

On the Optimal Design of Wall-to-Wall Heat Transport

CHARLES R. DOERING

University of Michigan

IAN TOBASCO

University of Michigan

Abstract

We consider the problem of optimizing heat transport through an incompressible fluid layer. Modeling passive scalar transport by advection-diffusion, we maximize the mean rate of total transport by a divergence-free velocity field. Subject to various boundary conditions and intensity constraints, we prove that the maximal rate of transport scales linearly in the r.m.s. kinetic energy and, up to possible logarithmic corrections, as the one-third power of the mean enstrophy in the advective regime. This makes rigorous a previous prediction on the near optimality of convection rolls for energy-constrained transport. On the other hand, optimal designs for enstrophy-constrained transport are significantly more difficult to describe: we introduce a “branching” flow design with an unbounded number of degrees of freedom and prove it achieves nearly optimal transport. The main technical tool behind these results is a variational principle for evaluating the transport of candidate designs. The principle admits dual formulations for bounding transport from above and below. While the upper bound is closely related to the “background method,” the lower bound reveals a connection between the optimal design problems considered herein and other apparently related model problems from mathematical materials science. These connections serve to motivate designs. © 2019 Wiley Periodicals, Inc.

Contents

| | |
|--|------|
| 1. Introduction | 2386 |
| 2. A Priori Bounds on Wall-to-Wall Optimal Transport | 2396 |
| 3. Optimal Design of Steady Wall-to-Wall Transport | 2405 |
| 4. Energy-Constrained Transport and Convection Roll Designs | 2413 |
| 5. Enstrophy-Constrained Transport and Branched Flow Designs | 2417 |
| 6. Implications for the Analysis of Turbulent Heat Transport | 2431 |
| 7. Optimal Transport as Energy-Driven Pattern Formation | 2441 |
| Bibliography | 2446 |

1 Introduction

1.1 The Wall-to-Wall Optimal Transport Problem

This paper concerns a class of optimal design problems from fluid dynamics that asks to maximize the overall transport of heat through an incompressible fluid layer. Passive scalar transport by an incompressible fluid is governed by the advection-diffusion equation

$$(1.1) \quad \partial_t T + \mathbf{u} \cdot \nabla T = \kappa \Delta T$$

where $T(\mathbf{x}, t)$ is the scalar field undergoing transport, referred to as temperature throughout, $\mathbf{u}(\mathbf{x}, t)$ is the velocity vector field of the fluid, and κ is the coefficient of molecular diffusivity. In general, the velocity field \mathbf{u} and temperature T may depend on both space $\mathbf{x} = (x, y, z)$ and time t . Due to incompressibility, \mathbf{u} must remain divergence-free. Thinking of \mathbf{u} as being in our control, we set ourselves the task of choosing it to maximize the overall transport of heat determined by (1.1).

This is a rich class of optimal design problems and we are interested in the dependence of any solutions, i.e., *optimal designs*, on various constraints that may be imposed. We discuss specific constraints for the velocities later on, but let us handle the temperature field first. Supposing the fluid is contained between two impenetrable parallel planar walls at a distance h , we fix the temperature at the walls by imposing the constant Dirichlet boundary conditions

$$(1.2) \quad T|_{z=0} = T_{\text{hot}} \quad \text{and} \quad T|_{z=h} = T_{\text{cold}}.$$

If the velocity field \mathbf{u} is regular enough—eventually our constraints on it will ensure this—the advection-diffusion equation (1.1) admits a unique solution T satisfying (1.2) for every essentially bounded initial temperature field $T|_{t=0} = T_0(\mathbf{x})$. We see, therefore, that the overall heat transport specified by (1.1) should depend in general on \mathbf{u} and T_0 . However, as the partial differential equation (PDE) (1.1) is dissipative, any dependence on the initial temperature T_0 is eventually lost as $t \rightarrow \infty$, and the resulting heat transport can be thought of as being set by \mathbf{u} alone.

In this paper, we study the optimal design of wall-to-wall heat transport in the long-time limit, subject to various boundary conditions and intensity constraints on the velocity field \mathbf{u} . To simplify matters, we consider all fields to be periodic in the wall-parallel variables x and y with periods l_x and l_y . That is, we take \mathbf{x} to belong to the domain $\Omega = \mathbb{T}_{xy}^2 \times [0, h]_z$, identifying \mathbb{T}_{xy}^2 with $[0, l_x] \times [0, l_y]$ in the usual way. We turn now to discuss the precise measure of overall heat transport that will be optimized throughout.

Finite-Time Wall-to-Wall Optimal Transport

According to the advection-diffusion equation (1.1) and the boundary conditions (1.2), the vertically averaged rate of heat transport per unit area up to time $t = \tau$ is

given by

$$\begin{aligned} J_\tau &= \frac{1}{\tau} \frac{1}{l_x l_y h} \int_0^\tau \int_\Omega \hat{\mathbf{k}} \cdot (\mathbf{u} T - \kappa \nabla T) d\mathbf{x} dt \\ &= \frac{\kappa}{h} (T_{\text{hot}} - T_{\text{cold}}) + \int_0^\tau \int_\Omega w T d\mathbf{x} dt. \end{aligned}$$

Here, $\mathbf{u} = u \hat{\mathbf{i}} + v \hat{\mathbf{j}} + w \hat{\mathbf{k}}$ and \int denotes an average over the integration domain. We are interested in determining those velocity fields that maximize the overall heat transport J_τ . Of course, unless \mathbf{u} is suitably constrained, the optimal transport $\sup J_\tau$ will be infinitely large. It is natural to prescribe the overall magnitude of \mathbf{u} , and to enforce suitable boundary conditions at the walls $\partial\Omega$. The resulting optimal design problems take the form

$$(1.3) \quad \sup_{\substack{\mathbf{u}(\mathbf{x},t) \\ \|\mathbf{u}\|=U \\ +\text{b.c.}}} J_\tau$$

where the parameter U sets the advective intensity of the admissible velocity fields.

We consider two classes of admissible velocity fields, which we refer to as being “energy-” or “enstrophy-constrained.” In the energy-constrained class, we take

$$\|\mathbf{u}\|^2 = \int_0^\tau \int_\Omega |\mathbf{u}|^2 d\mathbf{x} dt$$

in (1.3) so that the constraint $\|\mathbf{u}\| = U$ sets the average kinetic energy available for advection. As for boundary conditions, the no-penetration ones

$$w|_{\partial\Omega} = 0$$

are well-suited to this class. We call the problem that results the finite-time energy-constrained wall-to-wall optimal transport problem. The finite-time enstrophy-constrained¹ problem arises from taking

$$\|\mathbf{u}\|^2 = h^2 \int_0^\tau \int_\Omega |\nabla \mathbf{u}|^2 d\mathbf{x} dt$$

and enforcing the no-slip boundary conditions

$$\mathbf{u}|_{\partial\Omega} = \mathbf{0}$$

in (1.3). The essential results of this paper hold as well for the stress-free boundary conditions

$$w|_{\partial\Omega} = 0 \quad \text{and} \quad \partial_z u|_{\partial\Omega} = \partial_z v|_{\partial\Omega} = 0,$$

although our focus is mostly on the no-slip ones.

¹For various boundary conditions, including the ones considered here, the mean square rate of strain $\|\nabla \mathbf{u}\|_{L^2(\Omega)}^2$ and enstrophy $\|\nabla \times \mathbf{u}\|_{L^2(\Omega)}^2$ are the same.

Infinite-Time Wall-to-Wall Optimal Transport

Having introduced the finite-time energy- and enstrophy-constrained wall-to-wall optimal transport problems, we turn to discussing their infinite-time analogues, which are the focus of this paper.

Let $\langle \cdot \rangle$ denote the (limit superior) space and long-time average

$$\langle f \rangle = \limsup_{\tau \rightarrow \infty} \int_0^\tau \int_{\Omega} f(\mathbf{x}, t) d\mathbf{x} dt.$$

As an integration by parts shows, the space and long-time averaged heat transport determined by (1.1) satisfies

$$\limsup_{\tau \rightarrow \infty} J_\tau = \kappa \langle |\nabla T|^2 \rangle.$$

Note this depends on \mathbf{u} but not on the initial temperature T_0 so long as it is bounded. In direct analogy with the finite-time optimal transport problems, we define the infinite-time energy-constrained wall-to-wall optimal transport problem by

$$(1.4) \quad \sup_{\substack{\mathbf{u}(\mathbf{x}, t) \\ \langle |\mathbf{u}|^2 \rangle = U^2 \\ w|_{\partial\Omega} = 0}} \kappa \langle |\nabla T|^2 \rangle$$

and the infinite-time enstrophy-constrained problem by

$$(1.5) \quad \sup_{\substack{\mathbf{u}(\mathbf{x}, t) \\ \langle |\nabla \mathbf{u}|^2 \rangle = \frac{U^2}{h^2} \\ \mathbf{u}|_{\partial\Omega} = \mathbf{0}}} \kappa \langle |\nabla T|^2 \rangle.$$

It is these infinite-time optimal design problems that we study in the remainder of this paper. As we never return to the finite-time problems, we discontinue the use of the distinguishing phrases from now on.

A word is in order regarding the sense in which we consider (1.4) and (1.5) to be solved. We do not claim that there must exist maximizers for either problem. Although this certainly merits investigation, and is related to questions of Γ -convergence [4] of the finite-time problems to the infinite-time ones, we choose in this paper to focus instead on the maximum *value* of transport, which is always well-defined. To the maximum value is associated maximizing sequences, i.e., near optimizers that we may seek to describe. Even in the steady versions of (1.4) and (1.5)—where all fields are assumed to be independent of time and optimal designs are guaranteed to exist—determining the maximal transport is a nontrivial task.

The energy- and enstrophy-constrained wall-to-wall optimal transport problems (1.4) and (1.5) were introduced in [19] and studied further in [27, 39] by a combination of asymptotic and numerical methods. Similar methods have since been applied to study other related optimal transport problems [1, 25, 28]. A key question left unresolved by these works is whether the local maximizers constructed therein actually achieve heat transport comparable to that of global optimizers.

In this paper, we present a new mathematically rigorous approach to answering this question. Our methods do not rely on the use of Euler-Lagrange equations; as these are nonconcave maximization problems with many local maximizers, critical point conditions do not suffice to identify global optimizers. Rather, our starting point is a new variational formula for evaluating wall-to-wall heat transport, which is useful both for proving a priori upper bounds on optimal transport as well as lower bounds on the transport of candidate designs. For the energy-constrained problem, we prove that the convection roll designs from [19] achieve globally optimal heat transport up to a universal prefactor in the advection-dominated regime. For the enstrophy-constrained problem, we construct a new class of “branching” designs featuring a large and potentially unbounded number of degrees of freedom. A well-chosen branching design achieves optimal transport up to possible logarithmic corrections.

The wall-to-wall optimal transport problem is naturally related to the study of transport in turbulent fluids. One consequence of our results is a proof that any flows arising in Rayleigh’s original two-dimensional model of buoyancy-driven convection between stress-free walls [36] must achieve significantly suboptimal rates of heat transport in the large Rayleigh number regime $\text{Ra} \gg 1$. Indeed, while our results imply the existence of incompressible flows achieving transport consistent with the proposed “ultimate scaling” law $\text{Nu} \sim \text{Ra}^{1/2}$ (up to logarithmic corrections), such transport is impossible in Rayleigh’s original model [44]. In fact, our analysis leads us to wonder whether such logarithmic corrections to scaling should always hold, independent of dimension or boundary conditions. Behind these claims is a more or less explicit connection between the fluid dynamical optimal design problems considered herein and other apparently related model problems from the study of “energy-driven pattern formation” in materials science [23]. We discuss these considerations in detail at the end. A preliminary version of our methods and results was announced in [40].

1.2 Main Results and Methods

Nondimensionalization

We are concerned with the dependence of energy- and enstrophy-constrained wall-to-wall optimal transport (1.4) and (1.5) in their parameters. We make use of two standard nondimensional quantities. The Peclét number

$$\text{Pe} = \frac{Uh}{\kappa}$$

is a dimensionless measure of the intensity of advection relative to that of diffusion. Transport by (1.1) is dominated by advection when $\text{Pe} \gg 1$ and by diffusion when $\text{Pe} \ll 1$. The Nusselt number Nu is a dimensionless measure of the enhancement of heat transport by convection over that of pure conduction. In the fluid layer

geometry,

$$\text{Nu}(\mathbf{u}) = \frac{\kappa \langle |\nabla T|^2 \rangle}{\frac{\kappa}{h^2} (T_{\text{hot}} - T_{\text{cold}})^2} = 1 + \frac{h}{\kappa} \frac{1}{T_{\text{hot}} - T_{\text{cold}}} \langle wT \rangle.$$

Note this does not depend on the initial temperature T_0 .

Such nondimensionalization reduces the number of free parameters in (1.4) and (1.5) to three: the dimensionless group Pe and the aspect ratios of the domain $\frac{l_x}{h}$ and $\frac{l_y}{h}$. That is, it suffices to take

$$h = \kappa = T_{\text{hot}} = 1 \quad \text{and} \quad T_{\text{cold}} = 0$$

and study the dependence of the resulting nondimensionalized optimal transport problems

$$\sup_{\substack{\mathbf{u}(\mathbf{x},t) \\ \langle |\mathbf{u}|^2 \rangle = \text{Pe}^2 \\ w|_{\partial\Omega} = 0}} \text{Nu}(\mathbf{u}) \quad \text{and} \quad \sup_{\substack{\mathbf{u}(\mathbf{x},t) \\ \langle |\nabla \mathbf{u}|^2 \rangle = \text{Pe}^2 \\ \mathbf{u}|_{\partial\Omega} = \mathbf{0}}} \text{Nu}(\mathbf{u})$$

on Pe , l_x , and l_y . Henceforth, we understand the Nusselt number to be given by

$$(1.6) \quad \text{Nu}(\mathbf{u}) = \langle |\nabla T|^2 \rangle = 1 + \langle wT \rangle$$

where T is determined from \mathbf{u} by solving the advection-diffusion equation

$$(1.7) \quad \partial_t T + \mathbf{u} \cdot \nabla T = \Delta T$$

with Dirichlet boundary conditions

$$T|_{z=0} = 1 \quad \text{and} \quad T|_{z=1} = 0$$

and any essentially bounded initial data $T|_{t=0} = T_0$ (the choice of which is immaterial to our results). The domain $\Omega = \mathbb{T}_{xy}^2 \times I_z$ where \mathbb{T}_{xy}^2 is identified with $[0, l_x] \times [0, l_y]$ and $I_z = [0, 1]$. As always, \mathbf{u} is understood to be divergence-free.

Summary of Main Results

Our results concern the asymptotic dependence of optimal transport in the advective regime $\text{Pe} \gg 1$. Concerning energy-constrained transport, we find that the maximal transport rate scales linearly in the r.m.s. kinetic energy as $\text{Pe} \rightarrow \infty$. More precisely, we prove the following result:

THEOREM 1.1. *There exist positive constants C and C' so that*

$$C \text{Pe} \leq \sup_{\substack{\mathbf{u}(\mathbf{x},t) \\ \langle |\mathbf{u}|^2 \rangle = \text{Pe}^2 \\ w|_{\partial\Omega} = 0}} \text{Nu}(\mathbf{u}) \leq \frac{1}{2} \text{Pe}$$

for all $\text{Pe} \geq C'$. The constant C is independent of all parameters and C' depends only on the aspect ratios of the domain.

As noted in [19], the a priori upper bound $\text{Nu} \lesssim \text{Pe}$ can be proved by a quick application of the maximum principle and the Cauchy-Schwarz inequality. On the other hand, to prove the lower bound one must construct a certain family of admissible velocity fields $\{\mathbf{u}_{\text{Pe}}\}$ and prove that their Nusselt numbers scale linearly in Pe in the advective regime. Such a construction was described using methods of matched asymptotic analysis in [19] (albeit with no attempt to control the errors in the ensuing estimates). Our construction is inspired by that one: we consider a convection roll system as in Figure 1.1a and choose the number of rolls to scale optimally in Pe . Our approach to evaluating Nu allows us to rigorously justify the predictions from [19] regarding the (near) optimality of such flows.

The enstrophy-constrained problem turns out to be much more difficult to resolve. We prove that the maximal enstrophy-constrained transport rate scales, up to possible logarithmic corrections, as the two-thirds power of the r.m.s. rate-of-strain as $\text{Pe} \rightarrow \infty$. Furthermore, we obtain a bound on the size of any corrections to this scaling:

THEOREM 1.2. *There exist positive constants C , C' , and C'' so that*

$$C \frac{\text{Pe}^{2/3}}{\log^{4/3} \text{Pe}} \leq \sup_{\substack{\mathbf{u}(\mathbf{x},t) \\ \langle |\nabla \mathbf{u}|^2 \rangle = \text{Pe}^2 \\ \mathbf{u}|_{\partial\Omega} = 0}} \text{Nu}(\mathbf{u}) \leq C' \text{Pe}^{2/3}$$

for all $\text{Pe} \geq C''$. The constants C and C' are independent of all parameters and C'' depends only on the aspect ratios of the domain.

Remark 1.3. The same bounds apply to enstrophy-constrained optimal transport between no-penetration or stress-free walls. Indeed, by a simple inclusion argument, maximal transport between impenetrable walls is never less than for stress-free walls, and both are bounded below by maximal transport between no-slip walls. Since the a priori upper bound $\text{Nu} \lesssim \text{Pe}^{2/3}$ applies so long as $w|_{\partial\Omega} = 0$ (this is what is proved in Section 2), the result follows.

This result concerning the two-thirds-scaling law of enstrophy-constrained wall-to-wall optimal transport—modulo logarithms—was first announced in our paper [40]. The present paper provides all the mathematical details of the analysis outlined there, as well as a much more complete discussion of our general approach to the optimal design of heat transport. The bulk of it is devoted to motivating and evaluating the branching flows depicted in Figure 1.1b, which are the key to proving the logarithmically corrected lower bound from Theorem 1.2.

After this work was completed, a computational study of the Euler-Lagrange equations for the enstrophy-constrained wall-to-wall optimal transport problem reported convincing numerical evidence for velocity fields that produce $\text{Nu} \sim \text{Pe}^{2/3}$ in three dimensions [27]. Interestingly, numerical studies of the two-dimensional problem have thus far failed to produce heat transport scaling of this sort [19, 28,

39]. Whereas the velocity fields produced in these two-dimensional studies feature near-wall “recirculation zones,” which serve to enhance heat transport at moderate Pe , they come nowhere near the complexity of our branching flows. The three-dimensional computations, however, do exhibit branching of a fully three-dimensional character. Whether such three-dimensional branching flows can be constructed so as to eliminate the logarithmic gap in Theorem 1.2 as $\text{Pe} \rightarrow \infty$ remains to be seen.

Outline of the Approach

Theorem 1.1 and Theorem 1.2 contain two types of statements: a priori upper bounds on the Nusselt number Nu that hold for all velocity fields, and matching lower bounds on Nu for suitable designs. Methods to establish rigorous upper bounds on convective transport go back at least to Howard in the context of turbulent buoyancy-driven convection [21], and Constantin and one of the authors who developed the “background method” to prove upper bounds on Nu (albeit absent Howard’s hypothesis of statistical stationarity) [11–13].

Although a suitably adapted background method can be applied here [39], we do not proceed in this way. Instead, we present a new method for establishing upper bounds based on the fact that, for steady velocity fields, there exists a variational principle for evaluating heat transport. In the time-dependent case, this leads to new variational bounds on Nu that imply the background method. The bound we obtain is as follows:

$$(1.8) \quad \text{Nu}(\mathbf{u}) \leq \inf_{\eta} \langle |\nabla \eta|^2 + |\nabla \Delta^{-1}[\partial_t \eta + \text{div}(\mathbf{u} \eta)]|^2 \rangle$$

where η must satisfy

$$\eta|_{z=0} = 1 \quad \text{and} \quad \eta|_{z=1} = 0.$$

Here and throughout Δ^{-1} denotes the inverse Laplacian operator with vanishing Dirichlet boundary conditions. The bound (1.8) is sharp for steady flows; in that case (1.8) becomes an equality and η need not depend on time.

In contrast, methods to establish rigorous lower bounds on Nu are far and few between. The right-hand side of (1.8) is a convex minimization. Therefore, on general grounds, there should exist a concave maximization that is its dual. We find that

$$(1.9) \quad \text{Nu}(\mathbf{u}) - 1 \geq \sup_{\xi} \langle 2w\xi - |\nabla \xi|^2 - |\nabla \Delta^{-1}[\partial_t \xi + \text{div}(\mathbf{u} \xi)]|^2 \rangle$$

where ξ must satisfy

$$\xi|_{z=0} = 0 \quad \text{and} \quad \xi|_{z=1} = 0.$$

As with (1.8), the bound (1.9) becomes sharp for \mathbf{u} that do not depend on time.

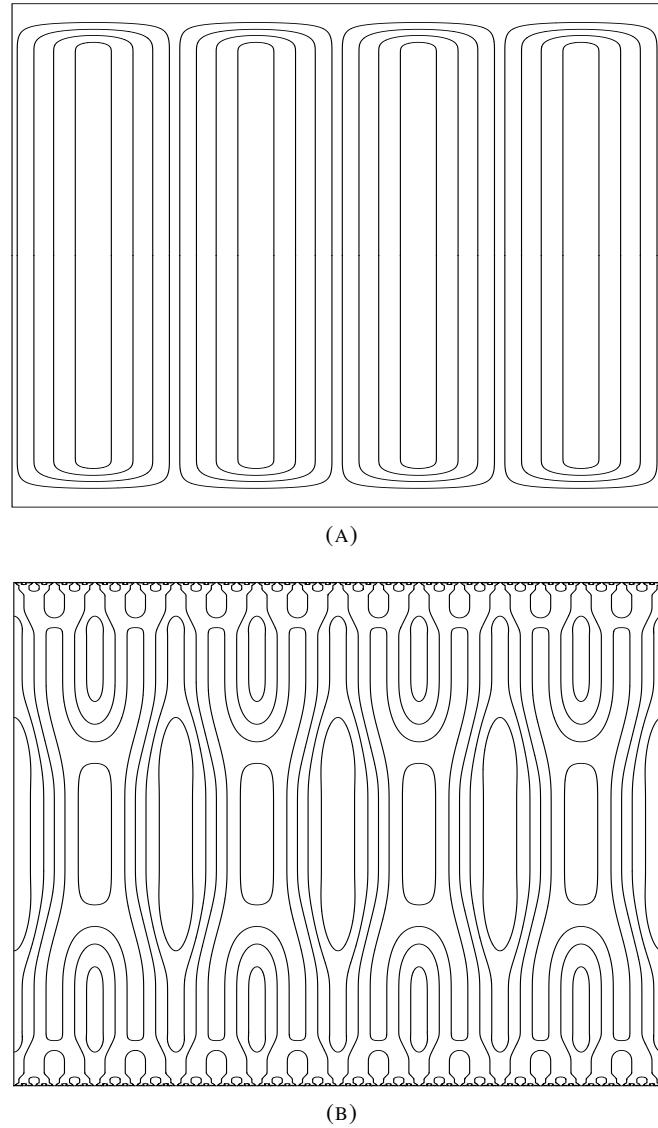


FIGURE 1.1. Streamlines from two families of velocity fields considered in this paper: (A) the convection roll construction and (B) the branching construction. The former involves a single horizontal wavenumber while the latter involves multiple horizontal wavenumbers, the total number of which is allowed to diverge in the advective limit $\text{Pe} \rightarrow \infty$. Such constructions are useful for establishing (nearly) sharp lower bounds on wall-to-wall optimal transport.

Armed with these observations, we describe a new duality-based approach to producing candidate designs. Consider the steady enstrophy-constrained wall-to-wall optimal transport problem

$$(1.10) \quad \max_{\substack{\mathbf{u}(\mathbf{x}) \\ f_{\Omega} |\nabla \mathbf{u}|^2 = \text{Pe}^2 \\ \mathbf{u}|_{\partial\Omega} = 0}} \text{Nu}(\mathbf{u}),$$

whose optimal value bounds the unsteady maximum from below. Appealing to the steady version of (1.9), we find that (1.10) can be rewritten as

$$(1.11) \quad \min_{\substack{\mathbf{u}(\mathbf{x}), \xi(\mathbf{x}) \\ f_{\Omega} w \xi = 1 \\ \mathbf{u}|_{\partial\Omega} = \mathbf{0}, \xi|_{\partial\Omega} = 0}} \int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \xi)|^2 + \varepsilon \int_{\Omega} |\nabla \mathbf{u}|^2 \cdot \int_{\Omega} |\nabla \xi|^2$$

where $\varepsilon = \text{Pe}^{-2}$. Indeed, the optimal values of (1.10) and (1.11) are reciprocals and their optimizers are in correspondence. Thus, solving the steady enstrophy-constrained problem (1.10) for $\text{Pe} \gg 1$ is equivalent to solving (1.11) for $\varepsilon \ll 1$. We refer to (1.11) as an “integral” formulation of wall-to-wall optimal transport.

The family of variational problems (1.11) is nonconvex and singularly perturbed. The situation shares important similarities with other model problems from the field of “energy-driven pattern formation” in materials science [23]. These include the study of branching patterns in micromagnetics [8, 9] and wrinkling cascades in thin elastic sheets [3, 22, 31]. For such problems, it is known that certain patterns which, at a glance, look like Figure 1.1b provide nearly optimal ways of matching low-energy states that are geometrically incompatible but forced to co-exist. We discuss such connections further in Section 7.

Of course, (1.11) does not derive from materials science but instead from fluid dynamics. We note the striking similarities between it and Howard’s variational problem, the latter of which gave birth to the field of variational bounds on turbulent transport [21]. It was recognized by Busse [5] that Howard’s problem should admit multiply scaled optimizers. The resulting construction is known as Busse’s “multi- α ” technique. After suitable modifications (wall-to-wall optimal transport and Howard’s problem are in the end quite distinct) Busse’s techniques can also be used to study (1.11). We consider these connections further in Section 6.

By either analogy, we are led to construct self-similar branching flows as candidates for (1.11). The streamlines depicted in Figure 1.1b are symmetric about $z = \frac{1}{2}$; each half of the domain is made up of n convection roll systems coupled through $n - 1$ transition layers. In the bulk there are large anisotropic convection rolls at some horizontal length scale l_{bulk} . Streamlines refine away from the bulk until there results an isotropic convection roll system at some much smaller length scale l_{bl} . The entire construction can be modeled by a single length scale function $\ell(z)$ that interpolates through the layers. In terms of ℓ , we find the optimal

branching construction to be picked out by the solution of

$$(1.12) \quad \min_{\substack{\ell(z) \\ \ell(z_{\text{bulk}}) = l_{\text{bulk}} \\ \ell(z_{\text{bl}}) = l_{\text{bl}}}} l_{\text{bl}} + \int_{z_{\text{bulk}}}^{z_{\text{bl}}} (\ell')^2 dz + \varepsilon \left(\frac{1}{l_{\text{bulk}}^2} + \int_{z_{\text{bulk}}}^{z_{\text{bl}}} \frac{1}{\ell^2} dz + \frac{1}{l_{\text{bl}}} \right)^2,$$

which satisfies

$$\ell(z) \sim \varepsilon^{1/6} \log^{1/6} \frac{1}{\varepsilon} \sqrt{1-z}, \quad z \in [z_{\text{bulk}}, z_{\text{bl}}],$$

$$l_{\text{bulk}} \sim \varepsilon^{1/6} \log^{1/6} \frac{1}{\varepsilon}, \quad l_{\text{bl}} \sim \varepsilon^{1/3} \log^{1/3} \frac{1}{\varepsilon}.$$

Although this analysis does not prove that optimal designs must exhibit fluctuations according to these rules, it does yield designs sufficient to obtain the asserted lower bounds from Theorem 1.2. The lower bounds from Theorem 1.1 on energy-constrained optimal transport are much simpler to obtain and serve as a test case for our approach.

1.3 Outline of the Paper

Section 2 proves the a priori upper bounds from Theorem 1.1 and Theorem 1.2 and establishes the variational principles and bounds on Nu alluded to above. The proof of the lower bounds from Theorems 1.1 and 1.2 is spread across Sections 3, 4, and 5. In Section 3 we describe our general approach to the optimal design of heat transport. In Section 4 we test our methods on the steady energy-constrained problem and obtain a proof of the lower bound part of Theorem 1.1. In Section 5 we consider the steady enstrophy-constrained problem and prove the lower bound part of Theorem 1.2. We conclude in Sections 6 and 7 with a discussion of bounds on turbulent heat transport, and a discussion of wall-to-wall optimal transport as a problem of energy-driven pattern formation.

1.4 Notation

Having nondimensionalized, we employ the domain $\Omega = \mathbb{T}_{xy}^2 \times I_z$ where \mathbb{T}_{xy}^2 is identified with $[0, l_x] \times [0, l_y]$ and $I_z = [0, 1]$. The spatial average of an integrable function f on Ω is denoted by

$$\mathbf{f}_{\Omega} f = \frac{1}{|\Omega|} \int_{\Omega} f(\mathbf{x}) d\mathbf{x}$$

where $|\Omega| = |\mathbb{T}_{xy}^2| = l_x l_y$. Generally speaking, \mathbf{f} indicates a well-defined average over the indicated domain of integration. Some distinguished averages used in this paper include

$$\bar{f} = \mathbf{f}_{\mathbb{T}_{xy}^2} f = \frac{1}{|\mathbb{T}_{xy}^2|} \int_{\mathbb{T}_{xy}^2} f(x, y, \cdot) dx dy,$$

which averages over the periodic variables x and y ; the (limit superior) space and long-time average

$$\langle f \rangle = \limsup_{\tau \rightarrow \infty} \mathbf{f} \int_0^\tau \int_{\Omega} f = \limsup_{\tau \rightarrow \infty} \frac{1}{\tau} \frac{1}{|\Omega|} \int_0^\tau \int_{\Omega} f(\mathbf{x}, t) d\mathbf{x} dt;$$

and the truncated space and time average

$$\langle f \rangle_\tau = \mathbf{f} \int_0^\tau \int_{\Omega} f = \frac{1}{\tau} \frac{1}{|\Omega|} \int_0^\tau \int_{\Omega} f(\mathbf{x}, t) d\mathbf{x} dt.$$

We use the standard L^2 - and \dot{H}^1 -norms for functions on Ω ,

$$\|f\|_{L^2(\Omega)} = \sqrt{\int_{\Omega} |f|^2} \quad \text{and} \quad \|f\|_{\dot{H}^1(\Omega)} = \sqrt{\int_{\Omega} |\nabla f|^2}.$$

The set of smooth and compactly supported functions on Ω is $C_c^\infty(\Omega)$. The Sobolev space $H_0^1(\Omega)$ is its completion in the norm $\|\cdot\|_{\dot{H}^1(\Omega)}$. We use (\cdot, \cdot) to denote the duality pairing of H^{-1} with H_0^1 . We denote by Δ^{-1} the inverse Laplacian operator with vanishing Dirichlet boundary conditions, which is well-defined from $H^{-1}(\Omega) \rightarrow H_0^1(\Omega)$.

The notation $X \lesssim Y$ means that there exists a positive constant C not depending on any parameters such that $X \leq CY$. We use the notations $X \wedge Y = \min\{X, Y\}$ and $X \vee Y = \max\{X, Y\}$.

2 A Priori Bounds on Wall-to-Wall Optimal Transport

We begin our analysis of wall-to-wall optimal transport by proving the a priori upper bounds from Theorems 1.1 and 1.2. Unless otherwise explicitly stated, we consider throughout that $\langle |\mathbf{u}|^2 \rangle < \infty$ so that (1.7) is well-posed.

The upper bound from Theorem 1.1 on energy-constrained transport is straightforward to prove, and we dispatch it first.

PROPOSITION 2.1. *We have that*

$$\text{Nu}(\mathbf{u}) \leq 1 + \frac{1}{2} \langle |w|^2 \rangle^{1/2}$$

whenever $w|_{\partial\Omega} = 0$.

PROOF. Let us recall the argument from [19]. First, note that Nu does not depend on the initial temperature T_0 . Thus, we can take $T_0 = 1 - z$ and conclude by the maximum principle that the associated solution of (1.7) satisfies

$$0 \leq T \leq 1 \quad \text{a.e.}$$

Note also that because w vanishes at $\partial\Omega$,

$$\langle w \rangle = 0.$$

Combining this with Jensen's inequality and the definition of the Nusselt number (1.6), we have that

$$\begin{aligned} \text{Nu} - 1 &= \langle wT \rangle = \left\langle w \left(T - \frac{1}{2} \right) \right\rangle \\ &\leq \langle |w| \rangle \|T - \frac{1}{2}\|_{L_x^\infty} \leq \frac{1}{2} \langle |w|^2 \rangle^{1/2}. \end{aligned} \quad \square$$

The remainder of this section is on upper bounds for enstrophy-constrained transport. We prove the following bound:

PROPOSITION 2.2. *There exists a positive constant C such that*

$$\text{Nu}(\mathbf{u}) \leq 1 + C \langle |\nabla w|^2 \rangle \wedge \langle |\nabla w|^2 \rangle^{1/3}$$

whenever $w|_{\partial\Omega} = 0$. This constant is independent of all parameters.

To the authors' knowledge, there are at least three proofs of Proposition 2.2. For one, it can be obtained via an application of the background method [39]. It can also be seen as a consequence of Seis' arguments from [37]. Our proof of Proposition 2.2 is different from either of these: we obtain it via a new approach using a Dirichlet-type variational principle for the functional $\text{Nu}(\mathbf{u})$.

It should be mentioned that we are not the first to notice the variational structure of the advection-diffusion equation. The existence of a variational principle for advection-diffusion in bounded domains appears to have been first reported in [30], where it was used to systematically derive “best approximation” finite element schemes. Around the same time, as described in [26], variational principles for computing effective complex conductivities in periodic homogenization were discovered by Gibiansky and Cherkaev (the relevant corrector equation is again divergence-form but not self-adjoint). We learned about the existence of such principles from the papers [2, 17], whose formulas for computing effective diffusivities in periodic homogenization inspired the formulas for Nu obtained below. Let us also mention the related work [18], which discusses nonstandard variational principles for PDEs at large. It was a pleasant surprise to learn that the seemingly ad hoc change of variables introduced in [19] for handling the Euler-Lagrange equations of wall-to-wall optimal transport turn out to be similar to those employed in previous works, and that behind it all is a variational principle for Nu.

The remainder of this section is organized as follows. First we establish a variational principle for Nu in the steady case where the reasoning is most transparent. We then extend the arguments to general unsteady flows, where the variational principle turns into a variational bound as anticipated in (1.8). Proposition 2.2 follows immediately thereafter. Later on in Section 3, we obtain the dual formula to bound Nu from below. In order to highlight the key step in the proof—a certain symmetrizing change of variables for the advection-diffusion equation—we refer to this as the “symmetrization method.”

2.1 Symmetrization Method for Steady Velocity Fields

We start with the case where \mathbf{u} is an arbitrary divergence-free vector field belonging to $L^2(\Omega; \mathbb{R}^3)$. In this case,

$$\text{Nu}(\mathbf{u}) = 1 + \int_{\Omega} w\theta = 1 + \int_{\Omega} |\nabla\theta|^2$$

where $\theta = T - (1 - z)$ is the deviation of the temperature field from the conductive state. That is, θ is the unique (essentially bounded) weak solution of

$$\mathbf{u} \cdot \nabla\theta = \Delta\theta + w$$

with zero Dirichlet boundary data $\theta|_{\partial\Omega} = 0$. To change variables, we let θ_{\pm} be the unique weak solutions of the pair of formally adjoint PDEs

$$(2.1) \quad \pm \mathbf{u} \cdot \nabla\theta_{\pm} = \Delta\theta_{\pm} + w$$

with $\theta_{\pm}|_{\partial\Omega} = 0$, and observe that $\theta = \theta_+$. Then, we define $\eta, \xi \in H_0^1(\Omega)$ by

$$\eta = \frac{1}{2}(\theta_+ - \theta_-) \quad \text{and} \quad \xi = \frac{1}{2}(\theta_+ + \theta_-)$$

and observe they satisfy the equivalent system of PDEs

$$(2.2) \quad \begin{cases} \mathbf{u} \cdot \nabla\eta = \Delta\xi + w, \\ \mathbf{u} \cdot \nabla\xi = \Delta\eta. \end{cases}$$

We claim the change of variables $(\theta_+, \theta_-) \leftrightarrow (\eta, \xi)$ yields a variational formula for Nu.

Testing the second equation in (2.2) against ξ and integrating by parts shows that $\nabla\xi \perp \nabla\eta$ in $L^2(\Omega)$, since

$$\int_{\Omega} \nabla\eta \cdot \nabla\xi = - \int_{\Omega} \mathbf{u} \cdot \nabla\xi \xi = 0.$$

Therefore,

$$\text{Nu} - 1 = \int_{\Omega} |\nabla\theta_+|^2 = \int_{\Omega} |\nabla\eta|^2 + |\nabla\xi|^2$$

or, using the first PDE in (2.2),

$$(2.3) \quad \text{Nu} - 1 = \int_{\Omega} |\nabla\eta|^2 + |\nabla\Delta^{-1}[\mathbf{u} \cdot \nabla\eta - w]|^2.$$

Consider the right-hand side of (2.3) as it depends on η . Since $\mathbf{u} \in L^2$ and is divergence-free, the right-hand side is well-defined for $\eta \in H_0^1(\Omega) \cap L^\infty(\Omega)$. Now consider the variational problem

$$(2.4) \quad \inf_{\eta \in H_0^1(\Omega) \cap L^\infty(\Omega)} \int_{\Omega} |\nabla\eta|^2 + |\nabla\Delta^{-1}[\mathbf{u} \cdot \nabla\eta - w]|^2,$$

which is strictly convex so that any minimizer must be unique. By a first-variation argument, we see that η is a minimizer of (2.4) if and only if it satisfies the Euler-Lagrange equation

$$\Delta\eta = \mathbf{u} \cdot \nabla \Delta^{-1}[\mathbf{u} \cdot \nabla \eta - w].$$

This is a rewrite of the system (2.2) with ξ defined by

$$\xi = \Delta^{-1}[\mathbf{u} \cdot \nabla \eta - w].$$

Since that system possesses solutions in the class $H_0^1 \cap L^\infty$, it immediately follows that (2.4) has a minimizer. It also follows that the minimal value in (2.4) is $\text{Nu} - 1$.

Relabeling η as $\eta + 1 - z$ and using that \mathbf{u} is divergence-free yields the following variational principle for heat transport:

THEOREM 2.3. *Let $\mathbf{u}(\mathbf{x})$ be a divergence-free vector field in $L^2(\Omega; \mathbb{R}^3)$. Then,*

$$\text{Nu}(\mathbf{u}) = \min_{\substack{\eta \in H^1(\Omega) \cap L^\infty(\Omega) \\ \eta|_{z=0}=1, \eta|_{z=1}=0}} \int_{\Omega} |\nabla \eta|^2 + |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \eta)|^2.$$

2.2 Upper Bounds on Unsteady Transport by Symmetrization

With some care, the previous argument can be adapted to the time-dependent case. Here the useful change of variables arises from the pair of PDEs

$$\pm(\partial_t + \mathbf{u} \cdot \nabla) \theta_{\pm} = \Delta \theta_{\pm} + w,$$

which are formally adjoint in space and time. These are in obvious analogy with (2.1) from the steady case. However, since parabolic PDEs are generically only well-posed forward in time, making sense of the “ $-$ ” equation presents an added difficulty. To deal with this, we will reverse the sense of time between the equations, performing the change of variables $t \rightarrow \tau - t$ for appropriately chosen $\tau \gg 1$. In the limit $\tau \rightarrow \infty$, we recover the unsteady variational bound.

Define the admissible set of test functions

$$\begin{aligned} \mathcal{A} = \{&\eta: \eta \in L^\infty([0, \infty); L^2(\Omega))\} \\ &\cap \{\eta(t) \in H^1(\Omega) \cap L^\infty(\Omega), \partial_t \eta(t) \in H^{-1}(\Omega) \text{ a.e. } t\}. \end{aligned}$$

THEOREM 2.4. *Let $\mathbf{u}(\mathbf{x}, t)$ be a divergence-free vector field with bounded mean energy $\langle |\mathbf{u}|^2 \rangle < \infty$. Then,*

$$\text{Nu}(\mathbf{u}) \leq \inf_{\substack{\eta \in \mathcal{A} \\ \eta|_{z=1}=0, \eta|_{z=0}=1}} \langle |\nabla \eta|^2 + |\nabla \Delta^{-1}[\partial_t \eta + \operatorname{div}(\mathbf{u} \eta)]|^2 \rangle.$$

PROOF. We begin by introducing the (approximately) symmetrized variables. Let $\theta_+ = T - (1 - z)$ and note it solves

$$\partial_t \theta_+ + \mathbf{u} \cdot \nabla \theta_+ = \Delta \theta_+ + w$$

on $[0, \infty) \times \Omega$ and vanishes at $\partial\Omega$. Let θ_-^τ be the unique essentially bounded weak solution of

$$-\partial_t \theta_- - \mathbf{u} \cdot \nabla \theta_- = \Delta \theta_- + w$$

on $\Omega_\tau = [0, \tau] \times \Omega$ that vanishes at $\partial\Omega$ and has final time data $\theta_-^\tau(\tau) = 0$. Next, define the symmetrized variables η^τ and ξ^τ by

$$\eta^\tau = \frac{1}{2}(\theta_+ - \theta_-^\tau) \quad \text{and} \quad \xi^\tau = \frac{1}{2}(\theta_+ + \theta_-^\tau).$$

Observe that η^τ and ξ^τ solve

$$(2.5) \quad \begin{cases} \partial_t \eta^\tau + \mathbf{u} \cdot \nabla \eta^\tau = \Delta \xi^\tau + w, \\ \partial_t \xi^\tau + \mathbf{u} \cdot \nabla \xi^\tau = \Delta \eta^\tau, \end{cases}$$

on Ω_τ , vanish at $\partial\Omega$, and remain bounded in L_{tx}^∞ uniformly in time. In particular, by the maximum principle,

$$\|\eta^\tau\|_{L_{tx}^\infty} \vee \|\xi^\tau\|_{L_{tx}^\infty} \lesssim \|\theta_+\|_{L_{tx}^\infty} \vee \|\theta_-^\tau\|_{L_{tx}^\infty} \leq C(T_0).$$

We proceed as in the steady case, accumulating errors that vanish as $\tau \rightarrow \infty$. From the second PDE in (2.5) we find that

$$\begin{aligned} \langle \nabla \eta^\tau \cdot \nabla \xi^\tau \rangle_\tau &= \int_0^\tau \frac{1}{|\Omega|} (-\Delta \eta^\tau, \xi^\tau) dt = \int_0^\tau \frac{1}{|\Omega|} (-\partial_t \xi^\tau - \operatorname{div}(\mathbf{u} \xi^\tau), \xi^\tau) dt \\ &= - \int_0^\tau \frac{1}{|\Omega|} \frac{d}{dt} \frac{1}{2} \|\xi^\tau\|_{L^2(\Omega)}^2 + \int_0^\tau \int_\Omega \mathbf{u} \xi^\tau \cdot \nabla \xi^\tau \\ &= \frac{1}{2\tau} \frac{1}{|\Omega|} (\|\xi^\tau(0)\|_{L^2(\Omega)}^2 - \|\xi^\tau(\tau)\|_{L^2(\Omega)}^2). \end{aligned}$$

Therefore,

$$|\langle \nabla \eta^\tau \cdot \nabla \xi^\tau \rangle_\tau| \leq C(T_0) \frac{1}{\tau}.$$

Since

$$\operatorname{Nu} - 1 = \langle |\nabla \theta|^2 \rangle = \langle |\nabla \theta_+|^2 \rangle$$

and

$$\langle |\nabla \theta_+|^2 \rangle_\tau = \langle |\nabla \eta^\tau|^2 + |\nabla \xi^\tau|^2 \rangle_\tau + 2 \langle \nabla \eta^\tau \cdot \nabla \xi^\tau \rangle_\tau,$$

we conclude that

$$(2.6) \quad \operatorname{Nu} - 1 = \langle |\nabla \eta^\tau|^2 + |\nabla \Delta^{-1}[\partial_t \eta^\tau + \mathbf{u} \cdot \nabla \eta^\tau - w]|^2 \rangle_\tau + O\left(\frac{1}{\tau}\right).$$

Next, we prove that η^τ is approximately minimal, with an error that vanishes as $\tau \rightarrow \infty$. Let $\eta \in \mathcal{A}$ vanish at $\partial\Omega$ but be otherwise arbitrary, and consider the difference

$$\begin{aligned} A_\tau &= \langle |\nabla \eta|^2 + |\nabla \Delta^{-1}[\partial_t \eta + \mathbf{u} \cdot \nabla \eta - w]|^2 \rangle_\tau \\ &\quad - \langle |\nabla \eta^\tau|^2 + |\nabla \Delta^{-1}[\partial_t \eta^\tau + \mathbf{u} \cdot \nabla \eta^\tau - w]|^2 \rangle_\tau. \end{aligned}$$

Using the convexity of $|\cdot|^2$, we can expand around η^τ and use (2.5) to arrive at the lower bound

$$\begin{aligned} \frac{A_\tau}{2} &\geq \langle \nabla(\eta - \eta^\tau) \cdot \nabla \eta^\tau \\ &\quad + \nabla \Delta^{-1}(\partial_t + \mathbf{u} \cdot \nabla)(\eta - \eta^\tau) \cdot \nabla \Delta^{-1}[(\partial_t + \mathbf{u} \cdot \nabla)\eta^\tau - w] \rangle_\tau \\ &= -\frac{1}{|\Omega|} \int_0^\tau ((\partial_t + \mathbf{u} \cdot \nabla)\xi^\tau, \eta - \eta^\tau) + ((\partial_t + \mathbf{u} \cdot \nabla)(\eta - \eta^\tau), \xi^\tau) dt \\ &= -\frac{1}{|\Omega|} \int_0^\tau \frac{d}{dt} \left[\int_\Omega (\eta - \eta^\tau) \xi^\tau \right] dt \\ &\gtrsim -\frac{1}{\tau} \frac{1}{|\Omega|} (\|\eta\|_{L_t^\infty L_x^2} \vee \|\eta^\tau\|_{L_t^\infty L_x^2}) \cdot \|\xi^\tau\|_{L_t^\infty L_x^2} \geq -C(T_0, \eta, \Omega) \frac{1}{\tau}. \end{aligned}$$

Combining this with (2.6), we find that

$$\text{Nu} - 1 \leq \langle |\nabla \eta|^2 + |\nabla \Delta^{-1}[\partial_t \eta + \mathbf{u} \cdot \nabla \eta - w]|^2 \rangle_\tau + C(T_0, \eta, \Omega) \frac{1}{\tau}.$$

Taking $\tau \rightarrow \infty$ yields the inequality

$$\text{Nu} - 1 \leq \langle |\nabla \eta|^2 + |\nabla \Delta^{-1}[\partial_t \eta + \mathbf{u} \cdot \nabla \eta - w]|^2 \rangle.$$

This holds for all $\eta \in \mathcal{A}$ that vanish at $\partial\Omega$. Changing variables by $\eta \rightarrow \eta + 1 - z$ and optimizing yields the result. \square

Even if \mathbf{u} depends on time, η can be taken to be independent of time and still used to bound Nu. The simplified version of Theorem 2.4 that results is analogous to Theorem 2.3, but for unsteady heat transport.

COROLLARY 2.5. *Let $\mathbf{u}(\mathbf{x}, t)$ be a divergence-free vector field with bounded mean energy $\langle |\mathbf{u}|^2 \rangle < \infty$. Then,*

$$(2.7) \quad \text{Nu}(\mathbf{u}) \leq \inf_{\substack{\eta \in H^1(\Omega) \cap L^\infty(\Omega) \\ \eta|_{z=1}=0, \eta|_{z=0}=1}} \int_\Omega |\nabla \eta|^2 + \langle |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \eta)|^2 \rangle.$$

2.3 Proof of Proposition 2.2

We prove Proposition 2.2 by deducing it from Corollary 2.5 with a good choice of test function η . Evidently, we must find a convenient way to bound the nonlocal term appearing there. Since $\nabla \Delta^{-1} \operatorname{div}$ is an L^2 -orthogonal projection, one has the bound

$$\langle |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \eta)|^2 \rangle \leq \langle |\mathbf{u}(\eta - c)|^2 \rangle$$

for an arbitrary constant c . In the case that \mathbf{u} satisfies no-slip boundary conditions, one can deduce Proposition 2.2 by choosing $\eta \approx c$, thereby localizing the right-hand side to a small neighborhood of $\partial\Omega$. Then, a straightforward application of Poincaré's inequality yields the result.

The final step in the preceding argument requires all components of \mathbf{u} to vanish at $\partial\Omega$. This is not useful for dealing with no-penetration boundary conditions.

Nevertheless, Proposition 2.2 holds in this more general case. The key is to approach the nonlocal term from Corollary 2.5 by duality. Observe that

$$\int_{\Omega} |\nabla \Delta^{-1} \operatorname{div} \mathbf{m}|^2 = \sup_{\theta \in H_0^1(\Omega)} \int_{\Omega} 2\nabla \theta \cdot \mathbf{m} - |\nabla \theta|^2$$

whenever $\mathbf{m} \in L^2(\Omega)$. Thus, the inequality

$$\int_{\Omega} |\nabla \Delta^{-1} \operatorname{div} \mathbf{m}|^2 \leq C$$

and the statement that

$$\int_{\Omega} 2\nabla \theta \cdot \mathbf{m} \leq C + \int_{\Omega} |\nabla \theta|^2 \quad \forall \theta \in H_0^1$$

are one and the same. Taking $\mathbf{m} = \mathbf{u}\eta$ where η depends only on z , we conclude it will be useful to have bounds of the form

$$\int_{\Omega} 2\overline{w\theta} \eta'(z) \leq C + \int_{\Omega} |\nabla \theta|^2 \quad \forall \theta \in H_0^1.$$

The following preliminary result allows us to establish bounds of this type.

LEMMA 2.6. *Let $w, \theta \in H_0^1(\Omega)$. Then,*

$$|\overline{w\theta}(z)| \lesssim \frac{|z \wedge (1-z)|}{|\mathbb{T}_{xy}^2|} \|\partial_z \theta\|_{L^2(\Omega)} \|\partial_z w\|_{L^2(\Omega)} \quad a.e.$$

Remark 2.7. The reader familiar with the background method may recognize that this inequality also plays a key role in carrying out that approach to a priori bounds. In particular, it is useful for verifying the spectral constraint. See Section 6.2 for more on the connection between the symmetrization method and the background method.

PROOF. By the usual approximation arguments, we can take \mathbf{u} and θ to be smooth. Differentiating and applying the Cauchy-Schwarz inequality, we find that

$$\left| \frac{d}{dz} \overline{w^2} \right| \leq \frac{2}{|\mathbb{T}_{xy}^2|} \|\partial_z w\|_{L^2(\mathbb{T}_{xy}^2)} \|w\|_{L^2(\mathbb{T}_{xy}^2)}.$$

Integrating this from 0 to z yields

$$\begin{aligned} \|w\|_{L^\infty([0,z];L_{xy}^2)}^2 &\leq \frac{2}{|\mathbb{T}_{xy}^2|} \int_0^z \|\partial_z w(z')\|_{L_{xy}^2} \|w(z')\|_{L_{xy}^2} dz' \\ &\leq \frac{2}{|\mathbb{T}_{xy}^2|} \|\partial_z w\|_{L^2(\Omega)} \|w\|_{L^2([0,z];L_{xy}^2)} \\ &\leq \frac{2}{|\mathbb{T}_{xy}^2|} \|\partial_z w\|_{L^2(\Omega)} \|w\|_{L^\infty([0,z];L_{xy}^2)} |z|^{1/2}. \end{aligned}$$

Therefore,

$$(2.8) \quad \|w\|_{L^\infty([0,z];L_{xy}^2)} \leq \frac{2}{|\mathbb{T}_{xy}^2|} |z|^{1/2} \|\partial_z w\|_{L^2(\Omega)}.$$

Similarly, we find that

$$(2.9) \quad \|\theta\|_{L^\infty([0,z];L_{xy}^2)} \leq \frac{2}{|\mathbb{T}_{xy}^2|} |z|^{1/2} \|\partial_z \theta\|_{L^2(\Omega)}.$$

Now consider the product $w\theta$. We have that

$$\left| \frac{d}{dz} \overline{w\theta} \right| \leq |\overline{\partial_z w\theta}| + |\overline{w\partial_z \theta}| \leq \frac{1}{|\mathbb{T}_{xy}^2|} (\|\partial_z w\|_{L_{xy}^2} \|\theta\|_{L_{xy}^2} + \|w\|_{L_{xy}^2} \|\partial_z \theta\|_{L_{xy}^2})$$

for all z . Therefore,

$$\begin{aligned} & \|\overline{w\theta}\|_{L^\infty([0,z])} \\ & \leq \frac{1}{|\mathbb{T}_{xy}^2|} \int_0^z \|\partial_z w(z')\|_{L_{xy}^2} \|\theta(z')\|_{L_{xy}^2} + \|w(z')\|_{L_{xy}^2} \|\partial_z \theta(z')\|_{L_{xy}^2} dz' \\ & \leq \frac{1}{|\mathbb{T}_{xy}^2|} (\|\partial_z w\|_{L^2(\Omega)} \|\theta\|_{L^2([0,z];L_{xy}^2)} + \|\partial_z \theta\|_{L^2(\Omega)} \|w\|_{L^2([0,z];L_{xy}^2)}) \\ & \leq \frac{|z|^{1/2}}{|\mathbb{T}_{xy}^2|} (\|\partial_z w\|_{L^2(\Omega)} \|\theta\|_{L^\infty([0,z];L_{xy}^2)} + \|\partial_z \theta\|_{L^2(\Omega)} \|w\|_{L^\infty([0,z];L_{xy}^2)}). \end{aligned}$$

Applying (2.8) and (2.9) we find that

$$\|\overline{w\theta}\|_{L^\infty([0,z])} \leq \frac{4|z|}{|\mathbb{T}_{xy}^2|} \|\partial_z w\|_{L^2(\Omega)} \|\partial_z \theta\|_{L^2(\Omega)}.$$

The argument above is symmetric under $z \rightarrow 1 - z$, so we immediately obtain the inequality

$$\|\overline{w\theta}\|_{L^\infty([1-z,1])} \leq \frac{4|1-z|}{|\mathbb{T}_{xy}^2|} \|\partial_z w\|_{L^2(\Omega)} \|\partial_z \theta\|_{L^2(\Omega)}.$$

These two combine to prove the result. \square

Now we are ready to prove Proposition 2.2, which immediately implies the upper bound part of Theorem 1.2. We follow the plan laid out above.

PROOF OF PROPOSITION 2.2. We apply Corollary 2.5 with an appropriate class of test functions $\{\eta_\delta\}$. Given $\delta \in (0, 1/2]$, we define η_δ by

$$(2.10) \quad \eta_\delta(z) = \begin{cases} 1 - \frac{1}{2\delta}z, & 0 \leq z \leq \delta, \\ \frac{1}{2}, & \delta \leq z \leq 1 - \delta, \\ \frac{1}{2\delta}(1-z), & 1 - \delta \leq z \leq 1. \end{cases}$$

Note these are admissible in (2.7). Thus,

$$(2.11) \quad \text{Nu}(\mathbf{u}) \leq \inf_{\delta \in (0, \frac{1}{2}]} \left\{ \int_{\Omega} |\nabla \eta_{\delta}|^2 + \langle |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \eta_{\delta})|^2 \rangle \right\}.$$

The first integral appearing above is simple to estimate and it satisfies

$$\int_{\Omega} |\nabla \eta_{\delta}|^2 \sim \frac{1}{\delta}.$$

So,

$$(2.12) \quad \text{Nu} \lesssim \inf_{\delta \in (0, \frac{1}{2}]} \left\{ \frac{1}{\delta} + \langle |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \eta_{\delta})|^2 \rangle \right\}.$$

Now we must estimate the nonlocal term. We claim that

$$(2.13) \quad \langle |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \eta_{\delta})|^2 \rangle \lesssim \delta^2 \langle |\nabla w|^2 \rangle$$

for all $\delta \in (0, \frac{1}{2}]$. We argue at a.e. time t . By duality,

$$\int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \eta_{\delta})|^2 = \sup_{\theta \in H_0^1(\Omega)} 2(\operatorname{div}(\mathbf{u} \eta_{\delta}), \theta) - \int_{\Omega} |\nabla \theta|^2.$$

Since η_{δ} only depends on z , $\operatorname{div}(\mathbf{u} \eta_{\delta}) = w \eta'_{\delta}$, and we have by Fubini that

$$\int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \eta_{\delta})|^2 = \sup_{\theta \in H_0^1(\Omega)} \int_{I_z} 2\bar{\theta}w \eta'_{\delta} - \int_{\Omega} |\nabla \theta|^2.$$

Therefore to show (2.13) it suffices to prove the inequality

$$\int_{I_z} 2\bar{\theta}w \eta'_{\delta} \leq C\delta^2 \int_{\Omega} |\nabla w|^2 + \int_{\Omega} |\nabla \theta|^2 \quad \forall \mathbf{u} \in H_0^1(\Omega; \mathbb{R}^d), \theta \in H_0^1(\Omega).$$

Recalling the formula for η_{δ} from (2.10), we see we must prove that

$$\int_0^{\delta} |\bar{\theta}w(z)| dz \lesssim \delta \left(\int_{\Omega} |\nabla w|^2 \right)^{1/2} \left(\int_{\Omega} |\nabla \theta|^2 \right)^{1/2} \\ \forall \mathbf{u} \in H_0^1(\Omega; \mathbb{R}^d), \theta \in H_0^1(\Omega).$$

Applying Lemma 2.6 proves this result and hence the desired estimate (2.13).

Assembling (2.12) and (2.13), we conclude that

$$\text{Nu} \lesssim \inf_{\delta \in (0, \frac{1}{2}]} \left\{ \frac{1}{\delta} + \delta^2 \langle |\nabla w|^2 \rangle \right\}.$$

If $\langle |\nabla w|^2 \rangle \geq 1$ we may choose $\delta \sim \langle |\nabla w|^2 \rangle^{-1/3}$ to conclude that

$$(2.14) \quad \text{Nu} \lesssim \langle |\nabla w|^2 \rangle^{1/3}.$$

To handle the case $\langle |\nabla w|^2 \rangle \leq 1$ we treat the choice $\delta = \frac{1}{2}$ more carefully in the above. Since $\eta_{1/2} = 1 - z$, (2.11) and (2.13) combine to prove that

$$(2.15) \quad \text{Nu} \leq 1 + \langle |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \eta_{1/2})|^2 \rangle \leq 1 + C \langle |\nabla w|^2 \rangle.$$

From (2.14) and (2.15) we conclude the result. \square

3 Optimal Design of Steady Wall-to-Wall Transport

This section begins the proof of the lower bounds from Theorem 1.1 and Theorem 1.2. While every admissible velocity field yields a lower bound on the maximal rate of heat transport, it is not at all clear what sorts of features are required for velocity fields to achieve maximal (or nearly maximal) transport. It is natural to wonder how the overall character of optimal designs depends on the intensity budget. One possibility is captured by the convection rolls pictured in Figure 1.1a. This design is a relatively simple one, as the number of length scales required to describe it remains independent of the intensity budget. A second, much more complicated possibility is captured by the branching designs from Figure 1.1b. There, the total number of length scales is allowed to depend on the intensity budget and can be unbounded as $\text{Pe} \rightarrow \infty$. In any case, one requires a general method by which to evaluate Nu to allow for comparison between candidate designs.

The best scenario would be to develop an ansatz-free approach to evaluating heat transport that, by its functional form, suggests optimal designs. In this section, we achieve this for a general class of steady (i.e., time-independent) wall-to-wall optimal transport problems, including the energy- and enstrophy-constrained ones as special cases. The class of problems we have in mind are of the form

$$(3.1) \quad \sup_{\substack{\mathbf{u}(\mathbf{x}) \\ \|\mathbf{u}\| = \text{Pe} \\ +\text{b.c.}}} \text{Nu}(\mathbf{u})$$

where $\|\cdot\|$ denotes any norm in which the advective intensity of \mathbf{u} may be measured. As described in the introduction, the steady energy-constrained problem arises from employing the (volume-averaged) L^2 -norm

$$\|\mathbf{u}\| = \left(\int_{\Omega} |\mathbf{u}|^2 \right)^{1/2}$$

to measure advective intensity, while the steady enstrophy-constrained one arises from the (volume-averaged) \dot{H}^1 -norm

$$\|\mathbf{u}\| = \left(\int_{\Omega} |\nabla \mathbf{u}|^2 \right)^{1/2}.$$

In any case, we require that $\mathbf{u} \in L^2$ in order that its heat transport be well-defined. As our aim in this section is to present a general approach to intensity-constrained optimal transport, we leave the boundary conditions unspecified. Of course, we do not claim that there exist optimizers at this level of generality.

The principal result of this section is that the general wall-to-wall optimal transport problem (3.1) can be reformulated as the double minimization

$$(3.2) \quad \inf_{\substack{\mathbf{u}(\mathbf{x}), \xi(\mathbf{x}) \\ \int_{\Omega} w \xi = 1 \\ +\text{b.c.}}} \int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \xi)|^2 + \frac{1}{\text{Pe}^2} \|\mathbf{u}\|^2 \cdot \int_{\Omega} |\nabla \xi|^2$$

in the velocity field \mathbf{u} and a new variable ξ . The boundary conditions for \mathbf{u} remain the same as for (3.1), while ξ is required to vanish at $\partial\Omega$. As will become clear, ξ plays a role in the analysis of Nu similar to that of η from the a priori bounds of Section 2—in fact, these variables are dual. The optimal values in (3.1) and (3.2) are reciprocals, and their optimizers are related through a certain change of variables. We refer the reader forward to Sections 4 and 5 for the application of these observations to energy- and enstrophy-constrained optimal transport. Presently, our goal is to establish the connection between (3.1) and (3.2), and to illustrate how the latter suggests optimal designs. As in Section 2, our approach centers on the existence of a variational principle for $\text{Nu}(\mathbf{u})$; it is dual to the one appearing there. After achieving this duality and using it to obtain (3.2), we proceed to make some general remarks on the construction of near optimal designs.

3.1 Dual Variational Formulations for Transport

Recall from the analysis of a priori bounds on transport in Section 2 that there is a variational principle for heat transport in the steady case, and that Nu can be written as the optimal value of a certain convex minimization problem:

$$\text{Nu}(\mathbf{u}) - 1 = \min_{\eta|_{\partial\Omega}=0} \mathbf{f} \int_{\Omega} |\nabla \eta|^2 + |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \eta)|^2.$$

For the precise statement, see Theorem 2.3. As this is convex it should, in principle, admit a dual formulation.

THEOREM 3.1. *Let \mathbf{u} be a divergence-free vector field in $L^2(\Omega; \mathbb{R}^3)$. Then,*

$$(3.3) \quad \text{Nu}(\mathbf{u}) - 1 = \max_{\xi \in H_0^1(\Omega) \cap L^\infty(\Omega)} \mathbf{f} \int_{\Omega} 2w\xi - |\nabla \xi|^2 - |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \xi)|^2.$$

PROOF. Recall from the proof of Theorem 2.3 the PDE system

$$(3.4) \quad \begin{cases} \mathbf{u} \cdot \nabla \eta = \Delta \xi + w, \\ \mathbf{u} \cdot \nabla \xi = \Delta \eta, \end{cases}$$

and the formula

$$\text{Nu} - 1 = \mathbf{f} \int_{\Omega} |\nabla \xi|^2 + |\nabla \eta|^2.$$

Testing the first equation in (3.4) with ξ , the second with η , and integrating by parts yields the string of equalities

$$\begin{aligned} \int_{\Omega} w\xi &= \int_{\Omega} |\nabla \xi|^2 + \int_{\Omega} \xi \mathbf{u} \cdot \nabla \eta \\ &= \int_{\Omega} |\nabla \xi|^2 - \int_{\Omega} \eta \mathbf{u} \cdot \nabla \xi = \int_{\Omega} |\nabla \xi|^2 + |\nabla \eta|^2. \end{aligned}$$

Thus,

$$\text{Nu} - 1 = \mathbf{f} \int_{\Omega} 2w\xi - |\nabla \xi|^2 - |\nabla \eta|^2 = \mathbf{f} \int_{\Omega} 2w\xi - |\nabla \xi|^2 - |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \xi)|^2.$$

Note in the last step we used the PDE system again.

Now consider the maximization

$$\sup_{\xi \in H_0^1(\Omega) \cap L^\infty(\Omega)} \int_{\Omega} 2w\xi - |\nabla \xi|^2 - |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u}\xi)|^2.$$

Reasoning with its Euler-Lagrange equation just as in the proof of Theorem 2.3, we deduce that this maximization problem is well-posed in the given admissible class, with optimal value $\text{Nu} - 1$. \square

As claimed in the introduction, there is a corresponding result holding for unsteady velocities that allows us to bound Nu from below, but not necessarily to evaluate it. This result was described in (1.9), and its proof is similar to that of Theorem 2.4. We remark that although the variational formulas for steady heat transport from Theorem 2.3 and Theorem 3.1 are strongly dual—they provide convex and concave alternatives for evaluating Nu whose optimal values agree—such strong duality need not hold for unsteady flows. More precisely, we note that for general velocity fields the bounds (1.8) and (1.9) need not coincide. In particular, there will be a duality gap for any velocity field that satisfies

$$\liminf_{\tau \rightarrow \infty} \langle wT \rangle_{\tau} < \limsup_{\tau \rightarrow \infty} \langle wT \rangle_{\tau}.$$

These fields have the peculiar property that the \limsup and \liminf alternatives for defining the space and long-time average heat transport Nu do not coincide.

3.2 An Integral Formulation of Wall-to-Wall Optimal Transport

Having written Nu for steady flows as the optimal value of the concave maximization problem (3.3), we can now give a useful reformulation of the entire class of steady wall-to-wall optimal transport problems from (3.1). This “integral” formulation of steady optimal transport will be used to design and evaluate nearly optimal flows in what follows.

Let $F(\text{Pe})$ denote the optimal value of the steady optimal transport problem (3.1),

$$F(\text{Pe}) = \sup_{\substack{\mathbf{u}(\mathbf{x}) \\ \|\mathbf{u}\| = \text{Pe} \\ +\text{b.c.}}} \text{Nu}(\mathbf{u}).$$

Applying Theorem 3.1, we find that

$$F(\text{Pe}) = \sup_{\substack{\mathbf{u}(\mathbf{x}), \xi(\mathbf{x}) \\ \|\mathbf{u}\| = \text{Pe} \\ +\text{b.c.}}} \int_{\Omega} 2w\xi - |\nabla \xi|^2 - |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u}\xi)|^2.$$

The boundary conditions for \mathbf{u} remain unchanged, while according to Theorem 3.1 we must require that $\xi|_{\partial\Omega} = 0$. Performing the substitution

$$\xi \rightarrow \lambda \xi, \quad \lambda \in \mathbb{R},$$

and maximizing over λ yields the equivalent variational problem

$$\frac{1}{F(\text{Pe})} = \inf_{\substack{\mathbf{u}(\mathbf{x}), \xi(\mathbf{x}) \\ \|\mathbf{u}\| = \text{Pe} \\ +\text{b.c.}}} \frac{\int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u}\xi)|^2 + |\nabla \xi|^2}{(\int_{\Omega} w\xi)^2}.$$

Changing variables via the substitutions

$$\mathbf{u} \rightarrow \text{Pe} \frac{\mathbf{u}}{\|\mathbf{u}\|} \quad \text{and} \quad \xi \rightarrow \frac{\|\mathbf{u}\|}{\text{Pe}} \xi$$

allows us to eliminate the intensity constraint on \mathbf{u} altogether so that

$$\frac{1}{F(\text{Pe})} = \inf_{\substack{\mathbf{u}(\mathbf{x}), \xi(\mathbf{x}) \\ +\text{b.c.}}} \frac{\int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u}\xi)|^2 + \frac{1}{\text{Pe}^2} \|\mathbf{u}\|^2 \int_{\Omega} |\nabla \xi|^2}{(\int_{\Omega} w\xi)^2}.$$

Given the scaling symmetries of the above, we may impose the constraint

$$\int_{\Omega} w\xi = 1$$

on the minimization without altering the result. This yields the promised integral reformulation of wall-to-wall optimal transport

$$(3.5) \quad \frac{1}{F(\text{Pe})} = \inf_{\substack{\mathbf{u}(\mathbf{x}), \xi(\mathbf{x}) \\ \int_{\Omega} w\xi = 1 \\ +\text{b.c.}}} \int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u}\xi)|^2 + \frac{1}{\text{Pe}^2} \|\mathbf{u}\|^2 \cdot \int_{\Omega} |\nabla \xi|^2,$$

and proves the equivalence between (3.1) and (3.2).

We turn to discuss the integral formulation of optimal transport just derived. Observe (3.5) consists of two types of terms, each of which prefers a different kind of design. The first term

$$(3.6) \quad \int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u}\xi)|^2$$

prefers $\mathbf{u}\xi$ to be divergence-free and we refer to it as the “advection term” throughout. This preference is strong in the advective regime $\text{Pe} \gg 1$, as it appears at leading order in Pe^{-1} in the functional above. The remaining terms

$$\|\mathbf{u}\|^2 \quad \text{and} \quad \int_{\Omega} |\nabla \xi|^2$$

contribute at higher order in Pe^{-1} , and act to regularize designs. Any admissible design must satisfy the “net flux” constraint

$$(3.7) \quad \int_{\Omega} w\xi = 1$$

as well as boundary conditions. While patterns such as the convection roll and branching ones depicted in Figure 1.1 can be easily made to satisfy such constraints—see Sections 4 and 5 for details—determining the optimal length scales for such designs requires performing an optimization as in (3.5).

Evidently, the most difficult term to evaluate is the advection one (3.6). Before turning to discuss its analysis in detail and what it implies for near optimal designs, we make two general remarks. In order to get a hint as to what designs (3.5) prefers in the advective limit $\text{Pe} \rightarrow \infty$, one can entertain the “limiting” wall-to-wall optimal transport problem

$$\inf_{\substack{\mathbf{u}(\mathbf{x}), \xi(\mathbf{x}) \\ \int_{\Omega} w\xi = 1 \\ +\text{b.c.}}} \oint_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u}\xi)|^2$$

obtained by formally taking $\text{Pe} = \infty$ in (3.5). This, however, is an ill-posed variational problem. Its optimal value is 0 as there exist admissible sequences $\{(\mathbf{u}_k, \xi_k)\}$ satisfying the net flux constraint (3.7) and achieving

$$\lim_{k \rightarrow \infty} \oint_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u}_k \xi_k)|^2 = 0.$$

Yet, no smooth enough admissible pair (\mathbf{u}, ξ) can satisfy the net flux constraint and simultaneously achieve

$$(3.8) \quad \operatorname{div}(\mathbf{u}\xi) = 0.$$

Indeed, if $\mathbf{u}\xi$ were divergence-free, then by averaging (3.8) in the periodic variables x and y we would find that the flux of ξ by \mathbf{u} through each slice $\{z = \text{const}\}$ is independent of the slice, i.e.,

$$\overline{w\xi}(z) = \oint_{\mathbb{T}_{x,y}^2} w\xi$$

is constant in z . By applying the boundary conditions that require at least that $\xi|_{\partial\Omega} = 0$, we conclude that $w\xi$ must vanish throughout the entire domain. This contradicts the net flux constraint (3.7). Therefore, wall-to-wall optimal transport is a singularly perturbed variational problem: the regularizing terms from (3.5), which at first glance appear to contribute at higher order in Pe^{-1} , are crucial for determining the character of optimal designs.

Our second observation is more straightforward: it is regarding the disappearance of the intensity constraint in the passage from (3.1) to its integral formulation (3.5). Since (3.5) is invariant under the rescaling

$$\mathbf{u} \rightarrow \lambda \mathbf{u} \quad \text{and} \quad \xi \rightarrow \frac{1}{\lambda} \xi, \quad \lambda \neq 0,$$

the magnitudes of any of its minimizers are not uniquely determined. Still, if $(\mathbf{u}_{Pe}, \xi_{Pe})$ achieves optimality in (3.5),

$$\mathbf{u} = \frac{Pe}{\|\mathbf{u}_{Pe}\|} \mathbf{u}_{Pe}$$

solves the wall-to-wall problem (3.1).

3.3 Analysis of the Advection Term

For a class of designs $\{(\mathbf{u}_\alpha, \xi_\alpha)\}_{\alpha \in \mathcal{I}}$ to compete in the minimization (3.5), it must at least achieve

$$\inf_{\alpha \in \mathcal{I}} \int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u}_\alpha \xi_\alpha)|^2 = 0.$$

How difficult is it for an admissible pair (\mathbf{u}, ξ) to make this advection term *nearly* zero? First, note that in such a situation, the vertical flux of ξ by \mathbf{u} through each slice $\{z = \text{const}\}$ must be nearly independent of the slice,

$$\frac{d}{dz} \overline{w\xi} \approx 0.$$

By the net flux constraint (3.7), it follows that

$$(3.9) \quad \overline{w\xi} \approx 1$$

in nearly all of the domain. This is an example of a “design principle” for wall-to-wall optimal transport: any nearly optimal design must achieve (3.9) with equality in the limit $Pe \rightarrow \infty$. Although (3.9) does not completely characterize optimal designs, it does give a necessary condition for constructing competitive ones. This will be particularly useful later on in Section 5, where we devise a functional form for the branching depicted in Figure 1.1b.

The advection term (3.6) contains a wealth of information for evaluating designs beyond (3.9), but to use it in practice one must confront its nonlocality. In the wall-to-wall domain $\Omega = \mathbb{T}_{xy}^2 \times [0, 1]_z \cong [0, l_x] \times [0, l_y] \times [0, 1]$, we may take advantage of periodicity to represent a function $f \in L^2(\Omega)$ via its Fourier series

$$f(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbb{Z}_{l_x, l_y}^2} \hat{f}_{\mathbf{k}}(z) e^{i\mathbf{k} \cdot \mathbf{x}'}$$

where $\mathbf{x}' = (x, y)$ and

$$\mathbb{Z}_{l_x, l_y}^2 = \left\{ (k_x, k_y) : \frac{l_x}{2\pi} k_x, \frac{l_y}{2\pi} k_y \in \mathbb{Z} \right\}.$$

We employ the Fourier transform

$$\hat{f}_{\mathbf{k}}(z) = \int_{\mathbb{T}_{xy}^2} e^{i\mathbf{k} \cdot \mathbf{x}'} f(\mathbf{x}', z) dx dy$$

and proceed to decompose the advection term mode by mode.

LEMMA 3.2. *Let $\mathbf{u} \in L^2(\Omega; \mathbb{R}^3)$ and $\xi \in L^\infty(\Omega)$. The advection term satisfies*

$$\int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \xi)|^2 = \int_{I_z} |\overline{w\xi} - \int_{\Omega} w\xi|^2 + \mathcal{Q}(\operatorname{div}(\mathbf{u} \xi))$$

where $\mathcal{Q} = \sum_{\mathbf{k} \neq \mathbf{0}} \mathcal{Q}_{\mathbf{k}}$ and $\mathcal{Q}_{\mathbf{k}}$ is the positive semidefinite quadratic form given by

$$\mathcal{Q}_{\mathbf{k}}(f) = \int_{I_z \times I_{z'}} G_{\mathbf{k}}(z, z') \hat{f}_{\mathbf{k}}(z) \hat{f}_{\mathbf{k}}^*(z') dz dz'$$

with kernel

$$G_{\mathbf{k}}(z, z') = \frac{\operatorname{csch}(|\mathbf{k}|)}{|\mathbf{k}|} \times \begin{cases} \sinh(|\mathbf{k}|z) \sinh(|\mathbf{k}|(1-z')), & z \leq z', \\ \sinh(|\mathbf{k}|z') \sinh(|\mathbf{k}|(1-z)), & z \geq z', \end{cases} \quad \mathbf{k} \neq 0.$$

PROOF. This follows from the Green's function representation for $-\Delta^{-1}$ on Ω with vanishing Dirichlet boundary conditions. Calling $J = \operatorname{div}(\mathbf{u} \xi)$ and applying Parseval's identity, we see that the advection term can be written as

$$\begin{aligned} \int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \xi)|^2 &= \frac{1}{|\Omega|} (J, -\Delta^{-1} J) = \sum_{\mathbf{k}} \int_{I_z} \left(-\frac{d^2}{dz^2} + |\mathbf{k}|^2 \right)^{-1} \hat{J}_{\mathbf{k}} \hat{J}_{\mathbf{k}}^* \\ &= \sum_{\mathbf{k}} \int_{I_z \times I_{z'}} G_{\mathbf{k}}(z, z') \hat{J}_{\mathbf{k}}(z) \hat{J}_{\mathbf{k}}^*(z') dz dz' \end{aligned}$$

where for $\mathbf{k} \neq \mathbf{0}$ the functions $G_{\mathbf{k}}$ are as defined above. For $\mathbf{k} = \mathbf{0}$, we have

$$G_{\mathbf{0}}(z, z') = \begin{cases} z(1-z'), & z \leq z', \\ z'(1-z), & z \geq z'. \end{cases}$$

We must only address the form of the zeroth term now.

We recognize that

$$\int_{I_z \times I_{z'}} G_{\mathbf{0}}(z, z') \hat{J}_{\mathbf{0}}(z) \hat{J}_{\mathbf{0}}^*(z') dz dz' = \int_{\Omega} |\nabla \Delta^{-1} \overline{J}(z)|^2.$$

By periodicity,

$$\overline{J} = \overline{\operatorname{div}(\mathbf{u} \xi)} = \frac{d}{dz} \overline{w\xi}.$$

Thus,

$$\begin{aligned} \int_{\Omega} |\nabla \Delta^{-1} \overline{J}(z)|^2 &= \int_{\Omega} \left| \frac{d}{dz} \left(\frac{d}{dz} \overline{w\xi}(z) \right) \right|^2 \\ &= \int_{I_z} \left| \overline{w\xi} - \int_{I_z} \overline{w\xi} \right|^2 = \int_{I_z} \left| \overline{w\xi} - \int_{\Omega} w\xi \right|^2 \end{aligned}$$

as required. \square

As the quadratic form \mathcal{Q} from Lemma 3.2 is nonnegative, the advection term satisfies the lower bound

$$\int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \xi)|^2 \geq \int_{I_z} \left| \overline{w\xi} - \int_{\Omega} w\xi \right|^2.$$

Note this quantifies the design principle (3.9). The appearance of \mathcal{Q} in Lemma 3.2 also makes clear why this principle alone does not suffice to characterize optimal designs.

We end this section by recording some useful estimates on the kernels $\{G_{\mathbf{k}}\}$ from the definition of \mathcal{Q} . These will be used later in Section 5.

LEMMA 3.3. *Let $A \subset I_z$ be Borel measurable. Then,*

$$\|G_{\mathbf{k}}\|_{L^1(A \times A)} \lesssim \frac{|A|}{|\mathbf{k}|} \left(|A| \wedge \frac{1}{|\mathbf{k}|} \right) \quad \mathbf{k} \neq 0$$

and

$$\|G_{\mathbf{0}}\|_{L^1(A \times A)} \lesssim |A|^2 \int_A z \wedge (1-z) dz.$$

PROOF. To see the first estimate, observe that $G_{\mathbf{k}}$ satisfies the pointwise estimate

$$|G_{\mathbf{k}}(z, z')| \leq \frac{1}{|\mathbf{k}|} e^{-|\mathbf{k}||z-z'|}, \quad \mathbf{k} \neq \mathbf{0}.$$

Now,

$$\begin{aligned} \int_{I_z \times I_z} e^{-|\mathbf{k}||z-z'|} \mathbb{1}_A(z) \mathbb{1}_A(z') dz dz' &\leq \int_{I_z} \left[\int_{\mathbb{R}} e^{-|\mathbf{k}||z-z'|} dz' \right] \mathbb{1}_A(z) dz \\ &= \frac{2}{|\mathbf{k}|} |A|. \end{aligned}$$

Therefore,

$$\|G_{\mathbf{k}}\|_{L^1(A \times A)} \lesssim \frac{|A|}{|\mathbf{k}|^2}.$$

On the other hand, we have that

$$\|G_{\mathbf{k}}\|_{L^1(A \times A)} \leq \|G_{\mathbf{k}}\|_{L^\infty} \|\mathbb{1}_{A \times A}\|_{L^1} \leq \frac{|A|^2}{|\mathbf{k}|}.$$

Combining these two bounds gives the first result.

Now we prove the second estimate. We need to show that

$$\int G_{\mathbf{0}}(z, z') \mathbb{1}_A(z) \mathbb{1}_A(z') dz dz' \lesssim |A| \int_A z \wedge (1-z) dz.$$

By symmetry,

$$\begin{aligned} \int G_0(z, z') \mathbb{1}_A(z) \mathbb{1}_A(z') dz dz' &= 2 \int_0^1 \int_0^{z'} z(1-z') \mathbb{1}_A(z) \mathbb{1}_A(z') dz dz' \\ &\leq 2 \int_0^1 \int_0^{z'} z'(1-z') \mathbb{1}_A(z) \mathbb{1}_A(z') dz dz' \\ &\leq 2|A| \int_0^1 z(1-z) dz. \end{aligned}$$

Since

$$z(1-z) \leq z \wedge (1-z) \quad \forall z \in [0, 1]$$

we conclude the desired result. \square

4 Energy-Constrained Transport and Convection Roll Designs

In the previous section, we considered the general class of steady wall-to-wall optimal transport problems (3.1) and produced their equivalent integral formulations (3.2). In this section and the next, we use these formulations to study the steady energy- and enstrophy-constrained problems. The subsequent analyses are largely independent. Nevertheless, the reader may find it helpful to study the energy-constrained problem first as its proof is much shorter and its technical details much less burdensome.

Here we discuss energy-constrained transport. The main result of this section is a proof of the lower bound from Theorem 1.1. Recall from Section 3 that the steady energy-constrained optimal transport problem

$$(4.1) \quad \max_{\substack{\mathbf{u}(\mathbf{x}) \\ \int_{\Omega} |\mathbf{u}|^2 = \text{Pe}^2 \\ w|_{\partial\Omega} = 0}} \text{Nu}(\mathbf{u})$$

admits the integral formulation

$$(4.2) \quad \min_{\substack{\mathbf{u}(\mathbf{x}), \xi(\mathbf{x}) \\ \int_{\Omega} w \xi = 1 \\ w|_{\partial\Omega} = \xi|_{\partial\Omega} = 0}} \int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \xi)|^2 + \varepsilon \int_{\Omega} |\mathbf{u}|^2 \cdot \int_{\Omega} |\nabla \xi|^2$$

with $\varepsilon = \text{Pe}^{-2}$. The optimal values of (4.1) and (4.2) are reciprocals and their optimizers are related through symmetrization.

Our goal now is to identify the scaling law of (4.2) in the advective regime $\varepsilon \ll 1$. Combined with the results of Section 3, this completes the proof of the lower bound half of Theorem 1.1.

PROPOSITION 4.1. *Let $\mathfrak{E}(\varepsilon)$ denote the optimal value of (4.2). Then,*

$$\mathfrak{E}(\varepsilon) \sim \varepsilon^{1/2}$$

whenever $\varepsilon \lesssim 1 \wedge l_x^4 \wedge l_y^4$.

The a priori lower bound $\mathfrak{E}(\varepsilon) \gtrsim \varepsilon^{1/2}$ is implied by the corresponding bound $\text{Nu} \lesssim \text{Pe}$ from Theorem 1.1. The remainder of this section is regarding the upper bound $\mathfrak{E}(\varepsilon) \lesssim \varepsilon^{1/2}$. We prove it by constructing the convection roll designs depicted in Figure 1.1a. Such designs can be parametrized using two variables: the number of rolls and their wall-to-wall extent. Carrying out the optimization from (4.2) with respect to these variables yields the desired upper bound. The condition that ε be small enough in the statement above is required to fit what would be, in the absence of an overall horizontal period, an optimal number of rolls inside the domain. Given the symmetry between x and y , we may suppose that $l_x \leq l_y$ in what follows.

4.1 Convection Roll Designs

The integral formulation (4.2) requires designing a velocity field \mathbf{u} and a test function ξ . For the velocity field, we introduce a family of streamfunctions of the form

$$\psi(x, z) = \chi(z)\Psi(x), \quad \chi \in C_c^\infty(I_z), \quad \Psi \in C^\infty(\mathbb{T}_x).$$

Each such ψ gives rise to a divergence-free velocity field by

$$\mathbf{u} = (-\partial_z \psi, 0, \partial_x \psi).$$

These are two-dimensional flows as their $\hat{\mathbf{j}}$ -component vanishes identically. Although we do not claim that optimizers must be of this form, we will prove that such a construction suffices to capture the optimal scaling law of (4.2).

Next we must describe test functions ξ well suited to the velocity fields. Recall the design principle (3.9), which states that for a design to be competitive it must satisfy

$$(4.3) \quad \overline{w\xi} \approx \int_{\Omega} w\xi = 1.$$

This rules out taking, for instance, $\xi = \psi$, as it would result in zero flux through each slice $\{z = \text{const}\}$. We can, however, choose ξ to depend only on x and z as does ψ . Then by Parseval's identity we can rewrite the flux as

$$\overline{w\xi} = \sum_{\mathbf{k}} \hat{w}_{\mathbf{k}}(z) \hat{\xi}_{\mathbf{k}}^*(z) = \sum_{\mathbf{k}} (ik_x^{1/2} \hat{\psi}_{\mathbf{k}})(k_x^{1/2} \hat{\xi}_{\mathbf{k}})^*.$$

Taking

$$|\partial_x^{1/2} \xi| \sim |\partial_x^{1/2} \psi|$$

allows us to satisfy (4.3).

Now we make the convection roll construction concrete. Given $\delta \in (0, 1/2)$ and l such that

$$\frac{1}{l} \in \frac{2\pi}{l_x} \mathbb{N},$$

we define

$$\psi_{\delta,l}(x, z) = \chi_{\delta}(z) \cdot l^{1/2} \Psi\left(\frac{x}{l}\right), \quad \xi_{\delta,l}(x, z) = \chi_{\delta}(z) \cdot l^{1/2} \Psi'\left(\frac{x}{l}\right),$$

where

$$\Psi(x) = c_0 \cos x$$

and c_0 is chosen so that $\overline{(\Psi')^2} = 1$. Here, the cutoff functions $\{\chi_\delta\} \subset C_c^\infty(I_z)$ are required to satisfy

- $\chi_\delta = 1$ on $[\delta, 1 - \delta]$,
- $|\chi_\delta| \lesssim 1$ and $|\chi'_\delta| \lesssim \frac{1}{\delta}$,
- $\int_{I_z} \chi_\delta^2 = 1$.

The constants in these assumptions are independent of all parameters. In what follows, we often neglect to record the subscripts δ and l as the meaning is clear.

First, we check admissibility.

LEMMA 4.2. *The convection roll construction described above is admissible for (4.2).*

PROOF. All conditions in admissibility are clear except for the net flux constraint, which we verify now. Given the above, we find that

$$\overline{w\xi} = \overline{\partial_x \psi \xi} = \chi_\delta^2 \overline{(\Psi')^2} = \chi_\delta^2$$

for all z , so that

$$\int_{\Omega} w\xi = \int_{I_z} \overline{w\xi} = \int_{I_z} \chi_\delta^2 = 1$$

as required. \square

Next, we estimate the advection term from (4.2).

LEMMA 4.3. *The convection roll construction satisfies*

$$\int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \xi)|^2 = \int_{I_z} |\overline{w\xi} - 1|^2 \lesssim \delta.$$

In particular, the quadratic form \mathcal{Q} from Lemma 3.2 vanishes on it.

PROOF. We apply Lemma 3.2. Note that since all functions entering into the construction are independent of y , we may work with k in place of $\mathbf{k} = (k_x, 0)$. We start with the $k = 0$ term from Lemma 3.2: it satisfies the estimate

$$\int_{I_z} |\overline{w\xi} - 1|^2 = \int_{I_z} |\chi_\delta^2 - 1|^2 \lesssim \delta.$$

Now we address the $k \neq 0$ terms. We must compute the quadratic form \mathcal{Q} from Lemma 3.2, and to do so we must compute \hat{J}_k for $k \neq 0$ where $J = \mathbf{u} \cdot \nabla \xi$. Note

that by the form of the convection roll construction,

$$\begin{aligned}
J &= \nabla^\perp \psi \cdot \nabla \xi = -\partial_z \psi \partial_x \xi + \partial_x \psi \partial_z \xi \\
&= -\left(\chi'_\delta(z) \cdot l^{1/2} \Psi\left(\frac{x}{l}\right)\right) \left(\chi_\delta(z) \cdot \frac{1}{l^{1/2}} \Psi''\left(\frac{x}{l}\right)\right) \\
&\quad + \left(\chi_\delta(z) \frac{1}{l^{1/2}} \Psi'\left(\frac{x}{l}\right)\right) \left(\chi'_\delta(z) \cdot l^{1/2} \Psi'\left(\frac{x}{l}\right)\right) \\
&= \frac{1}{2} (\chi_\delta^2)'(z) \cdot \Theta\left(\frac{x}{l}\right)
\end{aligned}$$

where $\Theta = (\Psi')^2 - \Psi \Psi''$. Since $\Theta = c_0^2 \sin^2 + c_0^2 \cos^2 = c_0^2$,

$$J = \frac{1}{2} (\chi_\delta^2)'(z),$$

which is entirely a function of z . This shows that $\hat{J}_k = 0$ for $k \neq 0$. Hence, \mathcal{Q} vanishes on the convection roll construction. \square

Continuing, we estimate the higher-order terms from (4.2).

LEMMA 4.4. *The convection roll construction satisfies*

$$\mathbf{f}_\Omega |\nabla \xi|^2 \vee \mathbf{f}_\Omega |\mathbf{u}|^2 \lesssim \frac{l}{\delta} + \frac{1}{l}.$$

PROOF. Clearly,

$$\mathbf{f}_\Omega |\nabla \xi|^2 \sim \mathbf{f}_\Omega |\nabla \psi|^2$$

since $\|\Psi^{(k)}\|_{L^2(\mathbb{T}_x)} \sim 1$ for all k . Noting that

$$\mathbf{f}_\Omega |\mathbf{u}|^2 = \mathbf{f}_\Omega |\nabla \psi|^2 = \mathbf{f}_\Omega (\chi'_\delta)^2 l \Psi^2 + \chi_\delta \frac{1}{l} (\Psi')^2 \lesssim \frac{l}{\delta} + \frac{1}{l},$$

we conclude the result. \square

Combining the above yields the following estimate on

$$E(\varepsilon; \delta, l) = \mathbf{f}_\Omega |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u}_{\delta, l} \xi_{\delta, l})|^2 + \varepsilon \mathbf{f}_\Omega |\mathbf{u}_{\delta, l}|^2 \cdot \mathbf{f}_\Omega |\nabla \xi_{\delta, l}|^2.$$

COROLLARY 4.5. *The convection roll construction satisfies*

$$E(\varepsilon; \delta, l) \lesssim \delta + \varepsilon \left(\frac{l}{\delta} + \frac{1}{l} \right)^2.$$

4.2 Proof of Proposition 4.1

Finally we choose δ and l to prove the desired bound.

PROOF OF PROPOSITION 4.1. Corollary 4.5 holds for all admissible δ and l , i.e., so long as $\delta \in (0, \frac{1}{2})$ and $l^{-1} \in 2\pi l_x^{-1} \mathbb{N}$. To optimize the result, we take

$$l \sim \sqrt{\delta} \quad \text{and} \quad \delta \sim \varepsilon^{1/2}$$

which we can do so as long as $\varepsilon \lesssim 1 \wedge l_x^4$. Under this condition,

$$\inf_{\delta, l} E(\varepsilon; \delta, l) \lesssim \varepsilon^{1/2}. \quad \square$$

5 Enstrophy-Constrained Transport and Branched Flow Designs

We now consider the steady enstrophy-constrained wall-to-wall optimal transport problem in the framework of Section 3. The main result of this section is a proof of the lower bound from Theorem 1.2. As in the previous section on energy-constrained transport, we exploit the fact that the enstrophy-constrained problem

$$(5.1) \quad \max_{\substack{\mathbf{u}(\mathbf{x}) \\ f_\Omega |\nabla \mathbf{u}|^2 = \text{Pe}^2 \\ \mathbf{u}|_{\partial\Omega} = \mathbf{0}}} \text{Nu}(\mathbf{u})$$

can be written in integral form as

$$(5.2) \quad \min_{\substack{\mathbf{u}(\mathbf{x}), \xi(\mathbf{x}) \\ f_\Omega w \xi = 1 \\ \mathbf{u}|_{\partial\Omega} = \mathbf{0}, \xi|_{\partial\Omega} = 0}} \int_\Omega |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \xi)|^2 + \varepsilon \int_\Omega |\nabla \mathbf{u}|^2 \cdot \int_\Omega |\nabla \xi|^2$$

where $\varepsilon = \text{Pe}^{-2}$. This form of the problem suggests the possibility of analyzing a certain multiple-scales ansatz for \mathbf{u} and ξ (we explain the intuition behind this further in Sections 6 and 7). As proved below, such an ansatz turns out to capture the scaling of the optimal value of (5.2) in ε up to possible logarithmic corrections. The precise statement is as follows:

PROPOSITION 5.1. *Let $\mathfrak{E}(\varepsilon)$ denote the optimal value of (5.2). Then,*

$$\varepsilon^{1/3} \lesssim \mathfrak{E}(\varepsilon) \lesssim \varepsilon^{1/3} \log^{4/3} \frac{1}{\varepsilon}$$

whenever $\varepsilon \lesssim 1$ and $\varepsilon \log \frac{1}{\varepsilon} \lesssim l_x^6 \wedge l_y^6$.

The a priori lower bound $\mathfrak{E}(\varepsilon) \gtrsim \varepsilon^{1/3}$ is implied by the upper bound $\text{Nu} \lesssim \text{Pe}^{2/3}$ from Theorem 1.2. (For a proof which is more self-contained, see the discussion surrounding (6.11).) To prove the upper bound $\mathfrak{E}(\varepsilon) \lesssim \varepsilon^{1/3} \log^{4/3} \frac{1}{\varepsilon}$, we must construct a suitable class of designs and estimate their heat transport. The successful ones are as depicted in Figure 1.1b. In contrast to the convection roll designs considered previously, such ‘branching’ designs are evidently more complicated to analyze. The main challenge of course lies with estimating the advection term.

Here, we make use of Lemmas 3.2 and 3.3. Note the requirement that ε be small enough is to ensure that our construction fits into the given domain. As in the previous section, we need only consider the case $l_x \leq l_y$ by symmetry.

Combined with the results of Section 3, Proposition 5.1 completes the proof of the lower bound from Theorem 1.1.

5.1 The Branching Construction

The integral formulation (5.2) requires the construction of a divergence-free velocity field \mathbf{u} and a test function ξ (the latter of which plays the role of temperature in this approach). For the velocity fields, we will use a streamfunction $\psi(x, z)$ whose streamlines are as in Figure 1.1b. That figure can be thought of as consisting of many individual convection roll systems that have been carefully fit together. In the bulk, there are large anisotropic convection rolls at some horizontal length scale l_{bulk} . At the walls, there are much smaller isotropic convection rolls at some other length scale $l_{\text{bl}} \ll l_{\text{bulk}}$. Between the bulk and the walls, streamlines branch and refine through several transition layers, a single one of which is shown in Figure 5.1. As the construction is symmetric about $z = \frac{1}{2}$, we only need describe it for $z \in [\frac{1}{2}, 1]$.

Counting upwards from the bulk, we understand by the j^{th} transition layer that part of the domain where $z \in [z_j, z_{j+1}]$. The points $\{z_j\}_{j=1}^n$ marking the edges of the layers satisfy

$$(5.3) \quad \frac{1}{2} < z_{\text{bulk}} = z_1 < \cdots < z_n = z_{\text{bl}} < 1.$$

At the horizontal slice $\{z = z_j\}$ the velocity components fluctuate at length scale l_j . These decrease monotonically according as

$$(5.4) \quad l_{\text{bulk}} = l_1 > \cdots > l_n = l_{\text{bl}}.$$

In what follows, we think of the parameters $\{z_j\}_{j=1}^n$ and $\{l_j\}_{j=1}^n$ as playing a distinguished role in specifying the branching design.

Given such a streamfunction ψ and its corresponding two-dimensional velocity field

$$\mathbf{u} = (-\partial_z \psi, 0, \partial_x \psi),$$

we must choose a “temperature” field ξ well suited to the minimization (5.2). Recall the design principle (3.9) discussed in Section 3, which requires that

$$\overline{w\xi} \approx \int_{\Omega} w\xi = 1$$

throughout the domain. For our purposes, it will suffice to set

$$(5.5) \quad \xi = w$$

and enforce that $\overline{w^2} \approx 1$. Such considerations significantly constrain the way that streamlines may branch. We note that while (5.5) may not necessarily hold for optimal designs, it greatly simplifies the ensuing analysis. And, as claimed in

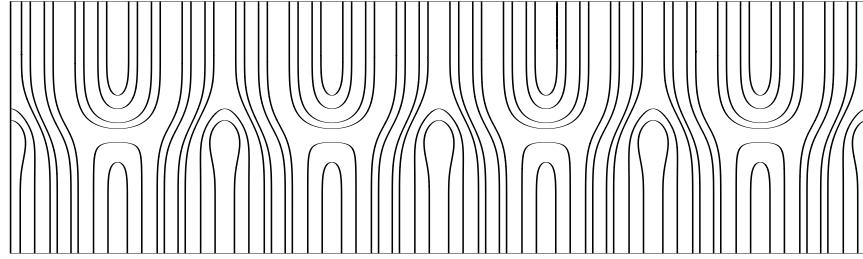


FIGURE 5.1. Streamlines from a period doubling transition layer.

Proposition 5.1 and proved below, such a choice introduces at most a logarithmic error in our estimates of enstrophy-constrained optimal transport.

We are now ready to give the precise functional form of our branching construction. Let points $\{z_j\}_{j=1}^n$ satisfying (5.3) and length scales $\{l_j\}_{j=1}^n$ satisfying

$$\frac{1}{l_j} \in \frac{2\pi}{l_x} \mathbb{N}$$

and (5.4) be given. Let

$$(5.6) \quad \Psi(x) = c_0 \cos x$$

where c_0 is chosen so that $\overline{(\Psi')^2} = 1$. We define $\{\psi_j\}_{j=1}^n$ by

$$\psi_j(x) = l_j \Psi\left(\frac{x}{l_j}\right)$$

and set

$$\psi(x, z) = \psi_1(x) \quad \text{for } z \in \left[\frac{1}{2}, z_1\right],$$

which corresponds to the bulk. In the boundary layer, we set

$$\psi(x, z) = g\left(\frac{z - z_n}{1 - z_n}\right) \psi_n(x) \quad \text{for } z \in [z_n, 1].$$

Here, $g \in C^\infty([0, 1])$ is a fixed cutoff function satisfying the matching conditions $g(0) = 1, g(1) = 0, g'(0) = g'(1) = 0$, as well as the integral condition

$$(5.7) \quad \int_0^1 (g(t))^2 dt = 1.$$

In the j^{th} transition layer we take

$$\psi(x, z) = f\left(\frac{z - z_j}{z_{j+1} - z_j}\right) \psi_j(x) + f\left(\frac{z_{j+1} - z}{z_{j+1} - z_j}\right) \psi_{j+1}(x) \quad \text{for } z \in [z_j, z_{j+1}]$$

where $f \in C^\infty([0, 1])$ is a second fixed cutoff function. We require it to satisfy the Pythagorean condition

$$(5.8) \quad (f(t))^2 + (f(1-t))^2 = 1 \quad \text{for } t \in [0, 1]$$

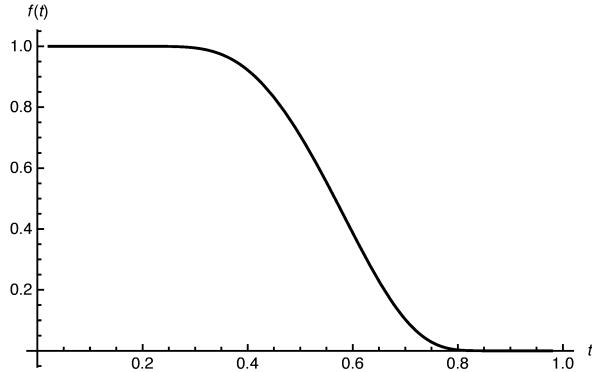


FIGURE 5.2. The cutoff function $f(t)$ used in the branching construction.

as well as the matching conditions $f(0) = 1$, $f(1) = 0$, and $f'(0) = f'(1) = 0$. One might choose, for instance,

$$f(t) = \sqrt{\frac{1}{2} - \frac{1}{2} \tanh\left(\tau \cdot \frac{t - \frac{1}{2}}{t^2(1-t)^2}\right)} \quad \text{for } 0 < t < 1$$

for some $\tau \in (0, \infty)$. Figure 5.2 shows such a cutoff function. This completes our construction of a general class of branching designs; we turn now to select an optimal one for use in (5.2).

5.2 Admissibility

Our first task is to check the admissibility of this construction. The Fourier series of ψ can be written in the form

$$(5.9) \quad \psi = \sum_{j=1}^n \chi_j(z) \psi_j(x),$$

where the amplitude functions $\{\chi_j\}_{j=1}^n$ belong to $H^2(I_z) \cap L^\infty(I_z)$. (Such regularity is guaranteed by the definition of the cutoff functions f and g .) The support of the j^{th} amplitude function satisfies

$$\text{supp } \chi_j \subset [z_{j-1}, z_{j+1}] \cup [1 - z_{j+1}, 1 - z_{j-1}]$$

and the expansion is nearly diagonal in the sense that

$$\chi_j \chi_{j'} \not\equiv 0 \iff |j - j'| \leq 1.$$

Thus,

$$(5.10) \quad \text{supp } \chi_j \chi_{j+1} \subset [z_j, z_{j+1}] \cup [1 - z_{j+1}, 1 - z_j] \quad \forall j,$$

and the same is true for products of derivatives thereof. Note also that by (5.8),

$$(5.11) \quad \sum_{j=1}^n \chi_j^2 = 1, \quad z \in [1 - z_{\text{bl}}, z_{\text{bl}}],$$

and by (5.7),

$$(5.12) \quad \int_{z_{\text{bl}}}^1 \chi_n^2 dz = 1.$$

LEMMA 5.2. *The branching construction defined above is admissible for (5.2).*

PROOF. That the boundary conditions for \mathbf{u} and ξ are met is clear. Here, we check that the net flux constraint $\int_{\Omega} w\xi = 1$ is satisfied. Since

$$\partial_x \psi = \sum_{j=1}^n \chi_j \psi'_j$$

and $\{\psi'_j\}_{j=1}^n$ forms an L^2 -orthonormal set, we can write that

$$\int_{\Omega} w\xi = \int_{\Omega} (\partial_x \psi)^2 = \int_{I_z} \sum_{j=1}^n \chi_j^2 dz.$$

By (5.11), the integrand is equal to 1 for $z \in [1 - z_{\text{bl}}, z_{\text{bl}}]$. Since

$$\int_{z_{\text{bl}}}^1 \chi_n^2 dz = 1 - z_{\text{bl}}$$

by (5.12), we conclude that

$$\int_{\Omega} w\xi = 2 \left(\int_{\frac{1}{2}}^{z_{\text{bl}}} + \int_{z_{\text{bl}}}^1 \right) \sum_{j=1}^n \chi_j^2 dz = 2 \left(z_{\text{bl}} - \frac{1}{2} + 1 - z_{\text{bl}} \right) = 1. \quad \square$$

Next, we record some technical requirements that will greatly simplify the identification of an optimal branching construction. These requirements are compatible with the upper bound from Proposition 5.1, and we have not been able to improve upon the scaling of this result by removing them. First, we require the transition layer thicknesses $\{\delta_k\}_{k=1}^n$, which are defined by

$$\delta_k = z_{k+1} - z_k, \quad 1 \leq k \leq n-1, \quad \text{and} \quad \delta_{\text{bl}} = \delta_n = 1 - z_n,$$

and the horizontal length scales $\{l_k\}_{k=1}^n$ to satisfy the relations

$$(5.13) \quad \delta_1 \gtrsim \cdots \gtrsim \delta_n$$

and

$$(5.14) \quad l_k \lesssim \delta_k, \quad 1 \leq k \leq n-1, \quad \text{and} \quad \delta_n \sim l_n.$$

The latter guarantees that a certain anisotropy is present throughout the construction that will simplify, amongst other things, the estimation of the higher-order terms from (5.2). Second, we require that

$$(5.15) \quad \delta_{\text{bulk}} = z_{\text{bulk}} - \frac{1}{2} \sim 1.$$

Note the constants implicit in (5.13)–(5.15) are not allowed to depend on any parameters. Third, we require that the refinement of length scale through each transition layer occur by period doubling, i.e.,

$$(5.16) \quad l_{k+1} = \frac{1}{2}l_k, \quad 1 \leq k \leq n-1.$$

This last requirement will serve to simplify the Fourier analysis involved in estimating the nonlocal advection term.

5.3 Estimating the Efficiency of Branching

In this section we estimate each of the terms from (5.2) for the branching construction. The requirements laid out in Section 5.1 and Section 5.2 are understood to hold. The constants appearing below are only allowed to depend on those implicit in (5.13)–(5.15), and so do not depend on any parameters.

First, we deal with the advection term. By Lemma 3.2,

$$\int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \xi)|^2 = \int_{I_z} |\overline{w\xi} - 1|^2 dz + \mathcal{Q}(\mathbf{u} \cdot \nabla \xi)$$

where $\mathcal{Q} = \sum_{\mathbf{k} \neq \mathbf{0}} \mathcal{Q}_{\mathbf{k}}$ and

$$\mathcal{Q}_{\mathbf{k}}(f) = \int_{I_z \times I_{z'}} G_{\mathbf{k}}(z, z') \hat{f}_{\mathbf{k}}(z) \hat{f}_{\mathbf{k}}^*(z') dz dz'.$$

As the construction is two-dimensional, k_y does not play a role. For ease of reading, we denote $\mathbf{k} = (k_x, 0)$ simply by k in what follows.

LEMMA 5.3. *The branching construction satisfies*

$$\mathcal{Q}(\mathbf{u} \cdot \nabla \xi) = \sum_{i=1}^{n-1} \mathcal{Q}_{k_i^{\text{sum}}}(\mathbf{u} \cdot \nabla \xi) + \mathcal{Q}_{k_i^{\text{diff}}}(\mathbf{u} \cdot \nabla \xi)$$

where

$$k_i^{\text{sum}} = \frac{1}{l_{i+1}} + \frac{1}{l_i} \quad \text{and} \quad k_i^{\text{diff}} = \frac{1}{l_{i+1}} - \frac{1}{l_i} \quad \text{for } 1 \leq i \leq n-1.$$

PROOF. Using the formula for the branching construction given in (5.9),

$$\begin{aligned} J &= \nabla^{\perp} \psi \cdot \nabla \partial_x \psi = \sum_{j,j'} \nabla^{\perp} (\chi_j \psi_j) \cdot \nabla \partial_x (\chi_{j'} \psi_{j'}) \\ &= \sum_{|j-j'| \leq 1} \nabla^{\perp} (\chi_j \psi_j) \cdot \nabla \partial_x (\chi_{j'} \psi_{j'}) = \end{aligned}$$

$$\begin{aligned}
&= \sum_{j=1}^n \nabla^\perp(\chi_j \psi_j) \cdot \nabla \partial_x(\chi_j \psi_j) \\
&\quad + \sum_{j=1}^{n-1} \nabla^\perp(\chi_j \psi_j) \cdot \nabla \partial_x(\chi_{j+1} \psi_{j+1}) + \nabla^\perp(\chi_{j+1} \psi_{j+1}) \cdot \nabla \partial_x(\chi_j \psi_j) \\
&= J_{\text{self}} + J_{\text{nbr}}.
\end{aligned}$$

Given our choice of fundamental streamfunction (5.6), these expressions can be made explicit and we do so now.

The general term in the first sum, J_{self} , satisfies

$$\begin{aligned}
\nabla^\perp(\chi_j \psi_j) \cdot \nabla \partial_x(\chi_j \psi_j) &= (-\chi'_j \psi_j, \chi_j \psi'_j) \cdot (\chi_j \psi''_j, \chi'_j \psi'_j) \\
&= -\chi'_j \chi_j \psi_j \psi''_j + \chi_j \chi'_j \psi'_j \psi'_j = \frac{1}{2} (\chi_j^2)'(z) \cdot \Theta\left(\frac{x}{l_j}\right)
\end{aligned}$$

where $\Theta = (\Psi')^2 - \Psi \Psi''$. Using (5.6), we see that $\Theta = c_0^2(\cos^2 + \sin^2) = c_0^2$ so that

$$(5.17) \quad J_{\text{self}} = c_0^2 \sum_{i=1}^n (\chi_i^2)'.$$

In particular, we find that J_{self} is constant in the periodic variable x , so that the Fourier coefficient $[J_{\text{self}}]_k^\wedge$ vanishes identically except for when $k = 0$. For $k = 0$, note that

$$[J_{\text{self}}]_0^\wedge = J_{\text{self}} = \bar{J} = \frac{1}{2} \frac{d}{dz} \overline{(\partial_x \psi)^2}.$$

Continuing, we see that the general term in the second sum, J_{nbr} , satisfies

$$\begin{aligned}
&\nabla^\perp(\chi_j \psi_j) \cdot \nabla \partial_x(\chi_{j+1} \psi_{j+1}) + \nabla^\perp(\chi_{j+1} \psi_{j+1}) \cdot \nabla \partial_x(\chi_j \psi_j) \\
&= -(\chi'_j \psi_j, \chi_j \psi'_j) \cdot (\chi_{j+1} \psi''_{j+1}, \chi'_{j+1} \psi'_{j+1}) \\
&\quad - (\chi'_{j+1} \psi_{j+1}, \chi_{j+1} \psi'_{j+1}) \cdot (\chi_j \psi''_j, \chi'_j \psi'_j) \\
&= -\chi'_j \psi_j \chi_{j+1} \psi''_{j+1} + \chi_j \psi'_j \chi'_{j+1} \psi'_{j+1} - \chi'_{j+1} \psi_{j+1} \chi_j \psi''_j \\
&\quad + \chi_{j+1} \psi'_{j+1} \chi'_j \psi'_j \\
&= \chi_j \chi'_{j+1} \Theta_{j,j+1} + \chi'_j \chi_{j+1} \Theta_{j+1,j}
\end{aligned}$$

where

$$\Theta_{j,j+1} = \psi'_j \psi'_{j+1} - \psi''_j \psi_{j+1}, \quad \Theta_{j+1,j} = \psi'_{j+1} \psi'_j - \psi''_{j+1} \psi_j.$$

Given (5.6), we find that

$$\begin{aligned}
\Theta_{j,j+1} &= \sin\left(\frac{x}{l_j}\right) \sin\left(\frac{x}{l_{j+1}}\right) + \frac{l_{j+1}}{l_j} \cos\left(\frac{x}{l_j}\right) \cos\left(\frac{x}{l_{j+1}}\right). \\
\Theta_{j+1,j} &= \sin\left(\frac{x}{l_j}\right) \sin\left(\frac{x}{l_{j+1}}\right) + \frac{l_j}{l_{j+1}} \cos\left(\frac{x}{l_j}\right) \cos\left(\frac{x}{l_{j+1}}\right).
\end{aligned}$$

Applying standard trigonometric identities,

$$\begin{aligned}\Theta_{j,j+1} &= \frac{1}{2} \left(1 + \frac{l_{j+1}}{l_j} \right) \cos(k_j^{\text{diff}} x) + \frac{1}{2} \left(\frac{l_{j+1}}{l_j} - 1 \right) \cos(k_j^{\text{sum}} x), \\ \Theta_{j+1,j} &= \frac{1}{2} \left(1 + \frac{l_j}{l_{j+1}} \right) \cos(k_j^{\text{diff}} x) + \frac{1}{2} \left(\frac{l_j}{l_{j+1}} - 1 \right) \cos(k_j^{\text{sum}} x),\end{aligned}$$

where k_j^{diff} and k_j^{sum} are as in the statement of the result. In sum,

$$(5.18) \quad J_{\text{nbr}} = \sum_{j=1}^n \chi_j \chi'_{j+1} \Theta_{j,j+1} + \sum_{j=1}^n \chi'_j \chi_{j+1} \Theta_{j+1,j}.$$

From (5.17), (5.18), and the decomposition $J = J_{\text{self}} + J_{\text{nbr}}$, it is clear which wavenumbers are present in \mathcal{Q} . We see that \hat{J}_k is not identically 0 if and only if

$$k \in \{0\} \cup \{k_j^{\text{diff}}: 1 \leq j \leq n-1\} \cup \{k_j^{\text{sum}}: 1 \leq j \leq n-1\}.$$

Since $l_{j+1} \neq l_j$, we see that

$$0 \notin \{k_j^{\text{diff}}: 1 \leq j \leq n-1\} \cup \{k_j^{\text{sum}}: 1 \leq j \leq n-1\}.$$

For general choices of length scales $\{l_j\}$ these two sets of wavenumbers may intersect; however, given our special choices of length scales in (5.16), we find that

$$(5.19) \quad \{k_j^{\text{diff}}: 1 \leq j \leq n-1\} \cap \{k_j^{\text{sum}}: 1 \leq j \leq n-1\} = \emptyset.$$

(If $k_i^{\text{sum}} = k_j^{\text{diff}}$ then $2^{i+1} + 2^i = 2^{j+1} - 2^j$ from which the contradiction $3 \cdot 2^i = 2^j$ follows.) Therefore,

$$\mathcal{Q}(\mathbf{u} \cdot \nabla \xi) = \sum_{k \neq 0} \mathcal{Q}_k(\mathbf{u} \cdot \nabla \xi) = \sum_{j=1}^{n-1} \mathcal{Q}_{k_j^{\text{sum}}}(\mathbf{u} \cdot \nabla \xi) + \mathcal{Q}_{k_j^{\text{diff}}}(\mathbf{u} \cdot \nabla \xi). \quad \square$$

Now we estimate each of the nonzero contributions to the advection term picked out by Lemma 5.3.

LEMMA 5.4. *The branching construction satisfies*

$$\int_{I_z} |\overline{w\xi} - 1|^2 dz \lesssim \delta_n$$

and

$$\mathcal{Q}_{k_j^{\text{sum}}}(\mathbf{u} \cdot \nabla \xi) \vee \mathcal{Q}_{k_j^{\text{diff}}}(\mathbf{u} \cdot \nabla \xi) \lesssim \frac{l_j^2}{\delta_j}, \quad 1 \leq j \leq n-1.$$

PROOF. We begin with the $k = 0$ term. Since

$$\overline{w\xi} = \sum_{j=1}^n \chi_j^2$$

we find that

$$\int_{I_z} |\overline{w\xi} - 1|^2 dz \lesssim \delta_n.$$

Next we wish to estimate $\mathcal{Q}_{k_j^{\text{sum}}}$ and $\mathcal{Q}_{k_j^{\text{diff}}}$. Recall that

$$\begin{aligned} \mathcal{Q}_{k_j^{\text{sum}}} &= \int_{I_z \times I'_z} G_{k_j^{\text{sum}}}(z, z') \hat{J}_{k_j^{\text{sum}}}(z) \hat{J}_{k_j^{\text{sum}}}^*(z') dz dz', \\ \mathcal{Q}_{k_j^{\text{diff}}} &= \int_{I_z \times I'_z} G_{k_j^{\text{diff}}}(z, z') \hat{J}_{k_j^{\text{diff}}}(z) \hat{J}_{k_j^{\text{diff}}}^*(z') dz dz'. \end{aligned}$$

By (5.17) and (5.18),

$$\hat{J}_k = [J_{\text{nbr}}]_k^\wedge = \sum_{j=1}^n \chi_j \chi'_{j+1} \widehat{\Theta_{j,j+1}}(k) + \chi'_j \chi_{j+1} \widehat{\Theta_{j+1,j}}(k)$$

for $k \neq 0$. It follows from (5.19) that

$$\begin{aligned} \hat{J}_{k_j^{\text{sum}}} &= \frac{1}{2} \left(\frac{l_{j+1}}{l_j} - 1 \right) \chi_j \chi'_{j+1} + \frac{1}{2} \left(\frac{l_j}{l_{j+1}} - 1 \right) \chi'_j \chi_{j+1}, \\ \hat{J}_{k_j^{\text{diff}}} &= \frac{1}{2} \left(1 + \frac{l_{j+1}}{l_j} \right) \chi_j \chi'_{j+1} + \frac{1}{2} \left(1 + \frac{l_j}{l_{j+1}} \right) \chi'_j \chi_{j+1}. \end{aligned}$$

In particular,

$$\begin{aligned} \text{supp } \hat{J}_{k_j^{\text{sum}}} \cup \text{supp } \hat{J}_{k_j^{\text{sum}}} &\subset \text{supp } \chi_j \chi'_{j+1} \cup \text{supp } \chi'_j \chi_{j+1} \\ &\subset [z_j, z_{j+1}] \cup [1 - z_{j+1}, 1 - z_j] = I_j \end{aligned}$$

by (5.10). Thus,

$$\begin{aligned} \mathcal{Q}_{k_j^{\text{sum}}} &= \int_{I_j \times I_j} G_{k_j^{\text{sum}}}(z, z') \hat{J}_{k_j^{\text{sum}}}(z) \hat{J}_{k_j^{\text{sum}}}^*(z') dz dz', \\ \mathcal{Q}_{k_j^{\text{diff}}} &= \int_{I_j \times I_j} G_{k_j^{\text{diff}}}(z, z') \hat{J}_{k_j^{\text{diff}}}(z) \hat{J}_{k_j^{\text{diff}}}^*(z') dz dz'. \end{aligned}$$

Now we estimate these quadratic forms. By Hölder's inequality,

$$\mathcal{Q}_{k_j^{\text{sum}}} \leq \|G_{k_j^{\text{sum}}}\|_{L^1(I_j \times I_j)} \|\hat{J}_{k_j^{\text{sum}}}\|_{L^\infty}^2, \quad \mathcal{Q}_{k_j^{\text{diff}}} \leq \|G_{k_j^{\text{diff}}}\|_{L^1(I_j \times I_j)} \|\hat{J}_{k_j^{\text{diff}}}\|_{L^\infty}^2.$$

Observe that

$$\|\chi_j \chi'_{j+1}\|_{L^\infty} \vee \|\chi'_j \chi_{j+1}\|_{L^\infty} \lesssim \frac{1}{\delta_j}$$

as a result of (5.13). Combining this with the first part of Lemma 3.3 applied with $A = I_j$, we find that

$$\begin{aligned}\mathcal{Q}_{k_j^{\text{sum}}} &\lesssim \frac{\delta_j}{k_j^{\text{sum}}} \left(\delta_j \wedge \frac{1}{k_j^{\text{sum}}} \right) \cdot \left(\frac{l_j}{l_{j+1}} \frac{1}{\delta_j} \right)^2, \\ \mathcal{Q}_{k_j^{\text{diff}}} &\lesssim \frac{\delta_j}{k_j^{\text{diff}}} \left(\delta_j \wedge \frac{1}{k_j^{\text{diff}}} \right) \cdot \left(\frac{l_j}{l_{j+1}} \frac{1}{\delta_j} \right)^2.\end{aligned}$$

It follows from (5.16) that

$$\frac{1}{l_j} \lesssim k_j^{\text{sum}} \wedge k_j^{\text{diff}},$$

so we can simplify these estimates to

$$\mathcal{Q}_{k_j^{\text{sum}}} \vee \mathcal{Q}_{k_j^{\text{diff}}} \lesssim \delta_j l_j (\delta_j \wedge l_j) \cdot \left(\frac{l_j}{l_{j+1}} \frac{1}{\delta_j} \right)^2.$$

Using the first part of (5.14) followed by (5.16), we conclude that

$$\mathcal{Q}_{k_j^{\text{sum}}} \vee \mathcal{Q}_{k_j^{\text{diff}}} \lesssim \frac{l_j^4}{\delta_j l_{j+1}^2} \lesssim \frac{l_j^2}{\delta_j}.$$

This completes the proof. \square

We turn to estimate the higher-order terms from (5.2).

LEMMA 5.5. *The branching construction satisfies*

$$\mathbf{f}_\Omega |\nabla \mathbf{u}|^2 \vee \mathbf{f}_\Omega |\nabla \xi|^2 \lesssim \frac{1}{l_1^2} + \sum_{j=1}^{n-1} \frac{\delta_j}{l_{j+1}^2} + \frac{l_n^2}{\delta_n^2}.$$

PROOF. Note that

$$\partial_{xx} \psi = \sum_{j=1}^n \chi_j \psi_j'', \quad \partial_{xz} \psi = \sum_{j=1}^n \chi_j' \psi_j', \quad \text{and} \quad \partial_{zz} \psi = \sum_{j=1}^n \chi_j'' \psi_j.$$

Therefore, by orthogonality,

$$\mathbf{f}_\Omega |\nabla \nabla \psi|^2 = \sum_{j=1}^n \mathbf{f}_\Omega |\chi_j \psi_j''|^2 + 2 \mathbf{f}_\Omega |\chi_j' \psi_j'|^2 + \mathbf{f}_\Omega |\chi_j'' \psi_j|^2.$$

For $j \neq 1$, we see from an application of (5.13) that

$$\mathbf{f}_\Omega |\chi_j \psi_j''|^2 = \|\chi_j\|_{L^2(I_z)}^2 \overline{(\psi_j'')^2} \lesssim \frac{\delta_j + \delta_{j-1}}{l_j^2} \lesssim \frac{\delta_{j-1}}{l_j^2}.$$

For $j = 1$, we have instead that

$$\mathbf{f}_\Omega |\chi_1 \psi_1''|^2 = \|\chi_1\|_{L^2(I_z)}^2 \overline{(\psi_1'')^2} \lesssim \frac{1}{l_1^2}.$$

Similarly, we find that

$$\begin{aligned} \int_{\Omega} |\chi'_j \psi'_j|^2 &= \|\chi'_j\|_{L^2(I_z)}^2 \overline{(\psi'_j)^2} \lesssim \frac{1}{\delta_{j-1}} + \frac{1}{\delta_j} \lesssim \frac{1}{\delta_j}, \\ \int_{\Omega} |\chi''_j \psi_j|^2 &= \|\chi''_j\|_{L^2(I_z)}^2 \overline{(\psi_j)^2} \lesssim \left(\frac{1}{\delta_{j-1}^3} + \frac{1}{\delta_j^3} \right) l_j^2 \lesssim \frac{l_j^2}{\delta_j^3}, \end{aligned}$$

for all j . Therefore,

$$\int_{\Omega} |\nabla \nabla \psi|^2 \lesssim \frac{1}{l_1^2} + \sum_{j=1}^{n-1} \left(\frac{\delta_j}{l_{j+1}^2} + \frac{1}{\delta_j} + \frac{l_j^2}{\delta_j^3} \right) + \frac{1}{\delta_n} + \frac{l_n^2}{\delta_n^3}.$$

By (5.14),

$$\frac{\delta_j}{l_{j+1}^2} \gtrsim \frac{1}{\delta_j} \vee \frac{l_j^2}{\delta_j^3} \quad \text{and} \quad \frac{l_n^2}{\delta_n^3} \gtrsim \frac{1}{\delta_n}.$$

The result follows. \square

We now assemble the previous estimates. Let

$$E(\varepsilon; \{z_k\}, \{l_k\}) = \int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \xi)|^2 + \varepsilon \int_{\Omega} |\nabla \mathbf{u}|^2 \cdot \int_{\Omega} |\nabla \xi|^2$$

where (\mathbf{u}, ξ) are constructed from $\{z_k\}_{k=1}^n$ and $\{l_k\}_{k=1}^n$ as described in Section 5.1. It will be convenient in what follows to think of estimating E in terms of some smoothly interpolated version of these parameters.

COROLLARY 5.6. *Let $\ell(z)$ be any smooth, monotonic function defined on $[z_{\text{bulk}}, z_{\text{bl}}]$ that satisfies*

$$\ell(z_k) = l_k, \quad 1 \leq k \leq n.$$

Then, the branching construction corresponding to $\{z_k\}_{k=1}^n$ and $\{l_k\}_{k=1}^n$ satisfies

$$E(\varepsilon; \{z_k\}, \{l_k\}) \lesssim l_{\text{bl}} + \int_{z_{\text{bulk}}}^{z_{\text{bl}}} (\ell'(z))^2 dz + \varepsilon \left(\frac{1}{l_{\text{bulk}}^2} + \int_{z_{\text{bulk}}}^{z_{\text{bl}}} \frac{1}{(\ell(z))^2} dz + \frac{1}{l_{\text{bl}}} \right)^2.$$

PROOF. Collecting the results above and using (5.16), we conclude that

$$\int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \xi)|^2 \lesssim l_{\text{bl}} + \sum_{j=1}^{n-1} \frac{l_j^2}{\delta_j} \sim l_{\text{bl}} + \sum_{j=1}^{n-1} \left| \frac{l_{j+1} - l_j}{\delta_j} \right|^2 \delta_j$$

and

$$\int_{\Omega} |\nabla \mathbf{u}|^2 \vee \int_{\Omega} |\nabla \xi|^2 \lesssim \frac{1}{l_1^2} + \sum_{j=1}^{n-1} \frac{\delta_j}{l_{j+1}^2} + \frac{l_n^2}{\delta_n^3} \sim \frac{\delta_{\text{bulk}}}{l_{\text{bulk}}^2} + \sum_{j=1}^{n-1} \frac{1}{l_j^2} \delta_j + \frac{1}{l_{\text{bl}}}.$$

By Jensen's inequality and the definition of $\ell(z)$,

$$\left| \frac{l_{j+1} - l_j}{\delta_j} \right|^2 = \left| \int_{z_j}^{z_{j+1}} \ell'(z) dz \right|^2 \leq \int_{z_j}^{z_{j+1}} |\ell'|^2 dz.$$

so that

$$\int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \xi)|^2 \lesssim l_{\text{bl}} + \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} |\ell'|^2 dz = l_{\text{bl}} + \int_{z_{\text{bulk}}}^{z_{\text{bl}}} (\ell')^2 dz.$$

Also, as $l_{j+1} \sim l_j$ by (5.16),

$$\frac{1}{l_j^2} \delta_j = \int_{z_j}^{z_{j+1}} \frac{1}{l_j^2} dz \sim \int_{z_j}^{z_{j+1}} \frac{1}{\ell^2} dz.$$

Therefore,

$$\int_{\Omega} |\nabla \mathbf{u}|^2 \vee \int_{\Omega} |\nabla \xi|^2 \lesssim \frac{\delta_{\text{bulk}}}{l_{\text{bulk}}^2} + \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \frac{1}{\ell^2} dz + \frac{1}{l_{\text{bl}}} = \frac{\delta_{\text{bulk}}}{l_{\text{bulk}}^2} + \int_{z_{\text{bulk}}}^{z_{\text{bl}}} \frac{1}{\ell^2} dz + \frac{1}{l_{\text{bl}}}.$$

□

5.4 Proof of Proposition 5.1

The result of the previous analysis is that the branching construction from Section 5.1 satisfies the efficiency estimate

$$E(\varepsilon; \{z_k\}, \{l_k\}) \lesssim l_{\text{bl}} + \int_{z_{\text{bulk}}}^{z_{\text{bl}}} (\ell')^2 dz + \varepsilon \left(\frac{1}{l_{\text{bulk}}^2} + \int_{z_{\text{bulk}}}^{z_{\text{bl}}} \frac{1}{(\ell)^2} dz + \frac{1}{l_{\text{bl}}} \right)^2$$

where $\ell(z)$ is obtained from $\{l_k\}_{k=1}^n$ by smooth and monotonic interpolation. Now to prove Proposition 5.1, we will optimize the right-hand side in the free parameters $\ell(z)$, l_{bulk} , and l_{bl} , and then back out admissible choices of $\{z_k\}_{k=1}^n$ and $\{l_k\}_{k=1}^n$ from the result. To ensure that the requirements from Section 5.1 and Section 5.2 hold, we must carry out this optimization under the constraint that

$$(5.20) \quad 0 \leq \ell'(z) \lesssim 1, \quad z \in [z_{\text{bulk}}, z_{\text{bl}}].$$

That the minimizer of

$$(5.21) \quad \min_{\substack{\ell(z) \\ l(z_{\text{bulk}}) = l_{\text{bulk}} \\ \ell(z_{\text{bl}}) = l_{\text{bl}}}} l_{\text{bl}} + \int_{z_{\text{bulk}}}^{z_{\text{bl}}} (\ell')^2 dz + \varepsilon \left(\frac{1}{l_{\text{bulk}}^2} + \int_{z_{\text{bulk}}}^{z_{\text{bl}}} \frac{1}{(\ell)^2} dz + \frac{1}{l_{\text{bl}}} \right)^2$$

satisfies (5.20) will be verified later on.

First, let us determine the optimal form of $\ell(z)$. We consider that $\varepsilon \ll 1$ throughout this preliminary discussion, which should serve to motivate the choices made in the formal proof that follows. Consider the contributions to (5.21) coming

from the transition layers where $z \in [z_{\text{bulk}}, z_{\text{bl}}]$. We can identify the scaling of their minimum value by balancing the corresponding integrands. This yields

$$(5.22) \quad \ell'(z) \sim \varepsilon^{1/2} \left(\int_{z_{\text{bulk}}}^{z_{\text{bl}}} \frac{1}{\ell^2} dz \right)^{1/2} \frac{1}{\ell(z)}.$$

It is natural to impose the boundary condition $\ell(1) = 0$ to determine ℓ . We find that

$$\ell(z) \sim c(\varepsilon)(1-z)^{1/2}$$

where $c(\varepsilon)$ must be determined by substitution into (5.22). Thus,

$$c(\varepsilon) \sim \varepsilon^{1/6} \left(\int_{z_{\text{bulk}}}^{z_{\text{bl}}} \frac{1}{1-z} dz \right)^{1/6} = \varepsilon^{1/6} \log^{1/6} \left(\frac{1-z_{\text{bulk}}}{1-z_{\text{bl}}} \right).$$

Since

$$l_{\text{bl}} \sim c(\varepsilon)(1-z_{\text{bl}})^{1/2} \quad \text{and} \quad l_{\text{bulk}} \sim c(\varepsilon)(1-z_{\text{bulk}})^{1/2}$$

we conclude that

$$\frac{1-z_{\text{bulk}}}{1-z_{\text{bl}}} \sim \frac{l_{\text{bulk}}}{l_{\text{bl}}}.$$

Anticipating that $l_{\text{bl}} \ll l_{\text{bulk}}$ for $\varepsilon \ll 1$, we conclude that the optimal form of the smooth length scale function $\ell(z)$ is given by

$$(5.23) \quad \ell(z) \sim \varepsilon^{1/6} \log^{1/6} \left(\frac{l_{\text{bulk}}}{l_{\text{bl}}} \right) (1-z)^{1/2}, \quad z \in [z_{\text{bulk}}, z_{\text{bl}}].$$

Such an ℓ yields the estimates

$$\begin{aligned} \int_{z_{\text{bulk}}}^{z_{\text{bl}}} (\ell')^2 dz &\sim \varepsilon \left(\int_{z_{\text{bulk}}}^{z_{\text{bl}}} \frac{1}{\ell^2} dz \right)^2 \sim \varepsilon^{1/3} \log^{1/3} \left(\frac{l_{\text{bulk}}}{l_{\text{bl}}} \right) \int_{z_{\text{bulk}}}^{z_{\text{bl}}} \frac{1}{1-z} dz \\ &= \varepsilon^{1/3} \log^{1/3} \left(\frac{l_{\text{bulk}}}{l_{\text{bl}}} \right) \log \left(\frac{1-z_{\text{bulk}}}{1-z_{\text{bl}}} \right) \sim \varepsilon^{1/3} \log^{4/3} \left(\frac{l_{\text{bulk}}}{l_{\text{bl}}} \right) \end{aligned}$$

for $\varepsilon \ll 1$.

Next, we determine the optimal choices for l_{bulk} and l_{bl} in this asymptotic regime. Plugging (5.23) back into (5.21) yields the resulting minimization

$$\min_{\substack{l(z_{\text{bulk}})=l_{\text{bulk}} \\ \ell(z_{\text{bl}})=l_{\text{bl}}}} l_{\text{bl}} + \varepsilon^{1/3} \log^{4/3} \left(\frac{l_{\text{bulk}}}{l_{\text{bl}}} \right) + \varepsilon \left(\frac{1}{l_{\text{bulk}}^2} + \frac{1}{l_{\text{bl}}^2} \right)^2.$$

Critical point tests yield the optimal scalings

$$(5.24) \quad l_{\text{bulk}} \sim \varepsilon^{1/6} \log^{1/6} \frac{1}{\varepsilon} \quad \text{and} \quad l_{\text{bl}} \sim \varepsilon^{1/3} \log^{1/3} \frac{1}{\varepsilon}$$

for $\varepsilon \ll 1$. Note this is consistent with the hypothesis that $l_{\text{bl}} \ll l_{\text{bulk}}$ in this regime. To summarize, the smooth length scale function $\ell(z)$ picked out by our analysis of (5.21) scales as

$$(5.25) \quad \ell(z) \sim \varepsilon^{1/6} \log^{1/6} \frac{1}{\varepsilon} (1-z)^{1/2}, \quad z \in [z_{\text{bulk}}, z_{\text{bl}}],$$

where

$$1 - z_{\text{bulk}} \sim 1 \quad \text{and} \quad 1 - z_{\text{bl}} \sim \varepsilon^{1/3} \log^{1/3} \frac{1}{\varepsilon}.$$

We are now ready to prove the upper bound from Proposition 5.1.

PROOF OF PROPOSITION 5.1. Our plan is to verify the existence of a branching construction, as described in Section 5.1, whose parameters $\{z_k\}_{k=1}^n$ and $\{l_k\}_{k=1}^n$ are consistent with the optimal smooth length scale function $\ell(z)$ from (5.25). Once we verify the requirements of Section 5.1 and Section 5.2 hold, the desired bound $E(\varepsilon; \{z_k\}, \{l_k\}) \lesssim \varepsilon^{1/3} \log^{4/3} \frac{1}{\varepsilon}$ follows as above. For the reader's convenience, we recall the requirements that must be checked: these are (5.3), (5.4), and (5.13)–(5.16).

We start by defining

$$(5.26) \quad \ell(z) = \varepsilon^{1/6} \log^{1/6} \frac{1}{\varepsilon} (1 - z)^{1/2} \quad z \in \left[\frac{1}{2}, 1 \right]$$

in obvious analogy with (5.25). To choose the horizontal length scales $\{l_k\}_{k=1}^n$, we set

$$l_{\text{bulk}} = \frac{l_x}{2\pi} \frac{1}{k_{\text{bulk}}}$$

where $k_{\text{bulk}} \in \mathbb{N}$ satisfies

$$(5.27) \quad k_{\text{bulk}} - 1 < \frac{l_x}{\pi} \frac{1}{\varepsilon^{1/6} \log^{1/6} \frac{1}{\varepsilon}} \leq k_{\text{bulk}},$$

and take

$$l_k = \frac{l_{\text{bulk}}}{2^{k-1}}, \quad k = 1, \dots, n.$$

Note to ensure $k_{\text{bulk}} \geq 1$ we must require that $\varepsilon^{1/6} \log^{1/6} \frac{1}{\varepsilon} \lesssim l_x$. This condition is given in the statement of Proposition 5.1. Note also that (5.4) and (5.16) hold.

Now as (5.26) is strictly decreasing, we may define the points $\{z_k\}_{k=1}^n$ by

$$(5.28) \quad \ell(z_k) = l_k, \quad k = 1, \dots, n.$$

This gives

$$z_k = 1 - c_1 \frac{1}{2^{2(k-1)}}, \quad k = 1, \dots, n,$$

where

$$c_1 = \left(\frac{l_{\text{bulk}}}{\varepsilon^{1/6} \log^{1/6} \frac{1}{\varepsilon}} \right)^2 = \frac{1}{4\pi^2} \frac{1}{\varepsilon^{1/3} \log^{1/3} \frac{1}{\varepsilon}} \frac{l_x^2}{k_{\text{bulk}}^2}.$$

By (5.27), $c_1 \leq \frac{1}{4}$ so that $z_1 = 1 - c_1 \geq \frac{3}{4}$ as required by (5.15). Note (5.3) and (5.13) are satisfied as well.

Finally, we fix $n \in \mathbb{N}$ by enforcing (5.24), which states here that

$$\frac{l_{\text{bulk}}}{2^{n-1}} \sim \varepsilon^{1/3} \log^{1/3} \frac{1}{\varepsilon}.$$

To achieve this, let us define $n \in \mathbb{N}$ via the inequalities

$$n - 1 < \log_2 \left(\frac{2\pi}{\varepsilon^{1/3} \log^{1/3} \frac{1}{\varepsilon} k_{\text{bulk}}} \right) \leq n.$$

Having chosen $\{z_k\}_{k=1}^n$ and $\{l_k\}_{k=1}^n$, we may invoke the definitions from Section 5.1 to produce a branching construction (\mathbf{u}, ξ) . Note we have checked each requirement from Sections 5.1 and 5.2 except for (5.14). That $l_n \sim \delta_n$ follows from (5.24) and (5.25).

Now we show that $l_k \lesssim \delta_k$ for all k . Since $\delta_k = z_{k+1} - z_k$ and $l_k \sim |l_{k+1} - l_k|$, this requires showing that

$$1 \lesssim \left| \frac{z_{k+1} - z_k}{l_{k+1} - l_k} \right|$$

for all k . Noting $z'(\ell) < 0$, we only need to show that

$$1 \lesssim |z'(\ell)|, \quad \ell \in [l_{\text{bl}}, l_{\text{bulk}}].$$

Differentiating (5.26) implicitly, we find that

$$|z'(\ell)| = 2 \frac{(1-z)^{1/2}}{\varepsilon^{1/6} \log^{1/6} \frac{1}{\varepsilon}} \gtrsim \frac{\delta_{\text{bl}}^{1/2}}{\varepsilon^{1/6} \log^{1/6} \frac{1}{\varepsilon}} \sim \frac{l_{\text{bl}}^{1/2}}{\varepsilon^{1/6} \log^{1/6} \frac{1}{\varepsilon}} \sim 1$$

as required.

In sum, we have produced a branching construction (\mathbf{u}, ξ) consistent with the requirements of Sections 5.1 and 5.2 whose parameters $\{z_k\}_{k=1}^n$ and $\{l_k\}_{k=1}^n$ interpolate the desired smooth length scale function $\ell(z)$ from (5.25). The estimates proved in Section 5.3 apply, and we may immediately conclude from Corollary 5.6 and the discussion surrounding (5.24) and (5.25) that

$$\begin{aligned} E(\varepsilon; \{z_k\}, \{l_k\}) &= \int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \xi)|^2 + \varepsilon \int_{\Omega} |\nabla \mathbf{u}|^2 \cdot \int_{\Omega} |\nabla \xi|^2 \\ &\lesssim l_{\text{bl}} + \int_{z_{\text{bulk}}}^{z_{\text{bl}}} (\ell')^2 dz + \varepsilon \left(\frac{1}{l_{\text{bulk}}^2} + \int_{z_{\text{bulk}}}^{z_{\text{bl}}} \frac{1}{(\ell')^2} dz + \frac{1}{l_{\text{bl}}} \right)^2 \\ &\lesssim \varepsilon^{1/3} \log^{1/3} \frac{1}{\varepsilon} + \varepsilon^{1/3} \log^{4/3} \left(\frac{\varepsilon^{1/6} \log^{1/6} \frac{1}{\varepsilon}}{\varepsilon^{1/3} \log^{1/3} \frac{1}{\varepsilon}} \right) + \varepsilon \left(\frac{2}{\varepsilon^{1/3} \log^{1/3} \frac{1}{\varepsilon}} \right)^2 \\ &\lesssim \varepsilon^{1/3} \log^{1/3} \frac{1}{\varepsilon} + \varepsilon^{1/3} \log^{4/3} \frac{1}{\varepsilon} + \frac{\varepsilon^{1/3}}{\log^{2/3} \frac{1}{\varepsilon}} \lesssim \varepsilon^{1/3} \log^{4/3} \frac{1}{\varepsilon} \end{aligned}$$

for $\varepsilon \lesssim 1$. Thus, Proposition 5.1 is proved. \square

6 Implications for the Analysis of Turbulent Heat Transport

There is a long history, originating in the works of Malkus [24] and Howard [21], of variational methods for the analysis of turbulent heat transport, the primary focus of which has been on absolute or a priori upper bounds. Consider

the usual setup of Rayleigh-Bénard convection (RBC), wherein an incompressible fluid layer is heated from below and cooled from above, and is subjected to a constant downwards-pointing gravitational force. The temperature field $T(\mathbf{x}, t)$ undergoes transport by means of advection-diffusion,

$$(6.1) \quad \partial_t T + \mathbf{u} \cdot \nabla T = \Delta T.$$

The advecting velocity $\mathbf{u}(\mathbf{x}, t)$ is coupled back to temperature field T through a suitable momentum equation. This could be, for instance, Darcy's law as it is for convection in a fluid-saturated porous layer. Here, we are concerned with convection in a fluid layer for which, in the Boussinesq approximation, (6.1) is supplemented with the buoyancy forced incompressible Navier-Stokes equations

$$(6.2) \quad \partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p = \text{Pr} \Delta \mathbf{u} + \text{Pr} \text{Ra} \hat{\mathbf{k}} T$$

and

$$(6.3) \quad \operatorname{div} \mathbf{u} = 0.$$

The two nondimensional parameters are the Prandtl number Pr , the ratio of the fluid's kinematic viscosity to its thermal diffusivity, and the Rayleigh number Ra , a ratio of the intensities of driving to damping forces that is proportional here to the bulk buoyancy force across the layer. Altogether, (6.1)–(6.3) constitute the equations of Rayleigh-Bénard convection in a fluid layer [36]. For boundary conditions we continue to assume that the temperature field is imposed at the top and bottom of the layer by

$$T|_{z=1} = 0 \quad \text{and} \quad T|_{z=0} = 1,$$

while the velocity field is taken to satisfy either the no-slip boundary conditions

$$\mathbf{u}|_{\partial\Omega} = \mathbf{0}$$

or the stress-free boundary conditions

$$w|_{\partial\Omega} = 0 \quad \text{and} \quad \partial_z u|_{\partial\Omega} = \partial_z v|_{\partial\Omega} = 0.$$

All fields are assumed to be periodic in the xy -plane.

The rate of heat transport in RBC can be measured by the Nusselt number Nu , which evidently depends on Pr and Ra in some unknown and complicated way. (It can also depend on the initial data, as well as on the aspect ratios of the fluid layer.) Determining this relationship and/or establishing absolute bounds on it continues to be the subject of numerous works across the physical and mathematical literatures. To date, the best known upper bound holding uniformly in Pr and for no-slip velocity boundary conditions states that

$$(6.4) \quad \text{Nu} \lesssim \text{Ra}^{1/2}$$

for $\text{Ra} \gg 1$ [13, 21, 37]. This bound also holds for stress-free velocity boundary conditions in the three-dimensional layer $\Omega = \mathbb{T}_{xy}^2 \times I_z$, but more is known in the two-dimensional case where $\Omega = \mathbb{T}_x \times I_z$: in two dimensions with stress-free boundary conditions, one has that $\text{Nu} \lesssim \text{Ra}^{5/12}$ uniformly in Pr for $\text{Ra} \gg 1$ [44].

(In the formal limit where $\text{Pr} = \infty$ and (6.2) is replaced with Stoke's equation, the situation is quite different [14, 16, 29, 33, 45].) There is little to no evidence, however, that any of these finite Pr bounds are in fact sharp, i.e., that there exist solutions of the equations of motion (6.1)–(6.3) satisfying $\text{Nu} \sim \text{Ra}^{1/2}$ as $\text{Ra} \rightarrow \infty$ (or $\text{Nu} \sim \text{Ra}^{5/12}$ for stress-free boundaries in two dimensions).

In light of all this, we note that the main fluid dynamical contribution of this paper is a proof that when the momentum equation (6.2) is replaced by the enstrophy-constraint

$$(6.5) \quad \langle |\nabla \mathbf{u}|^2 \rangle = \text{Ra}(\text{Nu} - 1),$$

which it implies, the upper bound $\text{Nu} \lesssim \text{Ra}^{1/2}$ becomes asymptotically sharp up to logarithmic corrections. That is, for all large enough Ra there exist velocity and temperature fields satisfying (6.1), (6.3), and (6.5) along with the requisite boundary conditions such that

$$(6.6) \quad \frac{\text{Ra}^{1/2}}{\log^2 \text{Ra}} \lesssim \text{Nu} \lesssim \text{Ra}^{1/2}.$$

This follows from Theorem 1.2 upon taking $\text{Pe}^2 = \text{Ra}(\text{Nu} - 1)$. Therefore, either the well-known bound (6.4) on RBC is asymptotically sharp as $\text{Ra} \rightarrow \infty$ and Pr is fixed, or details from the momentum equation (6.2) beyond the balance (6.5) are essential for determining the scaling law of maximal turbulent heat transport.

The remainder of this section places our analysis of wall-to-wall optimal transport into its proper fluid dynamical context. To keep the discussion at a reasonable length, we do not attempt to summarize the vast literature on the subject but instead focus on two of the most well-known methods for proving a priori bounds on transport: the variational approach of Howard, and the background method of Constantin-Doering. Our plan is to recall just enough about these methods to allow for comparison with the techniques developed in this paper. For Howard's approach, see Section 6.1, while for the background method, see Section 6.2. Section 6.3 concerns the role of the momentum equation.

Before we proceed, let us mention the existence of the recently developed “auxiliary functional” method for producing bounds on time-averaged quantities [7]. While the background method may ultimately be derived by a particular choice of auxiliary functional—the same is apparently true [6] for the recently proposed method of Seis [37]—it is not yet clear if there exists any auxiliary functional that yields an improvement to scaling beyond $\text{Nu} \lesssim \text{Ra}^{1/2}$. Although for ordinary differential equations the auxiliary functional method always yields sharp bounds on long-time averages [41], it remains to be seen if such a situation holds for general PDEs.

6.1 On the Variational Approach of Howard

Howard's Variational Problem

If RBC is to be taken as a predictive model for turbulent convection, one naturally asks: which of its solutions are realizable in experiments? Setting aside dynamical stability as a possible selection principle, Malkus introduced in [24] the idea that perhaps amongst all possible solutions of the equations of motion, those that are realized maximize their heat transport overall. An operational approach to establishing upper bounds inspired by Malkus' idea is to search for a larger admissible set of velocity and temperature fields, which contains all solutions of RBC, amongst which the maximal transport can analytically be determined. This is Howard's variational approach.

Following Howard [21], we observe that if \mathbf{u} and T arise in RBC, they must satisfy two identities known as the “power integrals.” To derive the first of these, dot the momentum equation (6.2) into \mathbf{u} , integrate by parts, and average in space and time. Changing variables by $\theta = T - (1 - z)$ yields the first of Howard's identities

$$(6.7) \quad \text{Ra}\langle w\theta \rangle = \langle |\nabla \mathbf{u}|^2 \rangle.$$

(Note that this is simply a restatement of (6.5) from above.) A similar manipulation involving the temperature equation (6.1) yields the second identity

$$(6.8) \quad \langle w\theta \rangle + \langle w\theta \rangle^2 - \langle |\overline{w\theta}|^2 \rangle = \langle |\nabla \theta|^2 \rangle.$$

Consider now the problem of maximizing Nu amongst all divergence-free vector fields \mathbf{u} and scalar fields θ that vanish at the walls and furthermore satisfy (6.7) and (6.8). Since the equations of motion of RBC imply these constraints, the resulting maximum sets an upper bound on Nu for RBC.

Setting aside matters of statistical stationarity [21], one can give an equivalent formulation of the variational problem described above that makes it tractable for analysis. Under certain further assumptions on the solutions of RBC (the “requirements of homogeneity” from [21]), Howard deduced that the minimization

$$(6.9) \quad \min_{\substack{\mathbf{u}(\mathbf{x}), \theta(\mathbf{x}) \\ \int_{\Omega} w\theta = 1 \\ \mathbf{u}|_{\partial\Omega} = \mathbf{0}, \theta|_{\partial\Omega} = 0}} \int_{\Omega} |\overline{w\theta} - 1|^2 + \varepsilon \int_{\Omega} |\nabla \mathbf{u}|^2 \int_{\Omega} |\nabla \theta|^2$$

is equivalent to the maximization $\sup \text{Nu}$ described above, and that its optimal value can be used to produce an a priori bound on RBC (the algebraic manipulations in the proof of this are like those performed in Section 3 in the derivation of the integral formulation of steady wall-to-wall optimal transport).

The minimization (6.9) is known as Howard's problem. It bears striking resemblance to our integral formulation of steady enstrophy-constrained wall-to-wall

transport

$$(6.10) \quad \min_{\substack{\mathbf{u}(\mathbf{x}), \xi(\mathbf{x}) \\ \int_{\Omega} w\xi = 1 \\ \mathbf{u}|_{\partial\Omega} = \mathbf{0}, \xi|_{\partial\Omega} = 0}} \int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u}\xi)|^2 + \varepsilon \int_{\Omega} |\nabla \mathbf{u}|^2 \int_{\Omega} |\nabla \xi|^2,$$

obtained in Section 3. To find the relationship between (6.9) and (6.10), we apply Lemma 3.2 along with the net flux constraint $\int_{\Omega} w\xi = 1$ and decompose the advection term as

$$\int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u}\xi)|^2 = \int_{\Omega} |\overline{w\xi} - 1|^2 + \mathcal{Q}(\operatorname{div}(\mathbf{u}\xi))$$

where \mathcal{Q} is the positive semidefinite quadratic form defined in Lemma 3.2.

This last equation reveals the precise distinction between Howard's problem (6.9) and our integral formulation in (6.10). Because \mathcal{Q} is positive semidefinite, it is evident that the minimum in (6.9) is not smaller than the minimum in (6.10). As a result, Howard's upper bound on heat transport is not lower than ours. Though the improvement in scaling in our approach is limited by (6.6) to at most a logarithmic correction, it remains to be seen whether such a correction holds as an absolute upper bound. We turn now to consider the difference between the optimizers of (6.9) and (6.10).

Busse's Multi- α Technique

As shown by Howard and Busse [5, 21], the optimal value of Howard's problem (6.9) scales as $\varepsilon^{1/3}$ for $\varepsilon \ll 1$. Thus, Howard's approach to bounds on RBC yields $\text{Nu} \lesssim \text{Ra}^{1/2}$ and no better. The a priori lower bound implicit in this result is due to Howard; the upper bound was obtained by Busse as an application of his “multi- α ” technique, which seeks to produce asymptotically valid solutions of the Euler-Lagrange equations of (6.9) involving multiple horizontal wave numbers. Busse's multi- α analysis turns out to share parallels with our construction of branching flows, which we would like to discuss now.

We start by recalling Howard's lower bound:

$$(6.11) \quad \min_{\substack{\mathbf{u}(\mathbf{x}), \theta(\mathbf{x}) \\ \int_{\Omega} w\theta = 1 \\ \mathbf{u}|_{\partial\Omega} = \mathbf{0}, \theta|_{\partial\Omega} = 0}} \int_{\Omega} |\overline{w\theta} - 1|^2 + \varepsilon \int_{\Omega} |\nabla \mathbf{u}|^2 \int_{\Omega} |\nabla \theta|^2 \gtrsim \varepsilon^{1/3}$$

for $\varepsilon \ll 1$. Let (\mathbf{u}, θ) be admissible, which we can take to be smooth. Let $\delta \in (0, \frac{1}{2})$ be such that

$$0 \leq |\overline{w\theta}| \leq \frac{1}{2} \quad \text{for } z \in [0, \delta] \quad \text{and} \quad \overline{w\theta}(\delta) = \frac{1}{2}.$$

(If there does not exist such a δ , then $\int_{\Omega} |\overline{w\theta} - 1|^2 \gtrsim 1 \gg \varepsilon^{2/3}$.) By its definition,

$$\int_{\Omega} |\overline{w\theta} - 1|^2 \geq \frac{1}{|I_z|} \int_0^{\delta} |\overline{w\theta} - 1|^2 \gtrsim \delta.$$

Lemma 2.6 states that

$$\frac{1}{|\mathbb{T}_{xy}^2|} \|\partial_z \theta\|_{L^2(\Omega)} \|\partial_z w\|_{L^2(\Omega)} \gtrsim \frac{|\overline{w\theta}(z)|}{|z \wedge (1-z)|} \quad \forall z.$$

Taking $z = \delta$ and squaring, we conclude that

$$\int_{\Omega} |\nabla \mathbf{u}|^2 \int_{\Omega} |\nabla \theta|^2 \gtrsim \frac{1}{\delta^2}.$$

Therefore, the optimal value in the left-hand side of (6.11) is bounded below by

$$\inf_{\delta \in (0, \frac{1}{2})} \left\{ \delta + \varepsilon \frac{1}{\delta^2} \right\} \sim \varepsilon^{1/3}$$

for $\varepsilon \ll 1$, and (6.11) is proved.

Now we discuss Busse's upper bound: it asserts the existence of admissible pairs $\{(\mathbf{u}_\varepsilon, \theta_\varepsilon)\}$ satisfying

$$(6.12) \quad \int_{\Omega} |\overline{w_\varepsilon \theta_\varepsilon} - 1|^2 + \varepsilon \int_{\Omega} |\nabla \mathbf{u}_\varepsilon|^2 \int_{\Omega} |\nabla \theta_\varepsilon|^2 \lesssim \varepsilon^{1/3}$$

for $\varepsilon \ll 1$. Busse's multi- α technique is analogous to our branching construction from Section 5. Arguing as in that section, we find that our branching construction with length scale $\ell(z)$ satisfies the estimates

$$\int_{\Omega} |\overline{w\theta} - 1|^2 \lesssim l_{\text{bl}} \quad \text{and} \quad \int_{\Omega} |\nabla \mathbf{u}|^2 \int_{\Omega} |\nabla \theta|^2 \lesssim \left(\frac{1}{l_{\text{bulk}}^2} + \int_{z_{\text{bulk}}}^{z_{\text{bl}}} \frac{1}{\ell^2} dz + \frac{1}{l_{\text{bl}}} \right)^2$$

so long as $0 \leq \ell'(z) \lesssim 1$. Since branching is admissible for Howard's problem, we find its optimal value is bounded above by

$$(6.13) \quad \min_{\substack{\ell(z) \\ \ell(z_{\text{bulk}}) = l_{\text{bulk}} \\ \ell(z_{\text{bl}}) = l_{\text{bl}} \\ 0 \leq \ell'(z) \lesssim 1}} l_{\text{bl}} + \varepsilon \left(\frac{1}{l_{\text{bulk}}^2} + \int_{z_{\text{bulk}}}^{z_{\text{bl}}} \frac{1}{\ell^2} dz + \frac{1}{l_{\text{bl}}} \right)^2.$$

Choosing

$$\ell(z) \sim 1 - z, \quad z \in [z_{\text{bulk}}, z_{\text{bl}}], \quad l_{\text{bulk}} \sim 1, \quad \text{and} \quad l_{\text{bl}} \sim \varepsilon^{1/3}$$

yields (6.12). Although Busse's construction is usually described in terms of discrete wavenumbers $\{\alpha_k\}_{k=1}^n$ and points $\{z_k\}_{k=1}^n$, for $\varepsilon \ll 1$ these can be seen to arise from interpolation of the continuous length scale $\ell(z) \sim 1 - z$, similar to the presentation in Section 5.

Coming back to wall-to-wall optimal transport, we can now discuss the difference between the optimizers of Howard's problem (6.9) and our integral formulation in (6.10). As the analysis in Section 5 indicates, adding \mathcal{Q} to Howard's problem (6.9) should change the preferred length scale for branching from Busse's

linear law $\ell \sim 1 - z$ to our square root law $\ell \sim c(\varepsilon)\sqrt{1-z}$. The estimates obtained there show that

$$(6.14) \quad \mathcal{Q} \sim \int_{z_{\text{bulk}}}^{z_{\text{bl}}} (\ell')^2 dz.$$

Thus, the one-dimensional problem (6.13) for selecting the length scale function ℓ turns into (1.12) for wall-to-wall optimal transport. It remains to be seen whether the true optimizers of (6.10) exhibit branching with these preferred length scales. Presumably, developing such fine detailed knowledge of the minimizers would help resolve the question of logarithmic corrections to scaling.

6.2 On the Background Method

Background Method for RBC

In [13], Constantin and one of the authors introduced an alternate method to Howard's for establishing a priori bounds on RBC, which can be applied without any assumptions of statistical stationarity or homogeneity. We recall the argument now, with the goal of connecting it to the symmetrization method from Section 2. We follow the presentation in [15].

Let \mathbf{u} and T arise from RBC and decompose the temperature field into the sum of stationary “background” and fluctuating parts,

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

where $\tau(0) = 1$ and $\tau(1) = 0$. Then,

$$\frac{1}{2} \frac{d}{dt} \left(\int_{\Omega} |\theta|^2 + \frac{1}{\text{PrRa}} \int_{\Omega} |\mathbf{u}|^2 \right) + \frac{1}{2} \int_{\Omega} |\nabla T|^2 = \frac{1}{2} \int_0^1 |\tau'|^2 - H_{\tau}(\mathbf{u}, \theta)$$

where H_{τ} is the quadratic form

$$H_{\tau}(\mathbf{u}, \theta) = \int_{\Omega} \frac{1}{\text{Ra}} |\nabla \mathbf{u}|^2 + \frac{1}{2} |\nabla \theta|^2 + w\theta(\tau' - 1).$$

Provided that $H_{\tau} \geq 0$ for all divergence-free vector fields $\mathbf{u}(\mathbf{x})$ and scalar fields $\theta(\mathbf{x})$ vanishing at $\partial\Omega$, we can drop the last term from the dissipation equation and take a long-time average to find the inequality

$$\langle |\nabla T|^2 \rangle \leq \int_0^1 |\tau'|^2.$$

This proves the following variational bound:

$$(6.15) \quad \text{Nu} \leq \inf_{\substack{\tau(z) \\ \tau(0)=1, \tau(1)=0 \\ H_{\tau} \geq 0}} \int_0^1 |\tau'|^2.$$

Those background fields τ that satisfy $H_{\tau} \geq 0$ are known as *spectrally stable*.

As proved in [15], there exist spectrally stable background fields $\{\tau_\delta\}$ satisfying

$$\int_0^1 |\tau'_\delta|^2 \sim \frac{1}{\delta}$$

for all $\delta \leq \text{Ra}^{-1/2}$. Minimizing the resulting bound $\text{Nu} \lesssim \frac{1}{\delta}$ over this range of δ proves that $\text{Nu} \lesssim \text{Ra}^{1/2}$. There is a remarkable similarity between the background fields τ_δ constructed in [15] and our fields η_δ constructed for the symmetrization method in (2.10). This hints that the upper bounds produced by these two approaches may in fact coincide. We turn to discuss this now.

Background Method for Optimal Transport

As observed in [39], one can obtain a priori bounds on optimal transport via a suitable modification of the background method. Here, our goal is to show that the symmetrization method from Section 2, when properly abstracted and optimized, yields an a priori bound on transport whose value is *exactly the same* as that obtained in [39]. This begs the question of whether better background fields might be constructed to improve upon the scaling $\text{Nu} \lesssim \text{Ra}^{1/2}$ (albeit by at most a logarithmic amount). Numerical evidence points in the opposite direction, as the optimal bounds found in [35] scale $\sim \text{Ra}^{1/2}$. We are not aware of a proof demonstrating this at the present time.

The modified background method from [39] is as follows. Let T solve the advection-diffusion equation (6.1). Performing the background decomposition

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

with $\tau(0) = 1$ and $\tau(1) = 0$ and introducing a Lagrange multiplier $\lambda \in \mathbb{R}$, we find that

$$(6.16) \quad \frac{1}{2} \frac{d}{dt} \int_{\Omega} |\theta|^2 + \frac{1}{2} \int_{\Omega} |\nabla T|^2 = \frac{1}{2} \int_0^1 |\tau'|^2 + \frac{\lambda}{2} \text{Pe}^2 - H_{\tau, \lambda}(\mathbf{u}, \theta)$$

where $H_{\tau, \lambda}$ is the quadratic form

$$H_{\tau, \lambda}(\mathbf{u}, \theta) = \int_{\Omega} \frac{\lambda}{2} |\nabla \mathbf{u}|^2 + \frac{1}{2} |\nabla \theta|^2 + w \theta \tau'.$$

If $H_{\tau, \lambda} \geq 0$ for all divergence-free vector fields $\mathbf{u}(\mathbf{x})$ and scalar fields $\theta(\mathbf{x})$ vanishing at $\partial\Omega$, the dissipation equation (6.16) implies that

$$\langle |\nabla T|^2 \rangle \leq \int_0^1 |\tau'|^2 + \lambda \text{Pe}^2.$$

Thus,

$$(6.17) \quad \text{Nu} \leq \inf_{\substack{\tau(z), \lambda \\ \tau(0)=1, \tau(1)=0 \\ H_{\tau, \lambda} \geq 0}} \left\{ \int_0^1 |\tau'|^2 + \lambda \text{Pe}^2 \right\}$$

In parallel with the background method discussed above, we refer to background fields τ satisfying $H_{\tau,\lambda} \geq 0$ as being *spectrally stable at Lagrange multiplier λ* .

On the other hand, the symmetrization method from Section 2 yields the bound

$$(6.18) \quad \sup_{\substack{\mathbf{u}(\mathbf{x},t) \\ \langle |\nabla \mathbf{u}|^2 \rangle = \text{Pe}^2 \\ \mathbf{u}|_{\partial\Omega} = 0}} \text{Nu}(\mathbf{u}) \leq \inf_{\substack{\eta(\mathbf{x}) \\ \eta|_{z=0} = 1 \\ \eta|_{z=1} = 0}} \left\{ \int_{\Omega} |\nabla \eta|^2 + \text{Pe}^2 \sup_{\substack{\mathbf{u}(\mathbf{x}) \\ \mathbf{u}|_{\partial\Omega} = \mathbf{0} \\ \int_{\Omega} |\nabla \mathbf{u}|^2 = 1}} \int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \eta)|^2 \right\}$$

when carried out optimally. As it turns out, these bounds are one and the same.

LEMMA 6.1. *Let $U_{\text{bm}}(\text{Pe})$ and $U_{\text{symm}}(\text{Pe})$ denote the optimal values appearing on the right-hand sides of (6.17) and (6.18), respectively. We have that $U_{\text{bm}} = U_{\text{symm}}$.*

Remark 6.2. As the following proof shows, the minimization in (6.18) can be performed over η depending on z alone without changing the resulting value.

PROOF. We prove this in two steps: first we show that $U_{\text{symm}} \leq U_{\text{bm}}$, and then we prove the reverse inequality. In both cases, we will use the fact that

$$(6.19) \quad \int_{\Omega} |\nabla \Delta^{-1} \operatorname{div} \mathbf{m}|^2 = \sup_{\substack{\theta(\mathbf{x}) \\ \theta|_{\partial\Omega} = 0}} \int_{\Omega} 2\mathbf{m} \cdot \nabla \theta - |\nabla \theta|^2$$

for all $\mathbf{m} \in L^2(\Omega; \mathbb{R}^3)$.

We begin by showing that $U_{\text{symm}} \leq U_{\text{bm}}$. Taking $\mathbf{m} = \mathbf{u} \tau$ in (6.19), we see that a background field $\tau(z)$ satisfies $H_{\tau,\lambda} \geq 0$ if and only if

$$\int_{\Omega} |\nabla \Delta^{-1} \operatorname{div} \mathbf{u} \tau|^2 \leq \lambda \int_{\Omega} |\nabla \mathbf{u}|^2$$

for all divergence-free \mathbf{u} that vanish at $\partial\Omega$. Therefore,

$$U_{\text{bm}} \geq \inf_{\substack{\tau(z), \lambda \\ \tau(0)=1, \tau(1)=0 \\ H_{\tau,\lambda} \geq 0}} \left\{ \int_0^1 |\tau'|^2 + \text{Pe}^2 \sup_{\substack{\mathbf{u}(\mathbf{x}) \\ \mathbf{u}|_{\partial\Omega} = \mathbf{0} \\ \int_{\Omega} |\nabla \mathbf{u}|^2 = 1}} \int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \tau)|^2 \right\} \geq U_{\text{symm}}$$

since enlarging the admissible set only decreases the resulting minimal value.

Now we prove that $U_{\text{symm}} \geq U_{\text{bm}}$. Parametrizing the admissible set from (6.18) via the level sets of

$$M(\eta) = \sup_{\substack{\mathbf{u}(\mathbf{x}) \\ \mathbf{u}|_{\partial\Omega} = \mathbf{0} \\ \int_{\Omega} |\nabla \mathbf{u}|^2 = 1}} \int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \eta)|^2,$$

we can write that

$$U_{\text{symm}} = \inf_{\lambda} \inf_{\substack{\eta(\mathbf{x}) \\ \eta|_{z=0}=1, \eta|_{z=1}=0 \\ M(\eta)=\lambda}} \left\{ \int_{\Omega} |\nabla \eta|^2 + \text{Pe}^2 \lambda \right\}.$$

Extend the definition of $H_{\tau, \lambda}$ to functions of \mathbf{x} by taking

$$H_{\eta, \lambda}(\mathbf{u}, \theta) = \int_{\Omega} \frac{\lambda}{2} |\nabla \mathbf{u}|^2 + \frac{1}{2} |\nabla \theta|^2 + \mathbf{u} \cdot \nabla \theta.$$

By (6.19),

$$M(\eta) = \lambda \iff \inf_{\substack{\mathbf{u}(\mathbf{x}), \theta(\mathbf{x}) \\ \mathbf{u}|_{\partial\Omega} = \mathbf{0}, \theta|_{\partial\Omega} = 0}} H_{\eta, \lambda}(\mathbf{u}, \theta) = 0,$$

and the latter happens if and only if $H_{\eta, \lambda} \geq 0$. Using that $\int_{\Omega} |\nabla \eta|^2$ is convex in η and that $\{\eta: H_{\eta, \lambda} \geq 0\}$ is also convex, we can replace η with its periodic average $\tau = \bar{\eta}$ to deduce that

$$U_{\text{symm}} \geq \inf_{\lambda} \inf_{\substack{\tau(z) \\ \tau|_{z=0}=1, \tau|_{z=1}=0 \\ H_{\tau, \lambda} \geq 0}} \left\{ \int_{\Omega} |\tau'|^2 + \text{Pe}^2 \lambda \right\} = U_{\text{bm}}$$

as desired. \square

While, in the end, the symmetrization method yields the same $\text{Pe}^{2/3}$ or $\text{Ra}^{1/2}$ upper bound as does the background method applied to optimal transport or RBC, its formulation is what ultimately led us to discover the relation between the optimal design of heat transport and Howard's variational approach as discussed in Section 6.1. This is not to say that a more careful analysis of the wall-to-wall optimal transport problem may not ultimately lead to new, logarithmically corrected upper bounds. Whether such corrections hold remains to be seen.

6.3 On the Realizability of Optimal Heat Transport by Buoyancy-Driven Convection

We return to the full system (6.1)–(6.3) to discuss the role of the momentum equation (6.2). One may wonder if buoyancy forces are capable of producing flows, time-dependent or steady, that realize near-optimal heat transport. The answer depends upon the way in which flow intensity is constrained.

First, we note that the energy-constrained wall-to-wall optimal transport problem corresponds to RBC in a fluid-saturated porous layer where the Navier-Stokes momentum equation (6.2) is replaced by Darcy's law. This implies the balance law $\langle |\mathbf{u}|^2 \rangle = \text{Ra}(\text{Nu} - 1)$, which, when combined with the result of Theorem 1.1, yields the optimal scaling $\text{Nu} \sim \text{Ra}$ in this setting. Direct numerical simulations of time-dependent high-Ra porous medium convection [20, 32] are consistent with this scaling, indicating that buoyancy forces *can* produce flows realizing optimal heat transport insofar as scaling is concerned. On the other hand, asymptotic and

numerical investigations indicate that the best possible transport by steady flows satisfies $\text{Nu} \sim \text{Ra}^{0.6}$ [43].

Second, we observe that the enstrophy-constrained optimal transport problem corresponds to Rayleigh's original model of buoyancy-driven convection in a fluid layer [36]. There, steady convection also appears to be strongly suboptimal with the highest computationally observed scaling being $\text{Nu} \sim \text{Ra}^{0.31}$ [38, 42]. To date, there are no turbulent high-Ra direct numerical simulations indicating heat transport scaling much higher than $\text{Nu} \sim \text{Ra}^{1/3}$.

We close our discussion of fluid dynamical implications by commenting on the certain suboptimality of heat transport in Rayleigh's original model. Rayleigh imposed (6.1)–(6.3) in two-dimensions with stress-free velocity boundary conditions and the usual Dirichlet temperature ones. Although RBC in a fluid layer must obey the bound $\text{Nu} \lesssim \text{Ra}^{1/2}$ in any dimension and for any boundary conditions, the result of [44] is that in two-dimensions and with stress-free boundaries $\text{Nu} \lesssim \text{Ra}^{5/12}$. Nevertheless, by combining the relevant balance law $\langle |\nabla \mathbf{u}|^2 \rangle = \text{Ra}(\text{Nu} - 1)$ implied by the Navier-Stokes momentum equation (6.2) with the result of Theorem 1.2 and the remark immediately thereafter, we conclude that optimal heat transport in the setting of Rayleigh's model must satisfy $\text{Nu} \sim \text{Ra}^{1/2}$ (up to logarithmic corrections). Our analysis is consistent with all the requirements of Rayleigh's model except for the Navier-Stokes momentum equation (6.2). Thus, buoyancy-driven convection in two-dimensions between stress-free boundaries must yield strongly suboptimal rates of heat transport as compared with what happens if (6.2) is not imposed. This underscores the importance of using the momentum equation—rather than only a balance law it implies—for determining the asymptotic heat transport of turbulent RBC.

7 Optimal Transport as Energy-Driven Pattern Formation

There is a second scientific context, other than the fluid dynamical one, in which the methods behind our analysis of wall-to-wall optimal heat transport have played a fundamental role. This is the subject of “energy-driven pattern formation” in mathematical materials science [23].

Perhaps the key methodological contribution of this paper is the reformulation of the general steady wall-to-wall optimal transport problem

$$(7.1) \quad \sup_{\substack{\mathbf{u}(\mathbf{x}) \\ \|\mathbf{u}\| = \text{Pe} \\ +\text{b.c.}}} \text{Nu}(\mathbf{u})$$

in its integral form

$$(7.2) \quad \inf_{\substack{\mathbf{u}(\mathbf{x}), \xi(\mathbf{x}) \\ \int_{\Omega} w \xi = 1 \\ +\text{b.c.}}} \int_{\Omega} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \xi)|^2 + \varepsilon \|\mathbf{u}\|^2 \int_{\Omega} |\nabla \xi|^2.$$

This change of viewpoint, accomplished in Section 3, hinges on the fact that the Nusselt number of a steady velocity field \mathbf{u} can be written as the maximal value of a certain nonlocal functional in ξ . The resulting problem (7.2) is equivalent to the original one (7.1), and optimizers correspond. In the examples of energy- and enstrophy-constrained optimal transport considered in Sections 4 and 5, where $\|\cdot\|$ is the (volume-averaged) L^2 - or \dot{H}^1 -norm, the integral formulation (7.2) plays a key role in the construction of divergence-free velocity fields that achieve nearly optimal transport. As that analysis shows, the complexity of the successful construction—whether it can be described using few length scales or many—depends strongly on the choice of norm.

Besides its practical use for the estimation of optimal transport, (7.2) shares striking similarities with other nonconvex and singularly perturbed variational problems from mathematical materials science. The study of patterns selected by energy minimization principles in this field is known as energy-driven pattern formation. It is important to note that the wall-to-wall optimal transport problem is variational by definition. Thus, our observation is not that there exists some variational formulation for it, but rather that the specific formulation (7.2) is reminiscent of various model problems from energy-driven pattern formation. From this point of view, it is no surprise that the (nearly) optimal patterns constructed in this paper for wall-to-wall transport—convection rolls and branching flows—bear similarities with other well appreciated patterns from materials science including domain branching in micromagnetics [8,9] and wrinkling cascades in thin elastic sheets [3,22,31]. What (7.2) offers is a functional analytic framework in which to make such connections precise.

We discuss below two model problems from energy-driven pattern formation and their connections to wall-to-wall optimal transport. We leave their general scientific introduction to the references therein, focusing instead on the salient features of their analysis. This discussion provides an alternate viewpoint on the role of branching patterns in the variational analysis of transport, which complements the older purely fluid dynamical arguments of Busse [5]. We hope these remarks prove useful to the reader interested in our approach.

7.1 Magnetic Domain Branching in a Uniaxial Ferromagnet

Our first example comes from micromagnetics and concerns the patterns formed by magnetic domains in a uniaxial ferromagnet. The energetic description is as follows. We take as the magnet the domain $\Omega = (-L, L)_x \times [0, 1]_{y,z}^2$, where x is the the preferred direction of magnetization and L is the magnet's (nondimensionalized) length. On Ω we define a magnetization vector field $\mathbf{m}(\mathbf{x}) = m_1 \hat{\mathbf{i}} + m_2 \hat{\mathbf{j}} + m_3 \hat{\mathbf{k}}$, which is required to be of unit size $|\mathbf{m}| = 1$ and is extended by zero to the rest of space $\mathbb{R}^3 \setminus \Omega$. The micromagnetic energy that results is

$$(7.3) \quad \int_{\text{all space}} |\nabla \Delta^{-1} \operatorname{div} \mathbf{m}|^2 + \int_{\text{magnet}} Q(1 - m_1^2) + \varepsilon |\nabla \mathbf{m}|$$

where the divergence is understood in the distributional sense. Strictly speaking, this is a “sharp interface” model in which the total variation norm

$$\int_{\Omega} |\nabla \mathbf{m}| = \sum_{i=1}^3 \int_{\Omega} |\nabla m_i| = \sum_{i=1}^3 \sup_{\substack{\mathbf{v} \in C_c^1(\Omega; \mathbb{R}^n) \\ \|\mathbf{v}\|_{L^\infty(\Omega)} \leq 1}} \int_{\Omega} m_i \operatorname{div} \mathbf{v}$$

features instead of the \dot{H}^1 -norm (for more on this reduction see [8]). The first term appearing in (7.3) is called the magnetostatic energy; it accounts for the cost of the magnetic field induced by \mathbf{m} in the ambient space. The second term is the anisotropy energy, and it arises from an underlying crystalline anisotropy that prefers \mathbf{m} to be $\pm \hat{\mathbf{i}}$. The third term is the interfacial energy. It permits \mathbf{m} to be discontinuous, but limits the total area of any interfaces across which \mathbf{m} jumps. The parameters Q and ε set the relative strengths of these effects. The magnetostatic and interfacial energies have direct analogues in the wall-to-wall problem (7.2); the anisotropy term does not. Note that, due to the constraint $|\mathbf{m}| = 1$, this functional is nonconvex.

There are various designs for \mathbf{m} one can entertain in minimizing (7.3). One is the so-called Kittel structure, in which \mathbf{m} is independent of x and $\pm \hat{\mathbf{i}}$ -valued throughout the magnet, alternating between these at some to be determined length scale l in the yz -plane. This design costs no anisotropic energy and the optimal l is selected by minimizing its magnetostatic and interfacial costs. Another important design is the Landau-Lifshitz structure, in which \mathbf{m} is independent of x and $\pm \hat{\mathbf{i}}$ -valued except for in a thin boundary layer near $x = \pm L$. There, it is taken to be perpendicular to $\hat{\mathbf{i}}$ in such a way as to eliminate the magnetostatic energy completely, thus coupling the thickness of the boundary layer to the length scale l of oscillations in the bulk. This is a sharp-interface version of the convection roll design described in Section 4.

Finally, there is the Privorotskii construction, which plays the role of the branching flows from Section 5. It too involves a very large number of distinct length scales that interpolate between a preferred length scale in the bulk l_{bulk} and a significantly smaller one at the boundary l_{bl} . We refer the reader for more details to [8, 9], including a description of the relevant regimes.

What can be proved regarding this nonconvex, nonlocal minimization problem? Following the reference [9] we assume that $\mathbf{m}(x, y, z)$ is periodic in (y, z) and identify $[0, 1]^2_{y,z}$ with $\mathbb{T}^2_{y,z}$. Then there exist positive constants C and C' such that the minimum micromagnetic energy satisfies

$$CQ^{1/3}\varepsilon^{2/3}L^{1/3} \leq \text{minimum micromagnetic energy} \leq C'Q^{1/3}\varepsilon^{2/3}L^{1/3}$$

for all sufficiently large Q and sufficiently small ε/L . The proof of this result requires two kinds of arguments. The upper bound comes from estimating the cost of an optimal Privorotskii construction (the conditions on Q , ε , and L ensure that the result is significantly less than those obtained by the Kittel and Landau-Lifshitz

structures). The lower bound asserts that the Prigorotskii construction cannot be beat as far as scaling is concerned. The original proof of it can be found in [9], but we note the existence of a second, more recent proof in [10], which utilizes the endpoint Gagliardo-Nirenberg interpolation inequality

$$\|f\|_{L^{4/3}(\mathbb{T}^2)} \lesssim \|\nabla f\|_{L^1(\mathbb{T}^2)}^{1/2} \|f\|_{H^{-1}(\mathbb{T}^2)}^{1/2}$$

holding for all mean-zero and periodic functions f .

7.2 Blistering Patterns in Thin Elastic Sheets

Our second example comes from elasticity theory. Consider a thin elastic sheet of (nondimensional) thickness h that is strongly bonded to the top of a large rubber block except for on some known subdomain $\Omega \subset \mathbb{R}^2$. Applying biaxial compression to the block causes the sheet to blister in the unbonded domain. The result is a complex pattern of wrinkles and folds whose details can be modeled through the minimization of a certain nonconvex and singularly perturbed variational problem. As in [3, 22], we consider minimization of the internal elastic energy under clamped boundary conditions. In the Föppl–von Karman model, the elastic energy (per unit thickness) is given by

$$(7.4) \quad \int_{\substack{\text{blistered} \\ \text{region}}} |e(\mathbf{v}) + \frac{1}{2} \nabla \phi \otimes \nabla \phi|^2 + h^2 |\nabla \nabla \phi|^2$$

where the “in-plane” displacement parallel to the top of the block is $\mathbf{v}(\mathbf{x})$ and the “out-of-plane” displacement perpendicular to it is $\phi(\mathbf{x})$. Here, $e(\mathbf{v})$ denotes the symmetric part of the in-plane displacement gradient $\nabla \mathbf{v}$. Taken together, the in- and out-of-plane displacements yield the map $(\mathbf{x}, 0) \mapsto (\mathbf{x} + \mathbf{v}(\mathbf{x}), \phi(\mathbf{x}))$ which describes the deformation of the blister. At the edge of the blister $\partial\Omega$ we impose the clamped boundary conditions

$$\mathbf{v}|_{\partial\Omega} = -\lambda \mathbf{x}, \quad \phi|_{\partial\Omega} = 0, \quad \text{and} \quad \partial_{\mathbf{v}} \phi|_{\partial\Omega} = 0.$$

The parameter λ is positive and sets the amount of overall compressive strain. The first term in the energy is called the membrane term. It prefers the in-plane strain $e(\mathbf{v}) + \frac{1}{2} \nabla \phi \otimes \nabla \phi$ to vanish. The second one is called the bending term, and it prefers the out-of-plane displacement to vary on longer length scales or not at all. The relative strength of these effects is determined by the parameter h , which is understood to be small.

There are significant parallels between the elastic energy functional (7.4) and the integral formulation of wall-to-wall transport (7.2). Of course, the bending term from (7.4) and the higher-order terms from (7.2) act to regularize designs. More interestingly, we observe a similarity between the membrane term from (7.4) and the advection term and net flux constraint from (7.2). Let us introduce a streamfunction ψ for the divergence-free velocity field \mathbf{u} (we work with a two-dimensional

fluid layer now) and rewrite the advection term as

$$(7.5) \quad \mathfrak{f}_{\text{fluid layer}} |\nabla \Delta^{-1} \operatorname{div}(\mathbf{u} \xi)|^2 = \mathfrak{f}_{\text{fluid layer}} |\nabla \Delta^{-1} J(\psi, \xi)|^2$$

where $J(\psi, \xi) = \nabla^\perp \psi \cdot \nabla \xi$. Recall also that the net flux constraint requires that

$$(7.6) \quad \mathfrak{f}_{\text{fluid layer}} w \xi = 1.$$

As pointed out in Section 3.2—see the discussion surrounding (3.8)—for smooth enough designs (\mathbf{u}, ξ) the advection term cannot vanish while the net flux constraint and boundary conditions $w|_{\partial\Omega} = \xi|_{\partial\Omega} = 0$ hold. As (7.2) makes clear, wall-to-wall optimal transport is precisely about balancing these competing effects. Regarding elasticity, we ask: what does it take for the membrane term to nearly vanish? This can be answered with the aid of the lower bound

$$(7.7) \quad \begin{aligned} \mathfrak{f}_{\text{blistered region}} \left| e(\mathbf{v}) + \frac{1}{2} \nabla \phi \otimes \nabla \phi \right|^2 &\gtrsim \mathfrak{f}_{\text{blistered region}} |\nabla \nabla (\Delta \Delta)^{-1} \det \nabla \nabla \phi|^2 \\ &\quad + \left| \mathfrak{f}_{\text{blistered region}} \frac{1}{2} \nabla \phi \otimes \nabla \phi - \lambda \operatorname{Id}_{2 \times 2} \right|^2, \end{aligned}$$

which we compare with (7.5) and (7.6). For the in-plane strain to nearly vanish, the bulk average of $\frac{1}{2} \nabla \phi \otimes \nabla \phi$ must be nearly constant and equal to a known multiple of the identity. At the same time, ϕ must nearly satisfy the degenerate Monge-Ampère equation $\det \nabla \nabla \phi = 0$. It follows from the results of [34] that these are incompatible constraints, i.e., the membrane term cannot vanish while the bending term remains finite. The situation is remarkably similar to that of wall-to-wall optimal transport.

The scaling law of the minimum energy for blistering is known. As proved in [3, 22], there exist constants C and C' depending only on Ω so that

$$(7.8) \quad Ch \leq \text{minimum elastic energy} \leq C'h$$

for small enough h . The upper bound comes from a branching construction involving finer and finer oscillations in $\nabla \phi$ at a certain length scale depending on the distance from the blister edge $\partial\Omega$. As opposed to the corresponding result for the wall-to-wall problem, there is no logarithmic correction to scaling in (7.8). This can be explained with the help of (7.5) and (7.7): whereas the advection term has a -1 scaling in its quadratic nonlinearity $J(\psi, \xi)$, the membrane term has a -2 scaling in $\det \nabla \nabla \phi$ and therefore permits much stronger oscillations. As a result, branching can be more easily accommodated in blistering than in optimal transport. The lower bound from (7.8) asserts that branching indeed achieves the minimum energy up to a prefactor depending only on the domain. Its proof reminds one of the proof of Howard's lower bound given after (6.11). For details we refer the reader to [22] for the case where Ω is a square with periodic boundary conditions

at opposite sides, and to [3] for the more general case of an arbitrary domain Ω with suitably smooth boundary.

Acknowledgment. We thank R. V. Kohn and A. N. Souza for helpful discussions. This work was supported by National Science Foundation Awards DGE-0813964 and DMS-1812831 (IT), DMS-1515161 and DMS-1813003 (CRD), a Van Loo Postdoctoral Fellowship (IT), and a Guggenheim Foundation Fellowship (CRD).

Bibliography

- [1] Alben, S. Improved convection cooling in steady channel flows. *Phys. Rev. Fluids* **2** (2017), 104501. doi:10.1103/PhysRevFluids.2.104501
- [2] Avellaneda, M.; Majda, A. J. An integral representation and bounds on the effective diffusivity in passive advection by laminar and turbulent flows. *Comm. Math. Phys.* **138** (1991), no. 2, 339–391.
- [3] Ben Belgacem, H.; Conti, S.; DeSimone, A.; Müller, S. Rigorous bounds for the Föppl-von Kármán theory of isotropically compressed plates. *J. Nonlinear Sci.* **10** (2000), no. 6, 661–683. doi:10.1007/s003320010007
- [4] Braides, A. *Γ -convergence for beginners*. Oxford Lecture Series in Mathematics and Its Applications, 22. Oxford University Press, Oxford, 2002. doi:10.1093/acprof:oso/9780198507840.001.0001
- [5] Busse, F. H. On Howard’s upper bound for heat transport by turbulent convection. *J. Fluid Mech.* **37** (1969), no. 3, 457–477. doi:10.1017/S0022112069000668
- [6] Chernyshenko, S. Relationship between the methods of bounding time averages. Preprint, 2017. arXiv:1704.02475 [physics.flu-dyn]
- [7] Chernyshenko, S. I.; Goulart, P.; Huang, D.; Papachristodoulou, A. Polynomial sum of squares in fluid dynamics: a review with a look ahead. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **372** (2014), no. 2020, 20130350, 18 pp. doi:10.1098/rsta.2013.0350
- [8] Choksi, R.; Kohn, R. V. Bounds on the micromagnetic energy of a uniaxial ferromagnet. *Comm. Pure Appl. Math.* **51** (1998), no. 3, 259–289. doi:10.1002/(SICI)1097-0312(199803)51:3<259::AID-CPA3>3.0.CO;2-9
- [9] Choksi, R.; Kohn, R. V.; Otto, F. Domain branching in uniaxial ferromagnets: a scaling law for the minimum energy. *Comm. Math. Phys.* **201** (1999), no. 1, 61–79. doi:10.1007/s002200050549
- [10] Cinti, E.; Otto, F. Interpolation inequalities in pattern formation. *J. Funct. Anal.* **271** (2016), no. 11, 3348–3392. doi:10.1016/j.jfa.2016.05.007
- [11] Constantin, P.; Doering, C. R. Variational bounds on energy dissipation in incompressible flows. II. Channel flow. *Phys. Rev. E (3)* **51** (1995), no. 4, part A, 3192–3198. doi:10.1103/PhysRevE.51.3192
- [12] Doering, C. R.; Constantin, P. Variational bounds on energy dissipation in incompressible flows: shear flow. *Phys. Rev. E (3)* **49** (1994), no. 5, part A, 4087–4099. doi:10.1103/PhysRevE.49.4087
- [13] Doering, C. R.; Constantin, P. Variational bounds on energy dissipation in incompressible flows. III. Convection. *Phys. Rev. E* **53** (1996), 5957–5981. doi:10.1103/PhysRevE.53.5957
- [14] Doering, C.; Constantin, P. On upper bounds for infinite Prandtl number convection with or without rotation. *J. Math. Phys.* **42** (2001), no. 2, 784–795. doi:10.1063/1.1336157
- [15] Doering, C. R.; Gibbon, J. D. *Applied analysis of the Navier-Stokes equations*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 1995. doi:10.1017/CBO9780511608803

- [16] Doering, C. R.; Otto, F.; Reznikoff, M. G. Bounds on vertical heat transport for infinite-Prandtl-number Rayleigh–Bénard convection. *J. Fluid Mech.* **560** (2006), 229–241. doi:10.1017/S0022112006000097
- [17] Fannjiang, A.; Papanicolaou, G. Convection enhanced diffusion for periodic flows. *SIAM J. Appl. Math.* **54** (1994), no. 2, 333–408. doi:10.1137/S0036139992236785
- [18] Ghoussoub, N. *Self-dual partial differential systems and their variational principles*. Springer Monographs in Mathematics, Springer, New York, 2009.
- [19] Hassanzadeh, P.; Chini, G. P.; Doering, C. R. Wall to wall optimal transport. *J. Fluid. Mech.* **751** (2014), 627–662. doi:10.1017/jfm.2014.306
- [20] Hewitt, D. R.; Neufeld, J. A.; Lister, J. R. Ultimate regime of high Rayleigh number convection in a porous medium. *Phys. Rev. Lett.* **108** (2012), 224503. doi:10.1103/PhysRevLett.108.224503
- [21] Howard, L. N. Heat transport by turbulent convection. *J. Fluid. Mech.* **17** (1963), no. 3, 405–432. doi:10.1017/S0022112063001427
- [22] Jin, W.; Sternberg, P. Energy estimates for the von Kármán model of thin-film blistering. *J. Math. Phys.* **42** (2001), no. 1, 192–199. doi:10.1063/1.1316058
- [23] Kohn, R. V. Energy-driven pattern formation. *International Congress of Mathematicians. Vol. I*, 359–383. European Mathematical Society, Zürich, 2007. doi:10.4171/022-1/15
- [24] Malkus, W. V. R. The heat transport and spectrum of thermal turbulence. *Proc. Roy. Soc. London. Ser. A.* **225** (1954), 196–212. doi:10.1098/rspa.1954.0197
- [25] Marcotte, F.; Doering, C. R.; Thiffeault, J.-L.; Young, W. R. Optimal heat transfer and optimal exit times. *SIAM J. Appl. Math.* **78** (2018), no. 1, 591–608. doi:10.1137/17M1150220
- [26] Milton, G. W. On characterizing the set of possible effective tensors of composites: the variational method and the translation method. *Comm. Pure Appl. Math.* **43** (1990), no. 1, 63–125. doi:10.1002/cpa.3160430104
- [27] Motoki, S.; Kawahara, G.; Shimizu, M. Maximal heat transfer between two parallel plates. *J. Fluid Mech.* **851** (2018), R4, 14 pp. doi:10.1017/jfm.2018.557
- [28] Motoki, S.; Kawahara, G.; Shimizu, M. Optimal heat transfer enhancement in plane Couette flow. *Journal of Fluid Mechanics* **835** (2018), 1157–1198. doi:10.1017/jfm.2017.779
- [29] Nobili, C.; Otto, F. Limitations of the background field method applied to Rayleigh–Bénard convection. *J. Math. Phys.* **58** (2017), no. 9, 093102, 46 pp. doi:10.1063/1.5002559
- [30] Ortiz, M. A variational formulation for convection-diffusion problems. *Internat. J. Engrg. Sci.* **23** (1985), no. 7, 717–731. doi:10.1016/0020-7225(85)90004-7
- [31] Ortiz, M.; Gioia, G. The morphology and folding patterns of buckling-driven thin-film blisters. *J. Mech. Phys. Solids* **42** (1994), no. 3, 531–559. doi:10.1016/0022-5096(94)90030-2
- [32] Otero, J.; Dontcheva, L. A.; Johnston, H.; Worthing, R. A.; Kurganov, A.; Petrova, G.; Doering, C. R. High-Rayleigh-number convection in a fluid-saturated porous layer. *J. Fluid Mech.* **500** (2004), 263–281. doi:10.1017/S0022112003007298
- [33] Otto, F.; Seis, C. Rayleigh–Bénard convection: improved bounds on the Nusselt number. *J. Math. Phys.* **52** (2011), no. 8, 083702, 24 pp. doi:10.1063/1.3623417
- [34] Pakzad, M. R. On the Sobolev space of isometric immersions. *J. Differential Geom.* **66** (2004), no. 1, 47–69.
- [35] Plasting, S. C.; Kerswell, R. R. 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.* **477** (2003), 363–379. doi:10.1017/S0022112002003361
- [36] Rayleigh, L. On convection currents in a horizontal layer of fluid, when the higher temperature is on the under side. *Philos. Mag.* **32** (1916), 529–546.
- [37] Seis, C. Scaling bounds on dissipation in turbulent flows. *J. Fluid. Mech.* **777** (2015), 591–603.
- [38] Sondak, D.; Smith, L. M.; Waleffe, F. Optimal heat transport solutions for Rayleigh–Bénard convection. *J. Fluid Mech.* **784** (2015), 565–595. doi:10.1017/jfm.2015.615

- [39] Souza, A. N. An optimal control approach to bounding transport properties of thermal convection. Ph.D. thesis, University of Michigan, 2016.
- [40] Tobasco, I.; Doering, C. R. Optimal wall-to-wall transport by incompressible flows. *Phys. Rev. Lett.* **118** (2017), 264502. doi:10.1103/PhysRevLett.118.264502
- [41] Tobasco, I.; Goluskin, D.; Doering, C. R. Optimal bounds and extremal trajectories for time averages in nonlinear dynamical systems. *Phys. Lett. A* **382** (2018), no. 6, 382–386. doi:10.1016/j.physleta.2017.12.023
- [42] Waleffe, F.; Boonkasame, A.; Smith, L. M. Heat transport by coherent Rayleigh-Bénard convection. *Physics of Fluids* **27** (2015), no. 5, 051702. doi:10.1063/1.4919930
- [43] Wen, B.; Corson, L. T.; Chini, G. P. Structure and stability of steady porous medium convection at large Rayleigh number. *J. Fluid Mech.* **772** (2015), 197–224. doi:10.1017/jfm.2015.205
- [44] Whitehead, J. P.; Doering, C. R. Ultimate state of two-dimensional Rayleigh-Bénard convection between free-slip fixed-temperature boundaries. *Phys. Rev. Lett.* **106** (2011), 244501. doi:10.1103/PhysRevLett.106.244501
- [45] Yan, X. On limits to convective heat transport at infinite Prandtl number with or without rotation. *J. Math. Phys.* **45** (2004), no. 7, 2718–2743. doi:10.1063/1.1763246

CHARLES R. DOERING
University of Michigan
530 Church St.
Ann Arbor, MI 48109
USA
E-mail: doering@umich.edu

IAN TOBASCO
University of Michigan
530 Church St.
Ann Arbor, MI 48109
USA
E-mail: itobasco@umich.edu

Received December 2017.