulent viscosity is driven by the parameters c_1 and c_2. However, it is important to notice that the total sensitivities are separated by an order of magnitude, with

$$\nabla_\nu \mathcal{J} \gg \nabla_{c_1} \mathcal{J} \gg \nabla_{c_2} \mathcal{J}, \qquad (83)$$

and so the turbulence modeling parameters actually have a relatively minor impact on the dynamics of the perturbation compared to modifications in the viscosity.

C. Discussion

Interestingly, although the conventional SVD analysis is not a formally correct method to study optimal energy growth in this system where the energy is a seminorm of the state vector, it still yields valuable insight, in particular that the frozen turbulent viscosity assumption is not appropriate. Indeed, we find a mode associated with a large turbulent viscosity perturbation amplitude which triggers, as we expected from the energy analysis carried out in Sec. III, a new energy production mechanism and gives rise to a very large transient growth of the energy. The interpretation and validation of the results of the SVD analysis however relies on the use of the seminorm framework presented in this paper. The seminorm framework allows us to conclude that we can linearly extract energy from the turbulent viscosity component, and the larger its amplitude, the larger the optimal energy at the end of the optimization interval. We can identify two qualitatively different types of behavior depending on the initial value C_0 of the ratio of the turbulent viscosity perturbation amplitude to the velocity perturbation amplitude. When this ratio is small, the energy production is small and is only due to the interaction between the mean flow velocity and the perturbation velocity, and the perturbation is essentially "laminar." However, for large values of this parameter, a new behavior appears, and such "turbulent" perturbations exhibit very large perturbation kinetic energy gain for which the gain evolves linearly with C_0 suggesting some universal behavior. Finally, from a sensitivity analysis, the influence of the modeling parameters can be shown to be very small compared to the influence of the laminar viscosity parameter on the flow evolution.

V. EXTENSIONS

This paper is, to the best of our knowledge, a first attempt to formally treat the seminorm gain optimization problem and, more generally, the inclusion of seminorm constraints. We address this problem by using a Lagrangian variational framework which, instead of having a single initial normalization constraint, has complementary seminorm constraints allowing us to control the relative contributions of the different components in the perturbation vector. The gain is then simply computed by calculating the ratio of the value of the seminorm at the final time (resulting from the optimization) to the value at the initial time (chosen through normalization). The optimization procedure can then be repeated for different values of the parameter quantifying the relative size of the initial amplitudes of the different components of the state vector.

This kind of multiscale nonmodal stability analysis is of a particular interest for systems where the definition of the energy arises from different physical contributions. Our framework then provides a systematic procedure to separate the different energy contributions and optimize any desired gain, even defined with seminorms. A few of the problems where nonmodal stability analysis is being applied, and where the multiscale stability analysis presented here could be appropriate, include turbulent mean flows [16,17], compressible flows and thermoacoustics [23,30–32], Rayleigh-Bénard type flows with thermally induced density gradients due to temperature effects [33], coupled fluid and electric field systems [24,25], magnetohydrodynamics [26], and irreversible mixing in density stratified flows [34,35].

Moreover, more than just being a way to calculate seminorm gains, the method can be used on other types of problems. Indeed, in flow control, the optimal placing and type of action of actuators could be derived from the optimization of conventional objective functionals (energy, drag, etc.) with constraints on seminorms defined on a compact support of the domain. By constraining the complementary part of the domain, and choosing a relatively large value for the seminorm defined on the area of interest with a relatively small value for the complementary seminorm, we would be able to identify the optimal localized forcing. We have developed our framework in terms of seminorms because of their relevance to fluid dynamics problems, but norm properties are not essential to the framework, and it is straightforward to generalize it for arbitrary functionals f, developing appropriate (physical) constraints to bound the magnitude of elements of the entire state vector space (in particular the kernel of f).

VI. CONCLUSION

In this paper, we develop a general Lagrangian variational framework for optimization problems using seminorm constraints. We present a systematic way to study optimal gains defined in terms of the state vector, with the introduction of new parameters setting the ratio between the different components of the perturbation state vector. This framework is a way to perform multiscale stability analysis, where the different components of the perturbation state vector do not have (necessarily) the same amplitude (which is relevant for multiphysics problems). To demonstrate the utility of this framework, we consider a simple idealized problem with a coupled set of two Burgers equations describing the evolution of flow velocity and a transport equation for a turbulent viscosity for which the production and destruction are controlled through two modeling parameters. This constitutes a minimal set to describe much of the key physics underlying the Reynolds-averaged Navier-Stokes equations and also some properties of mean flows in closed nonlinear dissipative systems. The Reynolds-averaged Burgers (RAB) equations are a simple one-dimensional approximation of a RANS equation for which a Boussinesq hypothesis of turbulence is considered, and a new turbulent viscosity is introduced, governed by a transport equation, reminiscent of the Spalart-Allmaras turbulence model [20]. The nonmodal stability analysis of this system also allows us to investigate

the usefulness of the assumption that the turbulent viscosity is "frozen" at a constant value.

After deriving the perturbation equations for this system, we perform the stability analysis in three different cases: the laminar case ("LAM") where the turbulent viscosity was zero, the frozen turbulent viscosity case ("FROZ") where the turbulent viscosity was set at a constant value, and the fully linearized analysis case ("FULL") where a turbulent viscosity perturbation was considered. The results obtained in the first two cases show that the only effect of the frozen turbulent viscosity is to add spatially varying damping in the system and thus a straightforward decrease of the energy gain. For the full linearized problem, the analysis of the total gain optimal (i.e., optimizing the 2-norm of the state vector) using SVD analysis establishes the presence of a new type of perturbation driven largely by turbulent viscosity perturbation effects, and associated with substantially larger gains than was found in the two simpler cases. The suboptimal perturbation identified by the SVD analysis is mainly driven by perturbation velocity and is very close in terms of perturbation structure and gain to the optimal perturbation identified for the frozen turbulent viscosity case.

The use of the seminorm gain optimization framework developed in this paper allows us to investigate the dependence of the different types of perturbations possible on the optimization method. Indeed, instead of obtaining the relative contribution of turbulent viscosity and mean flow in the perturbation vector as a result of the optimization of the most obvious 2-norm of the initial perturbations, we can, thanks to the new framework, consider this as an input of the optimization problem and then investigate it in a much deeper way.

In the limit of very low turbulent viscosity perturbation to mean flow perturbation ratio (i.e., $C_0 = K_0/E_0 \ll 1$), the results are (in order of magnitude) similar to the frozen turbulent viscosity case, though the fully linearized model never converges toward the frozen turbulent viscosity case's behavior, even as this ratio tends to zero, clearly illustrating the singularity of the problem. On the other hand, in the limit of very "turbulent" perturbations (i.e., when $C_0 \gg 1$), we show that a second type of behavior arises, leading to a large transfer of energy from the base mean flow to the perturbation velocity, catalyzed by substantial variation in the turbulent viscosity. This transfer of energy is possible because the turbulent viscosity perturbation adopts an optimal shape in order to extract energy from the mean flow. More precisely, we show that the perturbation energy gain evolves linearly with the ratio between the two components of the state vector which seems to be a universal behavior.

These results show that the seminorm framework is an interesting way to retrieve the physics given by the modal decomposition of a SVD analysis, through physical considerations, rather than through (at least potentially artificial) mathematical arrangements.

With a classical SVD approach, it is possible to make a decomposition of the linear mapping (in the time domain, or via a standard transfer function in the frequency domain) from any component of the state vector to any other, assuming that the initial perturbation (or the forcing, in the case of a transfer function) is only composed of a single component of

the state vector [36]. This decomposition is accomplished by performing an SVD analysis of the individual block matrix (of the full evolution operator) which connects the chosen input component to the chosen output component. In the above-presented seminorm framework, this component-wise analysis can be reproduced formally by choosing the ratio of relevant seminorms to be zero for the energy gain optimization problem (i.e., energy transport between the same state vector components over some finite optimization time interval). However, an equivalent analysis for the energy transport across state vector components would require a rearrangement of both the cost functional and the ratio of the relevant seminorms. In contrast, the above-presented seminorm variational framework allows us to consider all intermediate configurations of the initial state vector components that fall between the extremes given by the component-wise SVD analysis. This feature presents an opportunity to explore composite initial state vectors and assess their role in observed physical processes. To the authors' knowledge, such a compound state vector optimization problem cannot be handled within an SVD-based framework.

Besides, a sensitivity analysis of the system shows that the influence of modeling parameters (in particular the production and destruction of turbulent viscosity) is smaller than the influence of physical parameters (the viscosity). Moreover, for turbulent perturbations, we show that the mean and perturbation sensitivities with respect to modeling parameters have a different sign, meaning they are competing.

The main conclusion of our investigation of this model problem is that the frozen turbulent viscosity assumption might be relevant if the perturbation in turbulent viscosity is very low compared to the magnitude of the mean flow perturbation, although there still appears to be nontrivial quantitative differences. In the other limit, when the perturbation turbulent viscosity has significant initial magnitude, we clearly conclude that frozen turbulent viscosity is unable to describe the real dynamics of a perturbation governed by the full perturbation RAB equations (49). Therefore, we believe frozen turbulent viscosity is highly unlikely to describe correctly more complicated systems such as the Reynolds-averaged Navier-Stokes (RANS) equations since our one-dimensional model (47) possesses many of the key features of the RANS equations: i.e., time dependence, advection effects, a dissipative nature, and a closed nonlinearity. It is important to stress that the variational framework employed here also gives as an output the sensitivity with respect to the constraints of the problem, and thus offers a powerful analysis tool. Finally, we wish to reiterate that the problem chosen here (optimal perturbation gain defined in terms of a seminorm of the state vector) is one of the simplest we could have imagined and was chosen in order to present this method in a (hopefully) pedagogical way. However, nonlinearity, time averaging norms, nonautonomous operators, and more can be added to the framework with minor impact on the algorithmic approach.

APPENDIX A: LINEARIZED RAB OPERATORS

1. Direct operator

The direct linearized Reynolds-averaged Burgers operator defined in Sec. III in Eq. (47) is defined as follows:

$$\tilde{\mathbf{L}} = \begin{pmatrix} \tilde{L}_{11} & \tilde{L}_{12} \\ \tilde{L}_{21} & \tilde{L}_{22} \end{pmatrix}, \tag{A1}$$

with the following corresponding block matrices:

$$
\begin{aligned}
\tilde{L}_{11} &= -\overline{u}\partial_x - \partial_x\overline{u} + (\nu + \nu_t)\partial_{xx} + \partial_x\nu_t\partial_x, \\
\tilde{L}_{12} &= \partial_x\overline{u}\partial_x + \partial_{xx}\overline{u}, \\
\tilde{L}_{21} &= c_1\mathrm{sgn}(\partial_x\overline{u})\nu_t\partial_x - \partial_x\nu_t, \\
\tilde{L}_{22} &= -\overline{u}\partial_x + \partial_{xx}\nu_t + 2\partial_x\nu_t\partial_x + (\nu + \nu_t)\partial_{xx} \\
&\quad + c_1\,|\partial_x\overline{u}| - 2c_2\nu_t.
\end{aligned} \tag{A2}
$$

We notice that the matrix $\tilde{L}_{11}$ corresponds to the frozen turbulent viscosity equation.

2. Adjoint operator

The linear adjoint Reynolds-averaged Burgers operator defined in Sec. III in Eq. (68) is defined as follows:

$$\tilde{\mathbf{L}}^{\dagger} = \begin{pmatrix} \tilde{L}_{11}^{\dagger} & \tilde{L}_{12}^{\dagger} \\ \tilde{L}_{21}^{\dagger} & \tilde{L}_{22}^{\dagger} \end{pmatrix}, \tag{A3}$$

with the following corresponding block matrices:

$$
\begin{aligned}
\tilde{L}_{11}^{\dagger} &= \overline{u}\partial_x + (\nu + \nu_t)\partial_{xx} + \partial_x\nu_t\partial_x, \\
\tilde{L}_{12}^{\dagger} &= -\partial_x\nu_t - c_1\mathrm{sgn}(\partial_x\overline{u})(\nu_t\partial_x + \partial_x\nu_t), \\
\tilde{L}_{21}^{\dagger} &= -\partial_x\overline{u}\partial_x, \\
\tilde{L}_{22}^{\dagger} &= \overline{u}\partial_x + \partial_x\overline{u} + (\nu + \nu_t)\partial_{xx} + c_1\,|\partial_x\overline{u}| - 2c_2\nu_t.
\end{aligned} \tag{A4}
$$

APPENDIX B: EXPRESSION OF SENSITIVITY FUNCTIONS AND MATRICES

1. Base flow sensitivity functions

The sensitivity functions defined in (75), describing the sensitivity of the objective functional $\mathcal{J}$ defined in Eq. (56) to a change in the base mean flow $\overline{\mathbf{q}} = (\overline{u}, \overline{\nu}_t)$, are defined as

$$
\begin{aligned}
\mathbf{S}_{\overline{u}}(\tilde{\mathbf{q}}) &= \tilde{u}\partial_x\tilde{u}^{\dagger} + \tilde{v}\partial_{xx}\tilde{u}^{\dagger} + \partial_x\tilde{v}\partial_x\tilde{u}^{\dagger} - \tilde{v}^{\dagger}\partial_x\tilde{v} \\
&\quad + c_1\mathrm{sgn}\,(\partial_x\overline{u})\,(\tilde{v}^{\dagger}\partial_x\tilde{v} + \tilde{v}\partial_x\tilde{v}^{\dagger}),
\end{aligned} \tag{B1}
$$

$$
\begin{aligned}
\mathbf{S}_{\nu_t}(\tilde{\mathbf{q}}) &= -\partial_x\tilde{u}\partial_x\tilde{u}^{\dagger} + \tilde{v}\partial_{xx}\tilde{v}^{\dagger} - 2c_2\tilde{v}^{\dagger}\tilde{v} + \tilde{u}\partial_x\tilde{v}^{\dagger} \\
&\quad + \tilde{v}^{\dagger}\partial_x\tilde{u} - c_1\mathrm{sgn}\,(\partial_x\overline{u})\,\tilde{v}^{\dagger}\partial_x\tilde{u}.
\end{aligned} \tag{B2}
$$

2. Parameter sensitivity vectors and matrices

Base mean flow contribution. The sensitivity vectors defined in (79), describing the change of the objective functional $\mathcal{J}$ defined in Eq. (56) under a change of the parameters in the

base mean flow equations (47), are defined as

$$\overline{\mathbf{S}}_\nu(\overline{\mathbf{q}}) = \begin{pmatrix} \partial_{xx}\overline{u} \\ \partial_{xx}\nu_t \end{pmatrix}, \tag{B3}$$

$$\overline{\mathbf{S}}_{c_1}(\overline{\mathbf{q}}) = \begin{pmatrix} 0 \\ -|\partial_x\overline{u}|\,\nu_t \end{pmatrix}, \tag{B4}$$

$$\overline{\mathbf{S}}_{c_2}(\overline{\mathbf{q}}) = \begin{pmatrix} 0 \\ \nu_t^2 \end{pmatrix}. \tag{B5}$$

Perturbation contribution. The sensitivity vectors defined in (79), describing the change of the objective functional $\mathcal{J}$ defined in Eq. (56) under a change of the parameters in the perturbation equations (49), are defined as

$$\tilde{\mathbf{S}}_\nu = \begin{pmatrix} \partial_{xx} & 0 \\ 0 & \partial_{xx} \end{pmatrix}, \tag{B6}$$

$$\tilde{\mathbf{S}}_{c_1} = \begin{pmatrix} 0 & 0 \\ -\mathrm{sgn}\,(\partial_x\overline{u})\,\nu_t\partial_x & -|\partial_x\overline{u}| \end{pmatrix}, \tag{B7}$$

$$\tilde{\mathbf{S}}_{c_2} = \begin{pmatrix} 0 & 0 \\ 0 & 2\nu_t \end{pmatrix}. \tag{B8}$$

APPENDIX C: OPTIMAL PERTURBATION WITH SINGULAR VALUE DECOMPOSITION (SVD)

Given that the flow is stable for any value of the viscosity (the flow tends toward neutral stability when ν becomes small), we expect to be able to capture the properties of the dynamics by focusing on the transient growth mechanisms involved. For a general linear equation (discretized) of the type $d_t\mathbf{q} = \mathbf{L}q$, we can derive the exact solution at time T, which will from now on be called the horizon time. This solution can simply be expressed in term of the evolution operator $\mathbf{M}$ (which is the matrix exponential of $\mathbf{L}T$) and the initial condition $\mathbf{q}(0) = \mathbf{q}_0$; i.e.,

$$\mathbf{q}(T) = \mathbf{M}\mathbf{q}_0 = e^{T\mathbf{L}}\mathbf{q}_0. \tag{C1}$$

Let us now define the energy of a state vector as the following weighted scalar product:

$$\|\mathbf{q}\|_E^2 = \mathbf{q}^H\mathbf{W}\mathbf{q}. \tag{C2}$$

The maximum gain we can achieve for a time T is simply expressed as an optimization problem:

$$G(T) = \max_{\mathbf{q}_0\neq 0} \frac{\|\mathbf{q}(T)\|_E^2}{\|\mathbf{q}_0\|_E^2} = \max_{\mathbf{q}_0\neq 0} \frac{\|e^{T\mathbf{L}}\mathbf{q}_0\|_E^2}{\|\mathbf{q}_0\|_E^2}. \tag{C3}$$

Now, assuming that $\mathbf{W}$ is a symmetric positive definite matrix, we use a Cholesky decomposition to write

$$\mathbf{W} = \mathbf{F}^H\mathbf{F}. \tag{C4}$$

We then change the variable from $\mathbf{q}$ to $\mathbf{q}'$ defined as

$$\mathbf{q}' = \mathbf{F}\mathbf{q}. \tag{C5}$$

With this small transformation, we are able to write the gain in the following way:

$$G(T) = \max_{\mathbf{q}_0'\neq 0} \frac{\|\mathbf{F}e^{T\mathbf{L}}\mathbf{F}^{-1}\mathbf{q}_0'\|_2}{\|\mathbf{q}_0'\|_2}, \tag{C6}$$

which is simply the 2-norm of the matrix $\mathbf{M}' = \mathbf{F}e^{T\mathbf{L}}\mathbf{F}^{-1}$, which can be computed through a singular value decomposition. The output of the SVD is a diagonal matrix Σ, consisting of all the singular values which are real positive numbers, and two orthogonal matrices $\mathbf{U}'$ and $\mathbf{V}'$:

$$\mathbf{M}' = \mathbf{U}'\Sigma\mathbf{V}'^{\mathbf{H}}. \tag{C7}$$

The norm of the matrix (and as a consequence the gain) is given by the largest singular value (i.e., the first coefficient in Σ), while the first column of $\mathbf{V} = \mathbf{F}^{-1}\mathbf{V}'$ gives the optimal perturbation $\mathbf{q}_0$ and the first column of $\mathbf{U} = \mathbf{F}^{-1}\mathbf{U}'$ gives the optimal state at the horizon time T. With this method, we are able to find the optimal gain for different horizon times, as well as the associated optimal perturbation.

[1] O. Reynolds, Philos. Trans. R. Soc. London **174**, 935 (1883).
[2] S. Orszag, J. Fluid Mech. **50**, 689 (1971).
[3] V. Romanov, Functional Analysis and its Applications **7**, 137 (1973).
[4] S. J. Davies and C. M. White, Proc. R. Soc. London A **119**, 92 (1928).
[5] N. Tillmark and P. Alfredsson, J. Fluid Mech. **235**, 89 (1992).
[6] L. Trefethen, A. Trefethen, S. Reddy, and T. Driscoll, Science **261**, 578 (1993).
[7] L. Gustavsson, J. Fluid Mech. **224**, 241 (1991).
[8] K. Butler and B. Farrell, Phys. Fluids A **4**, 1637 (1992).
[9] S. Reddy and D. Henningson, J. Fluid Mech. **252**, 209 (1993).
[10] P. Schmid, Annu. Rev. Fluid Mech. **39**, 129 (2007).
[11] D. Hill, J. Fluid Mech. **292**, 183 (1995).
[12] M. Gunzburger, in *Inverse Design and Optimisation Methods*, Lecture Series 1997-05 (von Karman Institute for Fluid Dynamics, Rhode Saint Genese, Belgium, 1997).
[13] C. C. T. Pringle and R. R. Kerswell, Phys. Rev. Lett. **105**, 154502 (2010).
[14] O. Marquet, D. Sipp, and L. Jacquin, J. Fluid Mech. **615**, 221 (2008).
[15] A. Guégan, P. Schmid, and P. Huerre, J. Fluid Mech. **603**, 179 (2008).
[16] J. Crouch, A. Garbaruk, D. Magidov, and A. Travin, J. Fluid Mech. **628**, 357 (2009).
[17] C. Cossu, G. Pujals, and G. Depardon, J. Fluid Mech. **619**, 79 (2009).
[18] Y. Kuramoto, Prog. Theor. Phys. Suppl. **64**, 346 (1978).
[19] S. Zaleski, Physica D: Nonlinear Phenomena **34**, 427 (2002).
[20] P. Spalart and S. Allmaras, "A One-Equation Turbulence Model for Aerodynamic Flows," in Proceedings of the 30th AIAA Aerospace Sciences Meeting and Exhibit, 1992.
[21] L. Brandt, D. Sipp, J. Pralits, and O. Marquet, J. Fluid Mech. **687**, 503 (2011).
[22] F.-X. Le Dimet and O. Talagrand, Tellus, Ser. A **38**, 97 (1986).
[23] A. Hanifi, P. Schmid, and D. Henningson, Phys. Fluids **8**, 826 (1996).
[24] A. Castellanos and N. Agrait, IEEE Transaction on Industry Applications **28**, 513 (1992).

[25] F. Martinelli, M. Quadrio, and P. Schmid, "Stability of Planar Shear Flow in Presence of Electroconvection," in VII Int. Symp. on Turbulence and Shear Flow Phenomena, 28–31 July, 2011.

[26] J. Li, Y. Chen, and Z. Li, Phys. Plasmas **13**, 042101 (2006).

[27] A. Monokrousos, A. Bottaro, L. Brandt, A. Di Vita, and D. S. Henningson, Phys. Rev. Lett. **106**, 134502 (2011).

[28] T. Bewley, Prog. Aerospace Sci. **37**, 21 (2001).

[29] J. Boussinesq, *Essai sur la Théorie des Eaux Courantes* (Imprimerie Nationale, Paris, 1877).

[30] M. Juniper, J. Fluid Mech. **667**, 272 (2010).

[31] S. Nagaraja, K. Kedia, and R. Sujith, Proceedings of the Combustion Institute **32**, 2933 (2009).

[32] F. Selimefendigil, R. Sujith, and W. Polifke, Applied Mathematics and Computation **217**, 5134 (2011).

[33] J. J. S. Jerome, J.-M. Chomaz, and P. Huerre, Phys. Fluids **24**, 044103 (2012).

[34] G. Ivey, K. Winters, and J. Koseff, Annu. Rev. Fluid Mech. **40**, 169 (2008).

[35] W. Tang, C. P. Caulfield, and R. Kerswell, J. Fluid Mech. **634**, 487 (2009).

[36] M. R. Jovanovic and B. Bamieh, J. Fluid Mech. **534**, 145 (2005).