Time-Stepping Approach for Solving Upper-Bound Problems: Application to Two-Dimensional Rayleigh–Bénard Convection

Baole Wen,[1,2,3] Gregory P. Chini,[1,2,4,*] Rich R. Kerswell,[5] and Charles R. Doering[6,7,8]

[1] *Program in Integrated Applied Mathematics, University of New Hampshire, Durham, NH 03824 USA*
[2] *Center for Fluid Physics, University of New Hampshire, Durham, NH 03824 USA*
[3] *The Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX 78712 USA*
[4] *Department of Mechanical Engineering, University of New Hampshire, Durham, NH 03824 USA*
[5] *School of Mathematics, University of Bristol, University walk, Bristol, BS8 1TW, UK*
[6] *Department of Physics, University of Michigan, Ann Arbor, MI 48109-1043 USA*
[7] *Department of Mathematics, University of Michigan, Ann Arbor, MI 48109-1043 USA*
[8] *Center for the Study of Complex Systems, University of Michigan, Ann Arbor, MI 48109-1043 USA*

A new computational procedure for numerically solving a class of variational problems arising from rigorous upper bound analysis of forced-dissipative infinite-dimensional nonlinear dynamical systems, including the Navier–Stokes and Oberbeck–Boussinesq equations, is analyzed and applied to Rayleigh–Bénard convection. A proof that the only steady state to which this numerical algorithm can converge is the required global optimal of the relevant variational problem is given for three canonical flow configurations. In contrast with most other numerical schemes for computing the optimal bounds on transported quantities (e.g., heat or momentum) within the "background field" variational framework, which employ variants of Newton's method and hence require very accurate initial iterates, the new computational method is easy to implement and, crucially, does not require numerical continuation. The algorithm is used to determine the optimal background-method bound on the heat transport enhancement factor, i.e., the Nusselt number Nu, as a function of the Rayleigh number Ra, Prandtl number Pr, and domain aspect ratio L in two-dimensional Rayleigh–Bénard convection between stress-free isothermal boundaries (Rayleigh's original 1916 model of convection). The result of the computation is significant because analyses, laboratory experiments, and numerical simulations have suggested a range of exponents α and β in the presumed $\mathrm{Nu} \sim \mathrm{Pr}^{\alpha}\mathrm{Ra}^{\beta}$ scaling relation. The computations clearly show that for $\mathrm{Ra} \leq 10^{10}$ at fixed $L = 2\sqrt{2}$, $\mathrm{Nu} \leq 0.106\mathrm{Pr}^{0}\mathrm{Ra}^{5/12}$, which indicates that molecular transport cannot generally be neglected in the "ultimate" high-Ra regime.

* *Corresponding author*: greg.chini@unh.edu

I. INTRODUCTION

High-dimensional forced-dissipative nonlinear dynamical systems generally exhibit enhanced transport in turbulent regimes. For example, the transport of heat and momentum in turbulent shear- and buoyancy-driven flows, respectively, is increased by orders of magnitude over the transport achievable by molecular processes in the absence of fluid motion. *A priori* prediction or estimation of this transport enhancement is desirable in a wide range of geo-scientific and technological applications but is challenging owing to the complexity of the turbulent dynamics. One approach that has proved effective for certain problems is to derive, via rigorous analysis of – and without directly simulating – the governing nonlinear partial differential equations, bounds on the achievable transport that all flow solutions (whether steady, unsteady, turbulent, etc.) must satisfy. The derivation of upper bounds on flow quantities was first given by Howard [1] (originally motivated by the ideas of Malkus [2]) for Rayleigh–Bénard convection, the buoyancy-driven flow of a fluid heated from below and cooled from above [3, 4], and extended by Busse *et al.* to various other thermal convection processes as well as to shear flow turbulence [5–7]. The Malkus–Howard–Busse (MHB) variational formulation is strictly true for stationary flows but is presumed to be valid in an infinite (rather than finite) horizontal layer owing to the hypothesis of statistical stationarity, i.e., the technical assumption that horizontal averages, and thus also volume averages, are time independent [8]. Several decades later a "background field" method was proposed by Doering and Constantin [8–13] to produce rigorous upper bounds on energy dissipation and heat transport in a variety of turbulent flows without any statistical hypotheses, scaling assumptions, or closure approximations. This approach, which posits a decomposition of the flow variables into a steady background field plus arbitrarily large fluctuations, is based on Hopf's method for producing *a priori* estimates for solutions of the Navier–Stokes equations with inhomogeneous boundary conditions [14], and hereafter is referred to as the Constantin–Doering–Hopf (CHD) variational formalism. The link between these two variational schemes (MHB and CDH) has been discussed in detail by Kerswell [15, 16].

Strictly rigorous albeit sub-optimal bounds often can be obtained within the CDH framework by assuming simple (generally, piecewise-linear) functional forms for the background (e.g., temperature or streamwise velocity) profiles. To obtain the optimal achievable CDH bounds, the nonlinear, stationary Euler–Lagrange equations for the optimal fields must be solved numerically. The required computations are challenging because the solutions of these equations are non-unique; in contrast, the unique solution of the CDH problem, the global optimizer, not only satisfies the Euler–Lagrange equations but also a "spectral constraint" requiring a certain linear operator that depends on the global optimal background field to have strictly non-negative eigenvalues. Consequently, extreme care must be exercised to ensure that a candidate optimizer satisfies the spectral constraint.

Most prior approaches to numerically solving the CDH problem employ variants of Newton's method. In practice, to avoid numerous spurious optimizers, the use of sophisticated numerical continuation software packages is required [17] since Newton iterations generally will not converge to the *true* solution unless accurate initial iterates are provided. A major contribution of the present investigation is to demonstrate that a two-step algorithm recently developed by Wen *et al.* [18], in which the Euler–Lagrange equations are advanced in pseudo-time, does not require continuation even when so-called "balance parameters" are introduced into the upper bound analysis. Of equal importance, for three canonical turbulent dynamical systems, we prove that the only steady state to which our numerical method can converge is the true (globally optimal) solution of the variational CDH problem.

We apply our scheme to compute the optimal CDH upper bounds on the heat transport in two-dimensional (2D) Rayleigh–Bénard convection between stress-free isothermal boundaries, Rayleigh's original 1916 model [4]. Indeed, one of the central challenges in studies of Rayleigh–Bénard convection is the determination of the heat transport enhancement factor, i.e. the Nusselt number Nu, as a function of the Rayleigh number Ra, Prandtl number Pr, and domain aspect ratio L, representing, respectively, the ratios of driving to damping forces, the momentum to thermal diffusivity of the fluid, and the horizontal to vertical dimension of the container. The relationship between Nu, Pr, and Ra is often presumed to be $\mathrm{Nu} \sim \mathrm{Pr}^{\alpha}\mathrm{Ra}^{\beta}$ in the asymptotic high-Ra regime, but theoretical analyses, experiments, and simulations have yielded a variety of different scaling exponents α and β. For example, Malkus' marginally stable boundary layer theory [2, 19] yields a scaling $\mathrm{Nu} \sim \mathrm{Pr}^{0}\mathrm{Ra}^{1/3}$ wherein the heat flux is independent of the fluid layer height. However, theories by Spiegel [20] and Grossman & Lohse [21] suggest that $\mathrm{Nu} \sim \mathrm{Pr}^{1/2}\mathrm{Ra}^{1/2}$, in which case boundary layer effects are negligible and the heat flux becomes independent of the molecular transport coefficients as $\mathrm{Ra} \to \infty$. More recently, the study of 2D steady heat-flux-maximizing convective solutions with no-slip boundary conditions at $\mathrm{Pr} = 7$ by Waleffe [22] indicates $\mathrm{Nu} \sim 0.115\mathrm{Ra}^{0.31}$ for $10^7 < \mathrm{Ra} \leq 10^9$. Rigorous analyses of the three-dimensional (3D) Boussinesq equations governing Rayleigh–Bénard convection show $\mathrm{Nu} \leq c\mathrm{Ra}^{1/2}$ with prefactor $0 < c < \infty$ uniformly in Pr for no-slip and isothermal [8] or fixed heat flux [23] or mixed temperature [24] boundary conditions. This bound does not contradict the result $\mathrm{Nu} \sim \mathrm{Pr}^{0}\mathrm{Ra}^{1/3}$ obtained from marginally stable boundary layer theory, but it does rule out the prediction $\mathrm{Nu} \sim \mathrm{Pr}^{1/2}\mathrm{Ra}^{1/2}$ at large Prandtl numbers. And indeed, for convection between no-slip boundaries in the infinite-Pr limit, there exist rigorous upper bounds of the form

$\mathrm{Nu} \leq C\mathrm{Ra}^{1/3}$, where C depends on $\log \mathrm{Ra}$ [25, 26] or even only on $\log(\log \mathrm{Ra})$ [27]. On the other hand in Rayleigh's original 1916 model, the heat transport satisfies $\mathrm{Nu} < c\mathrm{Ra}^{5/12}$ [28, 29]. This 5/12 scaling exponent for the upper bound also holds for stress-free isothermal boundary conditions in 3D configurations when $\mathrm{Pr} = \infty$ [30, 31]. However, since piecewise linear functions were utilized as the background profiles, the upper bounds obtained by Otero [28] and Whitehead & Doering [29] are not the optimal ones within the CDH variational scheme. A second contribution of the present work is the computation of the optimal upper bounds on Nu as a function of Ra and L for 2D stress-free isothermal boundary conditions and arbitrary Pr. We obtain the optimal bounds by numerically solving the full background problem using the two-step algorithm developed by Wen *et al.* [18] and, as indicated above, find that in practice our scheme is much easier to implement and significantly more efficient than prior approaches employing Newton iteration and numerical continuation. Consequently, we hope that the computational methodology outlined here will enable a much broader community of researchers interested in transport in turbulent dynamical systems to employ the CDH formalism.

The reminder of this paper is organized as follows. In the next section, we derive the Euler–Lagrange equations for the relevant variational optimization problem, outline the numerical scheme to solve these equations, and prove that the only steady state to which the numerical algorithm will converge is the true solution. The upper bound computations for varying Rayleigh number and aspect ratio are described and analyzed in section III, and our conclusions are given in section IV.

II. PROBLEM FORMATION AND COMPUTATIONAL METHODOLOGY

The 2D (x–z) dimensionless Boussinesq equations for thermal convection are:

$$\frac{1}{\mathrm{Pr}}\left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u}\cdot\nabla\mathbf{u}\right) + \nabla p = \nabla^2\mathbf{u} + \mathrm{Ra}\hat{\mathbf{k}}T, \tag{1}$$

$$\nabla\cdot\mathbf{u} = 0, \tag{2}$$

$$\frac{\partial T}{\partial t} + \mathbf{u}\cdot\nabla T = \nabla^2 T, \tag{3}$$

where $\mathbf{u} = u\hat{\mathbf{i}} + w\hat{\mathbf{k}}$ is the velocity field, P is the pressure, and T is the temperature. At the upper and lower walls, located at $z = 0, 1$, the velocity field satisfies no-penetration and stress-free boundary conditions and the temperature is held fixed at the dimensional values 1 and 0, respectively; and all fields are L-periodic in x, as shown in Fig. 1. This system is governed by three control parameters: the domain aspect ratio L, the Prandtl number $\mathrm{Pr} = \nu/\kappa$, the ratio of the kinematic viscosity ν to the thermal diffusivity κ of the fluid, and the Rayleigh number $\mathrm{Ra} = \alpha g(T_{bot} - T_{top})H^3/(\nu\kappa)$, the normalized temperature drop across the layer, where α is the thermal expansion coefficient, g is the gravitational acceleration, $T_{bot} - T_{top}$ is the dimensional temperature change across the layer, and H is the layer depth. A primary quantity of interest in convection is the Nusselt number Nu, the ratio of the heat transport in the presence of convective motion to the conductive heat transport in the absence of fluid motion. One goal of this paper is to compute the optimal upper bounds on the Nusselt number,

$$\mathrm{Nu} = 1 + \lim_{\widetilde{t}\to\infty} \frac{1}{\widetilde{t}} \int_0^{\widetilde{t}} \langle wT\rangle dt,$$

within the CDH variational framework, where the angle brackets denote the spatial average, i.e. for some function f

$$\langle f\rangle = \frac{1}{L}\int_0^1\int_0^L f dx dz. \tag{4}$$

Taking the curl of Eq. (1) yields the evolution equation for the (negative) scalar vorticity compnent $\Omega = \partial w/\partial x - \partial u/\partial z$,

$$\frac{1}{\mathrm{Pr}}\left(\frac{\partial \Omega}{\partial t} + \mathbf{u}\cdot\nabla\Omega\right) = \nabla^2\Omega + \mathrm{Ra}\frac{\partial T}{\partial x}, \tag{5}$$

where $\Omega|_{z=0,1} = 0$ can be deduced from the no-penetration and stress-free boundary conditions on the velocity field. Note that homogenous boundary conditions on Ω are not realized for 3D stress-free Rayleigh–Bénard convection; consequently, the quadratic *enstrophy* constraint can only be imposed in the 2D stress-free convection problem to thereby reduce the upper bounds on the heat transport [28, 29, 31].

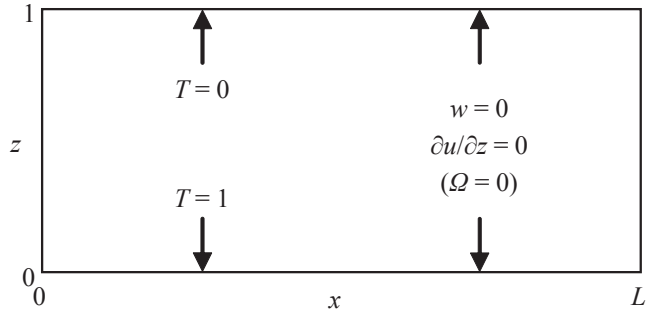

FIG. 1. Geometry and boundary conditions for the 2D stress-free convection problem.

A. Upper Bound Theory (CDH Formalism)

In the CDH upper bound theory the temperature $T(x,z,t)$ is decomposed into a time-independent background profile $\tau(z)$ carrying the inhomogeneous boundary conditions plus a nonlinear fluctuation $\theta(x,z,t)$ satisfying homogeneous boundary conditions:

$$T(x,z,t) = \tau(z) + \theta(x,z,t), \tag{6}$$

where $\tau(0) = 1$, $\tau(1) = 0$, and $\theta(x,0,t) = \theta(x,1,t) = 0$. Ensuring that appropriate test background profiles satisfy a certain spectral constraint produces rigorous upper bounds on global transport properties of the flow (whether laminar or turbulent).

By combining the energy and enstrophy balances [28, 29], the Nusselt number in this problem can be expressed as

$$\mathrm{Nu} \equiv nu - \frac{1}{1-b} \lim_{\tilde{t}\to\infty} \frac{1}{\tilde{t}} \int_0^{\tilde{t}} \mathcal{Q} dt, \tag{7}$$

where

$$nu = \frac{1}{1-b}\left(\int_0^1 (\tau')^2 dz - b\right), \tag{8}$$

$$\mathcal{Q} = \left\langle |\nabla\theta|^2 + \frac{a}{\mathrm{Ra}^{3/2}}|\nabla\Omega|^2 + \frac{b}{\mathrm{Ra}}\Omega^2 + 2\tau' w\theta - \frac{a}{\mathrm{Ra}^{1/2}}\Omega\theta_x \right\rangle, \tag{9}$$

primes denote ordinary differentiation with respect to z, a subscript denotes partial differentiation with respect to the given variable, and a and b are scalar "balance parameters" for the global energy and enstrophy constraints. Thus, if the background profile $\tau(z)$ and coefficients $a > 0$ and $0 < b < 1$ can be chosen so that $\mathcal{Q} \geq 0$ for all functions $\theta = \vartheta(x,z)$, $\Omega = \Omega(x,z)$, and $w = W(x,z)$ satisfying periodic boundary conditions in x and homogeneous Dirichlet conditions in z and the local constraint

$$\nabla^2 W - \Omega_x = 0, \tag{10}$$

which is derived from the definition of Ω and the continuity Eq. (2), then nu is an upper bound on Nu. Since W is a linear functional of Ω via Eq. (10), $\mathcal{Q}$ is indeed a quadratic form in terms of ϑ and Ω. Hence, the positivity constraint for this quadratic form is equivalent to a spectral constraint for the self-adjoint operator inside $\mathcal{Q}$, namely the non-negativity of the ground state eigenvalue λ^0 of the self-adjoint problem [28]:

$$-2\nabla^2\vartheta + 2W\tau' + \frac{a}{\mathrm{Ra}^{1/2}}\Omega_x = \lambda\vartheta, \tag{11}$$

$$-2\nabla^2\Omega_x + \frac{2b}{a}\mathrm{Ra}^{1/2}\Omega_x - \mathrm{Ra}\vartheta_{xx} + \frac{\mathrm{Ra}^{3/2}}{a}\gamma_{xx} = \frac{\mathrm{Ra}^{3/2}}{a}\lambda\Omega_x, \tag{12}$$

$$\nabla^2\gamma + 2\vartheta\tau' = 0, \tag{13}$$

$$\nabla^2 W - \Omega_x = 0, \tag{14}$$

where $\gamma(x,z)$ is the Lagrange-multiplier field enforcing the local constraint (i.e. Eq. (10)). To obtain the optimal upper bounds, we need to minimize nu subject to the spectral constraint $\lambda^0 \geq 0$.

B. Euler–Lagrange Equations

The optimal bounds are achieved when $\inf(\lambda^0) = 0$. To solve the background variational problem efficiently, we work directly with the corresponding Euler–Lagrange equations derived by identifying a Lagrange functional for this optimization problem:

$$\mathcal{L} = \frac{1}{1-b}\left(\int_0^1 (\tau')^2 dz - b\right) - \left\langle |\nabla\vartheta|^2 + \frac{a}{\mathrm{Ra}^{3/2}}|\nabla\Omega|^2 + \frac{b}{\mathrm{Ra}}\Omega^2 + 2\tau' W\vartheta - \frac{a}{\mathrm{Ra}^{1/2}}\Omega\vartheta_x \right\rangle - \left\langle \gamma(\nabla^2 W - \Omega_x)\right\rangle. \quad (15)$$

The first term in $\mathcal{L}$ is the objective functional to be extremized, and the second and third terms correspond to the spectral constraint and the local constraint Eq. (10), respectively. Note that in the second term, the scalar Lagrange multiplier has been absorbed into the spectral constraint by rescaling ϑ, Ω, W and γ, and in the third term the Lagrange-multiplier field $\gamma(x,z)$, like the direct field $\vartheta(x,z)$, satisfies periodic boundary conditions in x and homogenous Dirichlet conditions in z. The first variations (Frechet derivatives) of this functional with respect to τ, ϑ, Ω, W, γ, b, and a (i.e. $\delta\mathcal{L}/\delta\tau = 0$, $\delta\mathcal{L}/\delta\vartheta = 0$, etc.), respectively, yield the Euler–Lagrange equations

$$-\tau'' + (1-b)(\overline{W\vartheta})_z = 0, \quad (16)$$

$$-2\nabla^2\vartheta + 2W\tau' + \frac{a}{\mathrm{Ra}^{1/2}}\Omega_x = 0, \quad (17)$$

$$-\frac{2a}{\mathrm{Ra}^{3/2}}\nabla^2\Omega + \frac{2b}{\mathrm{Ra}}\Omega - \frac{a}{\mathrm{Ra}^{1/2}}\vartheta_x + \gamma_x = 0, \quad (18)$$

$$\nabla^2\gamma + 2\vartheta\tau' = 0, \quad (19)$$

$$\nabla^2 W - \Omega_x = 0, \quad (20)$$

$$b - 1 + \left\{ \frac{\mathrm{Ra}\left(\int_0^1 (\tau')^2 dz - 1\right)}{\langle\Omega^2\rangle} \right\}^{1/2} = 0, \quad (21)$$

$$\left\langle \frac{1}{\mathrm{Ra}^{3/2}}|\nabla\Omega|^2 - \frac{1}{\mathrm{Ra}^{1/2}}\Omega\vartheta_x \right\rangle = 0, \quad (22)$$

where $\overline{(\cdot)} = \frac{1}{L}\int_0^L dx(\cdot)$. After taking an x-derivative and rescaling, we can rewrite Eq. (18) as

$$-2\nabla^2\Omega_x + \frac{2b}{a}\mathrm{Ra}^{1/2}\Omega_x - \mathrm{Ra}\vartheta_{xx} + \frac{\mathrm{Ra}^{3/2}}{a}\gamma_{xx} = 0. \quad (23)$$

Since $\langle\Omega\cdot\delta\mathcal{L}/\delta\Omega\rangle = 0$, namely

$$\left\langle -\frac{2a}{\mathrm{Ra}^{3/2}}|\nabla\Omega|^2 - \frac{2b}{\mathrm{Ra}}\Omega^2 + \frac{a}{\mathrm{Ra}^{1/2}}\Omega\vartheta_x - \Omega\gamma_x \right\rangle = 0, \quad (24)$$

hence from Eq. (22), Eq. (24) becomes

$$a = \frac{\mathrm{Ra}^{3/2}\langle\gamma\Omega_x\rangle - 2b\mathrm{Ra}^{1/2}\langle\Omega^2\rangle}{\langle|\nabla\Omega|^2\rangle}. \quad (25)$$

Solving the Euler–Lagrange equations (16), (17), (23), (19)–(21), and (25) subject to the spectral constraint $\lambda^0 \geq 0$ yields the optimal background profile and upper bound.

Before solving these equations, it is instructive to analyze their structure. First, we observe that a subset of the Euler–Lagrange equations (17), (23), (19), and (20), is identical to the marginally stable eigenvalue system (i.e., Eqs. (11)–(14) for $\lambda^0 = 0$). Moreover, the Euler–Lagrange equations derived here have a similar mean-field (or quasi-linear) structure to those in Plasting & Kerswell [17] and Wen *et al.* [18]. In particular, using a Fourier series representation, the solution can be expressed as

$$\begin{bmatrix} \vartheta \\ \Omega_x \\ \gamma \\ W \end{bmatrix} = \sum_{n=1}^{N} \begin{bmatrix} \vartheta_n^*(z) \\ \Omega_{xn}^*(z) \\ \gamma_n^*(z) \\ W_n^*(z) \end{bmatrix} \cos(nkx), \quad (26)$$

where n and $k = 2\pi/L$ are the (integer) horizontal mode number and fundamental wavenumber, respectively, and N is the (generally) *finite* truncation mode number, i.e., the series in Eq. (26) generally terminates. The true solution,

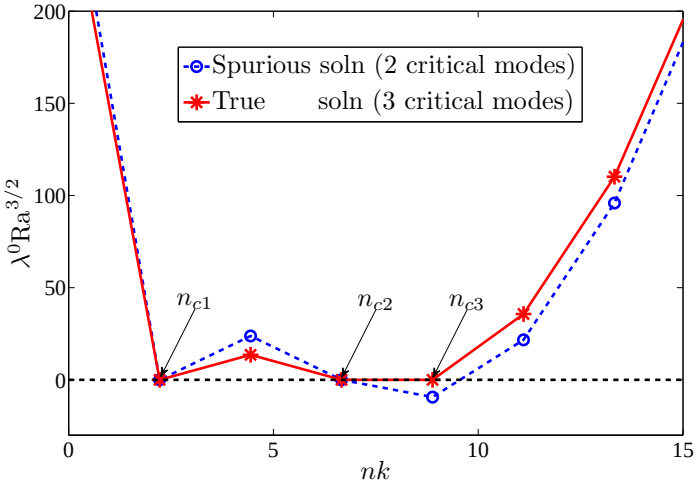

FIG. 2. Ground state eigenvalue distribution for the true (solid) and spurious (dashed) solutions of the Euler–Lagrange equations at $\mathrm{Ra} = 562341, L = 2\sqrt{2}$. Both solutions satisfy the Euler–Lagrange equations. However, the true solution, which includes all three critical modes (n_{c1}, n_{c2}, n_{c3}) for this case, also satisfies the spectral constraint $\lambda^0 \geq 0$; in contrast, the spurious solution, which captures only two of the critical modes (n_{c1}, n_{c2}), does not satisfy the spectral constraint, i.e., $\lambda^0 < 0$ at horizontal wavenumber $n_{c3}k$.

which not only satisfies the Euler–Lagrange equations but also the spectral constraint, is unique and determined solely by the critical modes $n_c k$, i.e. $\vartheta_n^* = \Omega_{xn}^* = \gamma_n^* = W_n^* \equiv 0$ for $n \neq n_c$. In contrast, the solution of the Euler–Lagrange equations is generally *not* unique: the true solution, or global optimal, includes all the critical modes and yields an *admissible* $\tau(z)$ satisfying the spectral constraint, while the spurious solutions, saddles which we refer to as local optimals, omit certain critical modes and/or include incorrect modes and yield an *inadmissible* $\tau(z)$ for which the ground state eigenvalue becomes negative at some horizontal wavenumber, as shown in Fig. 2. In the context of Rayleigh–Bénard convection, this sort of eigenstructure was first computed in early numerical work by Doering & Hyman [12], who used a finite-difference relaxation method to solve the *single*-critical-mode Euler–Lagrange equations; see in particular their Figure 2.

Thus, one of the central challenges in the numerical solution of the Euler–Lagrange equations is the determination of the *a priori* unknown critical modes, which usually requires the use of Newton iteration plus continuation [17]. Here we overcome this difficulty by utilizing a two-step algorithm we recently proposed in our previous computations of upper bounds on the heat transport in porous medium convection [18]. Crucially, the sole balance parameter arising in the porous medium problem can be scaled out of the Euler–Lagrange equations, and the two-step algorithm has been shown to be efficient and accurate in that case [18]. In the current problem, the balance parameters b and a cannot be scaled out of the Euler–Lagrange equations (21) and (25), introducing new sources of uncertainty regarding the efficacy of our algorithm. Nevertheless, in the following sections, we demonstrate that the two-step algorithm can indeed be successfully applied to background optimization problems even with balance parameters. Moreover, we prove that the only steady state to which this numerical approach can converge is the global optimal.

C. Two-Step Algorithm

The key idea of the first step is to convert the time-invariant Euler–Lagrange system into a *time-dependent* dynamical system by incorporating certain specific time derivatives into Eqs. (16)–(25). Then the solutions of the original Euler–Lagrange equations, which correspond exactly to the steady states of the "time-dependent" Euler–Lagrange equations, can be easily obtained by solving the extended equations numerically using a *time-marching* method with non-zero initial data for all horizontal modes $1 \leq n \leq N$. In this section, the behavior of two different types of time-dependent Euler–Lagrange systems, hereafter referred to as system A and system B, is investigated.

In system A, the terms $\partial_t \vartheta$ and $\partial_t \Omega_x$ are added, respectively, to Eqs. (17) and (23), as follows

$$\partial_t \vartheta - \frac{\delta \mathcal{L}}{\delta \vartheta} = 0 \quad \Rightarrow \quad \partial_t \vartheta - 2\nabla^2 \vartheta + 2W\tau' + \frac{a}{\mathrm{Ra}^{1/2}} \Omega_x = 0, \tag{27}$$

$$\partial_t \Omega_x - \frac{\delta \mathcal{L}}{\delta \Omega_x} = 0 \quad \Rightarrow \quad \partial_t \Omega_x - 2\nabla^2 \Omega_x + \frac{2b}{a} \mathrm{Ra}^{1/2} \Omega_x - \mathrm{Ra} \vartheta_{xx} + \frac{\mathrm{Ra}^{3/2}}{a} \gamma_{xx} = 0, \tag{28}$$

where τ, γ, W, b and a are determined from Eqs. (16), (19)–(21), and (25).

In system B, in addition to including the time-derivative terms in Eqs. (27) and (28), $\partial_t\tau$ is also added to Eq. (16) as follows

$$\partial_t\tau + \frac{\delta\mathcal{L}}{\delta\tau} = 0 \quad \Rightarrow \quad \partial_t\tau - \tau_{zz} + (1-b)(\overline{W\vartheta})_z = 0. \tag{29}$$

Moreover, the balance parameters can also be "freed up" by adding $\partial_t a$ and $\partial_t b$ to $\partial\mathcal{L}/\partial a = 0$ and $\partial\mathcal{L}/\partial b = 0$, respectively, that is

$$\partial_t a + \frac{\partial\mathcal{L}}{\partial a} = 0 \quad \Rightarrow \quad \partial_t a - \left\langle \frac{|\nabla\Omega|^2}{\mathrm{Ra}^{3/2}} - \frac{\Omega\vartheta_x}{\mathrm{Ra}^{1/2}} \right\rangle = 0, \tag{30}$$

$$\partial_t b + \frac{\partial\mathcal{L}}{\partial b} = 0 \quad \Rightarrow \quad \partial_t b + \frac{\int_0^1 (\tau')^2 dz - 1}{(1-b)^2} - \frac{\langle\Omega^2\rangle}{\mathrm{Ra}} = 0. \tag{31}$$

Then γ and W can be determined from Eqs. (19) and (20), respectively. The motivation for these choices is the observation that

$$\frac{d\mathcal{L}(\tau,\vartheta,\Omega;a,b,\mathrm{Ra})}{dt} = \left\langle \tau_t\frac{\delta\mathcal{L}}{\delta\tau} + \vartheta_t\frac{\delta\mathcal{L}}{\delta\vartheta} + \Omega_t\frac{\delta\mathcal{L}}{\delta\Omega} \right\rangle + a_t\frac{\partial\mathcal{L}}{\partial a} + b_t\frac{\partial\mathcal{L}}{\partial b}. \tag{32}$$

The signs of the diffusive terms in the Frechet derivatives for τ, ϑ and Ω dictate how the respective time derivatives are added but the situation is unclear for the balance parameters. A little experimentation, however, clearly indicates the "correct" sign to take so that Eq. (32) becomes

$$\frac{d\mathcal{L}(\tau,\vartheta,\Omega;a,b,\mathrm{Ra})}{dt} = \langle -\tau_t^2 + \vartheta_t^2 + \Omega_t^2 \rangle - a_t^2 - b_t^2. \tag{33}$$

By not adding τ_t, system A is an attempt to find an algorithm in which Nu increases monotonically with time to the global bound (for fixed balance parameters). System B includes τ_t to discern whether this makes the time-dependent system better behaved under temporal integration albeit at the expense of non-monotonic behavior in Nu. In the next section, we prove that if either of systems A or B (for fixed balance parameters) converges to a steady state then this state is the true solution, i.e., the global optimal.

In our computations, temporal discretization is achieved using the Crank–Nicolson method for the linear terms (i.e., terms linear in the variable being advanced) and a two-step Adams–Bashforth method for the nonlinear (i.e., remaining) terms, while a Chebyshev spectral collocation method is used for spatial discretization. As both systems were time-advanced, the balance parameters converged to specific values (see Fig. 3); Nu converged to the bound; and the ground state spectrum became marginally stable (see Fig. 4) with noncritical ϑ_n^* and Ω_{xn}^* converging to zero. It was found that although system A does generate a monotonically-increasing estimate for the global bound (even sometimes for free balance parameters), it required a good initial condition – which could be obtained by employing the time-marching solver for a short integration period with fixed a and b. In contrast, system B was more robust: it worked even for initial conditions distant from the global optimal and allowed larger time steps to be taken, although the inclusion of τ_t, a_t and b_t in Eqs. (29), (30) and (31), respectively, destroyed the monotonic-increasing property of Nu (see the left-hand plot of Fig. 4). Therefore, in the follow sections, system B is utilized for all the computations.

Generally, this first step of the two-step algorithm works efficiently and accurately for small and moderate Ra. However, for large Ra, e.g., $\mathrm{Ra} > 10^8$, the time-marching algorithm converges slowly owing to stringent restrictions on the allowable size of the numerical time step. Thus, for this range of Ra, we employ a second step: a Newton–Kantorovich iterative method [32]. As discussed previously, Newton iteration generally will converge to a spurious solution unless a *very* good initial condition that includes the correct critical modes is used. However, from the first step, both the critical modes for which $\lambda^0 \approx 0$ and the noncritical modes with $\hat{\vartheta}_n$ and $\hat{\Omega}_{xn}$ converging to zero can be accurately identified. Thus, the output from the first step can be used as the required (very) good initial condition for the second step. Moreover, during the first step we observe that the non-critical components of ϑ_n^* and Ω_{xn}^* require a comparably long time to converge to extremely small values even as the balance parameters have converged or oscillate in a very small range. Therefore, to improve the efficiency of the second step, we utilize the constant values of a and b as obtained from the first step. More details about the Newton–Kantorovich algorithm for this problem can be found in Appendix A.

Strictly speaking, after completing this two-step computation, the resulting background profile $\tau(z)$ should be substituted into the eigensystem, Eqs. (11)–(14), to verify that the spectral constraint is satisfied. Here we only verify the spectral constraint for $\mathrm{Ra} \leq 10^7$ owing to the numerical challenges associated with accurately solving the extremely ill-conditioned eigenvalue problem itself at large Ra. Nevertheless, for the cases with $\mathrm{Ra} > 10^7$, the critical modes can still be accurately predicted by identifying the noncritical modes for which ϑ_n^* and Ω_{xn}^* are damped as time evolves.

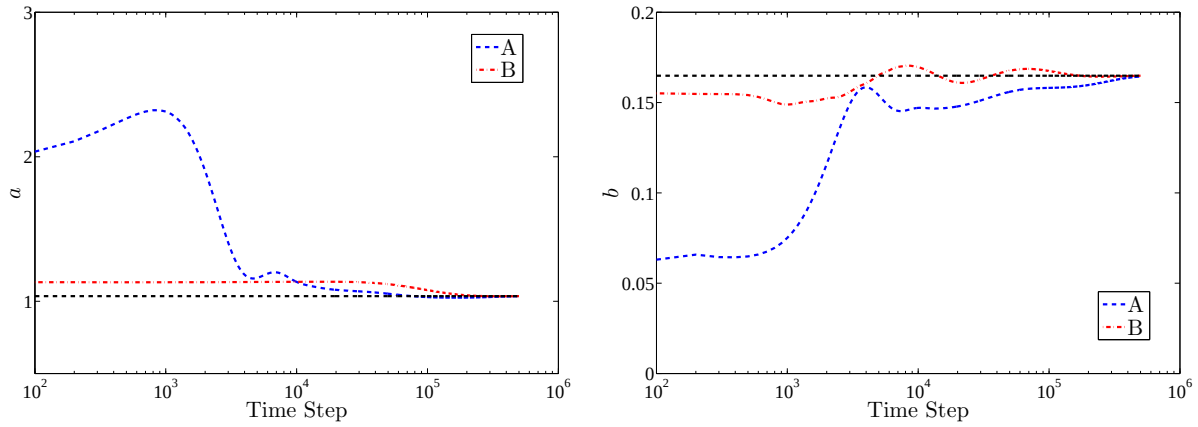

FIG. 3. Time-evolution, for Ra = 562341 and $L = 2\sqrt{2}$, of the balance parameters. As t increases, the balance parameters (Left: a; Right: b) converge to constant values (indicated by the black horizontal dashed lines). The same initial conditions and time steps are utilized for systems A and B. Note that the balance parameters in B exhibit far less variability than do those in A.

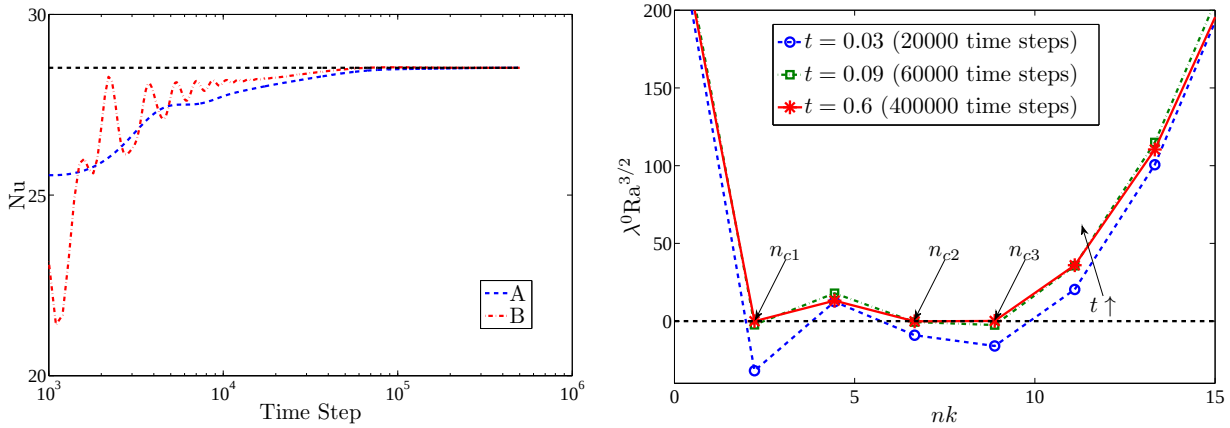

FIG. 4. Time-evolution, for Ra = 562341 and $L = 2\sqrt{2}$, of Nu (Left) and the ground state eigenvalue λ^0 (Right). As t increases, Nu converges to the bound (Left) and the ground state spectrum becomes marginally stable ($\lambda^0 \geq 0$), with the critical modes occurring at those wavenumbers for which $\lambda^0 = 0$ (Right). The same initial conditions and time steps are utilized for systems A and B, and the ground state eigenvalue in the right-hand plot is from B. Note that in the left-hand plot, Nu increases up to the bound monotonically for A.

D. The Global Optimal as the only Possible Steady Attractor

We next analyze the fixed-balance-parameter optimization problem to demonstrate that the global attractor is the only steady attractor of the time-augmented variational problem. Allowing the balance parameters to also vary can only further (temporally) destabilize states – so the global optimal is no longer guaranteed to be an attractor, but our numerical results suggest that it still is.

Let $(\tau, \vartheta, \Omega, \gamma, W)$ satisfy the Euler–Lagrange equations (16), (17), (19)–(20), and (23) and be made up of J x-wavenumbers $n_{cj}k$ so that

$$\tau = \tau(z), \qquad \begin{bmatrix} \vartheta \\ \Omega_x \\ \gamma \\ W \end{bmatrix} = \sum_{j=1}^{J} \begin{bmatrix} \vartheta^*_{n_{cj}}(z) \\ \Omega^*_{x\, n_{cj}}(z) \\ \gamma^*_{n_{cj}}(z) \\ W^*_{n_{cj}}(z) \end{bmatrix} \cos\left(n_{cj} k x\right). \tag{34}$$

As discussed previously, there are many such solutions but only one, the true solution or global optimal, satisfies the spectral constraint ($\lambda^0 \geq 0$ from the spectral problem, Eqs. (11)–(14)). These solutions – the global, true solution and the local, spurious optimals – are steady-state solutions of the time-derivative-enhanced equations: (19)–(20) and

(27)–(29). We now establish that all of the spurious (locally optimal) solutions are linearly unstable in the time-dependent system and so can never be attracting endstates of the time-dependent system. We further show that the globally optimal solution *is* linearly stable and therefore an attractor. The method of proof is relatively straightforward and, since it exploits the spectral constraint, generalizes easily to other canonical upper bound problems including plane Couette flow (e.g., [17]) and porous medium convection (e.g., [18]): see Appendix B. Extending the result that the global solution is an attractor when the balance parameters also vary does not, however, seem possible generally. Instead, this has to be verified on a case-by-case basis (via direct stability analysis) since the balance parameters enter into the optimization problems in rather arbitrary ways. Nevertheless, what can be proved is almost as strong as could be desired: *if* the time-dependent system converges to a steady state, then this steady state *is* the global optimal.

To establish this result, let $(\hat{\tau},\hat{\vartheta},\hat{\Omega},\hat{\gamma},\hat{W})$ be a possibly large perturbation of the steady state solution $(\tau,\vartheta,\Omega,\gamma,W)$ and consider the temporal evolution equations for this perturbation for *fixed* balance parameters

$$\partial_t\hat{\tau} = \hat{\tau}_{zz} - (1-b)\overline{(\hat{W}\vartheta + W\hat{\vartheta} + \hat{\boldsymbol{W}}\hat{\boldsymbol{\vartheta}})}_z, \tag{35}$$

$$\partial_t\hat{\vartheta} = 2\nabla^2\hat{\vartheta} - 2(\hat{W}\tau_z + W\hat{\tau}_z + \hat{\boldsymbol{W}}\hat{\boldsymbol{\tau}}_{\boldsymbol{z}}) - \frac{a}{\mathrm{Ra}^{1/2}}\hat{\Omega}_x, \tag{36}$$

$$\partial_t\hat{\Omega}_x = 2\nabla^2\hat{\Omega}_x - \frac{2b}{a}\mathrm{Ra}^{1/2}\hat{\Omega}_x + \mathrm{Ra}\hat{\vartheta}_{xx} - \frac{\mathrm{Ra}^{3/2}}{a}\hat{\gamma}_{xx}, \tag{37}$$

$$0 = \nabla^2\hat{\gamma} + 2(\hat{\vartheta}\tau_z + \vartheta\hat{\tau}_z + \hat{\boldsymbol{\vartheta}}\hat{\boldsymbol{\tau}}_{\boldsymbol{z}}), \tag{38}$$

$$0 = \nabla^2\hat{W} - \hat{\Omega}_x, \tag{39}$$

where the perturbation fields satisfy periodic boundary conditions in x and homogeneous boundary conditions in z (in particular, $\hat{\tau}(0)=\hat{\tau}(1)=0$). Then, the volume integrals $\langle\hat{\tau}\times\text{Eq.}(35)\rangle$, $\langle\hat{\vartheta}\times\text{Eq.}(36)\rangle$, $\langle-\int^x\hat{\Omega}(x')dx'\times\text{Eq.}(37)\rangle$, $\langle\hat{W}\times\text{Eq.}(38)\rangle$, and $\langle\hat{\gamma}\times\text{Eq.}(39)\rangle$ yield

$$\langle\,\tfrac{1}{2}\hat{\tau}^2\,\rangle_t = -\langle\,\hat{\tau}_z^2\,\rangle + (1-b)\langle\,(\hat{W}\vartheta + W\hat{\vartheta} + \hat{W}\hat{\vartheta})\hat{\tau}_z\,\rangle, \tag{40}$$

$$\langle\,\tfrac{1}{2}\hat{\vartheta}^2\,\rangle_t = -2\langle\,|\nabla\hat{\vartheta}|^2\,\rangle - 2\langle\,\hat{\vartheta}\hat{W}\tau_z + \hat{\vartheta}W\hat{\tau}_z + \hat{\vartheta}\hat{W}\hat{\tau}_z\,\rangle - \frac{a}{\mathrm{Ra}^{1/2}}\langle\,\hat{\vartheta}\hat{\Omega}_x\,\rangle, \tag{41}$$

$$\langle\,\tfrac{1}{2}\hat{\Omega}^2\,\rangle_t = -2\langle\,|\nabla\hat{\Omega}|^2\,\rangle - \frac{2b}{a}\mathrm{Ra}^{1/2}\langle\,\hat{\Omega}^2\,\rangle - \mathrm{Ra}\langle\,\hat{\vartheta}\hat{\Omega}_x\,\rangle + \frac{\mathrm{Ra}^{3/2}}{a}\langle\,\hat{\gamma}\hat{\Omega}_x\,\rangle, \tag{42}$$

$$0 = \langle\,\hat{W}\nabla^2\hat{\gamma}\,\rangle + 2\langle\,\hat{W}\hat{\vartheta}\tau_z + \hat{W}\vartheta\hat{\tau}_z + \hat{W}\hat{\vartheta}\hat{\tau}_z\,\rangle, \tag{43}$$

$$0 = \langle\,\hat{\gamma}\nabla^2\hat{W}\,\rangle - \langle\,\hat{\gamma}\hat{\Omega}_x\,\rangle. \tag{44}$$

Adding $2\times$Eq. (40), $(1-b)\times$Eq. (41), $a(1-b)/\mathrm{Ra}^{3/2}\times$Eq. (42), $(1-b)\times$Eq. (44) and subtracting $(1-b)\times$Eq. (43) gives

$$\left\langle \hat{\tau}^2 + \frac{1-b}{2}\hat{\vartheta}^2 + \frac{a(1-b)}{2\mathrm{Ra}^{3/2}}\,\hat{\Omega}^2 \right\rangle_t = -2\langle\,\hat{\tau}_z^2\,\rangle - 2(1-b)\left\langle\, |\nabla\hat{\vartheta}|^2 + \frac{a}{\mathrm{Ra}^{3/2}}|\nabla\hat{\Omega}|^2 + \frac{b}{\mathrm{Ra}}\hat{\Omega}^2 + \hat{W}\hat{\vartheta}(2\tau_z+\hat{\tau}_z) - \frac{a}{\mathrm{Ra}^{1/2}}\hat{\Omega}\hat{\vartheta}_x \right\rangle \tag{45}$$

Interestingly, the second term on the right-hand side is the spectral constraint based upon a linear combination of the base and perturbation background fields with only one higher-than-quadratic term, $\langle\hat{W}\hat{\vartheta}\hat{\tau}_z\rangle$, reflecting the fully nonlinear nature of the relation. Despite this relative simplicity, no global convergence result seems forthcoming so we instead *linearize* to consider local stability properties of the steady solutions. Now the second term is the spectral constraint on the steady background field τ, which is negative semidefinite if and only if τ satisfies the spectral constraint $\mathcal{Q}\geq 0$ in Eq. (9). At this point it is worth noting that eigenfunctions of the linearized evolution operator (Eqs. (35)–(39) with the nonlinear, bold terms dropped) take one of two distinct forms: either

$$\text{Type I}: \qquad \hat{\tau}=0, \qquad \begin{bmatrix}\hat{\vartheta}\\ \hat{\Omega}_x\\ \hat{\gamma}\\ \hat{W}\end{bmatrix} = \begin{bmatrix}\tilde{\vartheta}(z)\\ \tilde{\Omega}(z)\\ \tilde{\gamma}(z)\\ \tilde{W}(z)\end{bmatrix}\cos(nkx), \tag{46}$$

where $n\notin\{n_{c1},n_{c2},\ldots,n_{cJ}\}$, i.e., the perturbation shares no common wavenumber with the underlying steady state

(see Eq. (34)), or

$$\text{Type II}: \qquad \hat{\tau} = \hat{\tau}(z) \neq 0, \qquad \begin{bmatrix} \hat{\vartheta} \\ \hat{\Omega}_x \\ \hat{\gamma} \\ \hat{W} \end{bmatrix} = \sum_{j=1}^{J} \begin{bmatrix} \tilde{\vartheta}_{n_{cj}}(z) \\ \tilde{\Omega}_{x\,n_{cj}}(z) \\ \tilde{\gamma}_{n_{cj}}(z) \\ \tilde{W}_{n_{cj}}(z) \end{bmatrix} \cos(n_{cj} k x). \tag{47}$$

The crucial observation is that Type I eigenfunctions are also eigenfunctions of the spectral constraint operator defined by Eqs. (11)–(14). Hence if the spectral constraint is not satisfied as is the case for all spurious (local optimal) solutions, there is a temporally *unstable* Type I eigenfunction: hence the spurious solutions are not attracting states. Moreover, this linear instability persists when the balance parameters are also allowed to vary because Type I eigenfunctions cannot contain balance parameter perturbations (perturbations in the balance parameters are solely carried by Type II eigenfunctions).

In the case of the global optimal, the spectral constraint *is* satisfied and all Type I eigenfunctions are temporally damped (stable) since they have a spectral constraint eigenvalue λ^0 that is strictly positive. For perturbations spanned by Type II eigenfunctions, it is formally possible for the spectral constraint to be marginally satisfied *and* for $\langle\, \hat{\tau}_z^2 \,\rangle$ to vanish at some instant: i.e., strict monotonic decay of the functional on the left-hand side of Eq. (45) is not assured. However, if this happens (requiring

$$\begin{bmatrix} \tilde{\vartheta}_{n_{cj}}(z) \\ \tilde{\Omega}_{x\,n_{cj}}(z) \\ \tilde{\gamma}_{n_{cj}}(z) \\ \tilde{W}_{n_{cj}}(z) \end{bmatrix} = \alpha_j \begin{bmatrix} \vartheta^*_{n_{cj}}(z) \\ \Omega^*_{x\,n_{cj}}(z) \\ \gamma^*_{n_{cj}}(z) \\ W^*_{n_{cj}}(z) \end{bmatrix} \qquad \forall j \in \{1, \ldots, J\} \tag{48}$$

where α_j are some real scalars), Eq. (35) indicates that $\langle\, \hat{\tau}_z^2 \,\rangle_{tt} > 0$ so this situation cannot persist, i.e., there is no neutral Type II eigenfunction. Thus, all Type II eigenfunctions are also temporally damped and the global optimal is an attractor of the time-dependent problem with fixed balance parameters.

The overall conclusion is that the spurious (local optimal) solutions can never be attractors of the time-dependent system whereas the global optimal is for fixed balance parameters and may be for the full problem (with varying balance parameters), too. Put another way, if a steady solution is reached as the endstate of the time-dependent system it will be the global optimal, but a steady endstate is not guaranteed.

III. RESULTS AND DISCUSSION

A series of computations was performed for a discrete set of $\mathrm{Ra} = 100 \times 10^{(i-1)/4}$ (for integer i) from $\mathrm{Ra} = 100$ to $\mathrm{Ra} = 10^{10}$; for $\mathrm{Ra} \geq 10^7$, computations were also performed for a range of L values from $L = 0.01$ to $L = 2\sqrt{2}$ to study the aspect-ratio dependence of the upper bounds. As Ra was increased the number of Chebyshev modes used in the vertical discretization was increased from 65 to 401. Unlike for DNS, computational memory requirements are modest for upper bound computations; indeed, all the computations reported here were performed using laptop and desktop computers. The largest Rayleigh number for which we compute a bound, $\mathrm{Ra} = 10^{10}$, was selected based on the evident convergence of the large-Ra scaling of Nu (i.e., the exponent of Ra converges to 5/12), as shown below. Linear stability analysis of the conduction solution [4] indicates that stress-free Rayleigh–Bénard convection undergoes a stationary bifurcation at a critical Rayleigh number $\mathrm{Ra}_c = 27/4\pi^4$, above which the layer becomes convectively unstable, and that at Ra_c the wavelength of the corresponding marginal mode $L_c = 2\sqrt{2}$. Furthermore, as Ra increases, the wavelength of the shortest marginally stable mode decreases; specifically, for this mode, (abusing notation) $L_c \sim 2\pi \mathrm{Ra}^{-1/4}$ in the limit $\mathrm{Ra} \to \infty$.

Figure 5 compares our optimal upper bound, obtained by numerically solving the full background field variational problem, with previously obtained numerical and analytical upper bounds by Otero [23] and Whitehead & Doering [29], who employed a piecewise linear function as the background profile, and with the heat transport achieved by steady convection solutions obtained asymptotically by Chini & Cox [33] and numerically by Souza [34]. The new result indicates $\mathrm{Nu} \leq 0.106 \mathrm{Ra}^{5/12}$ in the asymptotic high-Ra regime at fixed $L = 2\sqrt{2}$ uniformly in Pr. Clearly, the new bound constitutes a quantitative improvement over the previous bounds, although in pre-factor only. For reference, we note that for stress-free Rayleigh–Bénard convection at *infinite* Pr, Plasting & Ierley [35] estimate $\mathrm{Nu} \lesssim 0.126 \mathrm{Ra}^{5/12}$. Moreover, all of these upper bounds lie above the predicted and computed *steady* heat transport, which indicate an approximate 1/3 scaling at large Ra, and the primary bifurcation at $\mathrm{Ra} = 27/4\pi^4$ is also clearly evident. Figure 6 shows the distribution of the critical modes predicted from the first step of our computations. At large Ra, the wavenumber of the critical mode with maximum wavenumber scales as $\mathrm{Ra}^{1/4}$, which agrees very well

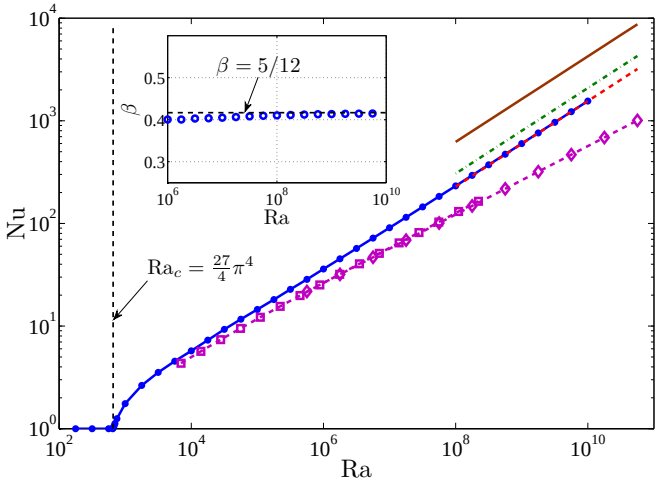

FIG. 5. Logarithmic Nu–Ra plot. The solid-dot line shows the optimal upper bound and the dashed line at large Ra is the best fit curve $\mathrm{Nu} \sim 0.106\mathrm{Ra}^{5/12}$. For reference, also included here are: the previous numerical bound from Otero [28] (dashed-dot line, $\mathrm{Nu} \sim 0.142\mathrm{Ra}^{5/12}$) and analytical bound from Whitehead & Doering [29] (solid line, $\mathrm{Nu} \sim 0.289\mathrm{Ra}^{5/12}$), and the Nu–Ra data from Chini & Cox [33] (dashed-diamond line) and Souza [34] (dashed-square line) corresponding, respectively, to the asymptotically-predicted and numerically-computed heat transport realized by *steady* cellular *solutions* of the Boussinesq equations (with $\mathrm{Pr} = 1$) maximized over the cell aspect ratio at each Ra. The inset shows the variation of the power-law scaling exponent β (recall $\mathrm{Nu} \sim \mathrm{Ra}^{\beta}$) of the new bound as a function of Ra, where β is computed using a central-differencing method. The scaling exponent of the optimal upper bound converges to 5/12 in the asymptotic high-Ra regime, consistent with the previous results from Otero [28] and Whitehead & Doering [29]. In contrast, for the steady solutions, $\beta \sim 1/3$ at large Ra in the analysis of Chini & Cox [33] and $\beta \approx 0.32$ at $\mathrm{Ra} = O(10^8)$ in the computations of Souza [34].

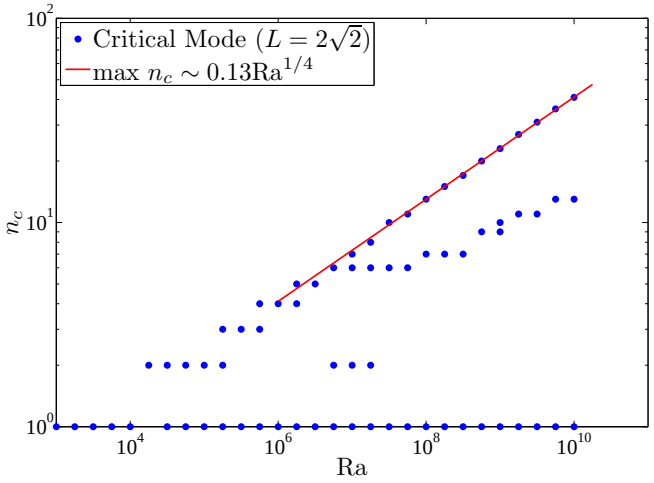

FIG. 6. Bifurcation diagram showing the number and values of the critical modes as a function of Ra for $L = 2\sqrt{2}$. The largest critical mode number $n_c \sim \mathrm{Ra}^{1/4}$ at large Ra, consistent with the analytical result from Whitehead & Doering [29].

with the analytical prediction by Whitehead & Doering [29]. Interestingly, this scaling is the same as that of the high-wavenumber marginal linear stability mode in the large-Ra limit.

The background profiles for different Ra and L are shown in Fig. 7. The optimal background profile shares many features exhibited by the horizontal and long-time mean temperature profile from DNS [36]: as Ra is increased, the boundary layers in $\tau(z)$ become thinner and the interior part of the profile becomes nearly z-independent. Moreover, for fixed Ra, the background profile becomes approximately invariant when L is large enough. After computing the optimal $\tau(z)$, a set of orthogonal eigenfunctions (see Fig. 8) can be obtained by solving the self-adjoint eigenvalue problem given by Eqs. (11)–(14). Since these eigenfunctions are extracted directly from the governing equations, they capture many features of the real flow (e.g., boundary layer structure) and should comprise a physically-preferred *a priori* basis for building reduced-order models via Galerkin projection [37]. The aspect-ratio dependence of the heat

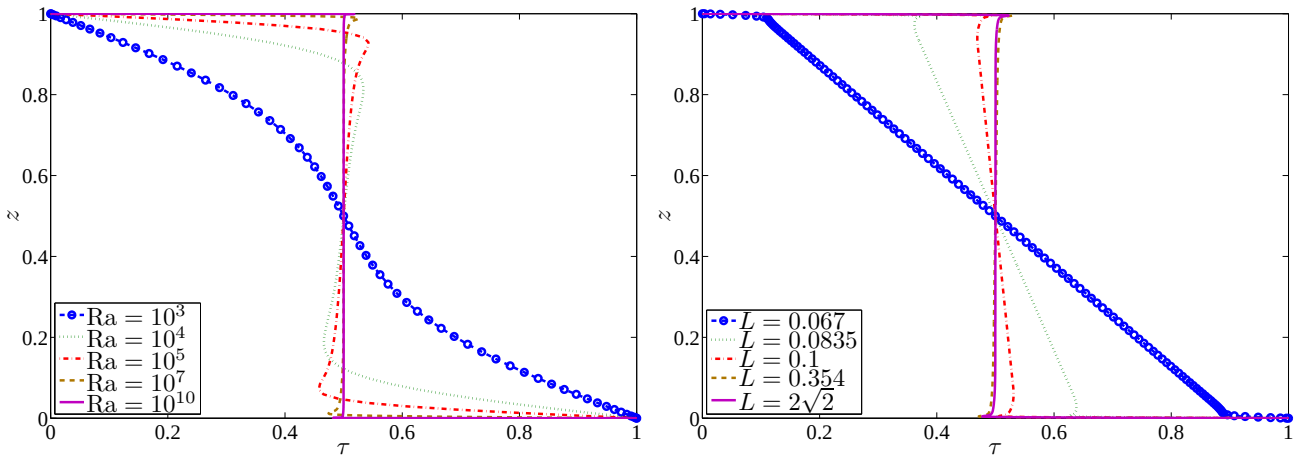

FIG. 7. Variation of $\tau(z)$ with Ra at $L = 2\sqrt{2}$ (Left) and with L at $\mathrm{Ra} = 10^8$ (Right).

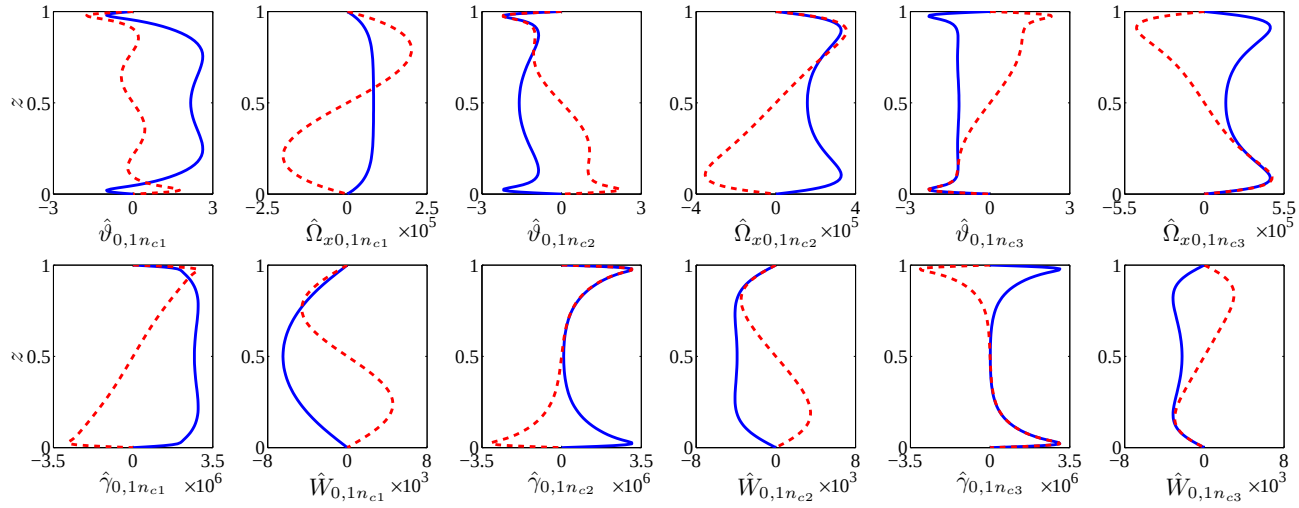

FIG. 8. Eigenfunctions, for $\mathrm{Ra} = 562341$ and $L = 2\sqrt{2}$, corresponding to the first and second lowest branches of eigenvalues (λ^0 and λ^1) at $n = n_{c1}$, n_{c2}, and n_{c3}. The ground state eigenfunctions (λ^0, solid curves) are even symmetric about $z = 1/2$, while the eigenfunctions corresponding to the first excited state (λ^1, dashed curves) are odd symmetric.

transport can also be explored through upper bound analysis, which has been shown for porous medium convection to accurately predict the $L(\mathrm{Ra})$ scaling relationship at large Ra corresponding to a given (re-normalized) Nusselt value computed using DNS [18]. Figure 9 shows the aspect-ratio dependence of the upper bounds for the Rayleigh–Bénard problem studied here for $10^7 \leq \mathrm{Ra} \leq 10^{10}$. Clearly, the upper bounds (again) capture the bifurcation occurring at the marginal stability boundary of the conduction state, corresponding to the dashed line in Fig. 9. To the left of this line, i.e. for $L < L_c$, there is no convection so that $\mathrm{Nu} \equiv 1$ and the the re-normalized Nu is constant. The re-normalized Nu quickly asymptotes to 1 as L increases, showing that for large Ra, the aspect ratio has little impact on the bounds until the domain becomes sufficiently small. This evidence suggests the existence of a "minimal flow unit" – the smallest domain for which the heat transport effectively remains constant with any further *increase* in L – although the precise size of this unit remains unclear.

IV. CONCLUSIONS

One possible explanation for the discrepancies among previously reported heat-transport scaling exponents in Rayleigh–Bénard convection is that many of these investigations did not reach the asymptotic high-Ra regime. Other explanations include uncertainties associated with non-Boussinesq effects and Pr-variations in experiments and the lack of sufficient spatiotemporal resolution in numerical simulations. Mathematical analysis of the governing Boussinesq

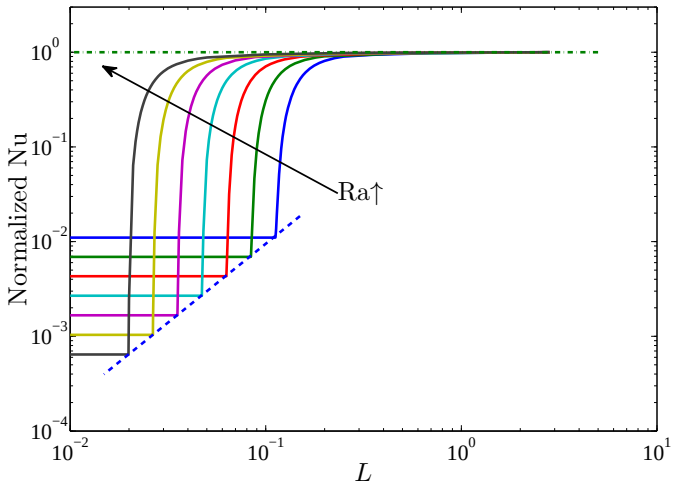

FIG. 9. Normalized Nu, $\mathrm{Nu}(\mathrm{Ra}, L)/\mathrm{Nu}(\mathrm{Ra}, 2\sqrt{2})$, obtained from the upper bound algorithm for $\mathrm{Ra} = 10^7$–10^{10}. There is no convection when L falls to the left of the dashed line, which corresponds to the marginal stability boundary of the conduction state L_c, i.e., $\mathrm{Nu}(\mathrm{Ra}, L < L_c) \equiv 1$. The differences in the constant values of normalized Nu for $L < L_c$ arise because $\mathrm{Nu}(\mathrm{Ra}, 2\sqrt{2})$ differs for each Ra.

equations both complements and informs experiments and DNS by providing provably true bounds on the achievable heat transport. Indeed, Whitehead & Doering [29, 31] used rigorous upper bound analysis to analytically prove that $\mathrm{Nu} \leq c\mathrm{Ra}^{5/12}$ as $\mathrm{Ra} \to \infty$ for both 2D Rayleigh–Bénard convection at arbitrary Pr and 3D infinite-Pr Rayleigh–Bénard convection, both with stress-free isothermal boundary conditions (albeit with slightly different pre-factors c). In this investigation, we have obtained the *optimal* upper bound – within the CDH framework – on the heat transport in 2D Rayleigh–Bénard convection with stress-free isothermal boundaries by solving the full background variational problem numerically. Moreover, we have proved that if the two-step numerical algorithm utilized here converges to a steady state then this state is the global optimal not only for stress-free Rayleigh–Bénard convection but also for other systems including, e.g., porous medium convection and plane Couette flow (see Appendix B). Since numerical continuation is *not* required using this approach, and because our algorithm can be readily programmed, thereby obviating the overhead associated with learning sophisticated commercial software, we hope that it should enable a broader community of researchers to employ computational upper bound theory.

For 2D stress-free Rayleigh–Bénard convection at large Ra and fixed $L = 2\sqrt{2}$, our results show that the optimal bound $\mathrm{Nu} \leq 0.106\mathrm{Ra}^{5/12}$ and the largest critical mode number $n_c \sim 0.13\mathrm{Ra}^{1/4}$, in close agreement with the analytical (sub-optimal) bounds obtained by Whitehead & Doering [29]. Crucially, the 5/12 exponent rules out general arguments (i.e., those not making specific reference to spatial dimension or boundary conditions) for the existence of an ultimate scaling regime with $\mathrm{Nu} \sim \mathrm{Pr}^{1/2}\mathrm{Ra}^{1/2}$ wherein the heat transport is independent of the molecular transport coefficients [20, 21]. The $\mathrm{Nu} \lesssim \mathrm{Ra}^{5/12}$ bound reveals that molecular transport processes are not generally negligible as $\mathrm{Ra} \to \infty$, and certainly not for the 2D stress-free problem. Furthermore, as a by-product, the upper bound analysis furnishes many useful features of the dynamical system including an approximation to the mean temperature profile, bifurcations of Nu with Ra and L, and an *a priori* orthogonal eigenbasis that should be well-suited for reduced-order modeling, *without* directly solving the governing equations. Our investigation constitutes the first systematic study of the influence of the domain aspect-ratio on the upper bounds on the heat transport in turbulent stress-free Rayleigh–Bénard convection. As for porous medium convection, our results imply that at large Ra there exists a minimal flow unit above which size the heat transport is effectively independent of L. The specific aspect ratio of this unit $L(\mathrm{Ra})$ remains an open question.

Finally, it is worthwhile to mention the remarkable quantitative correspondence between the best background asymptotic high Rayleigh number upper bound computed here, $\mathrm{Nu} \leq 0.106\mathrm{Ra}^{5/12}$, and the optimal (steady) high-Ra transport limit computed by explicitly constructing steady incompressible flows satisfying the same boundary conditions and intensity constraints – namely $\langle|\nabla u|^2\rangle = \mathrm{Ra}(\mathrm{Nu} - 1)$ – that passively advect the maximal heat flux between parallel isothermal boundaries, i.e., $\mathrm{Nu} \lesssim 0.115\mathrm{Ra}^{5/12}$ [38]. It is natural to wonder if the prefactors might actually be identical; the coincidence of their magnitudes is certainly notable and the 8.6% discrepancy might be the result of the numerical estimation of the relation over only a limited Rayleigh number range in Hassanzadeh *et al.* [38]. In any case the flows constructed there indicate that the background bounds computed here are essentially sharp in the sense that there exist "admissible" flows that transport heat very close to the estimated rate. This is in accord with

recent observations that the background method produces bounds that correspond precisely to optimally transporting flows in truncated versions of Rayleigh's model, specifically the Lorenz equations [39] and some distinguished higher order truncations that respect energy and enstrophy conservation in the inviscid limit [40]. We would emphasize that this correspondence seems to be special to Rayleigh's model with stress-free boundaries and, for the background method, in 2D explicitly exploiting the enstrophy balance. Indeed; we do not necessarily anticipate the same scaling for no-slip boundaries or for optimal transport in 3D.

ACKNOWLEDGMENTS

This research was supported in part by NSF Awards DMS-0928098 (GPC) and PHY-0855335, DMS-0927587, and PHY-1205219 (CRD). CRD is grateful for generous support from the Simons Foundation, and all the authors acknowledge the hospitality of the Geophysical Fluid Dynamics Program at Woods Hole Oceanographic Institution, supported by NSF and ONR, where some of this work was completed.

Appendix A: Newton–Kantorovich Algorithm for Solving the Euler–Lagrange Equations

For fixed a and b, we rewrite the Euler–Lagrange equations as

$$\tau_{zz} = F^{\tau}(\vartheta, W, \vartheta_z, W_z), \tag{A1}$$
$$2\nabla^2\vartheta = F^{\vartheta}(\Omega_x, W, \tau_z), \tag{A2}$$
$$2\nabla^2\Omega_x + \mathrm{Ra}\vartheta_{xx} - \frac{\mathrm{Ra}^{3/2}}{a}\gamma_{xx} = F^{\Omega_x}(\Omega_x), \tag{A3}$$
$$\nabla^2\gamma = F^{\gamma}(\vartheta, \tau_z), \tag{A4}$$
$$\nabla^2 W = F^{W}(\Omega_x). \tag{A5}$$

Suppose the iterates $\tau^i(z)$, $\vartheta^i(x,z)$, $\Omega^i_x(x,z)$, $\gamma^i(x,z)$, and $W^i(x,z)$ are good approximations to the true solution $\tau(z)$, $\vartheta(x,z)$, $\Omega_x(x,z)$, $\gamma(x,z)$, and $W(x,z)$. Taylor expansion of the functionals F^{τ}, F^{ϑ}, F^{Ω_x}, F^{γ} and F^{W} in Eqs. (A1)–(A5) about the i^{th} iterate yields

$$\begin{aligned}\tau_{zz} = {} & (F^{\tau})^i + (F^{\tau}_{\vartheta})^i[\vartheta - \vartheta^i] + (F^{\tau}_{W})^i[W - W^i] + (F^{\tau}_{\vartheta_z})^i[\vartheta_z - \vartheta^i_z] + \\ & (F^{\tau}_{W_z})^i[W_z - W^i_z] + O([\vartheta - \vartheta^i]^2, [W - W^i]^2, [\vartheta_z - \vartheta^i_z]^2, [W_z - W^i_z]^2),\end{aligned} \tag{A6}$$
$$\begin{aligned}2\nabla^2\vartheta = {} & (F^{\vartheta})^i + (F^{\vartheta}_{\Omega_x})^i[\Omega_x - \Omega^i_x] + (F^{\vartheta}_{W})^i[W - W^i] + (F^{\vartheta}_{\tau_z})^i[\tau_z - \tau^i_z] \\ & + O([\Omega_x - \Omega^i_x]^2, [W - W^i]^2, [\tau_z - \tau^i_z]^2),\end{aligned} \tag{A7}$$
$$2\nabla^2\Omega_x + \mathrm{Ra}\vartheta_{xx} - \frac{\mathrm{Ra}^{3/2}}{a}\gamma_{xx} = (F^{\Omega_x})^i + (F^{\Omega_x}_{\Omega_x})^i[\Omega_x - \Omega^i_x] + O([\Omega_x - \Omega^i_x]^2), \tag{A8}$$
$$\nabla^2\gamma = (F^{\gamma})^i + (F^{\gamma}_{\vartheta})^i[\vartheta - \vartheta^i] + (F^{\gamma}_{\tau_z})^i[\tau_z - \tau^i_z] + O([\vartheta - \vartheta^i]^2, [\tau_z - \tau^i_z]^2), \tag{A9}$$
$$\nabla^2 W = (F^{W})^i + (F^{W}_{\Omega_x})^i[\Omega_x - \Omega^i_x] + O([\Omega_x - \Omega^i_x]^2), \tag{A10}$$

where, for example, F^{τ}_{ϑ} denotes the Frechet derivative of the function $F^{\tau}(\vartheta, W, \vartheta_z, W_z)$ with respect to ϑ. By defining correction terms

$$\triangle^{\tau} = \tau^{i+1} - \tau^i, \triangle^{\vartheta} = \vartheta^{i+1} - \vartheta^i, \triangle^{\Omega_x} = \Omega^{i+1}_x - \Omega^i_x, \triangle^{\gamma} = \gamma^{i+1} - \gamma^i, \triangle^{W} = W^{i+1} - W^i \tag{A11}$$

and computing the Frechet derivatives, the linear differential equations for the corrections can be expressed as

$$\triangle^{\tau}_{zz} - (F^{\tau}_{\vartheta})^i\triangle^{\vartheta} - (F^{\tau}_{\vartheta_z})^i\triangle^{\vartheta}_z - (F^{\tau}_{W})^i\triangle^{W} - (F^{\tau}_{W_z})^i\triangle^{W}_z = (1-b)(\overline{W\vartheta})^i_z - \tau^i_{zz}, \tag{A12}$$
$$-(F^{\vartheta}_{\tau_z})^i\triangle^{\tau}_z + 2\nabla^2\triangle^{\vartheta} - (F^{\vartheta}_{\Omega_x})^i\triangle^{\Omega_x} - (F^{\vartheta}_{W})^i\triangle^{W} = 2W^i\tau^i_z + \frac{a}{\mathrm{Ra}^{1/2}}\Omega^i_x - 2\nabla^2\vartheta^i, \tag{A13}$$
$$\begin{aligned}\mathrm{Ra}\triangle^{\vartheta}_{xx} + [2\nabla^2 - (F^{\Omega_x}_{\Omega_x})^i]\triangle^{\Omega_x} - \frac{\mathrm{Ra}^{3/2}}{a}\triangle^{\gamma}_{xx} = {} & \frac{2b}{a}\mathrm{Ra}^{1/2}\Omega^i_x - \mathrm{Ra}\vartheta^i_{xx} \\ & -2\nabla^2\Omega^i_x + \frac{\mathrm{Ra}^{3/2}}{a}\gamma^i_{xx},\end{aligned} \tag{A14}$$
$$-(F^{\gamma}_{\tau_z})^i\triangle^{\tau}_z - (F^{\gamma}_{\vartheta})^i\triangle^{\vartheta} + \nabla^2\triangle^{\gamma} = -\nabla^2\gamma^i - 2\vartheta^i\tau^i_z, \tag{A15}$$
$$-(F^{W}_{\Omega_x})^i\triangle^{\Omega_x} + \nabla^2\triangle^{W} = -\nabla^2 W^i + \Omega^i_x. \tag{A16}$$

Let

$$\triangle^{\vartheta^*} = (\vartheta^*)^{i+1} - (\vartheta^*)^i, \;\; \triangle^{\Omega_x^*} = (\Omega_x^*)^{i+1} - (\Omega_x^*)^i, \;\; \triangle^{\gamma^*} = (\gamma^*)^{i+1} - (\gamma^*)^i, \;\; \triangle^{W^*} = (W^*)^{i+1} - (W^*)^i \quad \text{(A17)}$$

and $D \equiv \mathrm{d}/\mathrm{d}z$; then Eq. (A12) becomes

$$\begin{aligned} &D^2\triangle^\tau - \frac{1}{2}(1-b)\sum_{n=1}^{N}[(DW_n^{*i} + W_n^{*i}D)\triangle_n^{\vartheta^*}] - \frac{1}{2}(1-b)\sum_{n=1}^{N}[(D\vartheta_n^{*i} + \vartheta_n^{*i}D)\triangle_n^{W^*}] \\ &= -D^2\tau^i + \frac{1}{2}(1-b)D\sum_{n=1}^{N}(W_n^{*i}\vartheta_n^{*i}). \end{aligned} \quad \text{(A18)}$$

For a given wavenumber nk, Eqs. (A13)–(A16) become, respectively,

$$\begin{aligned} -2W_n^{*i}D\triangle^\tau + 2[D^2-(nk)^2]\triangle_n^{\vartheta^*} - \frac{a}{\mathrm{Ra}^{1/2}}\triangle_n^{\Omega_x^*} - 2D\tau^i\triangle_n^{W^*} &= -2[D^2-(nk)^2]\vartheta_n^{*i} \\ &\quad +2W_n^{*i}D\tau^i + \frac{a}{\mathrm{Ra}^{1/2}}\Omega_{xn}^{*i}, \end{aligned} \quad \text{(A19)}$$

$$\begin{aligned} -\mathrm{Ra}(nk)^2\triangle_n^{\vartheta^*} + [2(D^2-(nk)^2) - \frac{2b}{a}\mathrm{Ra}^{1/2}]\triangle_n^{\Omega_x^*} + \frac{\mathrm{Ra}^{3/2}}{a}(nk)^2\triangle_n^{\gamma^*} &= [-2(D^2-(nk)^2) + \frac{2b}{a}\mathrm{Ra}^{1/2}]\Omega_{xn}^{*i} \\ &\quad +\mathrm{Ra}(nk)^2\vartheta_n^{*i} - \frac{\mathrm{Ra}^{3/2}}{a}(nk)^2\gamma_n^{*i}, \end{aligned} \quad \text{(A20)}$$

$$2\vartheta_n^{*i}D\triangle^\tau + 2D\tau^i\triangle_n^{\vartheta^*} + [D^2-(nk)^2]\triangle_n^{\gamma^*} = -2\vartheta_n^{*i}D\tau^i - [D^2-(nk)^2]\gamma_n^{*i}, \quad \text{(A21)}$$

$$-\triangle_n^{\Omega_x^*} + [D^2-(nk)^2]\triangle_n^{W^*} = \Omega_{xn}^{*i} - [D^2-(nk)^2]W_n^{*i}. \quad \text{(A22)}$$

Although the Newton-Kantorovich method is only locally convergent, we can expand the basin of attraction (in the space of initial iterates) by updating the variables for each iterate using

$$\begin{bmatrix} \tau \\ \vartheta_n^* \\ \Omega_{xn}^* \\ \gamma_n^* \\ W_n^* \end{bmatrix}^{i+1} = \begin{bmatrix} \tau \\ \vartheta_n^* \\ \Omega_{xn}^* \\ \gamma_n^* \\ W_n^* \end{bmatrix}^{i} + d\begin{bmatrix} \triangle^\tau \\ \triangle_n^{\vartheta^*} \\ \triangle_n^{\Omega_x^*} \\ \triangle_n^{\gamma^*} \\ \triangle_n^{W^*} \end{bmatrix}, \quad \text{(A23)}$$

where $0 \le d \le 1$. The step length is reduced whenever $F_{res}^{i+1} > hF_{res}^i$, where F_{res}^i is the norm of the residual of the Euler–Lagrange equations at the i^{th} iterate, and $h \approx 1$ is an adjustable parameter.

Appendix B: Temporal Convergence in other Upper Bound Problems

Here we show that the analysis presented in section II D carries over to the porous medium convection problem where the time-stepping approach was first used [18] and to the plane Couette flow problem [17].

1. 2D Porous Medium Convection

Our starting point is equations (10)–(13) from Wen *et al.* [18],

$$\partial_t\tau = \tau_{zz} - \tfrac{1}{2}(\overline{W\vartheta})_z, \quad \text{(B1)}$$

$$\partial_t\vartheta = 2\nabla^2\vartheta - W\tau_z + \gamma_{xx}, \quad \text{(B2)}$$

$$0 = \nabla^2 W - \mathrm{ra}\,\vartheta_{xx}, \quad \text{(B3)}$$

$$0 = \nabla^2\gamma + \mathrm{ra}\,\tau_z\vartheta, \quad \text{(B4)}$$

where the one balance parameter a has been absorbed into a rescaled Rayleigh number $\mathrm{ra} \equiv \mathrm{Ra}/a$. Let $(\tau, \vartheta, W, \gamma)$ be a steady solution and consider a small disturbance $(\hat\tau, \hat\vartheta, \hat W, \hat\gamma)$ away from this solution. The (linearized) temporal

evolution equations for this disturbance are

$$\partial_t \hat{\tau} = \hat{\tau}_{zz} - \tfrac{1}{2}(\overline{\hat{W}\vartheta + W\hat{\vartheta}})_z, \tag{B5}$$

$$\partial_t \hat{\vartheta} = 2\nabla^2\hat{\vartheta} - \hat{W}\tau_z - W\hat{\tau}_z + \hat{\gamma}_{xx}, \tag{B6}$$

$$0 = \nabla^2\hat{W} - \mathrm{ra}\,\hat{\vartheta}_{xx}, \tag{B7}$$

$$0 = \nabla^2\hat{\gamma} + \mathrm{ra}\,(\hat{\tau}_z\vartheta + \tau_z\hat{\vartheta}). \tag{B8}$$

Adding $\langle\, \hat{\tau} \times$ Eq.(B5)$\,\rangle$ to $\frac{1}{2}\langle\, \hat{\vartheta} \times$ Eq.(B6)$\,\rangle$ produces

$$\left\langle\, \tfrac{1}{2}\hat{\tau}^2 + \tfrac{1}{4}\hat{\vartheta}^2 \,\right\rangle_t = -\langle\, \hat{\tau}_z^2 \,\rangle - \langle\, |\nabla\hat{\vartheta}|^2 \,\rangle + \tfrac{1}{2}\langle\, \hat{W}\vartheta\hat{\tau}_z - \hat{W}\hat{\vartheta}\tau_z + \hat{\vartheta}\hat{\gamma}_{xx} \,\rangle, \tag{B9}$$

which can be simplified by noticing that $\langle\, \hat{W} \times$ Eq.(B8) $- \hat{\gamma} \times$ Eq.(B7)$\,\rangle$ together with integration by parts and the homogeneous boundary conditions in z gives

$$\langle\, \hat{\vartheta}\hat{\gamma}_{xx} \,\rangle = -\langle\, \hat{W}\hat{\vartheta}\tau_z + \vartheta\hat{W}\hat{\tau}_z \,\rangle. \tag{B10}$$

Hence

$$\left\langle\, \tfrac{1}{2}\hat{\tau}^2 + \tfrac{1}{4}\hat{\vartheta}^2 \,\right\rangle_t = -\langle\, \hat{\tau}_z^2 \,\rangle - \left\langle\, |\nabla\hat{\vartheta}|^2 + \hat{W}\hat{\vartheta}\tau_z \,\right\rangle, \tag{B11}$$

where $\hat{W}$ is related to $\hat{\vartheta}$ via Eq. (B3). Mirroring the analysis of section II D, the second term on the right hand side is just the spectral constraint functional and hence all the same arguments carry over to establish that the global optimal is the only *steady* attractor for the time-stepping problem. This, of course, does not prove that the time-stepping approach will always converge to the global optimal as there may be other attractors (e.g., periodic orbits) but it does retrospectively prove that the steady state reached by the numerical computations reported in Wen *et al.* [18] must be the desired global optimal.

2. Plane Couette Flow

For plane Couette flow, we work with the Lagrangian defined in equation (2.8) of Plasting & Kerswell [17] or rather, as the optimal fields are steady, just

$$\mathcal{L} := \langle\, \phi_z^2 \,\rangle - \langle a\phi_z\nu_1\nu_2 + (a-1)|\nabla\boldsymbol{\nu}|^2 - (a-2)\nu_1\phi_{zz} \,\rangle. \tag{B12}$$

Let $(\phi, \boldsymbol{\nu})$ be a solution of the Euler–Lagrange equations (fixing the balance parameter a), so

$$\frac{\delta\mathcal{L}}{\delta\phi} := -2\phi_{zz} + a(\overline{\nu_1\nu_3})_z + (a-2)(\overline{\nu_1})_{zz} = 0, \tag{B13}$$

$$\frac{\delta\mathcal{L}}{\delta\boldsymbol{\nu}} := 2(a-1)\nabla^2\boldsymbol{\nu} - a\phi_z \begin{bmatrix} \nu_3 \\ 0 \\ \nu_1 \end{bmatrix} + \nabla p + (a-2)\phi_{zz}\hat{\boldsymbol{x}} = \mathbf{0}, \tag{B14}$$

(equations (2.9b) and (2.9c) of [17]) and let $(\hat{\phi}, \hat{\boldsymbol{\nu}})$ be a small perturbation superimposed on this steady solution. Adding time derivatives to the Euler–Lagrange equations such that

$$\partial_t\phi = -\frac{\delta\mathcal{L}}{\delta\phi} \quad \& \quad 2(a-1)\partial_t\boldsymbol{\nu} = \frac{\delta\mathcal{L}}{\delta\boldsymbol{\nu}}, \tag{B15}$$

(so

$$\frac{d\mathcal{L}}{dt} = \left\langle\, 2(a-1)\boldsymbol{\nu}_t.\frac{\delta\mathcal{L}}{\delta\boldsymbol{\nu}} + \phi_t\frac{\delta\mathcal{L}}{\delta\phi} \,\right\rangle = \left\langle\, 4(a-1)^2\boldsymbol{\nu}_t^2 - \phi_t^2 \,\right\rangle \tag{B16}$$

reflecting the fact that $\mathcal{L}$ is a saddle point functional), the (linearized) temporal evolution of the perturbations is then determined by

$$\partial_t \hat{\phi} := 2\hat{\phi}_{zz} - a(\overline{\hat{\nu}_1\nu_3 + \nu_1\hat{\nu}_3})_z - (a-2)(\overline{\hat{\nu}_1})_{zz}, \tag{B17}$$

$$2(a-1)\partial_t \hat{\boldsymbol{\nu}} := 2(a-1)\nabla^2\hat{\boldsymbol{\nu}} - a\hat{\phi}_z \begin{bmatrix} \nu_3 \\ 0 \\ \nu_1 \end{bmatrix} - a\phi_z \begin{bmatrix} \hat{\nu}_3 \\ 0 \\ \hat{\nu}_1 \end{bmatrix} + \nabla\hat{p} + (a-2)\hat{\phi}_{zz}\hat{\boldsymbol{x}}. \tag{B18}$$

Adding the volume integrals $\langle\, \hat{\phi} \times \text{Eq. (B17)} \,\rangle$ and $\langle\, \hat{\boldsymbol{\nu}} \cdot \text{Eq. (B18)} \,\rangle$ gives

$$\left\langle\, \tfrac{1}{2}\hat{\phi}^2 + (a-1)\hat{\boldsymbol{\nu}}^2 \,\right\rangle_t = -2\langle\, \hat{\phi}_z^2 \,\rangle - 2\left\langle\, (a-1)|\nabla\hat{\boldsymbol{\nu}}|^2 + a\phi_z\hat{\nu}_1\hat{\nu}_3 \,\right\rangle, \tag{B19}$$

where the last term on the right hand side is again the appropriate spectral constraint integral (see the part quadratic in $\boldsymbol{\nu}$ in Eq. (B12) or (2.14) in [17]). The same conclusion then follows as for porous medium convection (section B 1) and Rayleigh–Bénard convection (section II D): the global optimal can be the only steady attractor for the full (free balance parameter) time-dependent problem.

[1] L. N. Howard, J. Fluid Mech. **17**, 405 (1963).
[2] W. V. R. Malkus, Proc. R. Soc. Lond. **A225**, 196 (1954).
[3] H. Bénard, Revue Générale des Sciences Pures et Appliquées **11**, 1261 (1900).
[4] L. Rayleigh, Phil. Mag. **32**, 529 (1916).
[5] F. Busse, J. Fluid Mech. **37**, 457 (1969).
[6] F. Busse, J. Fluid Mech. **41**, 219 (1970).
[7] F. Busse and D. Joseph, J. Fluid Mech. **54**, 521 (1972).
[8] C. R. Doering and P. Constantin, Phys. Rev. E **53**, 5957 (1996).
[9] C. R. Doering and P. Constantin, Phys. Rev. Lett. **69**, 1648 (1992).
[10] C. R. Doering and P. Constantin, Phys. Rev. E **49**, 4087 (1994).
[11] P. Constantin and C. R. Doering, Phys. Rev. E **51**, 3192 (1995).
[12] C. R. Doering and J. M. Hyman, Phys. Rev. E **55**, 7775 (1997).
[13] C. R. Doering and P. Constantin, J. Fluid Mech. **376**, 263 (1998).
[14] E. Hopf, Mathematische Annalen **117**, 764 (1941).
[15] R. R. Kerswell, Physica D **121**, 175 (1998).
[16] R. R. Kerswell, Phys. Fluids **13**, 192 (2001).
[17] S. C. Plasting and R. R. Kerswell, J. Fluid Mech. **477**, 363 (2003).
[18] B. Wen, G. P. Chini, N. Dianati, and C. R. Doering, Phys. Lett. A **377**, 2931 (2013).
[19] L. Howard, in *Applied Mechanics, Proceedings of the 11th International Congress of Applied Mechanics*, edited by H. Görtler (Munich, 1964) pp. 1109–1115.
[20] E. A. Spiegel, Ann. Rev. Astron. Astrophys. **9**, 323 (1971).
[21] S. Grossmann and D. Lohse, J. Fluid Mech. **407**, 27 (2000).
[22] F. Waleffe, A. Boonkasame, and L. M. Smith, Phys. Fluids **27**, 051702 (2015).
[23] J. Otero, R. W. Wittenberg, R. A. Worthing, and C. R. Doering, J. Fluid Mech. **473**, 191 (2002).
[24] R. W. Wittenberg, J. Fluid Mech. **665**, 158 (2010).
[25] P. Constantin and C. R. Doering, J. Stat. Phys. **94**, 159 (1999).
[26] C. R. Doering, F. Otto, and M. G. Reznikoff, J. Fluid Mech. **560**, 229 (2006).
[27] F. Otto and C. Seis, J. Math. Phys. **52**, 083702 (2011).
[28] J. Otero, *Bounds for the heat transport in turbulent convection* (Ph.D. thesis, University of Michigan, 2002).
[29] J. P. Whitehead and C. R. Doering, Phys. Rev. Lett. **106**, 244501 (2011).
[30] G. R. Ierley, R. R. Kerswell, and S. C. Plasting, J. Fluid Mech. **560**, 159 (2006).
[31] J. P. Whitehead and C. R. Doering, J. Fluid Mech. **707**, 241 (2012).
[32] J. P. Boyd, *Chebyshev and Fourier Spectral Methods*, 2nd ed. (Dover, New York, 2000).
[33] G. P. Chini and S. M. Cox, Phys. Fluids **21**, 083603 (2009).
[34] A. N. Souza, "These Nu–Ra data are heat transport in steady solutions of the boussinesq equations with $Pr = 1$ maximized over the cell aspect ratio at each value of Ra," personal communication (unpublished) (2015).
[35] S. C. Plasting and G. R. Ierley, J. Fluid Mech. **542**, 343 (2005).
[36] R. Verzicco and K. R. Sreenivasan, J. Fluid Mech. **595**, 203 (2008).
[37] G. P. Chini, N. Dianati, Z. Zhang, and C. R. Doering, Physica D **240**, 241 (2011).
[38] P. Hassanzadeh, G. P. Chini, and C. R. Doering, J. Fluid Mech. **751**, 627 (2014).
[39] A. Souza and C. R. Doering, Phys. Lett. A **379**, 518 (2015).
[40] A. Souza and C. R. Doering, Physica D **308**, 26 (2015).