Unraveling self-similar energy transfer dynamics: A case study for the one-dimensional Burgers system

Pritpal Matharu [*,†]
Max Planck Institute for Mathematics in the Sciences, Inselstraße 22, Leipzig 04103, Germany and Department of Mathematics, KTH Royal Institute of Technology, Lindstedtsvägen 25, 114 28 Stockholm, Sweden

Bartosz Protas [‡,§]
Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario L8S 4K1, Canada

Tsuyoshi Yoneda [‖,¶]
Graduate School of Economics, Hitotsubashi University, Tokyo 186-8601, Japan

(Received 18 July 2025; accepted 13 February 2026; published 23 March 2026)

In this work we consider the problem of constructing initial conditions for a flow model such that the resulting flow evolution leads to a self-similar energy cascade consistent with Kolmogorov's statistical theory of turbulence. As a first step in this direction, we focus on the one-dimensional viscous Burgers equation as a toy model. Its solutions exhibiting self-similar behavior, in a precisely defined sense, are found by framing this problem in terms of partial differential equation (PDE)-constrained optimization. The main physical parameters are the time window over which self-similar behavior is sought (equal to approximately one eddy turnover time), viscosity (inversely proportional to the "Reynolds number"), and an integer parameter characterizing the distance in the Fourier space over which self-similar interactions occur. Local solutions to this nonconvex PDE optimization problem are obtained with a state-of-the-art adjoint-based gradient method. Two distinct families of solutions, termed *viscous* and *inertial*, are identified and are distinguished primarily by the behavior of enstrophy which, respectively, uniformly decays and grows in the two cases. The physically meaningful and appropriately self-similar inertial solutions are found only when a sufficiently small viscosity is considered. These flows achieve the self-similar behavior by a uniform steepening of the wave fronts present in the solutions. The results obtained demonstrate that the proposed methodology may be used to search for self-similar behavior in more complex flow models, including shell models, two-dimensional turbulence, and, ultimately, three-dimensional turbulence.

DOI: 10.1103/mxk3-jx9f

[*]Contact author: matharu@mis.mpg.de
[†]https://pipmath.github.io/
[‡]Contact author: bprotas@mcmaster.ca
[§]https://ms.mcmaster.ca/bprotas/
[‖]Contact author: t.yoneda@r.hit-u.ac.jp
[¶]https://www1.econ.hit-u.ac.jp/yoneda/

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Open access publication funded by Max Planck Society.

I. INTRODUCTION

Self-similarity of the energy cascade is one of the most prominent statistical properties of hydrodynamic turbulence [1]. The energy cascade between a hierarchy of spatial and temporal scales was described first by Richardson [2], who provided a statistical account of this cascade in turbulent flows. This idea further leads to the concept of self-similar flow structures predicted by Kolmogorov [3] who proposed a similarity hypothesis to explain the $-5/3$ power law describing the energy cascade. However, this approach is based on dimensional analysis of quantities defined in statistical terms and therefore is not directly related to the mathematical structure of the solutions of the Navier-Stokes system. In other words, being statistical in nature, this theory does not offer any insights into what kind of fluid motions can generate an energy cascade with the spectrum characterized by the $-5/3$ power law.

In 2017, Goto, Saito, and Kawahara [4] used direct numerical simulation (DNS) to reveal that the three-dimensional (3D) Navier-Stokes turbulence in a periodic box has a hierarchical structure consisting of vortex tubes with a range of characteristic length scales. At each of these scales, the vortex tubes generate a strain field that contributes to the generation of smaller-scale vortex tubes. Based on that result, Yoneda, Goto, and Tsurhashi [5] reformulated Kolmogorov's $-5/3$ power law by defining an energy-transfer function in terms of vorticity [Ref. 5, Eq. (2.6)] and then using it to identify a clear local-in-scale energy transfer structure from instantaneous spatially averaged turbulence snapshots [Ref. 5, Fig. 1], [6]. However, that approach was based on spatial averaging and therefore it was not possible to identify the elementary processes, in the form of specific flow evolutions in time, that give rise to a self-similar energy cascade. The purpose of this study is to address this question in the context of a simple toy model of turbulent flows, namely, the one-dimensional (1D) Burgers equation. A key novelty of our approach is that solutions with a certain self-similar structure are sought systematically by solving suitably defined optimization problems for a partial differential equation (PDE). By focusing on a simplified model problem we aim to develop and test an approach that can be generalized to search for analogous self-similar flow phenomena in the context of more complex flow models, eventually including 3D turbulence.

While PDE-constrained optimization has had a long history in fluid mechanics [7], most of these studies have focused on solutions of practical problems often motivated by engineering applications [8–11]. It was only recently that such techniques were employed to investigate fundamental questions concerning the structure and properties of solutions to hydrodynamic equations in a way that complements rigorous mathematical analysis [12]. Problems tackled in this way include probing sharpness of a priori bounds on various energy-type quantities [13,14], nonlinear stability [15], dissipation anomaly [16], search for finite-time singularities [17–20], and optimization of mixing [21,22]. The present study builds on this line of research.

The main goal of this work is to find and analyze a family of freely decaying (unforced) 1D Burgers flows featuring a self-similar energy cascade and parametrized by the value of the viscosity coefficient (or, equivalently, the Reynolds number) and a positive integer number characterizing the nonlocality of interactions in the Fourier space. The initial conditions for these flows are found as solutions of PDE-constrained optimization problems and are obtained numerically using a state-of-the-art adjoint-based optimization method. Our main finding is the evidence that families of such Burgers flows do indeed exist and that their self-similar behavior is realized by a uniform steepening of the fronts in the solutions. We also provide an analysis of how quantitative details of this process depend on physical parameters.

The structure of the paper is as follows: in the next section we introduce the model and state the optimization problem; in Sec. III we describe the solution approach whereas the results are presented in Sec. IV; discussion and final conclusions are deferred to Sec. V while some technical material is collected in the Appendix.

II. GOVERNING EQUATION AND OPTIMIZATION FORMULATION

In this section we first introduce the 1D Burgers system and briefly discuss some properties of its solutions. It will be considered on the periodic domain $\Omega = \mathbb{T} := [0, 2\pi]$ and over the time window $[0, T]$, where the choice of the end time T will be discussed below. Then, we formulate a PDE optimization problem designed to find solutions exhibiting a self-similar evolution.

A. One-dimensional Burgers equation

The Burgers system in 1D has the following form subject to the periodic boundary conditions

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} - \nu\frac{\partial^2 u}{\partial x^2} = 0 \quad \text{in } (0, T] \times \Omega, \tag{1a}$$

$$u(t = 0) = \phi \quad \text{in } \Omega, \tag{1b}$$

where $\nu > 0$ is the viscosity and ϕ is the initial condition. It is known to admit globally defined smooth solutions for a broad class of initial data ϕ [23]. There are three physical parameters in the problem, namely, the size $|\Omega|$ of the domain, the "size" U of the initial data ϕ and the viscosity coefficient ν, which can be combined to form the dimensionless Reynolds number $\mathrm{Re} = U|\Omega|/\nu$. In our study we will assume the first two parameters are fixed and will vary the viscosity coefficient. Anticipating the formulation of the optimization problem below, we will assume here that $\phi \in H^1(\Omega)$, where $H^1(\Omega)$ is the Sobolev space of functions with square-integrable weak derivatives endowed with the inner product [24]

$$\langle p_1, p_2\rangle_{H^1(\Omega)} := \int_\Omega p_1 p_2 + \ell^2 \left(\frac{\partial p_1}{\partial x}\,\frac{\partial p_2}{\partial x}\right) dx, \quad \forall\, p_1, p_2 \in H^1(\Omega), \tag{2}$$

where ":=" means "equal to by definition" and $0 < \ell < \infty$ is an adjustable parameter (its significance will become evident in Sec. III).

Various stochastic variants of system Eqs. (1) have been used as models of turbulence [25], typically in the forced setting realized by including a source term on the right-hand side in Eq. (1a). On a phenomenological level, equation (1a) describes the competition between a nonlinear amplification in the form of wave steepening, which represents the forward energy cascade with energy transferred from large to small scales, and viscous dissipation [26]. The wave steepening is characterized by the evolution of the enstrophy

$$\mathcal{E}(u(t,\cdot)) := \int_0^{2\pi} \left|\frac{\partial u(t,x)}{\partial x}\right|^2 dx = \|u(t,\cdot)\|^2_{\dot{H}^1}, \tag{3}$$

which in the unforced case is well understood [13,27,28]. In particular, sharp a priori bounds on the maximum growth of the enstrophy are now available [29]. The triadic interactions responsible for the amplification of the enstrophy in 1D Burgers flows were analyzed in Refs. [30–32].

B. Optimization problem

Before we can state the optimization problem, we need to define a suitable notion of "self-similar evolution." While there is more than one way in which this property can be defined, we focus here on a formulation suitable for the freely decaying problem we consider in which there is no statistical equilibrium. Let $\widehat{u}(t, k)$, $k \in \mathbb{N}$, denote the Fourier coefficient of the solution $u(t, x)$ at some time $t \geqslant 0$. Given a positive integer $\lambda \in \mathbb{N}^+$, we have the following:

Definition 1. A solution $u = u(t, x)$ of Eq. (1) is said to be self-similar with an index λ in terms of spectral energy transfer on $[t, t + T]$ if it satisfies the relation

$$|\widehat{u}(t, k)|^2 = \lambda\, |\widehat{u}(t + T, \lambda k)|^2, \quad k \in \mathbb{N}, \quad t > 0. \tag{4}$$

For brevity, we will refer to solutions satisfying this definition as simply "self-similar." Intuitively, relation (4) can be interpreted to mean that during a self-similar evolution over the time T a fraction $(1/\lambda) \in (0, 1)$ of the kinetic energy initially contained in the large-scale motions is transferred between the Fourier modes with wave numbers k and λk. While existence of such solutions in a mathematically precise sense remains an open question, we will attempt here to construct them by finding initial conditions ϕ that minimize the functional $\mathcal{J}_{\nu,\lambda} : H^1(\Omega) \to \mathbb{R}$ defined as

$$\mathcal{J}_{\nu,\lambda}(\phi) := \frac{1}{2} \sum_{k=0}^{\infty} \underbrace{|\widehat{\phi}(k)|^2 - \lambda|\widehat{u}(T, \lambda k; \phi)|^2|}_{f(k;\phi)}^2 \tag{5}$$

in which summation is restricted to positive wave numbers k only due to the conjugate symmetry $\widehat{u}(\cdot, k) = \overline{\widehat{u}(\cdot, -k)}$, where the overbar denotes complex conjugation, characterizing the Fourier coefficients of real-valued functions. Notions of self-similarity other than given by Definition 1 may also be considered, especially in the context of forced flows where some form of statistical equilibrium holds. In such situations self-similarity can be defined by requiring, e.g., that energy spectra at different times collapse or that the energy flux exhibits some form of scale invariance.

Clearly, in the absence of any constraints, the global minimum of $\mathcal{J}_{\nu,\lambda}(\phi)$ is achieved by $\phi_0 \equiv 0$, since then $u(t; \phi_0) \equiv 0$, $t > 0$, and $\mathcal{J}_{\nu,\lambda}(\phi_0) = 0$. To avoid such uninteresting trivial solutions, we need to constrain the initial condition to have a nonvanishing magnitude and we choose to do this by prescribing its enstrophy as $\mathcal{E}(\phi) = \mathcal{E}_0 > 0$. We thus have the following optimization problem:

Problem 1. For fixed T, $\mathcal{E}_0$, $\nu \in \mathbb{R}^+$ and $\lambda \in \mathbb{N}^+$ in Eq. (5), find

$$\varphi_{\nu,\lambda} = \arg\min_{\phi \in \mathcal{S}} \mathcal{J}_{\nu,\lambda}(\phi), \quad \text{where} \quad \mathcal{S} := \left\{ \phi \in H^1(\Omega) \ : \ \int_0^{2\pi} \phi(x)\, dx = 0, \quad \mathcal{E}(\phi) = \mathcal{E}_0 \right\}.$$

In the above $\mathcal{S} \subset H^1(\Omega)$ represents the constraint manifold which also includes the zero-mean condition [which is a property preserved in the evolution governed by Eq. (1a)]. The time window T will be taken on the order of an eddy turnover time t_e defined as

$$t_e := \sqrt{\frac{\nu}{\varepsilon}}, \tag{6}$$

where ε is the space-time averaged energy dissipation $\varepsilon := \frac{2\nu}{T|\Omega|} \int_0^T \mathcal{E}(u(t, \cdot))\, dt$ which is a well-defined quantity in freely decaying Burgers flows. We add that the constraints defining the manifold $\mathcal{S}$ in Problem 1 are independent of the problem parameters and also remain well-defined when $\nu \to 0$.

III. SOLUTION APPROACH

While methods of adjoint analysis are now well known [7], we introduce our approach in some detail below given the somewhat nonstandard structure of the objective functional which is defined in the Fourier space (5). Problem 1 is Riemanian in the sense that its solutions are constrained to a Riemannian manifold $\mathcal{S} \subset H^1(\Omega)$ [33]. It is also nonconvex and its local minimizers are approximated as $\varphi_{\nu,\lambda} = \lim_{n\to\infty} \phi^{(n)}$ using a retracted discrete gradient flow

$$\begin{aligned} \phi^{(n+1)} &= \mathcal{R}_{\mathcal{S}}(\phi^{(n)} + \tau_n \nabla \mathcal{J}_{\nu,\lambda}(\phi^{(n)})), \quad n = 1, 2, \ldots, \\ \phi^{(1)} &= \phi_1, \end{aligned} \tag{7}$$

where $\phi^{(n)}$ is an approximation of the minimizer obtained at the nth iteration, $\phi_1 \in \mathcal{S}$ is the initial guess, $\mathcal{R}_{\mathcal{S}} : H^1(\Omega) \to \mathcal{S}$ is the retraction operator defined below and τ_n is the length of the step in the direction of the gradient $\nabla \mathcal{J}_{\nu,\lambda}(\phi^{(n)})$. To exploit the mathematical structure of Problem 1, we will

adopt the "optimize-then-discretize" approach [7] where expressions for the different elements in Eq. (7) are derived in the continuous setting and discretized for the purpose of numerical evaluation at the final stage only; these steps are documented in Sec. III A below.

To obtain an expression for the gradient $\nabla\mathcal{J}_{\nu,\lambda}(\phi)$ of the objective functional $\mathcal{J}_{\nu,\lambda}(\phi)$, we begin by defining its Gâteaux (directional) differential $\mathcal{J}_{\nu,\lambda}{}'(\phi;\phi') := \lim_{\epsilon\to 0}\epsilon^{-1}[\mathcal{J}_{\nu,\lambda}(\phi+\epsilon\phi') - \mathcal{J}_{\nu,\lambda}(\phi)]$ for some arbitrary perturbation $\phi' \in H^1(\Omega)$ of the initial condition. Given the definition of the functional in Eq. (5), the Gâteaux differential is

$$
\begin{aligned}
\mathcal{J}_{\nu,\lambda}{}'(\phi;\phi') &= \sum_{k=0}^{\infty}[|\widehat{\phi}(k)|^2 - \lambda\,|\widehat{u}(T,\lambda k;\phi)|^2]\,[\widehat{\phi}'(k)\overline{\widehat{\phi}}(k) + \widehat{\phi}(k)\overline{\widehat{\phi}}'(k)] \\
&\quad - \sum_{k=0}^{\infty}\lambda[|\widehat{\phi}(k)|^2 - \lambda\,|\widehat{u}(T,\lambda k;\phi)|^2]\,[\widehat{u}'(T,\lambda k;\phi,\phi')\overline{\widehat{u}}(T,\lambda k;\phi)] \\
&\quad - \sum_{k=0}^{\infty}\lambda[|\widehat{\phi}(k)|^2 - \lambda\,|\widehat{u}(T,\lambda k;\phi)|^2]\,[\widehat{u}(T,\lambda k;\phi)\overline{\widehat{u}}'(T,\lambda k;\phi,\phi')], \\
&= 2\sum_{k=0}^{\infty}[|\widehat{\phi}(k)|^2 - \lambda\,|\widehat{u}(T,\lambda k;\phi)|^2]\widehat{\phi}(k)\,\overline{\widehat{\phi}}'(k) \\
&\quad \underbrace{- 2\sum_{k=0}^{\infty}\lambda[|\widehat{\phi}(k)|^2 - \lambda\,|\widehat{u}(T,\lambda k;\phi)|^2]\widehat{u}(t,\lambda k;\phi)\,\overline{\widehat{u}}'(t,\lambda k;\phi,\phi')}_{\mathcal{A}(\phi;\phi')},
\end{aligned}
\tag{8}
$$

where we used the identity $\sum_{k=0}^{\infty}\widehat{v}_1(k)\overline{\widehat{v}}_2(k) = \sum_{k=0}^{\infty}\overline{\widehat{v}}_1(k)\widehat{v}_2(k)$ valid for real-valued functions $v_1(x)$, $v_2(x)$, and $u'(t,x;\phi,\phi')$ satisfies the perturbation system

$$
\frac{\partial u'}{\partial t} + u'\frac{\partial u}{\partial x} + u\frac{\partial u'}{\partial x} - \nu\frac{\partial^2 u'}{\partial x^2} = 0 \quad \text{in } (0,T]\times\Omega, \tag{9a}
$$

$$
u'(t=0) = \phi' \quad \text{in } \Omega, \tag{9b}
$$

obtained by linearizing system Eqs. (1) around the solution $u(\cdot,\cdot;\phi)$. Noting that when viewed as a function of its second argument (ϕ') and for a fixed ϕ, the Gâteaux differential (8) is a bounded linear functional on both $H^1(\Omega)$ and $L^2(\Omega)$, and we can invoke the Riesz representation theorem [34],

$$
\mathcal{J}_{\nu,\lambda}{}'(\phi;\phi') = \left\langle \nabla^{H^1}\mathcal{J}_{\nu,\lambda},\phi'\right\rangle_{H^1(\Omega)} = \left\langle \nabla^{L^2}\mathcal{J}_{\nu,\lambda},\phi'\right\rangle_{L^2(\Omega)}, \tag{10}
$$

where the inner product $\langle\cdot,\cdot\rangle_{L^2(\Omega)}$ is obtained from Eq. (2) by setting $\ell = 0$. It is convenient to first obtain an expression for the L^2 gradient $\nabla^{L^2}\mathcal{J}_{\nu,\lambda}$ and then use identity (10) together with (2) to deduce the required Sobolev gradient $\nabla\mathcal{J}_{\nu,\lambda} = \nabla^{H^1}\mathcal{J}_{\nu,\lambda}$ [35]. We note that the last term in Eq. (8) for the Gâteaux differential, $\mathcal{A}(\phi;\phi')$, is not consistent with the Riesz representation (10) since the arbitrary perturbation ϕ' does not appear in it explicitly as a linear factor and is instead "hidden" as the initial condition in the perturbation system Eqs. (9). To transform this term to the required form, we use the adjoint calculus. Multiplying (9a) by the adjoint state $u^* : [0,T]\times\Omega\to\mathbb{R}$, integrating the resulting expression over $[0,T]\times\Omega$ and performing integration by parts with respect to t and

x, we obtain

$$0 = \int_0^T \int_0^{2\pi} \left(\frac{\partial u'}{\partial t} + u' \frac{\partial u}{\partial x} + u \frac{\partial u'}{\partial x} - \nu \frac{\partial^2 u'}{\partial x^2} \right) u^* \, dx \, dt$$
$$= \int_0^{2\pi} \left[u^*(T,x)u'(T,x) - u^*(0,x)u'(0,x) + \int_0^T \left(-\frac{\partial u^*}{\partial t} - u \frac{\partial u^*}{\partial x} - \nu \frac{\partial^2 u^*}{\partial x^2} \right) u' \, dt \right] dx, \quad (11)$$

where all the boundary terms resulting from integration by parts in space vanish due to the periodic boundary conditions. Defining the adjoint system as

$$-\frac{\partial u^*}{\partial t} - u \frac{\partial u^*}{\partial x} - \nu \frac{\partial^2 u^*}{\partial x^2} = 0 \quad \text{in } (0,T] \times \Omega, \quad (12a)$$

$$u^*(t = T) = W(\phi) \quad \text{in } \Omega, \quad (12b)$$

with the terminal condition stated in Fourier space as

$$\widehat{W}(k;\phi) = \begin{cases} -\frac{2\lambda}{\pi} [|\widehat{\phi}(k/\lambda)|^2 - \lambda \, |\widehat{u}(T,k;\phi)|^2] \widehat{u}(T,k;\phi), & k = \lambda n, \quad n \in \mathbb{N} \\ 0, & \text{otherwise} \end{cases} \quad (13)$$

reduces Eq. (11) to

$$\mathcal{A}(\phi;\phi') = \int_0^{2\pi} u^*(T,x)u'(T,x) \, dx = \int_0^{2\pi} u^*(0,x)\phi'(x) \, dx = \langle u^*(0,\cdot), \phi' \rangle_{L^2(\Omega)}, \quad (14)$$

where Parseval's theorem was used to express integrals of products of functions in terms of sums of products of their Fourier coefficients. The adjoint system Eqs. (12) is a terminal-value problem with the terminal condition (13) determined by the form of the Gâteaux differential (8), cf. (14), and hence also by the objective functional (5). Therefore, the system needs to be integrated backwards in time and its coefficients $u(t,x;\phi)$ are given by the solution of the governing system Eqs. (1) around which linearization is performed.

Combining Eqs. (8) and (14), we obtain the following expression for the L^2 gradient of the functional (5) given in terms of its Fourier-space representation as

$$\widehat{\nabla^{L^2} \mathcal{J}_{\nu,\lambda}}(k) = \frac{2}{\pi} [|\widehat{\phi}(k)|^2 - \lambda \, |\widehat{u}(T, \lambda k; \phi)|^2] \widehat{\phi}(k) + \widehat{u^*}(0,k), \quad k \in \mathbb{N}. \quad (15)$$

Since the L^2 gradient does not possess the regularity needed to construct solutions of Problem 1, we must determine the corresponding Sobolev H^1 gradient. Using the second equality in the Riesz identity (10) together with the definition (2) of the H^1 inner product, performing integration by parts and noting the arbitrariness of ϕ', we obtain the required H^1 gradient $\nabla \mathcal{J}_{\nu,\lambda}$ as a solution of the elliptic boundary-value problem

$$\left[\mathrm{Id} - \ell^2 \frac{\partial^2}{\partial x^2} \right] \nabla \mathcal{J}_{\nu,\lambda} = \nabla^{L^2} \mathcal{J}_{\nu,\lambda} \quad \text{in } \Omega, \quad (16)$$

subject to the periodic boundary conditions. As shown in Ref. [35], extraction of gradients in spaces of smoother functions such as $H^1(\Omega)$ can be interpreted as low-pass filtering of the L^2 gradients with the parameter ℓ acting as the cutoff length-scale.

The retraction operator in Eq. (7) is defined in terms of the normalization

$$\mathcal{R}_{\mathcal{S}}(\phi) := \sqrt{\frac{\mathcal{E}_0}{\mathcal{E}(\phi)}} \, \phi. \quad (17)$$

An optimal step size τ_n can be determined by solving the minimization problem

$$\tau_n = \arg\min_{\tau > 0} \{ \mathcal{J}_{\nu,\lambda} (\mathcal{R}_{\mathcal{S}} (\phi^{(n)} + \tau \, \nabla \mathcal{J}_{\nu,\lambda} (\phi^{(n)}))) \}, \quad (18)$$

which can be interpreted as a modification of a standard line search problem [36] with optimization performed following an arc (a geodesic) lying on the constraint manifold $\mathcal{S}$, rather than a straight line. A numerical implementation of the approach described here is presented below.

A. Numerical implementation

The governing system and the corresponding adjoint system, respectively, Eqs. (1) and (12), are discretized in space using a Fourier pseudo-spectral method. Nonlinear products are evaluated in physical space and appropriately dealiased using the 2/3 rule combined with a Gaussian filter introduced in Ref. [37]. Time integration is performed using a four-step, implicit/explicit Runge-Kutta scheme which yields globally third-order accuracy [38]. Since the number of spatial discretization points N and temporal step-size Δt heavily depend on the selected values of ν and λ, we defer the discussion of selection of these parameters to the next section. Nonetheless, validation of the gradient Eq. (15), which is a key element of the optimization approach and requires the numerical solution of both system Eqs. (1) and (12), is given in the Appendix.

The Sobolev gradient is determined using $\ell = 10$ in Eq. (2) and the corresponding elliptic boundary-value problem Eq. (16) is solved using the Fourier-Galerkin spectral method. The line minimization problem Eq. (18) is solved using Brent's algorithm [39]. Iterations (7) were declared converged when a stopping criterion given in terms of the relative decrease of the cost functional, $|\mathcal{J}_{\nu,\lambda}(\phi^{(n)}) - \mathcal{J}_{\nu,\lambda}(\phi^{(n-1)})|/\mathcal{J}_{\nu,\lambda}(\phi^{(n-1)}) < 1 \times 10^{-8}$, was satisfied or the number of iteration exceeded the limit of $n = 1000$. To accelerate convergence of the iterations in Eq. (7), the gradient $\nabla\mathcal{J}_{\nu,\lambda}(\phi^{(n)})$ in Eq. (7) was replaced with a conjugate gradient constructed using the Polak-Ribière formula [36]. Our approach is implemented in MATLAB and the code is available in Ref. [40].

IV. RESULTS

In this section we present results obtained by solving Problem 1 with $\mathcal{E}_0 = \pi$ and with λ and ν varying over a broad range of values, namely, $\lambda = 2, \ldots, 7$ and $\nu = 2.5 \times 10^{-3}, 2.5 \times 10^{-4}, 2.5 \times 10^{-5}, 2.5 \times 10^{-6}, 2.5 \times 10^{-7}$. We recall that, with the initial enstrophy $\mathcal{E}_0$ and the domain Ω both fixed, the characteristic "size" of the initial data can be expressed as $U = \sqrt{|\Omega|\mathcal{E}_0}$ and decreasing the viscosity coefficient ν can be interpreted as increasing the "Reynolds number" of the flow. When λ increases and ν decreases, the numerical resolution both in space and in time needs to be refined for the solutions $u(t, x; \phi)$ to be well resolved, which can be checked by examining their Fourier spectra. The values of N and Δt used for different λ and ν are summarized in Table I. As regards the choice of the time window T, our goal is for it to be on the order of an eddy turnover time (6), i.e., $T = C_T\, t_e$, where $C_T = \mathcal{O}(1)$. However, t_e depends on the solution $u(t, x; \varphi_{\nu,\lambda})$ of Problem 1 and therefore the actual value of the constant C_T can only be determined a posteriori, i.e., after Problem 1 has been solved. Hence, in our computations we use a fixed time window $T = 0.6$ and the resulting values of the constant C_T are reported in Table II. We have experimented with many other time windows T, both shorter and longer, and the results obtained were qualitatively similar to those reported below.

To illustrate the performance of algorithm (7), in Fig. 1(a) we show the dependence of the objective functional $\mathcal{J}_{\nu,\lambda}(\phi^{(n)})$ on iterations n for the solutions of Problem 1 obtained with $\nu = 2.5 \times 10^{-5}$ and $\lambda = 2, 4, 6$. We see that the values of the objective functional drop significantly, down to $\mathcal{O}(10^{-8} - 10^{-9})$ with smaller values achieved for smaller λ. To verify that the flows $u(t, x; \varphi_{\nu,\lambda})$ do indeed satisfy the definition of self-similarity Eq. (4) with a good accuracy, expressions $f(k, \varphi_{\nu,\lambda})$, cf. Eq. (5), corresponding to the optimal solutions $\varphi_{\nu,\lambda}$ obtained for different λ are shown as functions of k in Fig. 1(b). We see that in all cases these expressions are quite small, not exceeding $\mathcal{O}(10^{-5})$, and are generally decreasing functions of the wave number k.

To verify that the optimal flows we found are not numerical artifacts resulting from a finite numerical resolution, we solve Problem 1 with $\nu = 2.5 \times 10^{-5}$ and $\lambda = 2, 4, 6$ again, but with a refined numerical resolution. More specifically, the number of grid points N is increased by a factor

TABLE I. The number of spatial discretization points N and the temporal step size Δt used to solve Problem 1 for different values of the parameters ν and λ. Pink background in a cell indicates that no inertial solutions were found for the given values of ν and λ. If a cell is left blank, then no attempt was made to solve Problem 1 with the corresponding values of ν and λ.

ν \ λ	2	3	4	5	6	7
2.5×10^{-3}	$N = 4096$ $\Delta t = 1 \times 10^{-3}$	$N = 8192$ $\Delta t = 1 \times 10^{-3}$	$N = 8192$ $\Delta t = 1 \times 10^{-3}$	$N = 16384$ $\Delta t = 1 \times 10^{-3}$		
2.5×10^{-4}	$N = 8192$ $\Delta t = 5 \times 10^{-4}$	$N = 16384$ $\Delta t = 5 \times 10^{-4}$	$N = 16384$ $\Delta t = 5 \times 10^{-4}$	$N = 32768$ $\Delta t = 5 \times 10^{-4}$	$N = 65536$ $\Delta t = 5 \times 10^{-4}$	$N = 65536$ $\Delta t = 5 \times 10^{-4}$
2.5×10^{-5}	$N = 16384$ $\Delta t = 5 \times 10^{-4}$	N=131072 $\Delta t = 2 \times 10^{-4}$	N=131072 $\Delta t = 2 \times 10^{-4}$	N=131072 $\Delta t = 2 \times 10^{-4}$	N=131072 $\Delta t = 2 \times 10^{-4}$	N=131072 $\Delta t = 2 \times 10^{-4}$
2.5×10^{-6}	$N = 16384$ $\Delta t = 5 \times 10^{-4}$					
2.5×10^{-7}	$N = 16384$ $\Delta t = 5 \times 10^{-4}$					

of two as compared to the values indicated in Table I and the time step fixed to $\Delta t = 2 \times 10^{-4}$. In all cases, the solution $\varphi_{\nu,\lambda}$ obtained by solving Problem 1 with the original numerical parameters was used as the initial guess ϕ_1 in Eq. (7). The dependence of the objective functional $\mathcal{J}_{\nu,\lambda}(\phi^{(n)})$ normalized with respect to $\mathcal{J}_{\nu,\lambda}(\phi_1)$ on iterations n in these calculations is shown in Fig. 2(a). We see that this resolution refinement produces only a minuscule (less than 15%) further reduction of the objective functional. This should be contrasted with the decrease of the objective functional,

TABLE II. Values of the constant $C_T = T/t_e$ relating the optimization time window T and the eddy turnover time t_e, cf. (6), for different values of ν and λ. In addition, for inertial solutions, we show the number of wave fronts P present in the solution at final time. Pink background in a cell indicates that no inertial solutions where found for the given values of ν and λ (and hence P was not determined). If a cell is left blank, then no attempt was made to solve Problem 1 with the corresponding values of ν and λ.

ν \ λ	2	3	4	5	6	7
2.5×10^{-3}	$C_T = 0.70$ $P = 9$	$C_T = 0.09$	$C_T = 0.09$	$C_T = 0.21$		
2.5×10^{-4}	$C_T = 0.75$ $P = 10$	$C_T = 0.81$ $P = 19$	$C_T = 0.85$ $P = 15$	$C_T = 0.72$ $P = 27$	$C_T = 0.51$	$C_T = 0.53$
2.5×10^{-5}	$C_T = 0.75$ $P = 10$	$C_T = 1.44$ $P = 19$	$C_T = 0.90$ $P = 20$	$C_T = 1.35$ $P = 28$	$C_T = 0.67$ $P = 40$	$C_T = 0.86$ $P = 48$
2.5×10^{-6}	$C_T = 0.74$ $P = 10$					
2.5×10^{-7}	$C_T = 0.74$ $P = 10$					

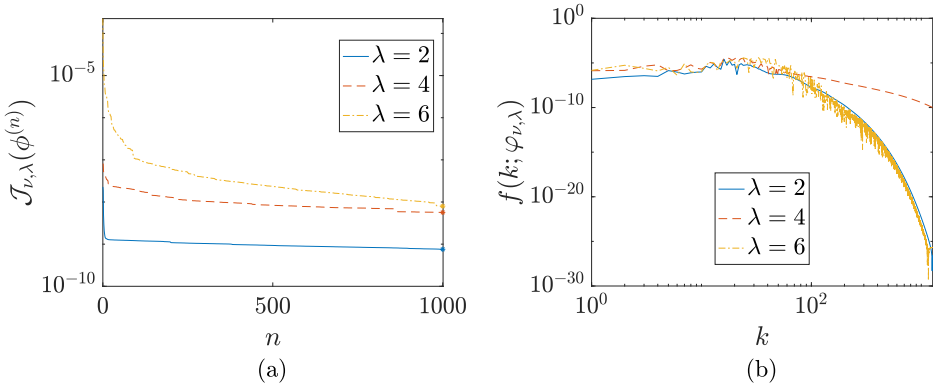


FIG. 1. (a) Decrease of the objective functional Eq. (5) with iterations n and (b) the corresponding distributions $f(k;\varphi_{\nu,\lambda})$ obtained at the minima when solving Problem 1 with $\nu = 2.5 \times 10^{-5}$ and (blue solid line) $\lambda = 2$, (red dashed line) $\lambda = 4$, and (yellow dot-dashed line) $\lambda = 6$.

typically involving several orders of magnitude [cf. Fig. 1(a)], which occurs when Problem 1 is solved "from scratch" with a fixed numerical resolution. These observations thus allow us to conclude that solutions of Problem 1 obtained with refined numerical resolutions produce convergent

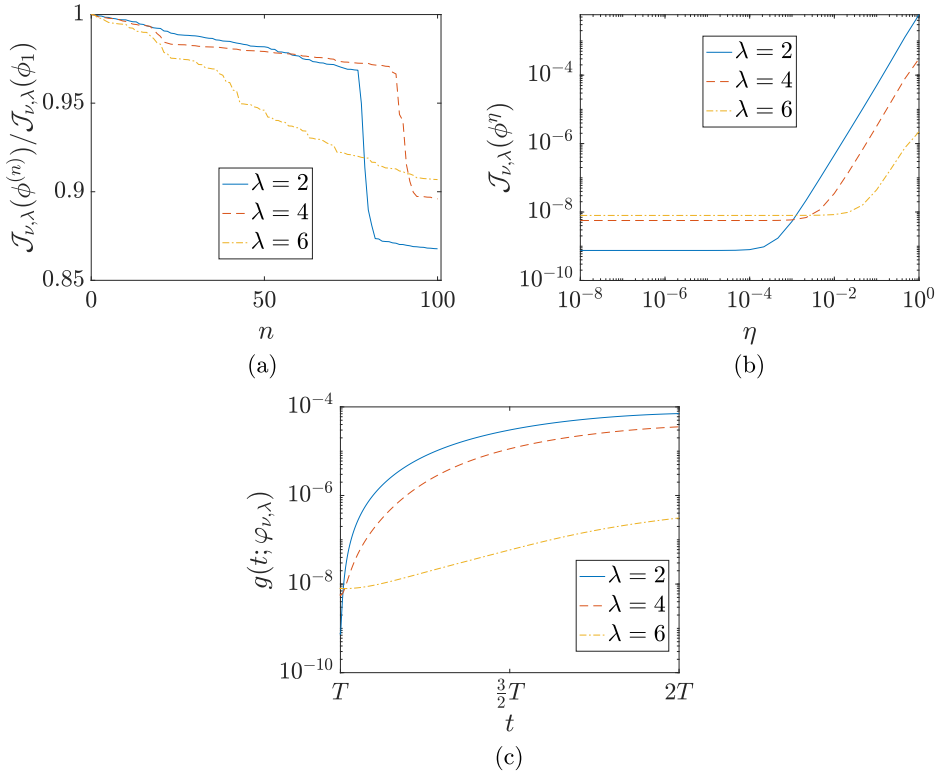


FIG. 2. (a) Decrease of the objective functional Eq. (5) with iterations n after refinement of the numerical resolution, (b) dependence of the objective functional $\mathcal{J}_{\nu,\lambda}(\phi^{\eta})$ on the noise magnitude η in the perturbed initial condition (19), and (c) the error function $g(t;\varphi_{\nu,\lambda})$ defined over a sliding time window for $t \in [T, 2T]$ for $\nu = 2.5 \times 10^{-5}$ and (blue solid line) $\lambda = 2$, (red dashed line) $\lambda = 4$, and (yellow dot-dashed line) $\lambda = 6$.

approximations of well-defined optimizers $\varphi_{\nu,\lambda}$. Moreover, since the obtained values of the objective functional are very small and approach zero (i.e., the global minimum) as the numerical resolution is refined, this suggests that these optimizers are in fact global.

Due to the nonlinearity of the governing system Eqs. (1), Problem 1 is nonconvex and as such may admit multiple local minimizers. This is indeed the situation we encounter at least for some combinations of ν and λ. More specifically, we have found evidence for the existence of two distinct classes of solutions of Problem 1 that we refer to as *viscous* and *inertial*. Their main features are contrasted in Fig. 3 and are summarized below.

(1) *Viscous solutions* involve highly oscillatory optimal initial conditions $\varphi_{\nu,\lambda}$, cf. Fig. 3(a), with energy concentrated at high wave numbers k in the dissipative subrange, cf. Fig. 3(e). Consequently, the resulting evolution does not involve any inertial transfer of energy between different Fourier modes and instead the solutions are quickly attenuated by the viscous dissipation, cf. Fig. 3(c) (which is the origin of the name we gave these solutions). The enstrophy $\mathcal{E}(u(t,\cdot;\varphi_{\nu,\lambda}))$ of viscous solutions is a rapidly decreasing function of time t, cf. Fig. 3(h). As they do not feature any large-scale structure and have a small magnitude, the viscous solutions resemble the minimizer $\phi_0 \equiv 0$ of functional (5) in the absence the constraint $\mathcal{E}(\phi) = \mathcal{E}_0 > 0$ (see the discussion before the statement of Problem 1).

(2) *Inertial solutions* minimize the objective functional (5) via inertial transfer of energy to Fourier modes with higher wave numbers k, cf. Fig. 3(f), which in the physical space is achieved by a self-similar steepening of the wave fronts, cf. Figs. 3(b) and 3(d). As a result, the enstrophy $\mathcal{E}(u(t,\cdot;\varphi_{\nu,\lambda}))$ of inertial solutions is a rapidly increasing function of time t, cf. Fig. 3(h). In contrast to the viscous solutions, the magnitude of the inertial solutions does not change much over the time window T.

Viscous solutions are "generic" as they can be found by solving Problem 1 using essentially any initial guess ϕ_1. However, inertial solutions are "rare," in the sense that they can only be found if a "good" initial guess ϕ_1 is used in Eq. (7). The final values of the objective functional (5) are in both cases comparable, cf. Fig. 1(a), although the expressions $f(k,\varphi_{\nu,\lambda})$ tend to have a different dependence on k for the viscous and inertial solutions, cf. Fig. 3(g). Moreover, for a given value of λ the inertial solutions can only be found provided the viscosity coefficient ν is sufficiently small. This is because the inertial solutions involve a self-similar energy transfer between Fourier modes with wave numbers k and λk, which can only happen if the inertial range is sufficiently wide or, equivalently, if the Reynolds number is sufficiently high. A key qualitative feature differentiating viscous and inertial solutions is the time evolution of the enstrophy (3) which is, respectively, a decreasing and an increasing function of time for these two families of solutions, cf. Fig. 3(h). Table II provides information about the values of λ for which inertial solutions could be found for decreasing values of ν, and also includes the number of wave fronts, P, present in the solution at the final time (during the evolution from $t = 0$ to $t = T$ this number typically decreases by no more than 15%). Needless to say, only the inertial solutions are self-similar in the sense of Definition 1 and hence we will exclusively focus on them in the remainder of this work.

The observation that inertial solutions are "rare," in the sense discussed above, raises the question about their robustness understood as persistence of self-similar evolution under perturbations of the optimal initial conditions $\varphi_{\nu,\lambda}$. To address this question, we solve the governing system Eqs. (1) over the time interval $[0,T]$ for a family of initial data ϕ^η obtained by adding random noise of magnitude $\eta \in \mathbb{R}^+$ to the optimal initial condition $\varphi_{\nu,\lambda}$ and then evaluating the objective functional (5) as function of η. The perturbed initial condition is obtained by adding random perturbations to the Fourier coefficients of $\varphi_{\nu,\lambda}$, i.e., as

$$\phi^\eta(x) = \sum_{k=-N/2+1}^{N/2} [\widehat{\varphi_{\nu,\lambda}}]_k (1+\eta\theta_k) e^{ikx}, \tag{19}$$

where θ_k are independent samples of a normally distributed random variable with zero mean and unit variance, and then normalizing ϕ^η such that the constraint $\mathcal{E}(\phi^\eta) = \mathcal{E}_0$ is satisfied. The

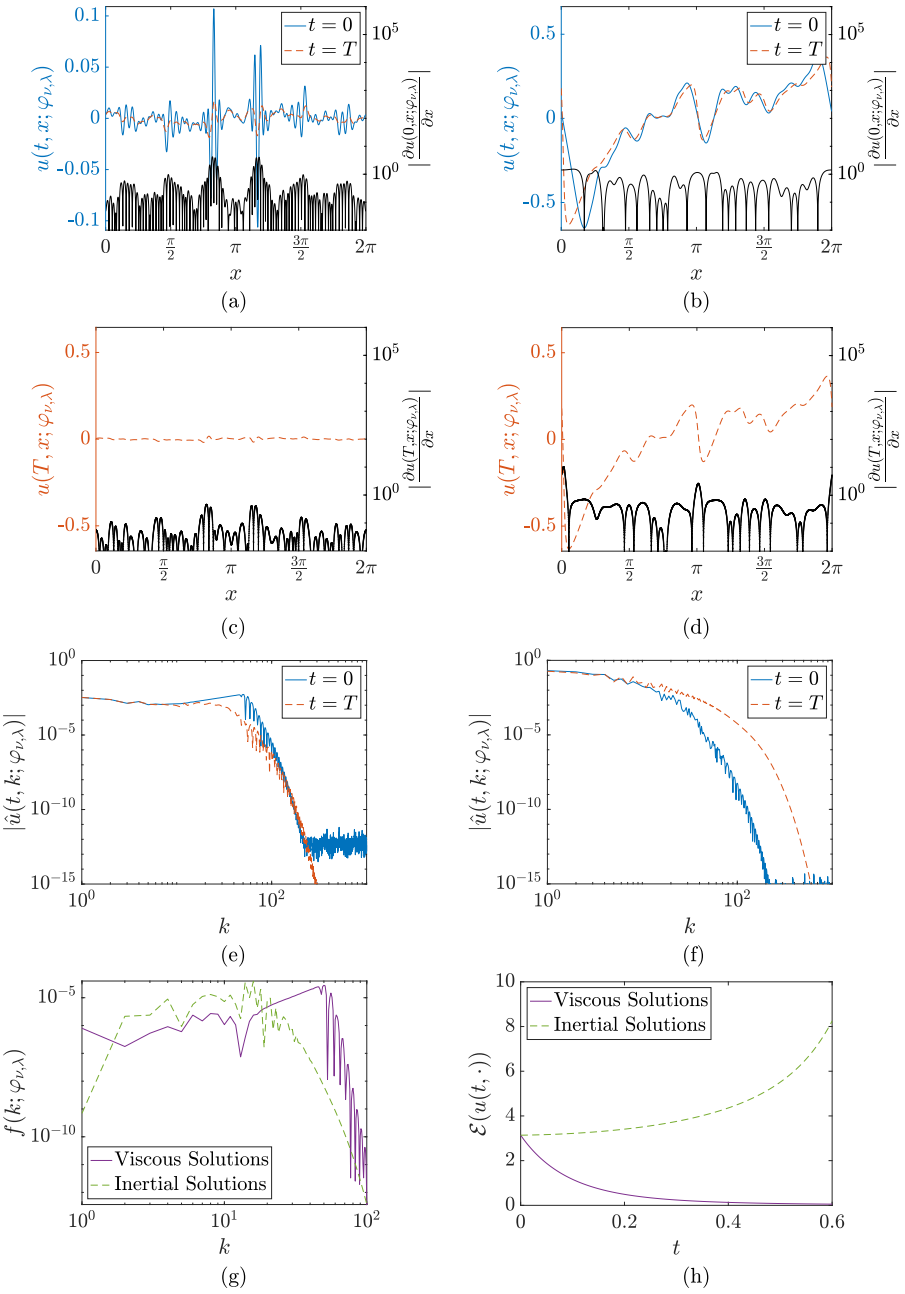


FIG. 3. Comparison of the properties of (left column) the viscous solutions and (right column) the inertial solutions obtained by solving Problem 1 with $\nu = 2.5 \times 10^{-3}$ and $\lambda = 2$: (a), (b) optimal initial conditions $\varphi_{\nu,\lambda}(x)$ (blue, left axis), the magnitude of their gradients $|d\varphi_{\nu,\lambda}(x)/dx|$ (black, right axis) and the corresponding solutions $u(T, x; \varphi_{\nu,\lambda})$ at the final time (red, left axis); (c), (d) the solutions at the final time $u(T, x; \varphi_{\nu,\lambda})$ (red, left axis) and their gradients $|\partial u(T, x; \varphi_{\nu,\lambda})/\partial x|$ (black, right axis); (e), (f) the Fourier spectra $\widehat{\varphi_{\nu,\lambda}}(k)$ (blue) and $\widehat{u}(T, k; \varphi_{\nu,\lambda})$ (red) of the optimal initial conditions and of the corresponding final states; (g), (h) the functions $f(k; \varphi_{\nu,\lambda})$ vs k and $\mathcal{E}(u(t, \cdot; \varphi_{\nu,\lambda}))$ vs t, cf. Eqs. (5) and (3), respectively, at the minimum for the viscous solution (purple) and the inertial solution (green).

results obtained for $\nu = 2.5 \times 10^{-5}$ and $\lambda = 2, 4, 6$ are shown in Fig. 2(b) where we see that flow evolutions cease to be self-similar, which is marked by the increase of $\mathcal{J}_{\nu,\lambda}(\phi^\eta)$ above its baseline value corresponding to $\eta = 0$, only provided the perturbations are sufficiently large, i.e., $\eta \gtrapprox 10^{-4}$ or $\eta \gtrapprox 10^{-2}$ depending on λ. For values of η larger than this threshold, we have $\mathcal{J}_{\nu,\lambda}(\phi^\eta) = \mathcal{O}(\eta^2)$ which, since the functional (5) is quartic in ϕ, suggests that the departure from self-similarity is only sublinear in η. These findings indicate that while inertial solutions are "rare," they are in fact relatively robust.

Finding optimal Burgers flows $u(t, \cdot; \varphi_{\nu,\lambda})$ as solutions of Problem 1 ensures these flows are self-similar on the time interval $[0, T]$ and a natural question is what happens with this property at later times $t > T$. To answer this question, we consider the function $g(t; \varphi_{\nu,\lambda}) := \frac{1}{2} \sum_{k=0}^{\infty} ||\widehat{u}(t, \lambda k; \varphi_{\nu,\lambda})|^2 - \lambda |\widehat{u}(t + T, \lambda k; \varphi_{\nu,\lambda})|^2|$ which is the expression from the objective functional (5) evaluated over a sliding window $[t, t + T]$. We plot this function for $t \in [T, 2T]$ in Fig. 2(c) for the optimal Burgers flows obtained for $\nu = 2.5 \times 10^{-5}$ and $\lambda = 2, 4, 6$. Unsurprisingly, the property of self-similarity is quickly lost. However, it is interesting to observe that the loss of self-similarity is much slower for larger values of λ.

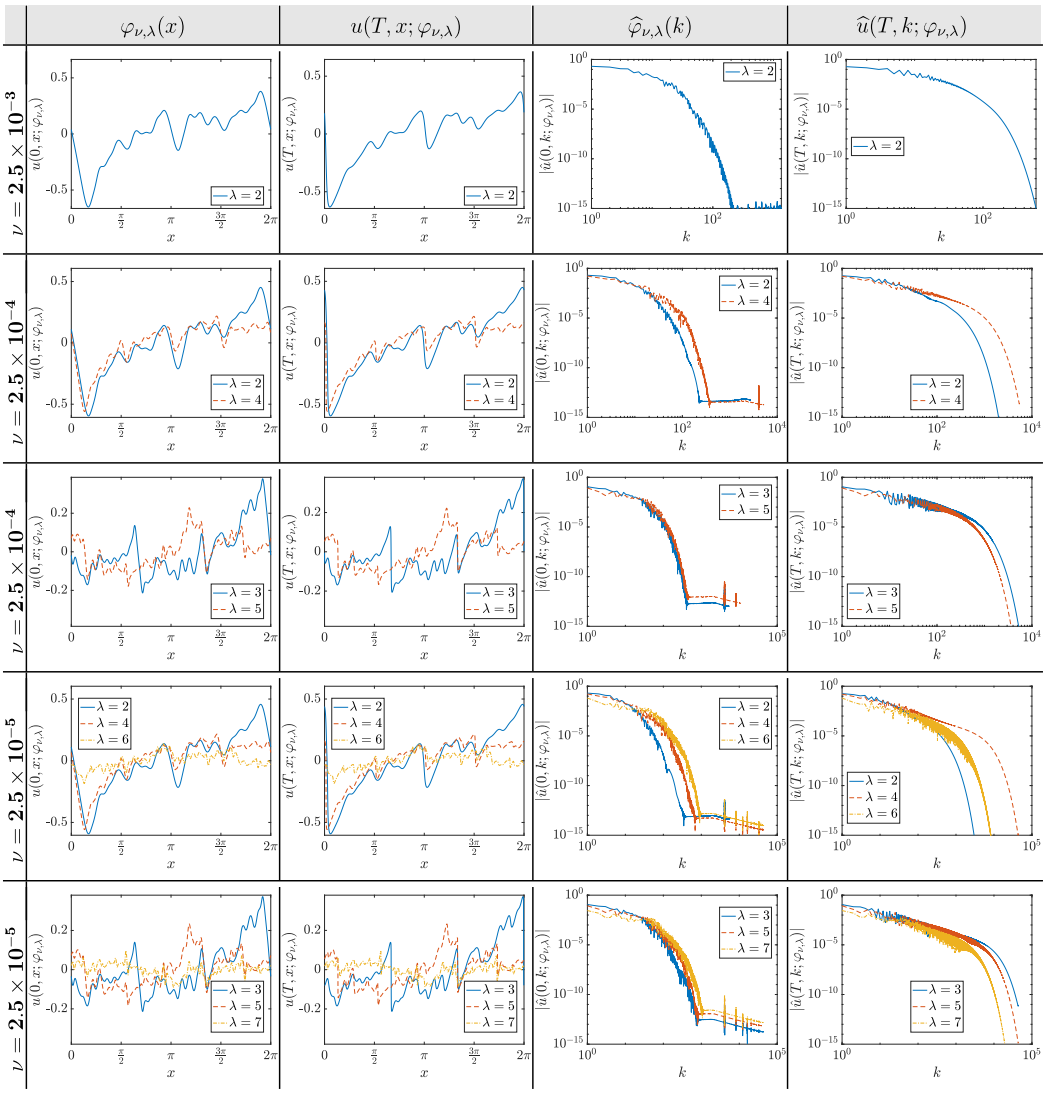
We now examine how properties of the inertial solutions vary in function of the parameters ν and λ. Families of optimal initial conditions $\varphi_{\nu,\lambda}$ were obtained using a continuation approach where Problem 1 was solved for some ν_1 and λ_1 using either $\varphi_{\nu_2,\lambda_1}$ or $\varphi_{\nu_1,\lambda_2}$ for some $\mathbb{R}^+ \ni \nu_2 > \nu_1$ and $\mathbb{N}^+ \ni \lambda_2 < \lambda_1$ as the initial guess ϕ_1 in Eq. (7). The solutions found for fixed viscosity coefficients ν and increasing values of λ are summarized in Table III, where we see that when $\lambda > 2$ inertial solutions can be found only provided ν is sufficiently small. The reason for this is clear when we examine the Fourier spectra of the optimal initial conditions $\varphi_{\nu,\lambda}$ and of the corresponding final states $u(T, \cdot; \varphi_{\nu,\lambda})$ (the third and fourth columns in Table III)—we see that during its evolution the solution develops a well-defined inertial range where energy is cascaded in a self-similar manner minimizing the objective functional (5). When the viscosity is large, this inertial range does not persist long enough to accommodate energy transfer between Fourier coefficients with a large separation in the wave-number space, which needs to happen when λ is large.

In terms of the physical-space representation, the self-similar energy cascade is realized by the steepening of the wave fronts which is clearly visible in the first and second columns of Table III. In other words, during the evolution over $[0, T]$ the global structure of the solutions remains essentially unchanged, but all the wave fronts steepen in almost the same manner. For beyond the time window T the steepened wave fronts form shocks, for which afterwards diffusion dominates and dissipates the flow. There are some apparent similarities in how this process unfolds in solutions obtained with even and odd values of λ. When ν is fixed, the number of distinct fronts in the solution is an increasing function of λ which is illustrated in Fig. 4. Table II provides more details about this trend, although additional data would be needed to determine the precise form of the dependence of P on λ for a fixed ν. Finally, in Table IV we characterize solutions obtained by solving Problem 1 with $\lambda = 2, 5$ and different values of ν (this is a subset of the results already shown in Table III, but organized in a different manner). We see that, interestingly, both the optimal initial conditions $\varphi_{\nu,\lambda}$ and the corresponding final states $u(T, x; \varphi_{\nu,\lambda})$ appear to converge to well-defined limits as $\nu \to 0$ while their Fourier spectra become more developed.

V. CONCLUSIONS

This study was motivated by an open question in turbulence research concerning the nature of fluid motions that can give rise to a self-similar energy cascade as predicted by Kolmogorov's statistical theory of turbulence [1,3]. While the real problem concerns flows in 3D, here we consider a simple toy model, namely, the 1D viscous Burgers equation, and provide a proof of concept for a novel approach to this problem, where solutions with a self-similar structure consistent with Definition 1 are constructed systematically. This is done by solving a family of PDE optimization problems, cf. Problem 1, where we seek initial conditions for system Eqs. (1) such that deviations from the self-similar behavior are minimized by the corresponding flow evolution over a period

TABLE III. Summary information about the inertial solutions found by solving Problem 1 with the values of ν indicated on the left and the values of λ given in the individual legends. For a better comparison the solutions obtained with even and odd λ are grouped together. The first and the second columns represent the optimal initial conditions $\varphi_{\nu,\lambda}$ and the final states $u(T, x; \varphi_{\nu,\lambda})$, with the corresponding spectra given in the third and fourth columns.



of time approximately equal to one eddy turnover time (6). This somewhat nonstandard PDE optimization problem is solved numerically for different values of ν and λ using a state-of-the-art adjoint-based gradient method described in Sec. III. The optimal solutions $\varphi_{\nu,\lambda}$ found in this way correspond to very small values of the objective functional (5), cf. Fig. 1(a), indicating that they satisfy Definition 1 of self-similarity down to numerical approximation and round-off errors. Moreover, since these values approach zero (i.e., the global minimum) as the numerical resolution is refined, cf. Fig. 2(a), this suggests that these optimizers are in fact global.

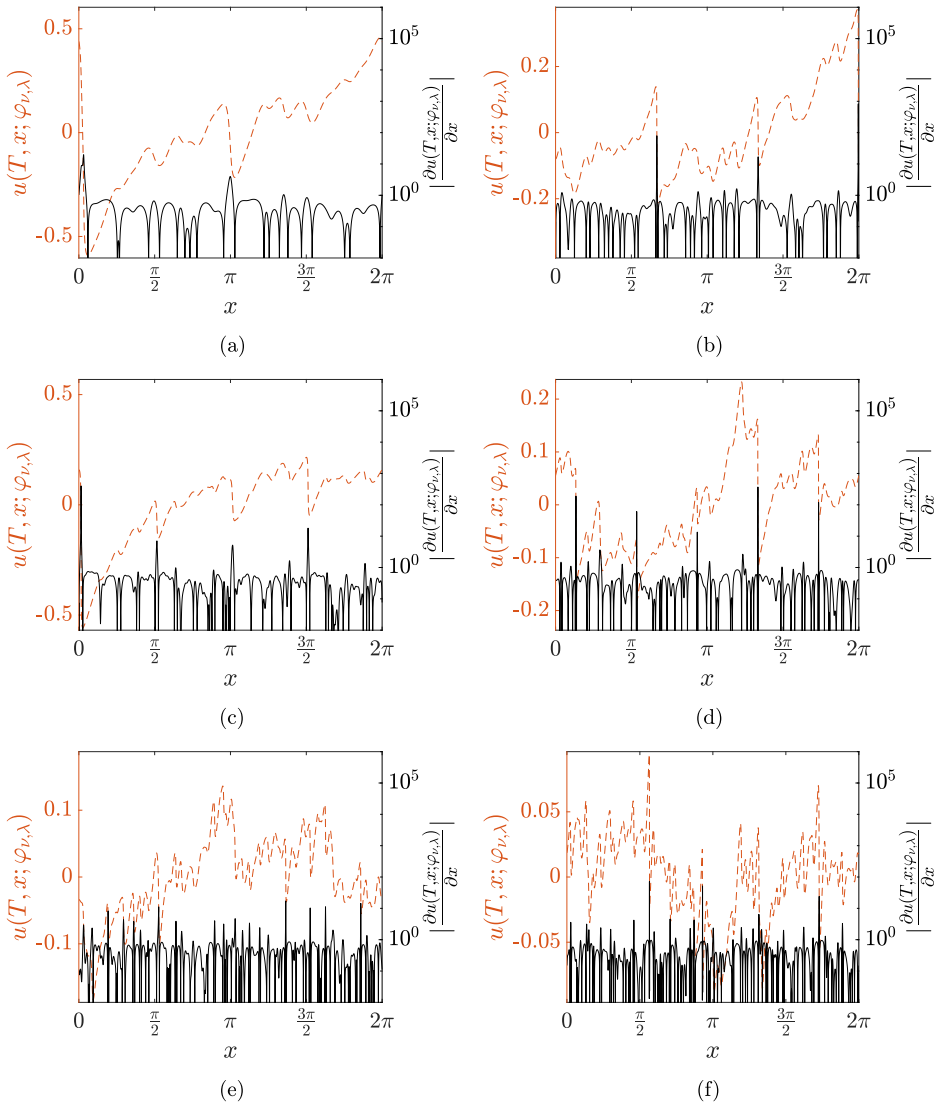
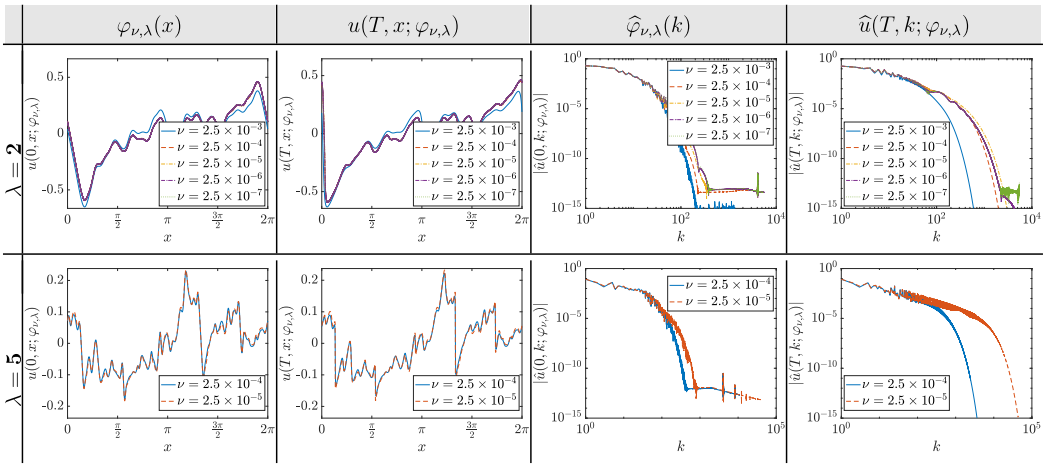


FIG. 4. Solutions at the final time $u(T, x; \varphi_{\nu,\lambda})$ (red, left axis) and their gradients $|\partial u(T, x; \varphi_{\nu,\lambda})/\partial x|$ (black, right axis) obtained by solving Problem 1 with $\nu = 2.5 \times 10^{-5}$ and (a) $\lambda = 2$, (b) $\lambda = 3$, (c) $\lambda = 4$, (d) $\lambda = 5$, (e) $\lambda = 6$, and (f) $\lambda = 7$.

The solutions found fall into two distinct categories referred to as viscous and inertial distinguished by the behavior of the enstrophy $\mathcal{E}(u(t, \cdot; \varphi_{\nu,\lambda}))$ which, respectively, decays and grows in these types of solutions, cf. Fig. 3(h). The former solutions are uninteresting as they consist of high-wave-number oscillations in the viscous subrange of the Fourier spectrum and as such are quickly attenuated by viscous dissipation without producing any significant energy transfer across scales. Therefore, they resemble the trivial solution $\phi_0 \equiv 0$ of the corresponding unconstrained problem, with the high-wave-number oscillation added to satisfy the constraint $\mathcal{E}(\phi) = \mathcal{E}_0$. The fact that such physically uninteresting, yet mathematically consistent, solutions have been found demonstrates the robustness of algorithm (7).

However, the inertial solutions lead to flow evolutions involving a self-similar, in the sense of Definition 1, transfer of energy in the Fourier space, which in the physical space is manifested

TABLE IV. Summary information about the inertial solutions found by solving Problem 1 with $\lambda = 2$ (top row) and $\lambda = 5$ (bottom row), and different values of ν indicated in the individual legends. The first and the second columns represent the optimal initial conditions $\varphi_{\nu,\lambda}$ and the final states $u(T, x; \varphi_{\nu,\lambda})$, with the corresponding spectra given in the third and fourth columns.



by a uniform steepening of the wave fronts, cf. Table III and Table IV. As expected, optimal solutions involving such energy transfer over a large distance in the Fourier space can only be found provided the viscosity ν is sufficiently small (or, equivalently, the "Reynolds number" is sufficiently large). Interestingly, the structure of these self-similar solutions is preserved as $\nu \to 0$, cf. Table IV, suggesting that it might persist in the limiting inviscid problem (for times before the inviscid solutions become singular). We add that front steepening in viscous Burgers flows, resulting from the competition between nonlinear amplification and viscous attenuation, is a well-studied and well-understood phenomenon [13,26–29]. While it occurs generically in such flows, our main contribution is to demonstrate that with a suitably chosen initial data, it can take place in a self-similar manner, in the sense of Definition 1.

To the best of our knowledge, the results reported here represent the first successful effort to construct time-dependent solutions of a hydrodynamic model characterized by self-similar energy transfer. Both the viscous and inviscid Burgers equations are in principle analytically solvable [23] and it is an interesting open question whether this fact could be used to rigorously justify the results presented here. Another interesting extension is to consider forced Burgers flows where notions of self-similarity other than the one given in Definition 1 would be applicable. In such case, the corresponding objective functional would be minimized with respect to some properties of the forcing term added on the right-hand side in Eq. (1a). While in this study we focused on a toy problem, the results obtained show promise for this approach to be applicable to more realistic flow models, including shell models of turbulence [41] and two-dimensional Navier-Stokes flows. Ultimately, the goal is to search for self-similar behavior in 3D Navier-Stokes flows which can provide new insights about hydrodynamic turbulence.

ACKNOWLEDGMENTS

B.P. was partially supported through an NSERC (Canada) Discovery Grant No. RGPIN-2020-05710. Research of T.Y. was partly supported by the JSPS Grants-in-Aid for Scientific Research under Grant No. 24H00186.

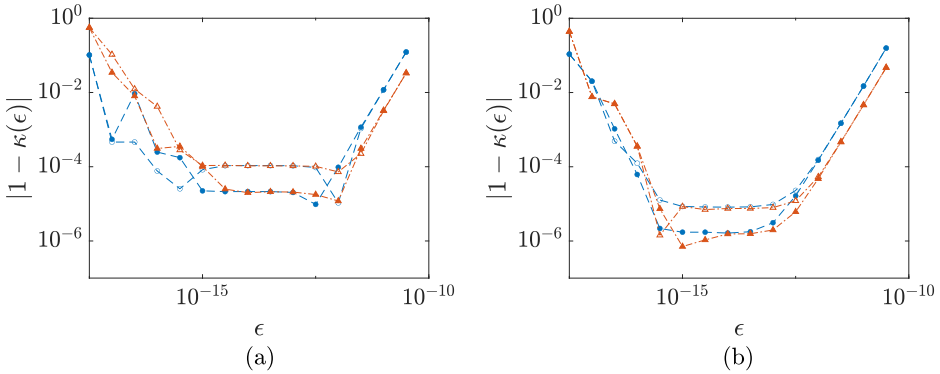


FIG. 5. Dependence of $|\kappa(\epsilon) - 1|$ on ϵ for (a) $\lambda = 2$ and (b) $\lambda = 7$. Equation (A1) was evaluated using the temporal discretizations (empty symbols) $\Delta t = 1 \times 10^{-3}$ and (filled symbols) $\Delta t = 2 \times 10^{-4}$, and the perturbations (blue circles) $\phi' = \sin(-x)$ and (red triangles) $\phi' = \frac{1}{2+\sin(x)}$.

DATA AVAILABILITY

The data that support the findings of this article are openly available [40].

APPENDIX: VALIDATION OF GRADIENT (15)

To validate the computation of gradient Eq. (15), which is obtained by solving the adjoint system Eqs. (12) with the rather nonstandard terminal condition (13) using the numerical approach described in Sec. III A, we compute the Gâteaux (directional) differential Eq. (8) in two ways: in terms of an approximation using a first-order forward finite difference formula with the step size ϵ and in terms of the Riesz formula Eq. (10) involving the gradient $\nabla^{L^2}\mathcal{J}_{\nu,\lambda}(\phi)$. The ratio of these two quantities is thus defined as

$$\kappa(\epsilon) := \frac{\epsilon^{-1}[\mathcal{J}_{\nu,\lambda}(\phi + \epsilon\phi') - \mathcal{J}_{\nu,\lambda}(\phi)]}{\langle \nabla^{L^2}\mathcal{J}_{\nu,\lambda}(\phi), \phi' \rangle_{L^2(\Omega)}}, \quad \epsilon > 0, \tag{A1}$$

where $\phi, \phi' \in L^2(\Omega)$ represent, respectively, the "point" where the Gâteaux differential is computed and the direction. Given the equivalence of the Riesz representations involving the L^2 and H^1 inner products, cf. (10), for simplicity we choose to use the former in the denominator of Eq. (A1).

For intermediate values of ϵ the dominant errors come from the discretization of system Eqs. (1) and (12), and are independent of ϵ; we thus expect that $\kappa(\epsilon) \to 1$ as the discretization is refined, i.e., as $N \to \infty$ and $\Delta t \to 0$. However, $\kappa(\epsilon)$ will deviate from 1 for both small and large ϵ due to, respectively, the subtractive cancellation (round-off) errors and the truncation errors in the finite-difference formula, both of which are well known effects. A detailed discussion of the interplay between these different types of errors is given in [Ref. 42, Fig. 2].

To fix attention, in the tests reported here, we set $T = 0.6$, $\nu = 2.5 \times 10^{-3}$ and $\lambda = 2, 7$, and use the initial condition $\phi(x) = -\sin(x)$. Since integration in time is usually the main source of errors when employing pseudospectral methods in space, Fig. 5 shows that indeed $\kappa(\epsilon) \to 1$ for intermediate $\epsilon \in [10^{-10}, 10^{-5}]$ as the temporal discretrization Δt is refined with a fixed spatial resolution of $N = 16,384$ and two different perturbations ϕ'. Similar trends were also observed in tests performed with other values of T, ν and λ (not shown here for brevity). This thus demonstrates the consistency of the computation of the gradient $\nabla^{L^2}\mathcal{J}_{\nu,\lambda}(\phi)$, both in terms of the derivation of the adjoint system Eqs. (12) and (13) and the numerical implementation.

[1] U. Frisch, *Turbulence*, 1st ed. (Cambridge University Press, Cambridge, UK, 1995).
[2] L. F. Richardson, *Weather Prediction by Numerical Process* (Cambridge University Press, Cambridge, UK, 1922).
[3] A. Kolmogorov, The local structure of turbulence in incompressible viscous fluid for very large Reynolds number, Dokl. Akad. Nauk SSSR **30**, 9 (1941); translated and reprinted in Proc. R. Soc. Lond. A **434**, 9 (1991).
[4] S. Goto, Y. Saito, and G. Kawahara, Hierarchy of antiparallel vortex tubes in spatially periodic turbulence at high Reynolds numbers, Phys. Rev. Fluids **2**, 064603 (2017).
[5] T. Yoneda, S. Goto, and T. Tsuruhashi, Mathematical reformulation of the Kolmogorov–Richardson energy cascade in terms of vortex stretching, Nonlinearity **35**, 1380 (2022).
[6] T. Tsuruhashi, S. Goto, S. Oka, and T. Yoneda, Self-similar hierarchy of coherent tubular vortices in turbulence, Philos. Trans. R. Soc. A **380**, 20210053 (2022).
[7] M. D. Gunzburger, *Perspectives in Flow Control and Optimization* (SIAM, Philadelphia, PA, 2003).
[8] A. Jameson, Aerodynamic design via control theory, J. Sci. Comput. **3**, 233 (1988).
[9] T. R. Bewley, P. Moin, and R. Temam, DNS–based predictive control of turbulence: An optimal benchmark for feedback algorithms, J. Fluid Mech. **447**, 179 (2001).
[10] P. Matharu and B. Protas, Optimal closures in a simple model for turbulent flows, SIAM J. Sci. Comput. **42**, B250 (2020).
[11] P. Matharu and B. Protas, Optimal eddy viscosity in closure models for two-dimensional turbulent flows, Phys. Rev. Fluids **7**, 044605 (2022).
[12] B. Protas, Systematic search for extreme and singular behaviour in some fundamental models of fluid mechanics, Philos. Trans. R. Soc. A **380**, 20210035 (2022).
[13] D. Ayala and B. Protas, On maximum enstrophy growth in a hydrodynamic system, Physica D **240**, 1553 (2011).
[14] D. Ayala, C. R. Doering, and T. M. Simon, Maximum palinstrophy amplification in the two-dimensional Navier-Stokes equations, J. Fluid Mech. **837**, 839 (2018).
[15] R. R. Kerswell, C. C. T. Pringle, and A. P. Willis, An optimization approach for analysing nonlinear stability with transition to turbulence in fluids as an exemplar, Rep. Prog. Phys. **77**, 085901 (2014).
[16] P. Matharu, B. Protas, and T. Yoneda, On maximum enstrophy dissipation in 2D Navier–Stokes flows in the limit of vanishing viscosity, Physica D **441**, 133517 (2022).
[17] D. Ayala and B. Protas, Extreme vortex states and the growth of enstrophy in 3D incompressible flows, J. Fluid Mech. **818**, 772 (2017).
[18] D. Kang, D. Yun, and B. Protas, Maximum amplification of enstrophy in three-dimensional Navier-Stokes flows, J. Fluid Mech. **893**, A22 (2020).
[19] D. Kang and B. Protas, Searching for singularities in Navier–Stokes flows based on the Ladyzhenskaya–Prodi–Serrin conditions, J. Nonlinear Sci. **32**, 81 (2022).
[20] X. Zhao and B. Protas, Systematic search for singularities in 3D Euler flows, J. Nonlinear Sci. **33**, 120 (2023).
[21] P. Hassanzadeh, G. Chini, and C. Doering, Wall to wall optimal transport, J. Fluid Mech. **751**, 627 (2014).
[22] C. J. Miles and C. R. Doering, A shell model for optimal mixing, J. Nonlinear Sci. **28**, 2153 (2018).
[23] H. Kreiss and J. Lorenz, *Initial-Boundary Value Problems and the Navier-Stokes Equations*, Classics in Applied Mathematics, Vol. 47 (SIAM, Philadelphia, PA, 2004).
[24] R. Adams, *Sobolev Spaces* (Academic Press, New York, NY, 2003).
[25] J. Bec and K. Khanin, Burgers turbulence, Phys. Rep. **447**, 1 (2007).
[26] J.-D. Fournier and U. Frisch, L'équation de Burgers déterministe et statistique (The deterministic and statistical Burgers equation), J. Méc. Théor. Appl. **2**, 699 (1983).
[27] D. Pelinovsky, Sharp bounds on enstrophy growth in the viscous Burgers equation, Proc. R. Soc. A **468**, 3636 (2012).
[28] D. Pelinovsky, Enstrophy growth in the viscous Burgers equation, Dynam. Partial Differ. Eqs. **9**, 305 (2012).
[29] D. Albritton and N. D. Nitti, Sharp bounds on enstrophy growth for viscous scalar conservation laws, Nonlinearity **36**, 7142 (2023).

[30] B. P. Murray and M. D. Bustamante, Energy flux enhancement, intermittency and turbulence via Fourier triad phase dynamics in the 1-D Burgers equation, J. Fluid Mech. **850**, 624 (2018).
[31] M. Buzzicotti, B. P. Murray, L. Biferale, and M. D. Bustamante, Phase and precession evolution in the Burgers equation, Eur. Phys. J. E **39**, 34 (2016).
[32] B. Protas, D. Kang, and M. D. Bustamante, Alignments of triad phases in extreme one-dimensional Burgers flows, Phys. Rev. E **109**, 055104 (2024).
[33] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds* (Princeton University Press, Princeton, NJ, 2008).
[34] M. S. Berger, *Nonlinearity and Functional Analysis* (Academic Press, San Diego, CA, 1977).
[35] B. Protas, T. R. Bewley, and G. Hagen, A computational framework for the regularization of adjoint analysis in multiscale PDE systems, J. Comput. Phys. **195**, 49 (2004).
[36] J. Nocedal and S. Wright, *Numerical Optimization*, 2nd ed., Springer Series in Operations Research and Financial Engineering (Springer, Berlin, 2006), pp. xxii+664.
[37] T. Y. Hou, Blow-up or no blow-up? A unified computational and analytic approach to 3D incompressible Euler and Navier-Stokes equations, Acta Numer. **18**, 277 (2009).
[38] R. Alimo, D. Cavaglieri, P. Beyhaghi, and T. R. Bewley, Design of IMEXRK time integration schemes via Delaunay-based derivative-free optimization with nonconvex constraints and grid-based acceleration, J. Glob. Optim. **79**, 567 (2020).
[39] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes*, 3rd ed. (Cambridge University Press, Cambridge, UK, 2007), pp. xxii+1235, the art of scientific computing.
[40] P. Matharu, Self-similar energy tranfer for 1D Burgers equation (2025), https://github.com/pipmath/BurgersEnergyTransfer.
[41] P. D. Ditlevsen, *Turbulence and Shell Models* (Cambridge University Press, Cambridge, UK, 2010).
[42] P. Matharu and B. Protas, Adjoint-based enforcement of state constraints in PDE optimization problems, J. Comput. Phys. **517**, 113298 (2024).