Exhausting the background approach for bounding the heat transport in Rayleigh-Bénard convection

Zijing Ding[1]† **and Rich R. Kerswell**[1]‡

Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, University of Cambridge, Cambridge, CB3 0WA, UK.

(Received ?; revised ?; accepted ?.)

We revisit the optimal heat transport problem for Rayleigh-Bénard convection in which a rigorous upper bound on the Nusselt number, Nu, is sought as a function of the Rayleigh number, Ra. Concentrating on the 2-dimensional problem with stress-free boundary conditions, we impose the time-averaged heat equation as a constraint for the bound using a novel 2-dimensional background approach thereby complementing the 'wall-to-wall' approach of Hassanzadeh *et al.* (*J. Fluid Mech.* **751**, 627-662, 2014). Imposing the same symmetry on the problem, we find correspondence with their maximal result for $Ra \leqslant Ra_c := 4468.8$ but, beyond that, the results from the two approaches diverge. The bound produced by the 2-dimensional background field approaches that produced by the 1-dimensional background field from below as the length of computational domain $L \to \infty$. On lifting the imposed symmetry, the optimal 2-dimensional temperature background field reverts to being 1-dimensional giving the best bound $Nu \leqslant 0.055Ra^{1/2}$ compared to $Nu \leqslant 0.026Ra^{1/2}$ in the non-slip case. We then show via an inductive bifurcation analysis that introducing 2-dimensional temperature *and* velocity background fields (in an attempt to impose the time-averaged Boussinesq equations) is also unable to lower the bound. This then exhausts the background approach for the 2-dimensional (and by extension 3-dimensional) Rayleigh-Bénard problem with the bound remaining stubbornly $Ra^{1/2}$ while data seems more to scale like $Ra^{1/3}$ for large Ra. Finally, we show that adding a velocity background field to the formulation of Wen *et al.* (*Phys. Rev. E.* **92**, 043012, 2015), which is able to use an extra vorticity constraint due to the stress-free condition to lower the bound to $Nu \leqslant O(Ra^{5/12})$, also fails to further improve the bound.

Key words: Upper bound, Rayleigh-Bénard convection

1. Introduction

In this paper we consider the fundamental problem of assessing how the heat flux behaves as a function of the Rayleigh number, Ra, in Rayleigh-Bénard convection where a layer of fluid is heated from below and cooled from above. This situation is ubiquitous in Nature and consequently the focus of a huge body of ongoing research work (e.g. Ahlers et al. 2009). The particular focus here is on the use of variational methods which seek an upper bound on the heat flux in the hope that this bound will capture the correct high-Ra scaling for turbulent convection. This approach involves constructing an optimisation

† z.ding@damtp.cam.ac.uk
‡ r.r.kerswell@damtp.cam.ac.uk

problem constrained by information gleaned from the governing equations. Inevitably, the constraints actually imposed form a strict subset of those implied by the governing equations so that any maximum which emerges is an upper bound on what can actually be realised. This approach has its roots in the work of Malkus (1954) who hypothesized that the fluid selects the flow state from all those possible states which maximises the heat transport. The subsequent mathematical formulation by Howard (1963) and Busse (1969) was as a maximization problem (see the early reviews by Howard (1972) and Busse (1978)). In the 1990s, an alternative complementary approach - the background method - was introduced by Doering & Constantin (1992,1994,1995,1996) which takes the form of a *minimization* problem. This has the considerable advantage that even a trial solution can yield an upper bound which, experience seems to indicate, yields the same scaling as the proper optimal (e.g. in shear flow and convection see Doering & Constantin 1992,1996 respectively compared to Plasting & Kerswell 2003, hereafter PK03).

In both approaches, however, the outstanding challenge has been to add further dynamical information to improve (lower) the scaling law (e.g. see Ierley & Worthing (2001) for efforts in the Howard-Busse maximization problem). The best current bound on the Nusselt number - the ratio of actual heat flux to the conductive value - for the case of non-slip boundary conditions on smooth walls is $Nu \leqslant 0.02634 Ra^{1/2}$ as $Ra \to \infty$ (PK03) whereas most of the current experimental data suggests $Nu \sim Ra^{0.31}$ (see the discussion in Waleffe et al. 2015) and so is more consistent with the simple theoretical prediction of $Nu \sim Ra^{1/3}$ (Malkus 1954, Priestley 1954) with some dependence on the Prandtl number also possible (Grossmann & Lohse 2000). A natural way of incorporating further information exists in the background method through simply extending the definitions of the background fields. To see this, recall that the Malkus-Howard-Busse (maximization) approach and the Doering-Constantin (minimization) approach are dual problems seeking to find an appropriate saddle point of a functional of the velocity and temperature fields (Kerswell 1998, 2001). To explain further we introduce the problem to be considered.

Let a Newtonian fluid be confined between two infinite isothermal plates at $z = 0$ and $z = d$ with the lower plate maintained at a constant temperature δT hotter than that of the upper plate. Using the gap width d, d^2/κ (κ is the thermal diffusivity) and δT as units of length, time and temperature together with adopting the Boussinesq approximation, the governing equations are

$$(\boldsymbol{\mathcal{N}}) := \frac{\partial \mathbf{u}}{\partial t} + \mathbf{u}\cdot\boldsymbol{\nabla}\mathbf{u} + \boldsymbol{\nabla} p - \sigma\nabla^2\mathbf{u} - \sigma Ra T\hat{\mathbf{z}} = \mathbf{0}, \tag{1.1}$$

$$(\mathcal{H}) := \frac{\partial T}{\partial t} + \boldsymbol{\nabla}\cdot(\mathbf{u}T - \boldsymbol{\nabla}T) = 0, \tag{1.2}$$

with $\boldsymbol{\nabla}\cdot\mathbf{u} = 0$ where

$$\sigma := \nu/\kappa \quad \& \quad Ra := g\beta\delta T d^3/\nu\kappa \tag{1.3}$$

are the Prandtl and Rayleigh numbers respectively (ν is the kinematic viscosity, β is the thermal expansion coefficient and $-g\hat{\mathbf{z}}$ is the acceleration due to gravity). The background method starts by writing down the functional

$$\mathscr{L} := \langle|\boldsymbol{\nabla}T|^2\rangle - \langle a\boldsymbol{\nu}\cdot(\boldsymbol{\mathcal{N}})\rangle - \langle b\theta(\mathcal{H})\rangle \tag{1.4}$$

where the first term on the right is the long-time-averaged Nusselt number Nu, $\boldsymbol{\nu}(\mathbf{x},t)$ and $\theta(\mathbf{x},t)$ are Lagrange multipliers imposing the momentum and heat equations as constraints respectively (the seemingly redundant extra scalars a and b play a key role

later) and

$$\langle\,(\ldots)\,\rangle := \lim_{\mathcal{T}\to\infty}\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}}\frac{1}{V}\int(\ldots)\,dV\,dt \tag{1.5}$$

is a spatial-temporal average. The crucial next step is to choose steady 'background' fields

$$\boldsymbol{\phi}(\mathbf{x}) := \mathbf{u}(\mathbf{x},t) - \boldsymbol{\nu}(\mathbf{x},t), \quad \tau(\mathbf{x}) := T(\mathbf{x},t) - \theta(\mathbf{x},t) \tag{1.6}$$

which connect the Lagrange multipliers with the physical fields and such that they carry any inhomogeneous boundary conditions (so here just those on the temperature field). In principle, time-dependence can be retained in the background fields but this leads to a substantially more complex problem beyond the scope of the current investigation (also it isn't clear this helps - see Souza & Doering (2015a,b) for calculations in reduced models). Changing variables from $(\mathbf{u}, T, \boldsymbol{\nu}, \theta)$ to $(\mathbf{u}, T, \boldsymbol{\phi}, \tau)$,

$$\begin{aligned}\mathscr{L} &= \langle|\boldsymbol{\nabla}T|^2\rangle - \langle a(\mathbf{u}-\boldsymbol{\phi})\cdot(\boldsymbol{\mathcal{N}})\rangle - \langle b(T-\tau)(\mathcal{H})\rangle,\\ &= \langle|\boldsymbol{\nabla}T|^2\rangle - a\langle\mathbf{u}\cdot(\boldsymbol{\mathcal{N}})\rangle + \frac{a}{V}\int\boldsymbol{\phi}\cdot\overline{(\boldsymbol{\mathcal{N}})}^t\,dV - b\langle T(\mathcal{H})\rangle + \frac{b}{V}\int\tau\overline{(\mathcal{H})}^t\,dV\end{aligned} \tag{1.7}$$

(where

$$\overline{(\,)}^t := \lim_{\mathcal{T}\to\infty}\sup\frac{1}{\mathcal{T}}\int_0^{\mathcal{T}}(\,)\,dt \tag{1.8}$$

is a long-time average) makes it clear that choosing the largest stationary value of $\mathscr{L}$ finds the largest long-time-averaged Nusselt number subject to the long-time-averaged power and entropy balances (Lagrange multipliers a and b respectively) and projected information from momentum and heat flux balances (Lagrange multipliers $\boldsymbol{\phi}$ and τ respectively). Since it can be shown that all the time derivative terms in these constraints vanish under long time averaging, the variational problem can be couched in terms of steady fields only. In particular, the goal is to evaluate the largest stationary value of the functional

$$\mathscr{L}_s := \langle|\boldsymbol{\nabla}T|^2\rangle - a\langle\mathbf{u}\cdot(\boldsymbol{\mathcal{N}})_s\rangle + a\langle\boldsymbol{\phi}\cdot(\boldsymbol{\mathcal{N}})_s\rangle - b\langle T(\mathcal{H})_s\rangle + b\langle\tau(\mathcal{H})_s\rangle \tag{1.9}$$

where the subscript s indicates the steady version of the unsubscripted quantity. So far only the minimal choice $(\tau,\boldsymbol{\phi}) = (\tau(z),\mathbf{0})$ has been explored (Doering & Constantin 1996) which leads to the simplified expression

$$\mathscr{L}_s := \langle|\boldsymbol{\nabla}T|^2\rangle - a\langle\mathbf{u}\cdot(\boldsymbol{\mathcal{N}})_s\rangle - b\langle T(\mathcal{H})_s\rangle + b\int_0^1\tau(z)\left[\lim_{L\to\infty}\frac{1}{L^2}\int_{-L/2}^{L/2}\int_{-L/2}^{L/2}(\mathcal{H})_s\,dx\,dy\right]dz \tag{1.10}$$

This choice turns out to give the dual problem to the Howard-Busse approach (Howard 1963, Busse 1969) and produces the same Nusselt number bound (Kerswell 2001, PK03). However, here, beyond the total power and entropy balances and insisting that the fluid is incompressible and the boundary conditions are satisfied, only the horizontally-averaged steady heat equation is imposed as a constraint. It seems reasonable to suppose that imposing further constraints from the governing equations by extending the definitions of the background fields should lower this current best bound for Boussinesq convection. Probing this hypothesis is the motivation for this paper.

The need to improve a bound so that it captures the observed scaling of a key flow quantity is quite general. Bounds developed on the energy dissipation rate for shear flows such as plane Couette flow (Doering & Constantin 1992), channel flow (Doering & Constantin 1994) and pipe flow (Plasting & Kerswell 2005) all seem to be too conservative

by a factor $1/(\log Re)^2$ (Re being the Reynolds number). There are special cases where the scaling is correctly captured - shear flow with suction (Doering *et al.*. 2000), porous media convection (Doering & Constantin 1998) and precessing flows (Kerswell 1996) - but generally some simplifying limit has to be taken to access further constraints (e.g. the momentum equation in infinite Prandtl number convection: Doering & Constantin 2001). Finite-Prandtl-number Rayleigh-Benard with non-slip smooth walls, however, is most studied partly because of its wide application and partly because the current best bound appears to have the wrong exponent and therefore calls for the most improvement. Of particular interest in recent efforts to lower the bound has been the introduction of the 'wall-to-wall' approach by Hassanzadeh et al. (2014) (see also Souza 2016 and Souza et al. 2019). Here the steady heat equation has been imposed as a constraint with some incompressible boundary-compliant flow field which, apart from an overall amplitude, is otherwise unconstrained and a *maximization* problem is solved. This appears to give a much improved (reduced) estimate of maximal flux with $Nu \sim Ra^{5/12}$ for stress-free boundary conditions in 2D convection with Souza (2016) finding a yet stronger (reduced) bound of $Nu \sim Ra^{0.371}$ for non slip boundary conditions. Later work by Tobasco & Doering (2017) (see also Doering & Tobasco 2019), however, has demonstrated through designing a sophisticated trial function that the upper bound must be at least $Nu \sim Ra^{1/2}$ up to logarithms for both stress-free and non-slip boundary conditions. This suggests that the non-slip maximal flux results of Souza (2016) and stress-free maximal results of Hassanzadeh et al. (2014) become only local maxima as $Ra \to \infty$. A good place to start our study is to try to shed some light on this by tackling the complementary background formulation - $(\tau, \boldsymbol{\phi}) = (\tau(x, z), \mathbf{0})$ which also imposes the steady heat equation in 2D convection and builds a *minimization* problem.

Concurrent work by Souza et al. (2019) has considered how the background method is connected to the wall-to-wall approach and speculated that there could be a 'duality gap' between them. Coming from a different perspective (the specific details of solving the variational equations), we share this speculation and confirm it here beyond a certain Rayleigh number. Motoki et al. (2018) have also built upon Hassanzadeh et al's work by extending the maximization search to 3 dimensions. Interestingly they find a 3D optimal solution which scales like $Ra^{1/2}$ with a numerical coefficient just 7.2% below the bound of PK03 (see their figure 2). This 3D result (using non-slip boundary conditions) and the 2D work of Tobasco & Doering (2017) (using non-slip boundary conditions) clearly beg the question whether further information from the momentum equation can be used to rule out the $Ra^{1/2}$ scaling which clearly persists despite imposing the steady heat equation. This also will be addressed here.

A further motivation for exploring the addition of further dynamical constraints is the hope that ultimately, the full governing equations can be imposed and then a direct connection forged between the optimal solution of an upper bounding variational problem and an actual solution of the governing equations. Realistically, this seems only likely if the optimal solution is indeed steady in which case only the time-averaged version of the equations need to be imposed. The possibility that steady solutions are close to if not capable of achieving maximal heat flux has been suggested by the asymptotic roll construction of Chini & Cox (2009) in 2D stress-free convection which attains $Nu \sim Ra^{1/3}$ and the recent computations tracking the (simple) 2D convection roll solution which initially bifurcates from the conductive state up to very high Rayleigh numbers of $O(10^9)$ in non-slip convection (Waleffe et al. 2015, Sondak et al. 2015). In this latter work, provided the aspect ratio of the rolls is optimized over, a heat flux relationship of $Nu \sim Ra^{0.31}$ is found which is intriguingly close to 3D turbulent convection measurements and

to $Nu \sim Ra^{1/3}$, the relationship many believe might be the ultimate scaling law although not all (e.g. Zhu et al. 2018).

A synopsis of the paper is as follows. Section 2 describes the set-up of 2D Boussinesq convection (§2.1), explains how a bound can be found using the background approach (§2.2) and then discusses the convexity of the optimization problem for a general temperature background field which ensures a unique optimal (§2.3). Section 2.4 explains how the numerical computations are performed with a choice having to be made between a branch continuation approach (PK03) and a time stepping method (Wen et al. 2013, 2015). Section 3 describes the results of tackling the upper bounding problem with the steady heat equation imposed in the presence of the same symmetry as used in Hassanzadeh et al. (2014). The appearance of a second fluctuation mode becoming 'spectrally unstable' at $Ra = Ra_c := 4468.8$ means: a) that there is a gap between Hassanzadeh et al's result and the background upper bound for $Ra > Ra_c$; and b) a new formulation for how the optimal is tracked needs to be introduced compared to previous work (e.g. PK03).

Section 4 discusses this new formulation which is significant because the various background and fluctuation optimal fields can no longer be used to define a set of physical temperature and velocity fields. In particular, the optimal fields do *not* satisfy the steady heat equation even though this is explicitly imposed as a constraint. Using this reformulation, section 5 shows how the optimal bound behaves for $Ra > Ra_c$. The size of the computational domain becomes important in the 2D background problem and it is found that the highest bound is only achieved in the infinite domain limit when the background field becomes increasingly 1D. Removing the symmetry used by Hassanzadeh et al. restores the translational invariance of the problem in which case the optimal has to be 1D and a bound of $Ra \leqslant 0.055Ra^{1/2}$ is found compared to the well-known result of $0.026Ra^{1/2}$ for non-slip walls (PK03).

Having found that imposing the steady heat equation does not improve the bound, we then consider adding extra information from the momentum equation by introducing a background velocity field $\boldsymbol{\phi}(x, z)$. Now the optimization problem is no longer convex and so we are unable to invoke uniqueness to dismiss non-vanishing $\boldsymbol{\phi}$. Instead we use an inductive bifurcation analysis to show that if $\boldsymbol{\phi} = \mathbf{0}$ before a bifurcation then it remains $\mathbf{0}$ after it too meaning that the continuous branch of optimals found by branch tracking out of the energy stability point always has $\boldsymbol{\phi} = \mathbf{0}$. Noting the one caveat that it's not impossible that there is an unconnected branch of optimals with $\boldsymbol{\phi} \neq \mathbf{0}$, this strongly suggests the surprising result that imposing the steady Boussinesq equations does not improve the bound over that obtained using the horizontally- and time-averaged heat equation and a global energy constraint from the momentum equation. Finally, section 7 observes that adding a velocity background temperature field to the formulation of Wen et al. (2015), which has an additional vorticity constraint, also fails to improve matters. A discussion follows in section 8.

2. Mathematical formulation

2.1. *Set-up*

We consider the 2 dimensional version of the Boussinesq equations (1.1) & (1.2) where $\boldsymbol{u} = u\hat{\boldsymbol{x}} + w\hat{\boldsymbol{z}}$ over a box $(x, z) \in [-\frac{1}{2}L, \frac{1}{2}L] \times [0, 1]$ together with the following stress-free

and isothermal boundary conditions

$$\frac{\partial u}{\partial z} = w = 0, \quad T = 1, \quad at \quad z = 0, \tag{2.1}$$

$$\frac{\partial u}{\partial z} = w = 0, \quad T = 0, \quad at \quad z = 1 \tag{2.2}$$

following Hassanzadeh et al. (2014). Applying the background method, we decompose the temperature field as

$$T = \tau(x, z) + \theta(x, z, t). \tag{2.3}$$

where the (steady) background temperature τ carries the boundary conditions of T (i.e. $\tau|_{z=0} = 1$ and $\tau|_{z=1} = 0$) so that the perturbation field θ vanishes at $z = 0, 1$. The time-averaged heat transport is characterized by the time-averaged Nusselt number Nu

$$Nu := \lim_{T\to\infty} \frac{1}{T} \int_0^T \frac{1}{L} \int_{-L/2}^{L/2} \left.\frac{\partial T}{\partial z}\right|_{z=1} dx\, dt = \langle |\boldsymbol{\nabla} T|^2 \rangle = 1 + \langle wT \rangle \tag{2.4}$$

where the spatial-temporal average defined in (1.5) becomes specifically

$$\langle\, (\ldots)\, \rangle := \lim_{T\to\infty} \frac{1}{T} \int_0^T \int_0^1 \frac{1}{L} \int_{-L/2}^{L/2} (\ldots)\, dx\, dz\, dt.$$

To find the maximum heat transport possible over all solutions to the Boussinesq equations, we construct the Lagrangian

$$\mathscr{L} = \langle |\boldsymbol{\nabla} T|^2 \rangle - \frac{a}{\sigma Ra} \langle \boldsymbol{u} \cdot \boldsymbol{\mathcal{N}} \rangle - b \langle \theta\, \mathcal{H} \rangle, \tag{2.5}$$

$$= \langle |\boldsymbol{\nabla} T|^2 \rangle - \frac{a}{\sigma Ra} \langle \boldsymbol{u} \cdot \boldsymbol{\mathcal{N}} \rangle - b \langle T\, \mathcal{H} \rangle + b \langle \tau\, \mathcal{H} \rangle. \tag{2.6}$$

where $a/\sigma Ra$ is the Langrange multiplier imposing the global constraint $\langle \boldsymbol{u} \cdot \boldsymbol{\mathcal{N}} \rangle = 0$, b is a Lagrange multiplier imposing the global constraint $\langle T\, \mathcal{H} \rangle = 0$ and $b\tau(x, z)$ is the Lagrange multiplier field imposing the time-averaged heat equation pointwise in the domain. The inclusion of b is actually redundant given the constraint imposed by τ implies $\langle T\, \mathcal{H} \rangle = 0$ so the value of b is chosen for convenience. Expression (2.6) can be rewritten using integration by parts and the fact that $\langle wT \rangle = \mathscr{L} - 1$ (see (2.4)) for solutions of the Boussinesq equations as

$$\mathscr{L} = \frac{1}{1-a} \left[\langle |\boldsymbol{\nabla} \tau|^2 \rangle - a \right] - \frac{1}{1-a} \mathscr{G} \tag{2.7}$$

where setting $b = 2$ makes

$$\mathscr{G} := \langle \frac{a}{Ra} |\boldsymbol{\nabla} \boldsymbol{u}|^2 + |\boldsymbol{\nabla} \theta|^2 + 2\theta \boldsymbol{u} \cdot \boldsymbol{\nabla} \tau \rangle \tag{2.8}$$

a purely quadratic form in θ and $\boldsymbol{u}$.

2.2. *Bounds on Nu*

The key realisation is that *if* $\mathscr{G} \geqslant 0$ for all $(\boldsymbol{u}, \theta) \in \Pi$ (the set of incompressible velocity and temperature fields which satisfy homogeneous versions of the boundary conditions (2.1) and (2.2)), which is a *spectral constraint* on τ, and $a \in (0, 1)$, we then have the bound

$$Nu \leqslant \frac{1}{1-a} \left[\langle |\boldsymbol{\nabla} \tau|^2 \rangle - a \right]. \tag{2.9}$$

The challenge is then to find the lowest such bound by minimizing over all (τ, a) which satisfy this spectral constraint, i.e.

$$(\tau, a) \in \Omega := \{(\tau, a) \mid \mathscr{G}(\boldsymbol{u}, \theta; \tau, a) \geqslant 0 \ \forall\, (\boldsymbol{u},\, \theta) \in \Pi\,\}. \tag{2.10}$$

After introducing a streamfunction, $(u, w) = (\partial\psi/\partial z, -\partial\psi/\partial x)$, the constraint that

$$\mathscr{G} = \langle \frac{a}{Ra}|\nabla^2\psi|^2 + |\boldsymbol{\nabla}\theta|^2 + 2\theta J(\tau, \psi)\rangle \geqslant 0 \tag{2.11}$$

where

$$J(A, B) := \frac{\partial A}{\partial x}\frac{\partial B}{\partial z} - \frac{\partial A}{\partial z}\frac{\partial B}{\partial x} \tag{2.12}$$

is equivalent to requiring that all of the eigenvalues λ of the linear problem

$$\lambda\theta \ = \nabla^2\theta - J(\tau, \psi), \tag{2.13}$$

$$\lambda\nabla^2\psi \ = \frac{a}{Ra}\nabla^4\psi - J(\tau, \theta) \tag{2.14}$$

(with boundary conditions $\psi = d^2\psi/dz^2 = \theta = 0$ for $z = \{0, 1\}$) are negative semi-definite.

2.3. *Convexity & Uniqueness*

The Euler-Lagrange equations for making the Lagrangian $\mathscr{L}$ in (2.7) stationary are

$$0 \ = \nabla^2\theta - J(\tau, \psi), \tag{2.15}$$

$$0 \ = \frac{a}{Ra}\nabla^4\psi - J(\tau, \theta), \tag{2.16}$$

$$0 \ = \nabla^2\tau - J(\theta, \psi), \tag{2.17}$$

$$\langle|\boldsymbol{\nabla}\tau|^2\rangle - 1 \ = \frac{(1-a)}{Ra}\langle|\nabla^2\psi|^2\rangle \tag{2.18}$$

and, as a nonlinear set of equations, can have many solutions. However, only solutions with $(\tau, a) \in \Omega$ yield a bound through the value of $\mathscr{L}$ generated. Due to the convexity of Ω (i.e. if (τ_1, a_1) and (τ_2, a_2) are in Ω then so is $\mu(\tau_1, a_1) + (1-\mu)(\tau_2, a_2)$ for $\mu \in (0, 1)$), and the fact that the objective functional

$$f(\tau, a) := \frac{1}{1-a}\left[\langle|\boldsymbol{\nabla}\tau|^2\rangle - a\right] \tag{2.19}$$

to be minimized is a strictly convex functional (the terms second order in $\delta\tau$ and δa in the difference $f(\tau + \delta\tau, a + \delta a) - f(\tau, a)$, specifically

$$\frac{1}{(1-a)^2}\langle\, \boldsymbol{\nabla}|(1-a)\delta\tau + (\tau + z - 1)\delta a\,|^2\,\rangle, \tag{2.20}$$

are positive definite), there is in fact at most one solution which satisfies the spectral constraint. This solution, hereafter referred to as the *optimal solution*, is what is sought.

2.4. *Numerical approach*

Recently, Wen *et al.* (2015) have proved that when τ is 1-dimensional, i.e. $\tau = \tau(z)$, appropriately augmenting the (steady) Euler-Lagrange equations with time derivatives leads to a system where the optimal solution is a unique attracting steady state. This proof carries over to 2-dimensional background fields $\tau = \tau(x, z)$ in the 3-dimensional Rayleigh-Benard problem but not in the 2-dimensional problem (see appendix A for details) where the dimensionality of the background field then matches that of the physical fields. This means that any steady attractor which emerges from time-stepping using

$\tau(x,z)$ in the 2D problem is not guaranteed to be the required optimal solution. The time stepping approach can still be used if it is married with a spectral constraint check but then there is always the prospect of rerunning with different initial conditions until the optimal solution is found. Given this, we chose instead to use the branch continuation approach - Newton's method with parametric continuation - starting from the energy stability bifurcation point as performed in PK03. While very robust, this has the general disadvantage of only being able to continuously trace optimal solutions from the energy stability bifurcation as Ra varies meaning that any new unconnected optimals cannot be found at a given Ra. This is not a problem here as the aforementioned uniqueness of the optimal solution means that no other optimal solution branches exist.

We consider periodic boundary conditions in x and, exactly as in Hassanzadeh et al. (2014), assume that the streamfunction ψ is odd (or antisymmetric), while θ and τ are even (or symmetric) about $x = 0$ by seeking the solution of (2.15)-(2.18) in the following form:

$$\psi = \sum_{m=1}^{M} \psi_m(z)\sin(m\alpha x), \quad \theta = \sum_{m=0}^{M} \theta_m(z)\cos(m\alpha x), \quad \tau = \sum_{m=0}^{M} \tau_m(z)\cos(m\alpha x). \tag{2.21}$$

We will find that this choice prevents a 1D background optimal even though this is allowed by the boundary conditions and imposed symmetry. Here $\alpha := 2\pi/L$ and ψ_m, θ_m, τ_m are expanded in Chebyshev polynomials, T_n,

$$[\psi_m(z), \theta_m(z), \tau_m(z)] = \sum_{n=0}^{N} [\psi_{mn}, \theta_{mn}, \tau_{mn}] T_n(z) \tag{2.22}$$

where $T_n(z) := \cos(n\cos^{-1}(2z-1))$. Resolution varies from $(N, M) = (30, 30)$ to $(80, 80)$ to ensure numerical accuracy as Ra increases and L changes .

3. Connecting to Hassanzadeh *et al.* (2014)

The conductive temperature profile $\tau = 1 - z$ is a spectrally-stable background field until $Ra = 27\pi^4/4$ in a domain of size $L = 2\sqrt{2}$ when energy instability first starts. Ensuring that the marginal fluctuation fields (θ, ψ) - hereafter a *mode* - stay marginal (called *pinning* as was done in PK03), the optimal solution was then tracked up to $Ra = 10^7$ with the domain size $L = L^*(Ra)$ simultaneously optimized to yield the highest heat flux at a given Ra: see figure 1. The calculated Nu values correspond exactly with those found by Hassanzadeh et al. (2014) (as do flow fields computed at $Ra = 10^5$ and 10^6; see the inset of figure 1). This indicates that Hassanzadeh et al.'s (2014) wall-to-wall transport approach is equivalent to the background method when a single mode is considered.

In their wall-to-wall optimal control approach, however, Hassanzadeh et al. (2014) had no way of identifying whether their local optimal was in fact the global optimal. It should be sufficiently close to the energy stability point but experience in other related problems (e.g. PK03) suggests that further modes in the spectral constraint eventually become marginal as Ra increases. The optimal solution should subsequently modify itself to keep these new modes marginal with concomitant adjustments in the Nu-scaling. Fortunately, in the background approach, the spectral constraint provides a check on whether a given Euler-Lagrange solution is the optimal solution. Solving the eigenvalue problem (2.13)-(2.14) for disturbances which are also periodic over $[0, L^*(Ra)]$ demonstrates that the eigenvalue (λ_1 in figure 2) of the first mode is pinned at 0, while a second mode becomes

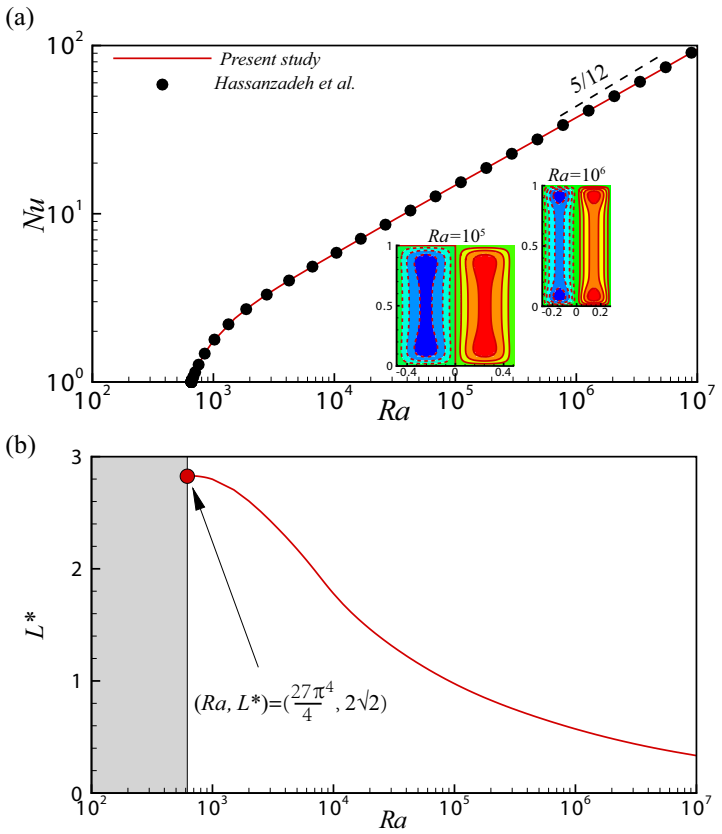

FIGURE 1. (a) The Nusselt number vs. the Rayleigh number tracked from the energy stability bifurcation point. The solution to the background upper bounding problem is optimized over the domain length and the data is in excellent agreement with Hassanzadeh *et al.*(2014) (their data courtesy of Dr A. Souza). Subplots show the flow streamfunctions at $Ra = 10^5$ and 10^6. (b) The optimal domain size L^* vs. the Rayleigh number. The bullet is the energy stability bifurcation point.

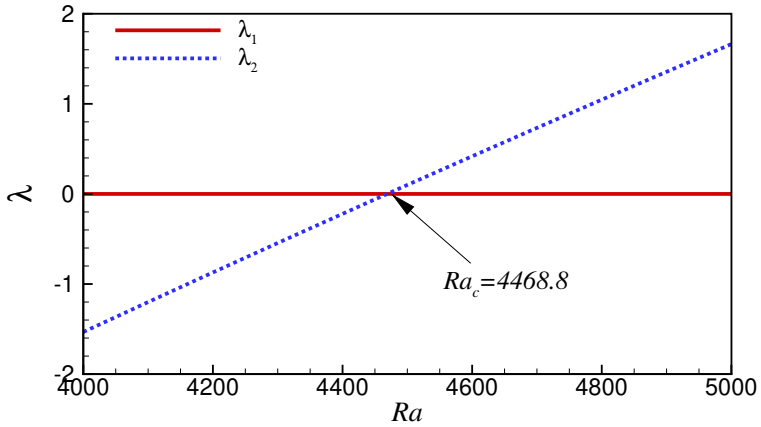

FIGURE 2. The first (λ_1) and second (λ_2) largest eigenvalues of the spectral constraint for $Ra < 4468.8$. At $Ra = 4468.8$, where they cross, the aspect ratio is $L^* = 2.234$.

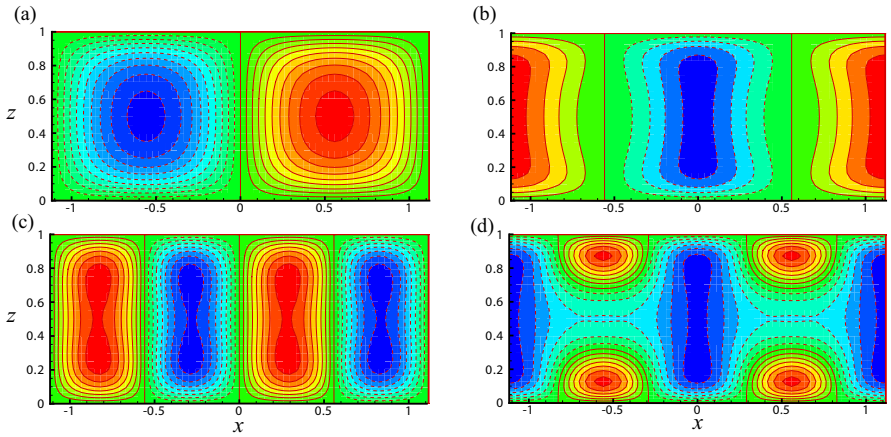

FIGURE 3. At $Ra = 4468.8$, the first mode (a) ψ_1 and (b) θ_1 and the new second mode (c) ψ_2 and (d) θ_2.

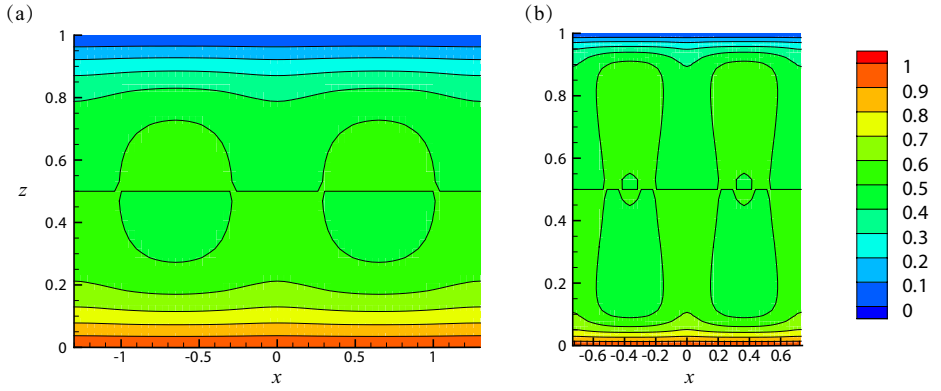

FIGURE 4. (a) The optimal background field τ plotted at $Ra = 2,000$ and (b) the non-optimal 1-mode solution at $Ra = 20,000$.

marginal at $Ra = 4468.8$ for an aspect ratio $L^* = 2.234$. This suggests that Hassanzadeh *et al.*'s (2014) result is either *not* a bound for $Ra > 4468.8$ or that the background bound has started to overestimate the actual maximal flux or, perhaps least likely, both. This divergence in results is apparently not because any further 2D bifurcations have been missed in the wall-to-wall calculations (Chini, *private communication*) but more a reflection of the 'duality gap' suggested by Souza et al. (2019) being realised. Figure 3(a,b) shows the first mode (ψ_1, θ_1) with wavenumber $\alpha_1 := 2\pi/L^*$ so the flow field contains one pair of convection cells. The second mode (ψ_2, θ_2) with $\alpha_2 = 2\alpha_1$ illustrated in figure 3(c,d) has two pairs of convection cells. The optimal background field at $Ra = 2000$ is shown in figure 4(a) and the now non-optimal 1-mode solution at $Ra = 20,000$ is shown in figure 4(b). In both cases the field is weakly 2-dimensional indicating that the first mode consists of non-monochromatic (i.e. non-single α) velocity and temperature fields. The emergence of the second mode at $Ra = 4468.8$ indicates that the background profile is now degenerate in a way which has important implications for solving the Euler-Lagrange equations for higher Ra while respecting the spectral constraint. We discuss this issue in the following section.

4. Multi-modal optimals

When a new mode becomes marginal in the spectral constraint as the background field τ evolves with Ra, a further 'pinning' constraint needs to be added to keep the new mode marginal in the spectral constraint as Ra increases further. This procedure is thoroughly discussed in Doering & Constantin (1996) and implemented in PK03 for a background field of lower dimensionality than the fluctuation field. In this situation, an example of which is using a 1D case $\tau = \tau(z)$ in the 2D Rayleigh-Bénard problem, the fluctuation field can be Fourier-transformed over the spatial dimension(s) across which τ is invariant and then considered parameterized by the Fourier wavenumber k. Different spectrally-marginal fluctuation fields have different k and are then naturally orthogonal under averaging over this spatial dimension. This means that the Euler-Lagrange equations (2.15)-(2.18),

$$0 = \nabla^2\theta_j - J(\tau,\psi_j), \qquad j = 1,\ldots,N \tag{4.1}$$

$$0 = \frac{a}{Ra}\nabla^4\psi_i - J(\tau,\theta_i), \qquad j = 1,\ldots,N \tag{4.2}$$

$$0 = \tau_{zz} - \sum_{j=1}^{N}\overline{J(\theta_j,\psi_j)}, \tag{4.3}$$

$$\langle|\tau_z|^2\rangle - 1 = \frac{(1-a)}{Ra}\sum_{j=1}^{N}\langle|\nabla^2\psi_j|^2\rangle, \tag{4.4}$$

(the overbar represents averaging over x) can simply be extended to include the new marginal mode

$$(\theta_{N+1},\psi_{N+1})(x,z) = (\hat{\theta}_{N+1},\hat{\psi}_{N+1})(z)e^{ik_{N+1}x} \tag{4.5}$$

when it appears. Equivalently, the Lagrangian is just

$$\mathscr{L} = \frac{1}{1-a}\left[\langle {\tau_z}^2\rangle - a\right] - \frac{1}{1-a}\mathscr{G} \tag{4.6}$$

where $\mathscr{G}$ naturally partitions into the contributions from the various marginal modes as follows

$$\mathscr{G} = \langle\frac{a}{Ra}|\nabla^2\psi|^2 + |\boldsymbol{\nabla}\theta|^2 + 2\theta J(\tau,\psi)\rangle = \sum_{j=1}^{N+1}\mathscr{G}_j \tag{4.7}$$

where

$$\mathscr{G}_j := \langle\frac{a}{Ra}|\nabla^2\psi_j|^2 + |\boldsymbol{\nabla}\theta_j|^2 + 2\theta J(\tau,\psi_j)\rangle. \tag{4.8}$$

The appearance of a new spectrally-marginal mode merely extends the set of wavenumbers contributing to the definition of the fluctuation field by one,

$$(\psi,\theta)(x,z) := \sum_{j=1}^{N+1}(\psi_j,\theta_j) = \sum_{j=1}^{N+1}(\hat{\psi}_j,\hat{\theta}_j)(z)e^{ik_jx}. \tag{4.9}$$

Importantly, this means it is possible to talk about the unique optimal solution of the variational problem which satisfies the imposed physical constraints as being

$$(\psi,T)(x,z) = (0,\tau)(z) + \sum_{j=1}^{N}(\hat{\psi}_j,\hat{\theta}_j)(z)e^{ik_jx}, \tag{4.10}$$

i.e. the spectral constraint is satisfied at a saddle point of $\mathscr{L}$.

This pleasing situation in which the marginal fluctuation fields have a physical interpretation changes, however, when the dimensionality of the background field equals the dimensionality of the problem (the case here), or, pathologically, there is more than one marginal mode for a *given* wavenumber (see chapter 3 of Fantuzzi 2018). In these scenarios, the natural orthogonality property of different marginal fluctuation fields disappears with the result that the physical meaning of the fluctuation fields is lost. To see this, the key is to realise that pinning the marginal fluctuation fields is done (Doering & Constantin 1996) as before by writing the Lagrangian as

$$\mathscr{L} = \frac{1}{1-a}(\langle|\boldsymbol{\nabla}\tau|^2\rangle - a) - \frac{1}{1-a}\sum_{j=1}\mathscr{G}_j. \tag{4.11}$$

The constraint that each $\mathscr{G}_j$ vanishes pins the j^{th} mode to be marginal (the Lagrange multiplier imposing this is absorbed into the amplitude of the j^{th} marginal fluctuation field) while $\mathscr{G} > 0$ for all other fluctuation fields. However, since the modes (ψ_j, θ_j) are not now orthogonal,

$$\sum_{j=1}^{N}\mathscr{G}_j \neq \mathscr{G} := \langle\frac{b}{Ra}|\nabla^2\psi|^2 + |\boldsymbol{\nabla}\theta|^2 + 2\theta J(\tau,\psi)\rangle \tag{4.12}$$

for $N \geqslant 2$ where

$$(\psi,\theta)(x,z) = \sum_{j=1}^{N}(\psi_j,\theta_j)(x,z) \tag{4.13}$$

is taken as the total optimal fluctuation field. In fact, $\mathscr{G} > 0$ and so this total optimal field is *not* a solution of the heat equation. The clear implication is that the spectral constraint is not satisfied for $N \geqslant 2$ at *any* saddle point of the Lagrangian (2.7) where the steady heat equation is imposed. Consequently, the optimization procedure is forced to find an optimal away from the saddle points of the Lagrangian (2.7) where the spectral constraint is satisfied to deliver a bound.

The fact that the form of the Lagrangian in (4.11) is different from that in (2.7) warrants further explanation. The Lagrangian in (2.7) is constructed in the usual way from the functional to be maximized (the heat flux) and the constraints to be applied with the highest stationary point being the maximal heat flux under the imposed constraints. The background method adds a *further* constraint by restricting the set of fields over which the Lagrangian is made stationary. Specifically, the allowed set (2.10) guarantees that any $(\tau, a) \in \Omega$ will give a value of the Lagrangian *greater or equal to* that at the highest stationary point: see the discussion in §2.2. When the background temperature field is of lower dimension than the physical problem - specifically here $\tau = \tau(z)$ in 2D Rayleigh-Benard convection, this extra requirement emerges naturally from the Lagrangian (2.7) since

$$\mathscr{G} = \sum_{j=1}^{N}\mathscr{G}_j \tag{4.14}$$

as the contributions to $\mathscr{G}$ from the various critical wavenumbers naturally separate. Put another way, the Lagrangians (4.11) and (2.7) are one and the same in this situation. There is, however, a further check that all other wavenumbers give $\mathscr{G} > 0$ which gets added onto the process of making the Lagrangian (2.7) stationary. Now, if τ has the same dimension as the physical problem, then the spectral problem also has the same dimension as the physical problem and (4.14) no longer holds (unless, trivially, $N = 1$). In this case the Lagrangian (2.7) has to be modified to that in (4.11) and the optimal

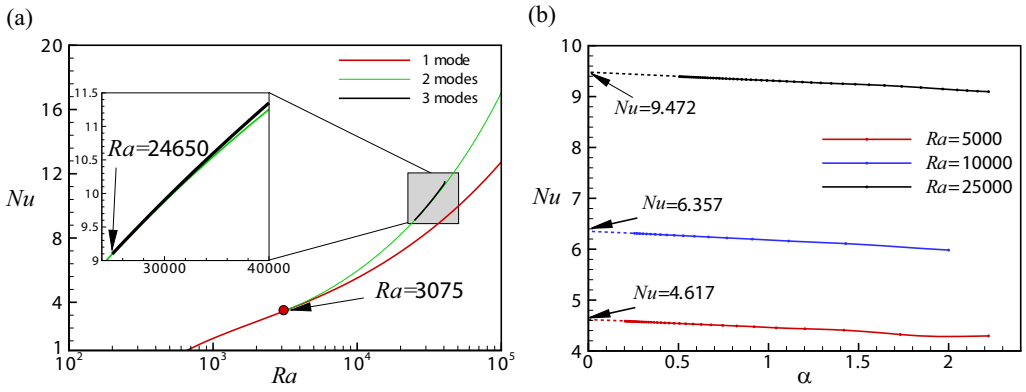

FIGURE 5. (a) The Nusselt number Nu vs. the Rayleigh number Ra at fixed aspect ratio $L = 2\sqrt{2}$. (b) Nu vs. $\alpha = 2\pi/L$.

solution no longer corresponds with the highest stationary point of (2.7). Instead, the highest stationary point of (4.11) will overestimate the highest stationary point of (2.7) as soon as $N \geqslant 2$ and in particular, the constraints imposed directly in (2.7) will not be satisfied.

From a different perspective, Souza et al. (2019) have also recently argued that this should happen when exploring the connection between the wall-to-wall approach (a max-min problem) with the associated background method (a min-max problem). A duality gap means that

$$\text{(wall}-\text{to}-\text{wall)} \qquad \sup_{\theta,\mathbf{u}} \inf_{\tau} \mathscr{L} \;<\; \inf_{\tau} \sup_{\theta,\mathbf{u}} \mathscr{L} \qquad \text{(background)} \tag{4.15}$$

(making the connection $\eta = \tau - (1 - z)$ and $\zeta = \theta$ with the variables used by Souza et al 2019) where the optimal solution to the wall-to-wall problem *is* achieved at a stationary point of $\mathscr{L}$ thereby implying that to the background method is not. They also supply a simple quadratic polynomial in 5 variables to illustrate the phenomenon. The calculations described in the next section confirm that this gap starts to exist as soon as $N = 2$.

5. Extending Hassanzedah *et al.* with a symmetric 2D background field $\tau(x, z)$

To explore multi-modal bounding solutions, a first series of computations were done in the fixed domain $L = 2\sqrt{2}$. In this geometry, the first mode appears at $Ra = 27\pi^4/4$ (the energy stability threshold), the second mode at $Ra = 3,075$ and the third mode at $Ra = 24,650$. The 1-mode and 2-mode optimal solution branches could be easily continued up to $Ra = 10^5$ whereas the 3-mode solution branch proved difficult to continue much beyond $Ra > 40,000$ due to numerical issues: see figure 5(a). The 3-mode solution, which provides an upper bound in this geometry over at least the range $24,650 \leqslant Ra \leqslant 40,000$, presents only a modest correction to the 2-mode optimal solution which is no longer a bound for these Ra.

A second series of computations were then carried out to investigate the dependence of the Nu-bound on the aspect ratio L. Three different Ra values were chosen to explore the dependence of the bound on L: $Ra = 5000$ and $10,000$ where the bound is given by a 2-mode solution, and $Ra = 25,000$ where the bound is given by a 3-mode solution. In all three cases, the largest bound is achieved as the aspect ratio $L \to \infty$: see figure 5(b).

This is very different from the optimal control results of Hassanzadeh *et al.*(2014) where the optimal aspect ratio scales like $Ra^{-1/4}$ and so vanishes as $Ra \to \infty$.

Figure 6 shows the structure of the two modes at $Ra = 10^4$. The fluctuation fields ψ_i and θ_i for both $i = 1$ and 2 have a convection roll structure and increasing L just means that more of the rolls fit into the domain. On closer inspection it is clear that the rolls are slightly different near to $x = 0$ and $x = \pm\frac{1}{2}L$ where they are forced to have a certain symmetry (symmetry around $x = 0$ and periodicity over a length L force symmetry about $x = \pm\frac{1}{2}L$ as well). When the domain is short, e.g. $L = \pi$, the background field is clearly two-dimensional as seen in figure 7. However, as L increases to $L = 8\pi$, the background field become predominantly one-dimensional (1-D) away from the imposed lines of symmetry at $x = 0$ and $x = \pm\frac{1}{2}L$ (the ends of the domain shown). Plotting the streamfunctions ψ_1 and ψ_2 over this long domain - see figure 8 - confirms that the convection cells are similar away from the symmetry lines ('zone 1' in figure 8) where τ is predominantly 1-D but are quite different close to the symmetry lines ('zone 2') where τ is clearly 2-D.

The structure of the optimal fields (both background and fluctuation) and the fact that the bound is maximised as $L \to \infty$ indicate that the optimal solution is trying to minimise the effect of the imposed symmetry requirements at $x = 0$ and $x = \pm\frac{1}{2}L$. Without this imposed symmetry, the problem becomes translationally invariant and the optimal solution must be 1-D by the convexity result in §2.3 (Doering & Constantin 1996). There is another simple way to see this. Since the bounding functional $f(\tau, a)$ (see (2.19)) is strictly convex in both $\tau(x, z)$ and a, any 2-D solution $(\tau_{2D}(x, z), a) \in \Omega$ satisfies

$$f\left(\frac{1}{N}\sum_{j=1}^{N}\tau_{2D}(x + \tfrac{jL}{N}, z), a\right) < \frac{1}{N}\sum_{j=1}^{N} f\left(\tau_{2D}(x + \tfrac{jL}{N}, z), a\right) \tag{5.1}$$

by Jensen's inequality. Taking the limit $N \to \infty$ in the left hand side and using translational invariance of the problem in the right hand side leads to

$$f\left(\tau_{1D}(z) := \frac{1}{L}\int_0^L \tau_{2D}(x, z)\, dx, a\right) < f(\tau_{2D}(x, z), a) \tag{5.2}$$

so that a 1-D background field always produces a better bound than a 2-D field. The results in figure 7 indicate that this is what the optimal solution to the background problem is trying to achieve.

5.1. *Lifting the symmetry: 1D background field*

Calculation of the optimal solution assuming from the onset that the background field is 1-D simplifies the computation since the fluctuation fields can then be parametrised by their single wavenumber in x (as in (4.5)). In this case, rather than setting a domain periodicity and insisting the fluctuation wavenumbers be consistent with this, the wavenumbers themselves can be optimised over as real continuous variables meaning, in effect, that L is infinite. For example, the Euler-Lagrange equation corresponding to the m^{th} wavenumber k_m is

$$\delta\mathscr{L}/\delta k_m := -2\int_0^1 a k_m(u_m^2 + w_m^2) + k_m\theta_m^2 dz + \int_0^1 p_m u_m dz = 0. \tag{5.3}$$

With this formulation, Newton's method with branch continuation proved much faster than the time-stepping approach. It took around 4 hours cputime on a 2.8Ghz laptop using Newton's method to obtain the optimal solution from $Ra = 27\pi^4/4$ up to $Ra = 5 \times 10^8$ while the time-stepping approach took at least a day to generate a single point

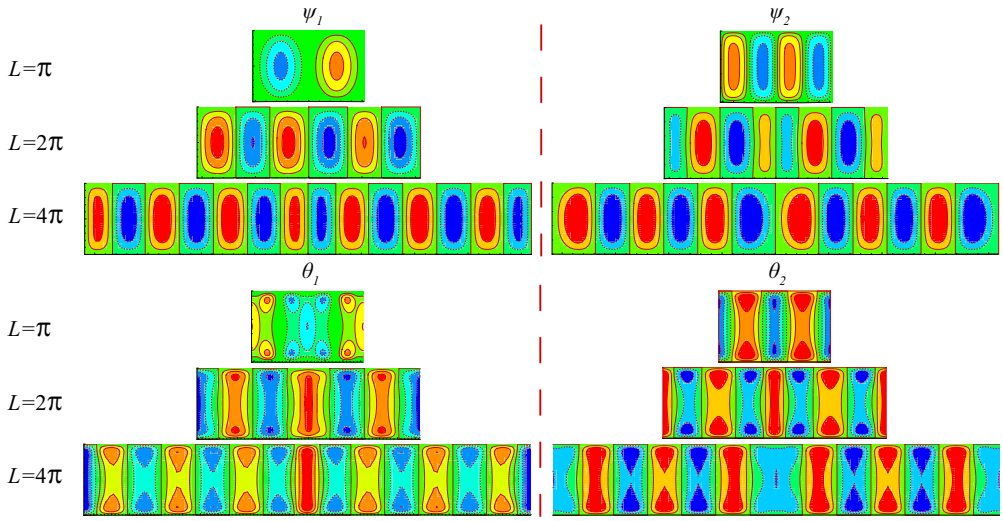

FIGURE 6. Left panel: The first mode (ψ_1, θ_1); right panel: the second mode (ψ_2, θ_2) at $Ra = 10^4$ (only two critical modes are present for this Ra).

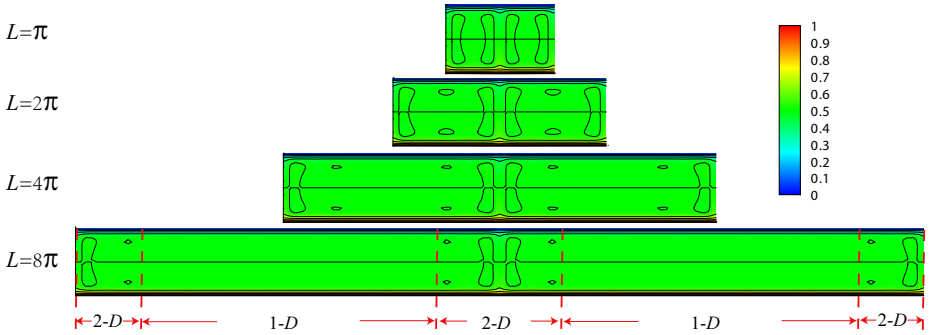

FIGURE 7. The background field plotted at different aspect ratios at $Ra = 10^4$.

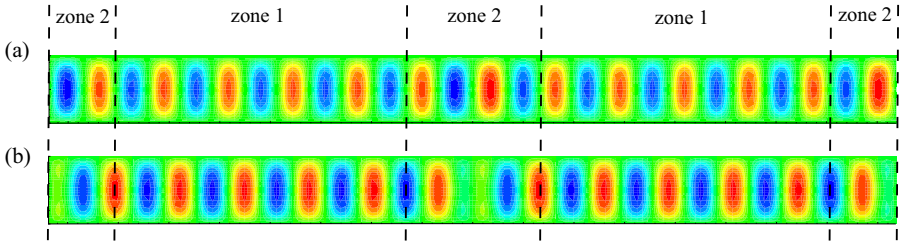

FIGURE 8. The two critical flow fields at $Ra = 10^4$ and $L = 8\pi$: (a) ψ_1 and (b) ψ_2.

at $Ra = 5 \times 10^8$. However, when the domain is fixed, Newton's method becomes very inefficient as the critical wavenumbers k_m are discrete and cannot be tracked using a (continuous) continuation method: in this case, time-stepping is the better choice. The numerical solution of the one-dimensional background problem gives the upper bound of $Nu \leqslant 0.055 Ra^{1/2}$ as shown in figure 9(a) with 5 critical modes present by $Ra = 10^9$ (see figure 9(b)). This result has the same scaling exponent as the non-slip result $Nu \leqslant 0.026 Ra^{1/2}$ of PK03 but with a larger numerical coefficent as should be expected for

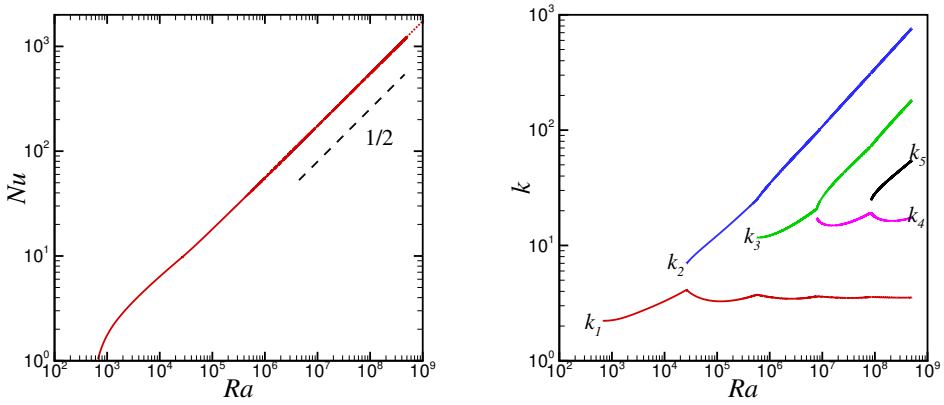

FIGURE 9. Left panel: the upper bound of Nu vs. the Rayleigh number Ra with $Nu \leqslant 0.055Ra^{1/2}$ in the asymptotic regime. right panel: the bifurcation diagram of critical wave number k_m vs. the Rayleigh number.

stress-free boundary conditions. The prior work of Wen *et al.* (2015) indicates that adding a further enstrophy constraint (possible only in stress-free 2D convection) significantly improves the bound obtained here down to $Nu \leqslant 0.106Ra^{5/12}$.

It is worth briefly discussing how the critical wavenumbers which appear in the 1-D and 2-D background field calculations are related. Figure 9(b) indicates that at $Ra = 10^4$ there is only one critical wavenumber $k_1 = 3.284$ for the 1-D background problem. However, for the 2-D (symmetric) background problem, there are two critical modes as seen in figure 6 and figure 8 with both having an approximate wavenumber ≈ 3.3. Both these modes are forced to be antisymmetric (and so in phase) about $x = 0$ and $x = \pm L/2$ (zone 2 in figure 8) but away from these points endeavour to be approximately $\pi/2$ out of phase (zone 1 in figure 8). With this phase difference together with matching amplitudes so

$$\psi_1 = f(z)\sin(k_1 x), \qquad \theta_1 = g(z)\cos(k_1 x), \tag{5.4}$$

$$\psi_2 = f(z)\sin(k_1 x + \pi/2), \quad \theta_2 = g(z)\cos(k_1 x + \pi/2), \tag{5.5}$$

the nonlinear term in Eq.(2.17)

$$\sum_{i=1}^{2}\left[\frac{\partial\psi_i}{\partial z}\frac{\partial\theta_i}{\partial x} - \frac{\partial\psi_i}{\partial x}\frac{\partial\theta_i}{\partial z}\right] = -k_1\left[\frac{df}{dz}g + \frac{dg}{dz}f\right] \tag{5.6}$$

is x independent and therefore can only drive a 1-D background field. From another perspective, the 1-D background field problem really has two modes with $k = 3.284$ but only one needs to be tracked as the nonlinear term is horizontally averaged (e.g. see (4.3)) ensuring that the background field stays 1-D.

At $Ra = 25,000$ there is even a third mode in the 2-D background problem compared to still only 1 mode in the 1-D background problem (the second wavenumber k_2 appears at $Ra \approx 26,450$). Figure 10 shows that in fact ψ_3 is only significant in zone 2 where the imposed symmetries dominate. In zone 1 where the background field is essentially 1-D profile, ψ_3 vanishes.

The conclusion of the computations so far is that using a 2D background temperature field does not improve the upper bound produced by a 1D background temperature field, a realisation also reached independently in Doering & Tobasco (2019) (see their lemma

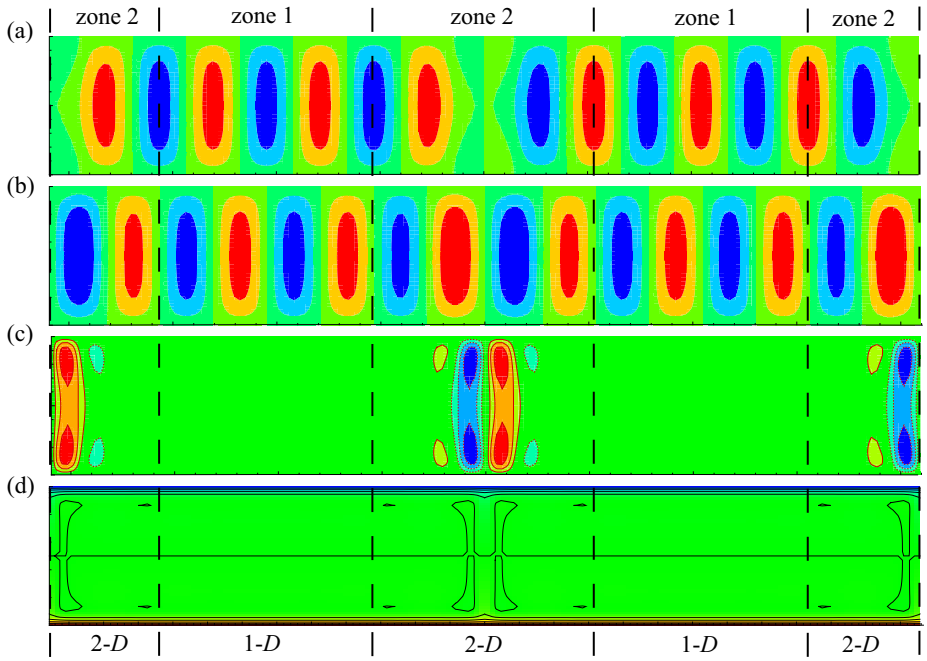

FIGURE 10. (a-c) The profiles of the three critical flow fields ψ_1, ψ_2, ψ_3 at $Ra = 25000$, $L = 4\pi$. (d) The profile of the two-dimensional optimal background field.

6.1). The next obvious question is whether adding a background velocity field helps either and we now turn our attention to this.

6. Imposing the steady momentum equation: $\boldsymbol{\phi} \neq \mathbf{0}$

In this section, we attempt to improve the bound by using a background temperature field *and* a background velocity field of the same dimension as the physical problem which means that the full, albeit steady, momentum and heat equations are imposed as constraints. Importantly, the optimization problem is no longer convex and so permitting $\boldsymbol{\phi} \neq \mathbf{0}$ could produce a better (reduced) upper bound. However, numerical calculations suggest otherwise with $\boldsymbol{\phi}$ remaining zero after every bifurcation. To attempt to explain this, we use an inductive bifurcation analysis to show that if $\boldsymbol{\phi} = \mathbf{0}$ before a bifurcation then it remains $\mathbf{0}$ after it too implying that the continuous branch of optimals found by branch tracking out of the energy stability point always has $\boldsymbol{\phi} = \mathbf{0}$.

The analysis begins by constructing the following Lagrangian:

$$\mathscr{L} = \langle |\boldsymbol{\nabla} T|^2 \rangle - \frac{a}{\sigma Ra} \langle \boldsymbol{v} \cdot \boldsymbol{\mathcal{N}} \rangle - 2 \langle \theta \, \mathcal{H} \rangle \tag{6.1}$$

which, after introducing the extended background decomposition

$$\boldsymbol{u} = \boldsymbol{\phi} + \boldsymbol{v}, \quad T = \tau + \theta, \tag{6.2}$$

(now both $\boldsymbol{\phi}$ and $\boldsymbol{v}$ vanish on the boundaries $z = 0, 1$) can be rewritten as

$$\mathscr{L} = \frac{\langle |\boldsymbol{\nabla} \tau|^2 \rangle - a}{1 - a} - \frac{a}{1 - a} \langle \phi_z \tau \rangle - \frac{1}{1 - a} \mathscr{G} \tag{6.3}$$

($\phi_z := \boldsymbol{\phi}\cdot\hat{\mathbf{z}}$) where

$$\mathscr{G}(\boldsymbol{v},\theta) := \langle 2\theta(\boldsymbol{\phi}+\boldsymbol{v})\cdot\boldsymbol{\nabla}\tau + |\boldsymbol{\nabla}\theta|^2\rangle + a\langle\phi_z\theta\rangle \\ + \langle\frac{a}{\sigma Ra}\boldsymbol{v}\cdot\boldsymbol{v}\cdot\boldsymbol{\nabla}\boldsymbol{\phi} + \frac{a}{\sigma Ra}\boldsymbol{v}\cdot\boldsymbol{\phi}\cdot\boldsymbol{\nabla}\boldsymbol{\phi} + \frac{a}{Ra}|\boldsymbol{\nabla}\boldsymbol{v}|^2 - \frac{a}{Ra}\boldsymbol{v}\cdot\nabla^2\boldsymbol{\phi}\rangle$$

(note $\mathscr{G}$ depends parametrically on τ, $\boldsymbol{\phi}$, a, σ and Ra but this is suppressed for clarity). If $\inf_{\boldsymbol{v},\theta}\mathscr{G}$ exists (and necessarily $0<a<1$), a bound is then given by

$$Nu \leqslant \frac{\langle|\boldsymbol{\nabla}\tau|^2\rangle - a}{1-a} - \frac{a}{1-a}\langle\phi_z\tau\rangle - \frac{1}{1-a}\inf_{\boldsymbol{v},\theta}\mathscr{G}(\tau,\boldsymbol{\phi}). \tag{6.4}$$

Minimization of $\mathscr{G}$ with respect to incompressible $\boldsymbol{v}$ and θ requires

$$-\frac{2a}{Ra}\nabla^2\boldsymbol{v} + \frac{a}{\sigma Ra}\boldsymbol{v}\cdot(\boldsymbol{\nabla}\boldsymbol{\phi}+\boldsymbol{\nabla}\boldsymbol{\phi}^{\mathrm{T}}) + \frac{a}{\sigma Ra}\boldsymbol{\phi}\cdot\boldsymbol{\nabla}\boldsymbol{\phi} - \frac{a}{\sigma Ra}\nabla^2\boldsymbol{\phi} + 2\theta\boldsymbol{\nabla}\tau + \boldsymbol{\nabla}p = 0, \tag{6.5}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{v} = 0, \tag{6.6}$$

$$-2\nabla^2\theta + a\phi_z + 2(\boldsymbol{v}+\boldsymbol{\phi})\cdot\boldsymbol{\nabla}\tau = 0 \tag{6.7}$$

the solution of which is denoted as $(\boldsymbol{v}_0,\theta_0)$. The Lagrangian can then be written as

$$\mathscr{L} = \frac{\langle|\boldsymbol{\nabla}\tau|^2\rangle - a}{1-a} - \frac{a}{1-a}\langle\phi_z\tau\rangle - \frac{1}{1-a}\left\{\mathscr{G}(\boldsymbol{v}_0,\theta_0;\tau,\boldsymbol{\phi}) + \sum_{i=1}^{N}\mathscr{H}(\boldsymbol{v}_i,\theta_i)\right\} \tag{6.8}$$

where $\boldsymbol{v}_0$, $\boldsymbol{v}_i$ ($i=1\ldots N$) and $\boldsymbol{\phi}$ are incompressible fields and

$$\mathscr{H}(\boldsymbol{v},\theta) := \langle 2\theta\boldsymbol{v}\cdot\boldsymbol{\nabla}\tau + |\boldsymbol{\nabla}\theta|^2 + \frac{a}{\sigma Ra}\boldsymbol{v}\cdot\boldsymbol{v}\cdot\boldsymbol{\nabla}\boldsymbol{\phi} + \frac{a}{Ra}|\boldsymbol{\nabla}\boldsymbol{v}|^2\rangle \tag{6.9}$$

is a purely quadratic functional of $(\boldsymbol{v},\theta)$ which must be positive semi-definite - the spectral constraint - for $\inf\mathscr{G}$ to exist. The fields $(\boldsymbol{v}_i,\theta_i)$ are marginal in that $\mathscr{H}(\boldsymbol{v}_i,\theta_i)=0$ and their number N increases with Ra. The aim is to minimize the upper bound over τ, $\boldsymbol{\phi}$ and a at fixed σ and Ra subject to this spectral constraint. The Euler-Lagrange equations are: the spectral constraint equations for $\boldsymbol{v}_i$ and θ_i

$$-(1-a)Ra\frac{\delta\mathscr{L}}{\delta\boldsymbol{v}_i} := -2a\nabla^2\boldsymbol{v}_i + \frac{a}{\sigma}\boldsymbol{v}_i\cdot(\boldsymbol{\nabla}\boldsymbol{\phi}+\boldsymbol{\nabla}\boldsymbol{\phi}^{\mathrm{T}}) + 2Ra\theta_i\boldsymbol{\nabla}\tau + \boldsymbol{\nabla}p_i = \mathbf{0}, \qquad i=1\ldots N \tag{6.10}$$

$$-(1-a)\frac{\delta\mathscr{L}}{\delta\theta_i} := -2\nabla^2\theta_i + 2\boldsymbol{v}_i\cdot\boldsymbol{\nabla}\tau = 0. \qquad i=1\ldots N; \tag{6.11}$$

the forced field equations for $\boldsymbol{v}_0$ and θ_0

$$-(1-a)Ra\frac{\delta\mathscr{L}}{\delta\boldsymbol{v}_0} := -2a\nabla^2\boldsymbol{v}_0 + \frac{a}{\sigma}\boldsymbol{v}_0\cdot(\boldsymbol{\nabla}\boldsymbol{\phi}+\boldsymbol{\nabla}\boldsymbol{\phi}^{\mathrm{T}}) + 2Ra\theta_0\boldsymbol{\nabla}\tau + \boldsymbol{\nabla}p_0 + \frac{a}{\sigma}\boldsymbol{\phi}\cdot\boldsymbol{\nabla}\boldsymbol{\phi} - a\nabla^2\boldsymbol{\phi} = \mathbf{0}; \tag{6.12}$$

$$-(1-a)\frac{\delta\mathscr{L}}{\delta\theta_0} := -2\nabla^2\theta_0 + 2\boldsymbol{v}_0\cdot\boldsymbol{\nabla}\tau + 2\boldsymbol{\phi}\cdot\boldsymbol{\nabla}\tau + a\phi_z = 0; \tag{6.13}$$

the background field equations

$$(1-a)Ra\frac{\delta\mathscr{L}}{\delta\boldsymbol{\phi}} := a\nabla^2\boldsymbol{v}_0 - aRa(\tau+\theta_0)\boldsymbol{e}_z - 2Ra\theta_0\boldsymbol{\nabla}\tau + \boldsymbol{\nabla}q + \frac{a}{\sigma}\boldsymbol{v}_0\cdot\boldsymbol{\nabla}\boldsymbol{v}_0 \\ - \frac{a}{\sigma}(\boldsymbol{v}_0\cdot\boldsymbol{\nabla}\boldsymbol{\phi}^{\mathrm{T}} - \boldsymbol{\phi}\cdot\boldsymbol{\nabla}\boldsymbol{v}_0) + \frac{a}{\sigma}\sum_{i=1}^{N}\boldsymbol{v}_i\cdot\boldsymbol{\nabla}\boldsymbol{v}_i = \mathbf{0}, \tag{6.14}$$

$$(1-a)\frac{\delta\mathscr{L}}{\delta\tau} := -2\nabla^2\tau - a\phi_z + 2(\boldsymbol{v}_0+\boldsymbol{\phi})\cdot\boldsymbol{\nabla}\theta_0 + 2\sum_{i=1}^{N}\boldsymbol{v}_i\cdot\boldsymbol{\nabla}\theta_i = 0; \tag{6.15}$$

and finally the balance parameter equation

$$(1-a)\frac{\delta\mathscr{L}}{\delta a} := \mathscr{L} - 1 - \langle\phi_z\tau\rangle - \Big\{\sum_{i=1}\langle\frac{1}{\sigma Ra}\boldsymbol{v}_i\cdot\boldsymbol{v}_i\cdot\boldsymbol{\nabla}\boldsymbol{\phi} + \frac{1}{Ra}|\boldsymbol{\nabla}\boldsymbol{v}_i|^2\rangle$$
$$+\langle\phi_z\theta_0 + \frac{1}{Ra}|\boldsymbol{\nabla}\boldsymbol{v}_0|^2 + \frac{1}{\sigma Ra}\boldsymbol{v}_0\cdot\boldsymbol{v}_0\cdot\boldsymbol{\nabla}\boldsymbol{\phi} + \frac{1}{\sigma Ra}\boldsymbol{v}_0\cdot\boldsymbol{\phi}\cdot\boldsymbol{\nabla}\boldsymbol{\phi} - \frac{1}{Ra}\boldsymbol{v}_0\cdot\nabla^2\boldsymbol{\phi}\rangle\Big\}. \tag{6.16}$$

The pressure-like quantities have been rescaled as follows $Ra\,p_i \to p_i$, $Ra\,p_0 \to p_0$ and $(1-a)Ra\,q \to q$) and incompressibility conditions on $\boldsymbol{v}_0$, $\boldsymbol{v}_i$ and $\boldsymbol{\phi}$ are left implicit. A key point here is that the forced field pair $(\boldsymbol{v}_0, \theta_0)$ is *not* marginal in the spectral constraint (otherwise equations (6.12) and (6.13) could not be satisfied) and since $\mathscr{H}$ is positive semi-definite, $\mathscr{H}(\boldsymbol{v}_0, \theta_0) > 0$.

6.1. *The first bifurcation point*

The solution at the first critical point $Ra_c = 27\pi^4/4$ is $\tau = 1 - z$ and $\boldsymbol{\phi} = 0$, $(\boldsymbol{v}_0, \theta_0) = (\boldsymbol{0}, 0)$ and $a = 1$. At $Ra = Ra_c$, the spectral constraint becomes marginal for the first time, i.e. there is a non-trivial solution to the spectral problem

$$-2\nabla^2\boldsymbol{v}_i + 2Ra\theta_i\boldsymbol{\nabla}\tau + \boldsymbol{\nabla}p_i = 0, \tag{6.17}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{v}_i = 0, \tag{6.18}$$

$$-\nabla^2\theta_i + \boldsymbol{v}_i\cdot\boldsymbol{\nabla}\tau = 0. \tag{6.19}$$

There are two different modes (using symmetries):

$$(\boldsymbol{v}_1, \theta_1) = A_1(\,U(z)\sin(kx)\boldsymbol{e}_x + W(z)\cos(kx)\boldsymbol{e}_z, \Theta(z)\cos(kx)\,) \tag{6.20}$$

$$(\boldsymbol{v}_2, \theta_2) = A_2(\,U(z)\cos(kx)\boldsymbol{e}_x - W(z)\sin(kx)\boldsymbol{e}_z, -\Theta(z)\sin(kx)\,) \tag{6.21}$$

Since $d\tau/dz = -1$, the structure in z is simple: $U := \pi\cos(\pi z)$, $W := -k\sin(\pi z)$ and $\Theta := -\frac{\sqrt{2}}{3\pi}\sin(\pi z)$ where $k = \pi/\sqrt{2}$. Slightly away from the critical point, $Ra = Ra_c + \varepsilon$, the fields need to be expanded as follows

$$\begin{aligned}
\tau &= \tau_0 + \varepsilon\tau_1 + \varepsilon^2\tau_2 + \ldots,\\
\boldsymbol{\phi} &= \quad \varepsilon\boldsymbol{\phi}_1 + \varepsilon^2\boldsymbol{\phi}_2 + \varepsilon^3\boldsymbol{\phi}_3\ldots,\\
\boldsymbol{v}_0 &= \quad \varepsilon\boldsymbol{v}_0^1 + \varepsilon^2\boldsymbol{v}_0^2 + \ldots,\\
\theta_0 &= \quad \varepsilon\theta_0^1 + \varepsilon^2\theta_0^2 + \ldots,\\
\boldsymbol{v}_i &= \varepsilon^{1/2}\boldsymbol{v}_i^0 + \varepsilon^{3/2}\boldsymbol{v}_i^1 + \varepsilon^{5/2}\boldsymbol{v}_i^2 + \ldots,\\
\theta_i &= \varepsilon^{1/2}\theta_i^0 + \varepsilon^{3/2}\theta_i^1 + \varepsilon^{5/2}\theta_i^2 + \ldots,\\
a &= \ a_0 + \varepsilon a_1 + \varepsilon^2 a_2 + \ldots
\end{aligned}$$

where $\tau_0 := 1 - z$ and $a_0 := 1$.

6.1.1. *Leading Order (e.g. the problem for v_i^0 and θ_i^0)*

To leading order, the spectral constraint is satisfied by $(\boldsymbol{v}_i^0, \theta_i^0)$ defined in (6.20-6.21) and these fields force the other leading order equations for the background fields

$$2\nabla^2\tau_1 + \phi_{1z} = 2\sum_{i=1,2} \boldsymbol{v}_i^0 \cdot \boldsymbol{\nabla}\theta_i^0, \tag{6.22}$$

$$\nabla^2\boldsymbol{v}_0^1 - Ra_c(\tau_1 + a_1\tau_0 - \theta_0^1)\boldsymbol{e}_z - \tau_0\boldsymbol{e}_z + \boldsymbol{\nabla}q = \underbrace{-\frac{1}{\sigma}\sum_{i=1,2} \boldsymbol{v}_i^0 \cdot \boldsymbol{\nabla}\boldsymbol{v}_i^0}_{balanced\ by\ pressure}. \tag{6.23}$$

($\phi_{1z} := \boldsymbol{\phi}_1 \cdot \hat{\mathbf{z}}$) which are coupled with the forced field equations

$$-2\nabla^2\boldsymbol{v}_0^1 + 2Ra_c\theta_0^1\boldsymbol{\nabla}\tau_0 + \boldsymbol{\nabla}p = \nabla^2\boldsymbol{\phi}_1, \tag{6.24}$$

$$-2\nabla^2\theta_0^1 + 2\boldsymbol{v}_0^1 \cdot \boldsymbol{\nabla}\tau_0 = \phi_{1z}. \tag{6.25}$$

The forcing term in (6.22) is

$$\sum_{i=1}^{2} \boldsymbol{v}_i^0 \cdot \boldsymbol{\nabla}\theta_i^0 = \frac{\pi(A_1^2 + A_2^2)}{6}\sin(2\pi z) \tag{6.26}$$

and in (6.23)

$$\sum_{i=1}^{2} \boldsymbol{v}_i^0 \cdot \boldsymbol{\nabla}\boldsymbol{v}_i^0 = \frac{k\pi^2}{2}(A_1^2 - A_2^2)\sin(2kx)\boldsymbol{e}_x + \frac{k^2\pi}{2}(A_1^2 + A_2^2)\sin(2\pi z)\boldsymbol{e}_z \tag{6.27}$$

(recall $k = \pi/\sqrt{2}$) so that simply

$$\boldsymbol{\phi}_1 = \boldsymbol{v}_0^1 = \mathbf{0}, \quad \theta_0^1 = 0, \quad q = c_1(A_1^2 - A_2^2)\cos(2kx) + c_2(A_1^2 + A_2^2)\cos(2\pi z) \tag{6.28}$$

$$\& \quad \tau_1 = c_3(A_1^2 + A_2^2)\sin(2\pi z) \tag{6.29}$$

where c_1, c_2 and c_3 are specific constants. Finally, the leading order balance (which is at $O(\varepsilon)$) in the balance parameter equation (6.16) is

$$-\frac{1}{a_1}\langle|\boldsymbol{\nabla}\tau|^2\rangle - \frac{1}{Ra_c}\sum_{i=1,2}\langle|\boldsymbol{\nabla}\boldsymbol{v}_i^0|^2\rangle = 0 \tag{6.30}$$

which relates $A_1^2 + A_2^2$ and a_1.

6.1.2. *Next Order (e.g. the problem for v_i^1 and θ_i^1)*

A further piece of information to identify the leading order fields comes from a solvability condition on the spectral constraint equations at next order ($(O(\varepsilon^{3/2}))$ which is

$$-2\nabla^2\boldsymbol{v}_i^1 + 2Ra_c\theta_i^1\boldsymbol{\nabla}\tau_0 + \boldsymbol{\nabla}p = 2a_1\nabla^2\boldsymbol{v}_i^0 - 2Ra_c\theta_i^0\boldsymbol{\nabla}\tau_1 - 2\theta_i^0\boldsymbol{\nabla}\tau_0, \tag{6.31}$$

$$-2\nabla^2\theta_i^1 + 2\boldsymbol{v}_i^1 \cdot \boldsymbol{\nabla}\tau_0 = -2\boldsymbol{v}_i^0 \cdot \boldsymbol{\nabla}\tau_1. \tag{6.32}$$

Formally, this has two solvability conditions:

$$\langle \boldsymbol{v}_i^0 \cdot (2a_1\nabla^2\boldsymbol{v}_i^0 - 2Ra_c\theta_i^0\boldsymbol{\nabla}\tau_1 - 2\theta_i^0\boldsymbol{\nabla}\tau_0) - 2\theta_i^0\boldsymbol{v}_i^0 \cdot \boldsymbol{\nabla}\tau_1\rangle = 0, \quad i = 1, 2 \tag{6.33}$$

but they are equivalent since τ_1 is 1-dimensional (i.e. solely a function of z) with the resulting condition providing a second equation linking a_1 and $A_1^2 + A_2^2$. These ((6.30) and (6.33)) can then be solved to give

$$A_1^2 + A_2^2 = 24, \quad a_1 = -\tfrac{9}{4} \tag{6.34}$$

and as a consequence $\tau_1 = -\frac{1}{\pi}\sin(2\pi z)$. The fields $(\boldsymbol{v}_i^1, \theta_i^1)$ depend linearly on $\boldsymbol{v}_i^0$, θ_i^0 and so can be written as

$$(\boldsymbol{v}_1^1, \theta_1^1) = A_1(\,\mathcal{U}(z)\sin(kx)\boldsymbol{e}_x + \mathcal{W}(z)\cos(kx)\boldsymbol{e}_z, \mathcal{T}(z)\cos(kx)\,) \tag{6.35}$$

$$(\boldsymbol{v}_2^1, \theta_2^1) = A_2(\,\mathcal{U}(z)\cos(kx)\boldsymbol{e}_x - \mathcal{W}(z)\sin(kx)\boldsymbol{e}_z, -\mathcal{T}(z)\sin(kx)\,) \tag{6.36}$$

where $\mathcal{U} \neq U$, $\mathcal{W} \neq W$. These fields along with $(\boldsymbol{v}_i^0, \theta_i^0)$ drive the higher order equations governing further corrections to the background fields and the forced fields. These are

$$2\nabla^2\tau_2 + \phi_{2z} = \underbrace{2\sum_{i=1,2} \boldsymbol{v}_i^0 \cdot \boldsymbol{\nabla}\theta_i^1 + \boldsymbol{v}_i^1 \cdot \boldsymbol{\nabla}\theta_i^0}_{driving\ term}. \tag{6.37}$$

where $\phi_{2z} := \boldsymbol{\phi}_2 \cdot \hat{\mathbf{z}}$ and

$$\begin{aligned}&\nabla^2\boldsymbol{v}_0^2 - Ra_c(a_0\tau_2 + a_2\tau_0 - \theta_0^2)\boldsymbol{e}_z + \boldsymbol{\nabla}q \\ &= \underbrace{([a_1 Ra_c + 1]\tau_1 + a_1\tau_0)\boldsymbol{e}_z - \frac{a_1}{\sigma}\sum_{i=1}^{2} \boldsymbol{v}_i^0 \cdot \boldsymbol{\nabla}\boldsymbol{v}_i^0}_{balanced\ by\ pressure} \underbrace{- \frac{1}{\sigma}\sum_{i=1}^{2} \boldsymbol{v}_i^0 \cdot \boldsymbol{\nabla}\boldsymbol{v}_i^1 + \boldsymbol{v}_i^1 \cdot \boldsymbol{\nabla}\boldsymbol{v}_i^0}_{driving\ term} = 0.\end{aligned} \tag{6.38}$$

$$-2\nabla^2\boldsymbol{v}_0^2 + 2Ra_c\theta_0^2\boldsymbol{\nabla}\tau_0 + \boldsymbol{\nabla}p_i = \nabla^2\boldsymbol{\phi}_2, \tag{6.39}$$

$$-2\nabla^2\theta_0^2 + 2\boldsymbol{v}_0^2 \cdot \boldsymbol{\nabla}\tau_0 = \phi_{2z}. \tag{6.40}$$

The apparent driving term $-a_1/\sigma\sum_{i=1}^{2} \boldsymbol{v}_i^0 \cdot \boldsymbol{\nabla}\boldsymbol{v}_i^0$ can be balanced by the pressure term (see (6.27)) as can $([a_1 Ra_c + 1]\tau_1 + a_1\tau_0)\boldsymbol{e}_z$. Also importantly for what follows, $Ra_c a_2\tau_0\boldsymbol{e}_z$ in (6.38) can also be absorbed into the pressure term which means that τ_2 and $\boldsymbol{\phi}_2$ do not depend on a_2 (this is crucial for the argument surrounding (6.50) below). This leaves the driving term for the 2D background temperature field

$$\begin{aligned}\sum_{i=1,2} \boldsymbol{v}_i^0 \cdot \boldsymbol{\nabla}\theta_i^1 + \boldsymbol{v}_i^1 \cdot \boldsymbol{\nabla}\theta_i^0 &= 12\left(-kU\mathcal{T} - k\mathcal{U}\Theta + W\frac{d\mathcal{T}}{dz} + \mathcal{W}\frac{d\Theta}{dz}\right) \\ &+ \frac{1}{2}(A_1^2 - A_2^2)\left(kU\mathcal{T} + k\mathcal{U}\Theta + W\frac{d\mathcal{T}}{dz} + \mathcal{W}\frac{d\Theta}{dz}\right)\cos(2kx).\end{aligned} \tag{6.41}$$

and the driving term in Eq.(6.38) for $\boldsymbol{v}_0^2$

$$\begin{aligned}\sum_{i=1}^{2} \boldsymbol{v}_i^0 \cdot \boldsymbol{\nabla}\boldsymbol{v}_i^1 + \boldsymbol{v}_i^1 \cdot \boldsymbol{\nabla}\boldsymbol{v}_i^0 &= \underbrace{12\left(-kU\mathcal{W} - k\mathcal{U}W + \frac{d(W\mathcal{W})}{dz}\right)\boldsymbol{e}_z}_{balanced\ by\ pressure} \\ &+ \frac{1}{2}(A_1^2 - A_2^2)\left(2kU\mathcal{U} + W\frac{d\mathcal{U}}{dz} + \mathcal{W}\frac{dU}{dz}\right)\sin(2kx)\boldsymbol{e}_x.\end{aligned} \tag{6.42}$$

Given the form of these driving terms, τ_2 can be split into two parts: a 1D part which depends only on z and a 2D part proportional to $A_1^2 - A_2^2$ which has both x and z dependence whereas the remaining corrections $\boldsymbol{\phi}_2$, $\boldsymbol{v}_0^2$ and θ_0^2 only have a 2D part proportional to $A_1^2 - A_2^2$, i.e.

$$\tau_2 = \tau_2^{1D}(z) + \tau_2^{2D}(x, z) := P_1(z) + (A_1^2 - A_2^2)P_2(z)\cos(2kx), \tag{6.43}$$

$$\boldsymbol{\phi}_2 = (A_1^2 - A_2^2)\Big[G_1(z)\sin(2kx)\boldsymbol{e}_x + G_2(z)\cos(2kx)\boldsymbol{e}_z\Big]. \tag{6.44}$$

(the expressions for $\boldsymbol{v}_0^2$ and θ_0^2 are not needed in what follows and hence suppressed). At this point $\boldsymbol{\phi}_2$ is now known as a function of $A_1^2 - A_2^2$. Further information about A_1 and A_2 comes from solvability conditions at the next order of the spectral constraint ($O(\varepsilon^{5/2})$).

Before pursuing this, we remark that the next order ($O(\varepsilon^2)$) of the balance equation (6.16) involves the higher order unknown $\boldsymbol{\phi}_3$ and so at this order a_2 is unspecified. In fact a_2 is also set by solvability conditions at $O(\varepsilon^{5/2})$ of the spectral constraint to which we now turn.

6.1.3. *Solvability at $O(\varepsilon^{5/2})$ in the problem for v_i^2 and θ_i^2*

The spectral constraint equations at $O(\varepsilon^{5/2})$ are

$$-2\nabla^2\boldsymbol{v}_i^2 + 2Ra_c\theta_i^2\boldsymbol{\nabla}\tau_0 + \boldsymbol{\nabla}p_i = 2a_1\nabla^2\boldsymbol{v}_i^1 + 2a_2\nabla^2\boldsymbol{v}_i^0 - 2Ra_c(\theta_i^0\boldsymbol{\nabla}\tau_2 + \theta_i^1\boldsymbol{\nabla}\tau_1) - 2(\theta_i^0\boldsymbol{\nabla}\tau_1 + \theta_i^1\boldsymbol{\nabla}\tau_0) - \frac{1}{\sigma}(\boldsymbol{\nabla}\boldsymbol{\phi}_2 + \boldsymbol{\nabla}\boldsymbol{\phi}_2^T)\cdot\boldsymbol{v}_i^0 \quad (6.45)$$

$$-2\nabla^2\theta_i^2 + 2\boldsymbol{v}_i^2\cdot\boldsymbol{\nabla}\tau_0 = -2\boldsymbol{v}_i^0\cdot\boldsymbol{\nabla}\tau_2 - 2\boldsymbol{v}_i^1\cdot\boldsymbol{\nabla}\tau_1. \quad (6.46)$$

The operator on the left hand side of (6.45)−(6.46) is self adjoint and annihilates the leading order fields $(\boldsymbol{v}_j^0, \theta_j^0)$ $j = 1, 2$. As a result, there are solvability conditions for $(\boldsymbol{v}_0^2, \theta_0^2)$ of the form

$$\langle \boldsymbol{v}_j^0\cdot[2a_1\nabla^2\boldsymbol{v}_i^1 + 2a_2\nabla^2\boldsymbol{v}_i^0 - 2Ra_c(\theta_i^0\boldsymbol{\nabla}\tau_2 + \theta_i^1\boldsymbol{\nabla}\tau_1) - 2(\theta_i^0\boldsymbol{\nabla}\tau_1 + \theta_i^1\boldsymbol{\nabla}\tau_0) - \frac{1}{\sigma}(\boldsymbol{\nabla}\boldsymbol{\phi}_2 + \boldsymbol{\nabla}\boldsymbol{\phi}_2^T)\cdot\boldsymbol{v}_i^0] - 2\theta_j^0(\boldsymbol{v}_i^0\cdot\boldsymbol{\nabla}\tau_2 + \boldsymbol{v}_i^1\cdot\boldsymbol{\nabla}\tau_1)\rangle = 0. \quad (6.47)$$

Taking $i = j$ (the $i \neq j$ conditions vanish trivially), this can be rearranged to

$$\mathrm{Term}_1(i) + (A_1^2 - A_2^2)\mathrm{Term}_2(i) = 0 \qquad i = 1, 2 \quad (6.48)$$

where

$$(A_1^2 - A_2^2)\mathrm{Term}_2(i) := -\frac{1}{A_i^2}\left\langle \frac{2}{\sigma}\boldsymbol{v}_i^0\cdot\boldsymbol{\nabla}\boldsymbol{\phi}_2\cdot\boldsymbol{v}_i^0 + 2(Ra_c+1)\theta_i^0\boldsymbol{v}_i^0\cdot\boldsymbol{\nabla}\tau_2^{2D}\right\rangle \quad (6.49)$$

Crucially $\mathrm{Term}_1(1)$=$\mathrm{Term}_1(2)$ whereas $\mathrm{Term}_2(1)$=−$\mathrm{Term}_2(2)$ so that (6.48) implies that

$$\mathrm{Term}_1(1) = (A_1^2 - A_2^2)\mathrm{Term}_2(1) = 0. \quad (6.50)$$

The unspecified coefficient a_2 only enters Term_1 and so is set by the condition this vanishes. In contrast, there are no free constants in Term_2 which is non-zero in our computations (although we have been unable to prove this is always the case). In this situation, $A_1^2 - A_2^2 = 0$ is forced instead which eliminates at a stroke all 2-dimensional fields in the bifurcation analysis. Consequently, the background flow field remains zero and the background temperature field stays 1D after the first bifurcation.

6.2. *Subsequent bifurcations*

Now we consider subsequent bifurcations to establish that if $\tau = \tau(z), \boldsymbol{\phi} = 0$ exists before then that situation persists after the bifurcation. The approach is inductive: assume $\tau = \tau(z), \boldsymbol{\phi} = 0$ after m bifurcations and consider the $(m+1)^{th}$ bifurcation at $Ra = Ra_c^{(m+1)}$ where two new neutral modes appear so that there are now $2(m+1)$ critical modes in

the spectral constraint. Defining $\varepsilon := Ra - Ra_c^{(m+1)}$ we expand:

$$(\tau, \boldsymbol{\phi}) = (\ \tau_0(z) + \varepsilon\tau_1(x,z) + \dots,\ \varepsilon\boldsymbol{\phi}_0(x,z) + \dots\ ) \tag{6.51}$$

$$(\boldsymbol{v}_0, \theta_0) = (\ \varepsilon\boldsymbol{v}_0^0 + \dots,\ \varepsilon\theta_0^0 + \dots\ ) \tag{6.52}$$

$$a = a_0 + \varepsilon a_1 + \dots \tag{6.53}$$

$$(\boldsymbol{v}_i, \theta_i) = \begin{cases} (\boldsymbol{v}_i^0 + \varepsilon\boldsymbol{v}_i^1 + \dots, & \theta_i^0 + \varepsilon\theta_i^1 + \dots) & i = 1, 2, \dots, 2m \\ (\varepsilon^{1/2}\boldsymbol{v}_i^0 + \varepsilon^{3/2}\boldsymbol{v}_i^1 + \dots, & \varepsilon^{1/2}\theta_i^0 + \varepsilon^{3/2}\theta_i^1 + \dots) & i = 2m+1, 2m+2. \end{cases} \tag{6.54}$$

where the leading fields $\tau_0(z)$, $(\boldsymbol{v}_i^0, \theta_i^0)$ $(i = 1, \dots 2m+2)$ and a_0 are all known. In particular, the i^{th} wavenumber k_i $(i = 1, 2, \dots, m)$, is associated with two modes which, to leading order, are

$$(\boldsymbol{v}_{2i-1}^0, \theta_{2i-1}^0) = A_i(\ U_i(z)\sin(k_i x)\boldsymbol{e}_x + W_i(z)\cos(k_i x)\boldsymbol{e}_z, \Theta_i(z)\cos(k_i x)\ ) \tag{6.55}$$

and

$$(\boldsymbol{v}_{2i}^0, \theta_{2i}^0) = B_i(\ U_i(z)\cos(k_i x)\boldsymbol{e}_x - W_i(z)\sin(k_i x)\boldsymbol{e}_z, -\Theta_i(z)\sin(k_i x)\ ) \tag{6.56}$$

where $A_i^2 = B_i^2$ for $i = 1, 2, ..., m$. The two new modes emerging at the $(m+1)^{th}$ bifurcation point can be assumed to have the following general 3D form:

$$\boldsymbol{v}_{2m+1}^0 = A_{m+1}\begin{pmatrix} U_{m+1}(z)\sin(\boldsymbol{k}_{m+1}\cdot\boldsymbol{x}) \\ V_{m+1}(z)\sin(\boldsymbol{k}_{m+1}\cdot\boldsymbol{x}) \\ W_{m+1}(z)\cos(\boldsymbol{k}_{m+1}\cdot\boldsymbol{x}) \end{pmatrix} \tag{6.57}$$

$$\theta_{2m+1}^0 = A_{m+1}\Theta_{m+1}(z)\cos(\boldsymbol{k}_{m+1}\cdot\boldsymbol{x}), \tag{6.58}$$

and

$$\boldsymbol{v}_{2m+2}^0 = B_{m+1}\begin{pmatrix} U_{m+1}(z)\cos(\boldsymbol{k}_{m+1}\cdot\boldsymbol{x}) \\ V_{m+1}(z)\cos(\boldsymbol{k}_{m+1}\cdot\boldsymbol{x}) \\ -W_{m+1}(z)\sin(\boldsymbol{k}_{m+1}\cdot\boldsymbol{x}) \end{pmatrix} \tag{6.59}$$

$$\theta_{2m+2}^0 = -B_{m+1}\Theta_{m+1}(z)\sin(\boldsymbol{k}_{m+1}\cdot\boldsymbol{x}), \tag{6.60}$$

where $\boldsymbol{k}_{m+1} = (k_x, k_y, 0)$. The objective in what follows is to show that $A_{m+1}^2 = B_{m+1}^2$ after the bifurcation so that the optimization problem remains 1D.

At leading order $(O(\epsilon))$ in the forced field and background field equations

$$-2a_0\nabla^2\boldsymbol{v}_0^0 + 2Ra_c^{m+1}\theta_0^0\boldsymbol{\nabla}\tau_0 + \boldsymbol{\nabla}p - a_0\nabla^2\boldsymbol{\phi}_0 = \boldsymbol{0}, \tag{6.61}$$

$$-2\nabla^2\theta_0^0 + 2\boldsymbol{v}_0^0\cdot\boldsymbol{\nabla}\tau_0 + 2\boldsymbol{\phi}_0\cdot\boldsymbol{\nabla}\tau_0 + a_0\phi_{0z} = 0 \tag{6.62}$$

where $\phi_{0z} := \boldsymbol{\phi_0}\cdot\hat{\mathbf{z}}$,

$$\begin{aligned} &a_0\nabla^2\boldsymbol{v}_0^0 - Ra_c^{m+1}(a_0\tau_1 + a_1\tau_0 + a_0\theta_0^0)\boldsymbol{e}_z - a_0\tau_0\boldsymbol{e}_z - 2Ra_c^{m+1}\theta_0^0\boldsymbol{\nabla}\tau_0 + \boldsymbol{\nabla}q \\ &+ \frac{a_0}{\sigma}\sum_{i=1}^{2m}\Big(\boldsymbol{v}_i^0\cdot\boldsymbol{\nabla}\boldsymbol{v}_i^1 + \boldsymbol{v}_i^1\cdot\boldsymbol{\nabla}\boldsymbol{v}_i^0\Big) = \underbrace{-\frac{a_1}{\sigma}\sum_{i=1}^{2m}\boldsymbol{v}_i^0\cdot\boldsymbol{\nabla}\boldsymbol{v}_i^0}_{1D} \underbrace{-\frac{a_0}{\sigma}\sum_{i=2m+1}^{2m+2}\boldsymbol{v}_i^0\cdot\boldsymbol{\nabla}\boldsymbol{v}_i^0}_{driving\ term}, \end{aligned} \tag{6.63}$$

$$-2\nabla^2\tau_1 - a_0\phi_{0z} + 2\sum_{i=1}^{2m}\Big(\boldsymbol{v}_i^0\cdot\boldsymbol{\nabla}\theta_i^1 + \boldsymbol{v}_i^1\cdot\boldsymbol{\nabla}\theta_i^0\Big) = \underbrace{-2\sum_{i=2m+1}^{2m+2}\boldsymbol{v}_i^0\cdot\boldsymbol{\nabla}\theta_i^0}_{driving\ term}. \tag{6.64}$$

and again it is implicit that $\boldsymbol{v}_0$, $\boldsymbol{v}_i$ and $\boldsymbol{\phi}$ are incompressible fields. The spectral constraint for modes $i = 1, 2, ..., 2m$ at $O(\varepsilon)$ and for modes $i = 2m+1, 2m+2$ at $O(\varepsilon^{3/2})$ is

$$-2a_0\nabla^2\boldsymbol{v}_i^1 + 2Ra_c^{m+1}\theta_i^1\boldsymbol{\nabla}\tau_0 + \boldsymbol{\nabla}p = 2a_1\nabla^2\boldsymbol{v}_i^0 - \frac{a_0}{\sigma}\boldsymbol{v}_i^0\cdot(\boldsymbol{\nabla}\boldsymbol{\phi}_0 + \boldsymbol{\nabla}\boldsymbol{\phi}_0^{\mathtt{T}}) - 2Ra_c^{m+1}\theta_i^0\boldsymbol{\nabla}\tau_1 - 2\theta_i^0\boldsymbol{\nabla}\tau_0, \quad (6.65)$$

$$-2\nabla^2\theta_i^1 + 2\boldsymbol{v}_i^1\cdot\boldsymbol{\nabla}\tau_0 = -2\boldsymbol{v}_i^0\cdot\boldsymbol{\nabla}\tau_1. \quad (6.66)$$

The system of equations (6.61)-(6.66) is linear in $\boldsymbol{v}_0^0$, θ_0^0, $\boldsymbol{\phi}_0$, $\boldsymbol{v}_i^1$ and θ_i^1 ($i = 1, 2, ..., 2m+2$). The emergent critical modes at Ra_c^{m+1} give rise to the new driving terms in (6.63) & (6.64)

$$\sum_{i=2m+1}^{2m+2}\boldsymbol{v}_i^0\cdot\boldsymbol{\nabla}\boldsymbol{v}_i^0 = \tfrac{1}{2}(A_{m+1}^2 + B_{m+1}^2)\boldsymbol{\nabla}[W_{m+1}^2(z)] + \tfrac{1}{2}(A_{m+1}^2 - B_{m+1}^2)\begin{bmatrix} -U_{m+1}\frac{dW_{m+1}}{dz} + \frac{dU_{m+1}}{dz}W_{m+1} \\ -V_{m+1}\frac{dW_{m+1}}{dz} + \frac{dV_{m+1}}{dz}W_{m+1} \\ 0 \end{bmatrix}\sin(2\boldsymbol{k}_{m+1}\cdot\boldsymbol{x}) \quad (6.67)$$

$$\sum_{i=2m+1}^{2m+2}\boldsymbol{v}_i^0\cdot\boldsymbol{\nabla}\theta_i^0 = \tfrac{1}{2}(A_{m+1}^2 + B_{m+1}^2)\frac{d(W_{m+1}\Theta_{m+1})}{dz} + \tfrac{1}{2}(A_{m+1}^2 - B_{m+1}^2)\left(W_{m+1}\frac{d\Theta_{m+1}}{dz} - \frac{dW_{m+1}}{dz}\Theta_{m+1}\right)\cos(2\boldsymbol{k}_{m+1}\cdot\boldsymbol{x}) \quad (6.68)$$

which have a 1D part proportional to $A_{m+1}^2 + B_{m+1}^2$ and a non-1D part proportional to $A_{m+1}^2 - B_{m+1}^2$. *If* $A_{m+1}^2 = B_{m+1}^2$ then $\tau_1 = \tau_1(z)$ and $\boldsymbol{\phi} = \boldsymbol{0}$ using the arguments of section 6.1. Hence for $A_{m+1}^2 \neq B_{m+1}^2$ and using equation (6.64), we can assume a solution structure of the form

$$\tau_1 = \underbrace{\sum_{i=1}^{m+1}(A_i^2 + B_i^2)P_i(z)}_{\tau_1^{1D}} + \underbrace{(A_{m+1}^2 - B_{m+1}^2)\Big(Q(z)\cos(2\boldsymbol{k}_{m+1}\cdot\boldsymbol{x}) + \tau_1^*(x,z)\Big)}_{\tau_1^{2D}}, \quad (6.69)$$

$$\boldsymbol{\phi}_0 = (A_{m+1}^2 - B_{m+1}^2)\left(\begin{bmatrix} G_1(z)\sin(2\boldsymbol{k}_{m+1}\cdot\boldsymbol{x}) \\ G_2(z)\sin(2\boldsymbol{k}_{m+1}\cdot\boldsymbol{x}) \\ G_3(z)\cos(2\boldsymbol{k}_{m+1}\cdot\boldsymbol{x}) \end{bmatrix} + \boldsymbol{\Phi}^*(x,z)\right) \quad (6.70)$$

where $\tau_1^*(x,z)$ and $\boldsymbol{\Phi}^*(x,z)$ collect all the other wavenumber dependence on x in τ_1 and $\boldsymbol{\phi}_0$ respectively (this is unimportant in what follows). The key is then examining the solvability conditions

$$\left\langle \boldsymbol{v}_j^0\cdot\left(2a_1\nabla^2\boldsymbol{v}_i^0 - \frac{a_0}{\sigma}\boldsymbol{v}_i^0\cdot(\boldsymbol{\nabla}\boldsymbol{\phi}_0 + \boldsymbol{\nabla}\boldsymbol{\phi}_0^{\mathtt{T}}) - 2Ra_c^{m+1}\theta_i^0\boldsymbol{\nabla}\tau_1 - 2\theta_i^0\boldsymbol{\nabla}\tau_0\right) - 2\theta_i^0\boldsymbol{v}_j^0\cdot\boldsymbol{\nabla}\tau_1\right\rangle = 0 \quad (6.71)$$

$(i, j, = 1, \ldots m+1)$ on the corrections $\boldsymbol{v}_i^1$ to all the critical modes $\boldsymbol{v}_i^0$ of the spectral constraint. These set the amplitudes $A_i^2\,(= B_i^2)$ $(i = 1, \ldots m)$ and (A_{m+1}, B_{m+1}) (a_1 is determined by the balance parameter equation at $O(\varepsilon)$).

To establish that the background fields stay 1D, it is sufficient to focus on the solvability conditions for the new critical modes ($i = 2m+1$ and $2m+2$). Here, the solvability condition is explicitly

$$-a_1\int_0^1 \boldsymbol{k}_{m+1}^2(U_{m+1}^2+V_{m+1}^2+W_{m+1}^2)+\left(\frac{dU_{m+1}}{dz}\right)^2+\left(\frac{dV_{m+1}}{dz}\right)^2+\left(\frac{dW_{m+1}}{dz}\right)^2 dz$$
$$-\int_0^1 W_{m+1}\Theta_{m+1}\frac{d\tau_0}{dz}dz+\sum_{j=1}^{m+1}\left[(A_j^2+B_j^2)(Ra_c^{m+1}+1)\int_0^1 W_j\Theta_j\frac{dP_j}{dz}dz\right]$$
$$+(A_{m+1}^2-B_{m+1}^2)\mathrm{Term}(i)=0 \quad (6.72)$$

where

$$(A_{m+1}^2-B_{m+1}^2)\mathrm{Term}(i):=-\frac{2}{c}\left\langle\frac{a_0}{\sigma}\boldsymbol{v}_i^0\boldsymbol{\nabla}\boldsymbol{\phi}_0\cdot\boldsymbol{v}_i^0+(Ra_c^{m+1}+1)\theta_i^0\boldsymbol{v}_i^0\cdot\boldsymbol{\nabla}\tau_1^{2D}\right\rangle, \quad (6.73)$$

with $c:=A_{m+1}^2$ for $i=2m+1$ or B_{m+1}^2 for $i=2m+2$. Crucially, it is straightforward to show that

$$\mathrm{Term}(2m+1)=-\mathrm{Term}(2m+2) \quad (6.74)$$

so that, as in subsection (6.1), we must have

$$(A_{m+1}^2-B_{m+1}^2)\mathrm{Term}(i)=0 \quad (6.75)$$

leaving equation (6.72) to determine the value of $A_{m+1}^2+B_{m+1}^2$. Our computations indicate $\mathrm{Term}(2m+1)\neq 0$ (although, as in subsection (6.1), we have been unable to prove this in general) implying that $A_{m+1}^2=B_{m+1}^2$. This forces $\tau_1=\tau_1(z)$ and $\boldsymbol{\phi}_0=\boldsymbol{0}$ so that the optimal solution stays 1D after the $(m+1)^{th}$ $(m\geqslant 0)$ bifurcation if it is of this form before.

Taken together with the first bifurcation analysis in subsection (6.1), the result of this section means that the optimal solution is 1D for all Ra and so, surprisingly, there is *no* benefit of imposing the full steady momentum and heat balances in the upper bound problem.

7. Adding a background velocity field to Wen *et al.* (2015)

Following the success of adding an enstrophy constraint in 2D stress-free convection Wen *et al.* (2015), an interesting question is whether adding a 1-D background velocity field by using the decomposition

$$\boldsymbol{u}(x,z)=\phi(z)\boldsymbol{e}_x+\boldsymbol{v}(x,z), \quad T(x,z)=\tau(z)+\theta(x,z) \quad (7.1)$$

would improve the bound further since this imposes additional information from the Navier-Stokes equations (this question is not covered by the conclusion in section 6 since the extra enstrophy constraint was not included there). To maximize the heat flux, the Lagrangian

$$Nu=\langle|\boldsymbol{\nabla}T|^2\rangle-\frac{a}{\sigma Ra}\langle\boldsymbol{v}\cdot\boldsymbol{\mathcal{N}}\rangle-\frac{b}{\sigma Ra}\langle\boldsymbol{\omega}\cdot\boldsymbol{\nabla}\times\boldsymbol{\mathcal{N}})\rangle-2\langle\theta\,\mathcal{H}\rangle \quad (7.2)$$

is constructed where $\boldsymbol{\omega}=\omega(x,z)\boldsymbol{e}_y:=(\boldsymbol{\nabla}\times\boldsymbol{v})$. After some integration by parts, judicious use of boundary conditions, and building in the fact that $Nu=1+\langle wT\rangle$, this leads to

the expression

$$(1-a)Nu + a = \langle|\tau'|^2\rangle - \left\langle |\nabla\theta|^2 + 2\theta v_z\tau' + \frac{a}{\sigma Ra}v_x v_z\phi' + \frac{b}{\sigma Ra}v_z\omega\phi'' - b\omega\frac{\partial\theta}{\partial x} + \frac{a}{Ra}\left(|\nabla\boldsymbol{v}|^2 - v_x\phi''\right) + \frac{b}{Ra}\left(|\nabla\omega|^2 - \omega\phi'''\right)\right\rangle. \tag{7.3}$$

where $\boldsymbol{v} = v_x\hat{\mathbf{x}} + v_z\hat{\mathbf{z}}$. The two linear terms in the second line of this expression - $v_x\phi''$ and $\omega\phi'''$ - mean that optimization over the fluctuation fields $\boldsymbol{v}$ and θ will give rise to a non-zero contribution to be added to $\langle|\tau'|^2\rangle$. This complication can be avoided (or rather made more explicit) by defining a shifted variable

$$\hat{\boldsymbol{v}} := \boldsymbol{v} + \tfrac{1}{2}\phi(z)\boldsymbol{e}_x \tag{7.4}$$

which is possible if $\boldsymbol{v}$ and ϕ are both assumed to satisfy (natural) homogeneous boundary conditions and allows both linear terms to be absorbed into perfect squares. As a result of this, the expression becomes

$$Nu = \frac{1}{1-a}(\langle|\tau'|^2\rangle - a) + \frac{a}{4(1-a)Ra}\langle|\phi'|^2\rangle + \frac{b}{4(1-a)Ra}\langle|\phi''|^2\rangle - \frac{1}{1-a}\mathscr{G}(\hat{\boldsymbol{v}},\omega,\theta;\tau,\phi,a,b,Ra,\sigma) \tag{7.5}$$

where

$$\mathscr{G} := a\langle\frac{1}{Ra}|\nabla\hat{\boldsymbol{v}}|^2 + \frac{1}{\sigma Ra}\hat{v}_x\hat{v}_z\phi'\rangle + b\langle\frac{1}{\sigma Ra}\omega\hat{v}_z\phi'' + \frac{1}{Ra}|\boldsymbol{\nabla}\hat{\omega}|^2 - \hat{\omega}\frac{\partial\theta}{\partial x}\rangle + \langle|\boldsymbol{\nabla}\theta|^2 + 2\theta\hat{v}_z\tau'\rangle$$

is a purely quadratic functional of $\hat{\boldsymbol{v}}$, $\hat{\omega} := \boldsymbol{e}_y\cdot\boldsymbol{\nabla}\times\hat{\boldsymbol{v}} = \omega + \frac{1}{2}\phi'$ and θ. Provided the background fields τ and ϕ are chosen such that $\mathscr{G} \geqslant 0$ for all permissable $\hat{\boldsymbol{v}}$, $\hat{\omega}$ and θ, then a bound follows on Nu. Now it is clear that: 1) the objective functional is strictly convex in the background fields and 2) the set of allowable background fields is convex (if (τ_1,ϕ_1) and (τ_2,ϕ_2) ensure $\mathscr{G} \geqslant 0$ so does $(\tau,\phi) = \mu(\tau_1,\phi_1) + (1-\mu)(\tau_2,\phi_2)$ with $0 \leqslant \mu \leqslant 1$). This implies that the optimizer is unique and is attained for $(\tau,\phi) = (\tau,0)$ i.e. the background velocity field vanishes indicating that the extra information this folds into the optimization is, in fact, unimportant. Physically, a bifurcation analysis shows that the fluctuation fields are always such as to produce zero Reynolds stress so that no background flow field is generated.

A numerical solution shown in figure 11 using Newton's method on the Euler-Lagrange equations in an infinitely long domain confirms that $\phi = 0$ as does a bifurcation analysis developed in the same way as the previous section (not shown). The bound compares well with the earlier results of Wen *et al.* (2015) who considered a fixed domain of $L = 2\sqrt{2}$ indicating further that the bound is not that sensitive to the domain size. The fashion in which the necessarily discretized critical modes found by Wen *et al.* (2015) cluster around the (continuous) optimal wave numbers in our study confirms this conclusion.

8. Discussion

This paper has revisited the optimal heat transport problem in two-dimensional Rayleigh-Bénard convection with stress-free boundary conditions using an extended background method. The key novelty has been to consider background temperature and velocity fields whose dimensional dependence matches that of the physical problem (so 2D here). This situation needs a reformulation in the way the variational equations are solved which has the significant consequence of breaking any link between the optimal fields which

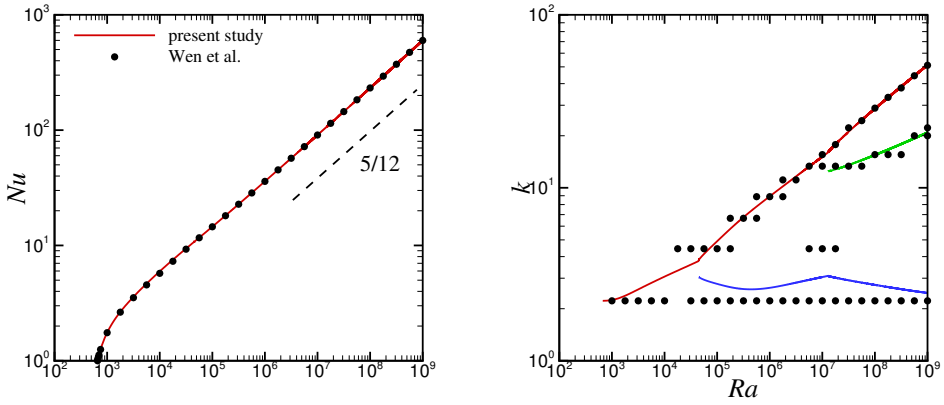

FIGURE 11. Left panel: the upper bound of Nu vs. the Rayleigh number Ra, and $Nu \leqslant 0.107a^{5/12}$ in the asymptotic ultimate regime. right panel: the bifurcation diagram of critical wavenumbers k_m vs. the Rayleigh number (solid found here for $L = \infty$; dots from Wen *et al.* (2015) for fixed $L = 2\sqrt{2}$). The prefactor predicted 0.107 is very slightly higher than Wen *et al.* (2015)'s 0.106 as the domain size L is also optimized over here (data courtesy of Baole Wen).

emerge and a single physical temperature and velocity field. In particular, this means that the optimal fields do not obviously satisfy the steady heat equation even though that is explicitly imposed when the background temperature field is allowed to be fully 2D. This is due to the spectral constraint (that ensures a bound) which means the optimal bound found does not correspond with the highest stationary point of the Lagrangian (i.e. the Euler-Lagrange equations are not all satisfied) but is strictly above it. In other words, there is a gap between the highest heat flux attained by a steady solution of the governing equations imposed and the best (lowest) bound because of the additional spectral constraint. Importantly, this means that there is no direct connection between the optimal solution in the background method built around the steady governing equations and a steady solution of the governing equations (here the Boussinesq equations but clearly more generally true). This realisation removes the possibility, for example, that the simple 2D roll solution computed by Waleffe et al. (2015) could actually be the optimal solution to the background bounding problem. It now seems clear that it would be spectrally unstable.

In revisiting the exact 2D Rayleigh-Benard problem treated by Hassanzadeh et al. (2014), we have shown that their maximal heat flux result is only guaranteed to be a global maximum up to $Ra \leqslant Ra_c := 4468.8$. Beyond this, a gap develops between the bound generated by the background method and the wall-to-wall maximal result. If the spatial domain is extended, the background method optimal becomes increasingly 1D. Removing the symmetry imposed by Hassanzadeh et al. (2014) and reinstating translational invariance in the horizontal direction by making the domain unbounded, the optimal background optimal solution is then provably just 1D and the classic scaling result of $Nu \sim Ra^{1/2}$ is recovered albeit with the larger numerical coefficient of 0.055 as opposed to the already known 0.026 (PK03) for non-slip boundary conditions. The conclusion is then that imposing the steady heat equation in the bounding calculation does *not* improve (reduce) the bound over that obtained using the horizontally-averaged steady heat equation. We then considered adding extra information from the momen-

tum equation to the upper bounding problem by introducing a background velocity field $\boldsymbol{\phi}(x,z)$. Now the optimization problem is no longer convex and we use an inductive bifurcation analysis to show that if $\boldsymbol{\phi}=\mathbf{0}$ before a bifurcation then it remains $\mathbf{0}$ after it too. This means that the continuous branch of optimals found by branch tracking out of the energy stability point always has $\boldsymbol{\phi}=\mathbf{0}$. Noting the caveats that a) it is not impossible that there is an unconnected branch of optimals with $\boldsymbol{\phi}\neq\mathbf{0}$ and b) $Term(2m+1)$ in (6.75) could serendipitously vanish at a subsequent bifurcation beyond our calculations, this strongly suggests the surprising result that imposing the steady Boussinesq equations does *not* improve the bound over that already obtained using the horizontally-averaged steady heat equation and an energy constraint derived from the steady momentum equation.

The 'take-home' message from this study is that the background method of seeking an upper bound on heat flux in Rayleigh-Benard convection has been exhausted (at least using steady fields) with disappointingly *no* improvement possible over the minimal choice of a 1D background temperature field originally made in 1996 by Doering and Constantin. It is hard not to imagine this realisation also generalising to the analogous background formulations for shear flows too, e.g. plane Couette flow (Doering & Constantin, 1992), channel flow (Constantin and Doering, 1995) and pipe flow (Plasting & Kerswell, 2005). Simply extending the definition of the background fields ostensibly folds in more information from the governing equations but not in a fruitful way. However, it seems generating extra information by *differentiating* the governing equations can help. Whitehead & Doering (2011) (see also Wen et al. 2015) used an extra vorticity constraint to significantly lower the bound from $Nu\sim Ra^{1/2}$ to $Nu\sim Ra^{5/12}$ but only in the 2D situation with stress-free boundary conditions. Interestingly, this approach can be inverse-engineered into the form of a background method by loosening the connection between the Lagrange multiplier $\boldsymbol{\nu}(\mathbf{x},t)$ and the velocity field $\mathbf{u}(\mathbf{x},t)$ from $\mathbf{u}(\mathbf{x},t)-\boldsymbol{\nu}(\mathbf{x},t)=\phi(z)\hat{\mathbf{x}}$ to

$$\mathbf{u}(\mathbf{x},t)-\boldsymbol{\nu}(\mathbf{x},t)=\phi(z)\hat{\mathbf{x}}+c\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\mathbf{u}(\mathbf{x},t)$$

where c is a new scalar Lagrange multiplier imposing the global vorticity constraint

$$\langle\boldsymbol{\nabla}\times\mathbf{u}\cdot\boldsymbol{\nabla}\times(\boldsymbol{\mathcal{N}})_s\rangle=0.$$

This clearly extracts something more from the governing equations than just taking projections. Maybe there is some mileage in exploring this but a shortage of boundary conditions is the usual impediment to this approach. It is also worth remarking here that Nobili & Otto (2017) have demonstrated that the background method has a fundamental limitation in the infinite-Prandtl number convection problem in that it can not produce the optimal bound established using other techniques.

Looking ahead, a recent generalisation of the background approach - the 'auxiliary function method' (Chernyshenko et al. 2014, Fantuzzi et al. 2016, Tobasco et al. 2018, Goluskin & Fantuzzi 2019) - appears to offer greater potential for progress since it extends the quadratic constraints used here to more general polynomials albeit at the expense of a fully numerical approach. Here the upper bound problem can be posed as a convex optimisation problem where the optimal value can provably be attained for a system governed by ordinary differential equations at least (Tobasco et al. 2018).

Acknowledgements: The authors are very grateful to Andre Souza and Charlie Doering for helpful discussions and sharing their recent preprint (Souza et al. 2019) as well as the 3 referees who provided insightful reviews. The authors also acknowledge the support

of EPSRC under grant EP/P001130/1. Declaration of Interests. The authors report no conflict of interest.

Appendix A. Time stepping for a 2D background temperature field in 2D

Here we show that the time-marching method of Wen et al. (2015) is not guaranteed to have the optimal solution as the unique steady attractor when $\tau = \tau(x, z)$ has the same *spatial* dimensionality as the physical temperature field $T(x, z, t)$. Time-stepping would work, however, for a 2-dimensional $\tau(x, z)$ in a 3-dimensional problem. To explain this, we revisit the proof of Wen et al. (2015). The time-stepping approach consists of adding time derivatives for θ, $\nabla^2\psi$ and τ to the left hand sides of (2.15)-(2.17) respectively. Small disturbances $(\theta', \boldsymbol{u}', \tau', p')$ on top of a solution to the Euler-Lagrange equations, $(\theta, \boldsymbol{u}, \tau, p)$, then evolve according to the following equations

$$\frac{\partial \theta'}{\partial t} = \nabla^2\theta' - J(\tau', \psi) - J(\tau, \psi'), \tag{A 1}$$

$$\frac{\partial \nabla^2\psi'}{\partial t} = \frac{a}{Ra}\nabla^4\psi' - J(\tau', \theta) - J(\tau, \theta'), \tag{A 2}$$

$$\frac{\partial \tau'}{\partial t} = \nabla^2\tau' - J(\theta', \psi) - J(\theta, \psi') \tag{A 3}$$

at fixed balance parameter a. Then $\langle\, \theta'(A\,1) - \psi'(A\,2) + \tau'(A\,3)\,\rangle$ gives

$$\frac{\partial}{\partial t}\frac{1}{2}\langle \theta'^2 + |\boldsymbol{\nabla}\psi'|^2 + \tau'^2\rangle = -\langle|\boldsymbol{\nabla}\tau'|^2\rangle - \underbrace{\langle \frac{a}{Ra}|\nabla^2\psi'|^2 + |\boldsymbol{\nabla}\theta'|^2 + 2\theta' J(\tau, \psi')\rangle}_{\mathscr{G}}. \tag{A 4}$$

In the 1-dimensional background field case, $\tau = \tau(z)$, (A 3) becomes

$$\frac{\partial \tau'}{\partial t} - \frac{\partial^2\tau'}{\partial z^2} = -\overline{J(\theta', \psi)} - \overline{J(\theta, \psi')} \tag{A 5}$$

(where the overbar represents averaging over x) and the possible fluctuation fields can, after a Fourier transform, be assumed to have a specific wavenumber in x. There are then two types of fluctuation fields: 1) those with wavenumbers which don't overlap with those in the optimal solution ($\lambda < 0$ in the spectral constraint) and therefore do not generate any concomitant disturbance τ', and 2) those which do have a non-vanishing τ' but necessarily have $\lambda = 0$ (the optimal solution is unique for any *given* balance parameter $a \in (0, 1)$ by the same arguments presented in the main text and, by construction, includes any fluctuation fields (θ, ψ) which are neutral in the spectral constraint). In both cases, the fluctuation fields have to decay, in the former case because $\lambda < 0$ and in the latter through the τ' component generated in (A 5). The unique solution is therefore an attractor but the key step is proving that it is the only such. This follows by realising that if a solution to the Euler-Lagrange equations does not satisfy the spectral constraint, then there is an unstable eigenfunction of the linear time-stepping operator defined in (A 1)-(A 3) which consists of the fluctuation field (θ', ψ') which makes $\mathscr{G} < 0$. This is because a fluctuation field with $\lambda \neq 0$ does not overlap under x-averaging with the underlying state and so does not generate a τ' component via (A 5). This argument can clearly be extended to 2-dimensional $\tau(x, z)$ in 3-dimensional Rayleigh-Benard convection since orthogonality in x is replaced by orthogonality of y but breaks down for 2-dimensional Rayleigh-Benard convection. In the latter situation, fluctuation fields which violate the spectral constraint will generate a τ' component via (A 3) and may not then represent

a growing eigenfunction for the time stepping procedure. The implication of this is that some saddles of $\mathscr{L}$ may also be local attractors so if the time-stepping procedure leads to a steady state it is not guaranteed to be the optimal solution. Preliminary numerical tests demonstrated this multistability with the final steady state depending on the initial condition used.

REFERENCES

Ahlers G., Grossmann S. and Lohse D. 2009 Heat transfer and large scale dynamics in turbulent Rayleigh-Benard convection. *Rev. Mod. Phys.* **81**, 503.

Busse F.H. 1969 On Howard's upper bound for heat transport by turbulent convection. *J. Fluid Mech.* **37**, 457.

Busse F.H. 1978 The optimum theory of turbulence *Adv. Appl. Mech.* **18**, 77.

Chernyshenko S.I., Goulart P., Huang D. and Papachristodoulou A. 2014 Polynomial sum of squares in fluid dynamics: a review with a look ahead. *Phil. Trans. Roy. Soc. A* **372**, 20130350.

Constantin P. and Doering, C.R. 1995 Variational bounds on energy dissipation in incompressible flows: II. channel flow. *Phys. Rev. E* **51**, 3192-3198.

Doering C.R. and Constantin P. 1992 Energy dissipation in shear driven turbulence. *Phys. Rev. Lett.* **69**, 1648.

Doering C.R. and Constantin P. 1994 Variational bounds on energy dissipation in incompressible flows: shear flow. *Phys. Rev. E* **49**, 4087.

Doering C.R. and Constantin P. 1996 Variational bounds on energy dissipation in incompressible flows. III. convection. *Phys. Rev. E* **53**, 5957.

Doering C.R. and Constantin P. 1998 Bounds for heat transport in a porous layer *J. Fluid Mech.* **376**, 263-296.

Doering C.R., Spiegel E.A and Worthing R.A. 2000 Energy dissipation in a shear layer with suction. *Phys. Fluids* **12**, 1955-1968.

Doering C.R. and Constantin P. 2001 On upper bounds for infinite Prandtle number convection with and without rotation *J. Math. Phys.* **42**, 784-795.

Doering C.R. and Tobasco 2019 On the optimal design of wall-to-wall heat transport *Commun. Pure. Appl. Math* **LXXII**, 2385-2448.

Fantuzzi, G. 2018 Construction of optimal background fields using semidefinite programming *PhD thesis Imperial College.*

Fantuzzi, G. , Goluskin, D., Huang, D. and Chernyshenko S.I. 2016 Bounds for deterministic and stochastic dynamical systems using sum-of-squares optimization *SIAM J. App. Dyn. Sys.* **15**, 1962-1988.

Goluskin, D. and Fantuzzi, G. 2019 Bounds on mean energy in the Kuramoto-Sivashinsky equation computed using semidefinite programming *Nonlinearity* **32**, 1705-1730.

Grossmann S. and Lohse, D. 2000 Scaling in thermal convection: A unifying theory. *J. Fluid Mech.* **407**, 27.

Hassanzadeh P., Chini G.P. and Doering C.R. 2014 Wall to wall optimal transport. *J. Fluid Mech.* **751**, 627-662.

Howard L.N. 1963 Heat transport by turbulent convection. *J. Fluid Mech.* **17**, 405.

Howard L.N. 1972 Bounds on flow quantities. *Ann. Rev. Fluid Mech.* **4**, 473-494.

Ierley G.R. and Worthing R.A. 2001 Bound to improve: a variational approach to convective heat transport. *J. Fluid Mech.* **441**, 223.

Kerswell R.R. 1996 Upper bounds on the energy dissipation in turbulent precession *J. Fluid Mech.* **321**, 335-370.

Kerswell R.R. 1998 Unification of variational principles for turbulent shear flows: the background method of Doering-Constantin and the mean-fluctuation formulation of Howard-Busse. *Physica D* **121**, 175-192.

Kerswell R.R. 2001 New results in the variational approach to turbulent Boussinesq convection. *Phys. Fluids* **13**, 192.

Malkus W.V.R. 1954 The heat transport and spectrum of thermal turbulence. *Proc. R. Soc. London A* **225**, 196.

Motoki S., Kawahara G. and Shimizu M. 2018 Maximal heat transfer between two parallel plates. *J. Fluid Mech.* **851**, R4.

Nobili, C. and Otto, F. 2017 Limitations of the background method applied to Rayleigh Benard convection *J. Math. Phys.* **58**, 093102.

Plasting S.C. and Kerswell R.R. 2003 Improved upper bound on the energy dissipation rate in plane Couette flow: the full solution to Busse's problem and the Constantin-Doering-Hopf problem with one-dimensional background field. *J. Fluid Mech.* **16**, 363-379 (**referred to as PK03 in the text**).

Plasting S.C. and Kerswell R.R. 2005 A friction factor bound for transitional pipe flow *Phys. Fluids* **17**, 011706.

Priestley C.H.B 1954 Convection from a large horizontal surface *Aust. J. Phys.* **7**, 176-201 .

Sondak D., Smith L. and Waleffe F. 2015 Optimal heat transport solutions for Rayleigh-Bénard convection. *J. Fluid Mech.* **784**, 565.

Souza A. and Doering C. 2015a Maximal transport in the Lorenz equations *Phys. Lett. A* **379**, 518-523.

Souza A. and Doering C. 2015b Transport bounds for a truncated model of Rayleigh-Benard convection *Physica D* **308**, 26-33.

Souza A. 2016 An optimal control approach to bounding transport properties of thermal convection. *Ph.D. thesis*, *University of Michigan*.

Souza A., Tobasco I. and Doering C. 2019 Wall-to-wall optimal transport: theory and 2D computations. *preprint*

Tobasco I. and Doering C.R. 2017 Optimal wall-to-wall transport by incompressible flows. *Phys. Rev. Lett.* **118**, 264502.

Tobasco I. Goluskin, D. and Doering C.R. 2018 Optimal bounds and extremal trajectories for time averages in nonlinear dynamical systems. *Phys. Lett. A* **382**, 382-386.

Waleffe F., Boonkasame A. and Smith L. 2015 Heat transport by coherent Rayleigh-Bénard convection. *Phys. Fluids* **27**, 051702.

Wen B., Chini G.P., Dianati N.and Doering C. 2013 Computational approaches to aspect-ratio-dependent upper bounds and heat flux in porous medium convection *Phys. Lett. A* **A 377**, 2931.

Wen B., Chini G.P., Kerswell R.R. and Doering C. 2015 Time-stepping approach for solving upper-bound problems: application to two-dimensional Rayleigh-Bénard convection. *Phys. Rev. E* **92**, 043012.

Whitehead J. and Doering C.R. 2011 Ultimate state of two-dimensional Rayleigh-Bénard convection between free-slip fixed-temperature boundaries. *Phys. Rev. Lett.* **106**, 244501.

Whitehead J. and Doering C.R. 2012 Rigid bounds on heat transport by a fluid between slippery boundaries. *J. Fluid Mech.* **707**, 241.

Zhu X., Mathai V., Stevens R.J.A.M., Verzicco R. and Lohse D. 2018 Transition to the ultimate regime in two-dimensional Rayleigh-Benard convection. *Phys. Rev. Lett.* **120**, 144502.