

Data Assimilation in Large Prandtl Rayleigh–Bénard Convection from Thermal Measurements*

A. Farhat[†], N. E. Glatt-Holtz[‡], V. R. Martinez[§], S. A. McQuarrie[¶], and J. P. Whitehead^{||}

Abstract. This work applies a continuous data assimilation scheme—a framework for reconciling sparse and potentially noisy observations to a mathematical model—to Rayleigh–Bénard convection at infinite or large Prandtl numbers using only the temperature field as observables. These Prandtl numbers are applicable to the earth’s mantle and to gases under high pressure. We rigorously identify conditions that guarantee synchronization between the observed system and the model, then confirm the applicability of these results via numerical simulations. Our numerical experiments show that the analytically derived conditions for synchronization are far from sharp; that is, synchronization often occurs even when sufficient conditions of our theorems are not met. We also develop estimates on the convergence of an infinite Prandtl model to a large (but finite) Prandtl number generated set of observations. Numerical simulations in this hybrid setting indicate that the mathematically rigorous results are accurate, but of practical interest only for extremely large Prandtl numbers.

Key words. data assimilation, Rayleigh–Bénard convection, large Prandtl limit

AMS subject classifications. 76E06, 62M20, 35Q35

DOI. 10.1137/19M1248327

1. Introduction. In order to make accurate predictions, numerical models for geophysical processes require establishing accurate initial conditions. Data assimilation is used to estimate weather or ocean (or any other geophysical) variables by incorporating the real world data into the mathematical system to obtain an accurate initialization. One of the classical methods of data assimilation (see, e.g., [46, 23, 58, 60, 69, 66, 79, 65, 70, 7, 15, 8]) is to insert observational measurements directly into a model as the latter is being integrated in time (also known as nudging or Newtonian relaxation). There is a significant amount of recent literature concerning the mathematically rigorous analysis of nudging algorithms for data assimilation developed for hydrodynamic equations with a particular focus on weather and climate systems. Recently, a variational scheme, known as 3DVAR, was studied in [15], in the case where observables are given as noisy Fourier modes, and in [8], which successfully accommodates a larger class of observables that, in particular, includes the more physically relevant cases of nodal values and

*Received by the editors March 6, 2019; accepted for publication (in revised form) December 5, 2019; published electronically February 19, 2020.

<https://doi.org/10.1137/19M1248327>

Funding: The second author would like to acknowledge the grants NSF-DMS-1816551 and NSF-DMS-1313272 and Simons Foundation 515990 which supported this work.

[†]Mathematics, Florida State University, Tallahassee, FL, 32306-4510 (afarhat@fsu.edu).

[‡]Tulane University, New Orleans, LA 70118 (negh@tulane.edu).

[§]Mathematics and Statistics, Hunter College, CUNY, New York, NY, 10065 (vmartinez@hunter.cuny.edu).

[¶]Institute for Computational Engineering Sciences, University of Texas, Austin, TX, 78712-1229 (shanemcq@utexas.edu).

^{||}Mathematics, Brigham Young University, Provo, UT, 84602-0002 (whitehead@mathematics.byu.edu).

volume elements; see also [9], where observational error is accounted for. In these articles, rigorous proofs are obtained for the synchronization of the approximating signal with the true signal that corresponds to the observations, using the two-dimensional (2D) incompressible Navier–Stokes equations (NSEs) as a model.

The data assimilation algorithm analyzed in [15, 8] can be described as follows: suppose that $u(t)$ represents a solution of some dynamical system governed by an evolution equation of the type

$$(1.1) \quad \frac{du}{dt} = F(u),$$

where the initial state of the system, $u(0) = u_0$, is *unknown*. We would like to accurately track this solution $u(t)$ as t increases notwithstanding our uncertainty in u_0 . Let $I_h(u(t))$ represent an interpolant operator based on the observations of the system at a coarse spatial resolution of size h , for $t \in [0, T]$. We then construct a solution $v(t)$ from the observations that satisfies the equations

$$(1.2a) \quad \frac{dv}{dt} = F(v) - \mu(I_h(v) - I_h(u)),$$

$$(1.2b) \quad v(0) = v_0,$$

where $\mu > 0$ is a relaxation (nudging) parameter and v_0 can be prescribed as an arbitrary initial condition. We then take $v(t)$ as prediction of $u(t)$ which we anticipate becomes more accurate as t (and therefore the amount of observed data $I_h(u(t))$) increases.

The algorithm designated by (1.2) was designed to work for dissipative dynamical systems of the form (1.1) that are known to have global-in-time solutions, a finite-dimensional global attractor, as well as a finite set of determining parameters (see, e.g., [38, 39, 40, 53, 52, 21, 47, 35] and references therein). Typically in these settings, following the ideas in [38], lower bounds on $\mu > 0$ and upper bounds on $h > 0$ can be derived such that the approximate solution $v(t)$ converges to the reference solution $u(t)$ as $t \rightarrow \infty$. This was initially proved for the 2D NSEs in [15, 8].

Numerous further studies, both analytical and numerical, have been carried out for the algorithm (1.2), illustrating its broad scope of applicability. For instance, the nudging approach has been validated for models including the 2D magnetohydrodynamic system [11], the 2D surface quasi-geostrophic equation [51], three-dimensional (3D) Brinkman–Forchheimer–extended Darcy model [61], and 3D simplified Bardina model [4]. The practically and physically relevant scenarios of discrete-time and time-averaged observables were studied in [37, 14, 49, 56]; more recently, it was shown in [10] that this nudging algorithm is capable of synchronizing the statistics propagated by the flow as they are observed only on a coarse-mesh scale; the efficacy of this algorithm for assimilating actual data sampled from a regional domain encompassing most of Northern Africa and the Middle East was recently tested in [24]. We refer the reader to [23] for a summary on the use of data assimilation in practical forecasting and [54] for a comprehensive text on numerical weather prediction where nudging has been employed.

Regarding related numerical studies, [42] demonstrated in the case of the 2D NSEs that the number of observables required for synchronization using (1.2) is much lower in practice

than what has been deemed sufficient by the rigorous analysis. In the setting of the 2D Rayleigh Bénard (RB) system, numerical studies were carried out in [6], and then in [28] for nearly turbulent flows using vorticity and local circulation measurements. We emphasize that the numerical experiments carried out in the present article are for moderately turbulent flows whose dynamics are significantly more complex than the regime of two-cell convection rolls that [6] was restricted to. Moreover, our studies are carried out to a similar high degree of numerical precision as found in [28]. We refer the reader to [5, 34, 25, 19] for various other studies in the context of turbulent flows such as how one can leverage the nudging scheme to infer unknown parameters of the flow. In [12, 48, 62], analytical studies on the various modes of synchronization of the algorithm (1.2) and on certain variants on its numerical discretization were carried out.

The earth system is heated from within and cooled by the atmosphere or ocean at the earth's surface. On geological time scales, the mantle's motion can be modeled as a fluid. The big difference between the temperature of the top mantle and the bottom mantle is a major source of the convective motion (fluid motion driven by temperature difference). The full compressible, temperature-dependent viscous equations of motion that ostensibly describe flow in the mantle [68] are currently beyond the reach of a rigorous mathematical analysis, but a first order approximation to this system is adequately described by taking the infinite Prandtl limit [71, 72, 73, 74, 75, 76, 41] of the RB system first described in [64] by Lord Rayleigh. We recall that the Prandtl number represents the ratio of the kinematic viscosity to the thermal diffusivity. Since the original formulation of the minimal mathematical model in [64], extensive research has sought to quantify its dynamical evolution [59, 1, 2, 16] and resultant large spatial and long temporal scale impact of convective flow; see [3, 57] for example. Despite the seeming simplicity of the RB system, there remain open questions regarding the exact nature of the convective heat transport and the impact and nature of boundary layers at physically relevant values (see, e.g., [3]). To further complicate matters, mantle convection is far more nuanced than RB convection, having several other unanswered questions, in addition to the well-known open problems in the latter setting. Unlike the low Prandtl number setting, experimental investigations of mantle convection are not practical, so numerical simulations provide one of the only avenues to investigate these issues. Of fundamental concern in such simulations is the dependence of the simulation on the initial condition and/or true physical setting, which would ideally be accompanied by physical observations. The collection of observational data from the mantle is an onerous inverse problem that obfuscates much of the desired resolution in time and space [68], both in the sense of physical accessibility and due to the relevant time scales involved. Indeed, the evolution of the mantle is on millennial time scales. Despite remarkable advances in imaging technologies, observations of the mantle are sparse and prone to noise and are insufficient to determine the mantle's state. Thus, the development of the advanced real-time prediction systems that are capable of depicting and predicting the mantle's state is necessary to gain insight into the dynamics in the earth's interior.

Access to observations from earth's mantle is limited. Geophysicists have decent observations at the surface (top layer) of plate velocities and heat flow distribution and have probes of mantle temperature where volcanism occurs. Since the motion of the tectonic plates is very slow (the relative movement of the plates typically ranges from zero to 100 mm annually; see [68], for example), we can assume that the top plate is stationary and with fixed temperature.

Thus, a realistic forecast model for the dynamics of earth’s mantle will employ *sparse thermal observations only*. In the context of atmospheric and oceanic physics, data assimilation algorithms where some state variable observations are not available as an input have been studied in [20, 27, 44, 43] for simplified numerical forecast models. Particularly related to the study carried out in this article, it was shown in the case of the 2D NSEs [30] and the 2D RB system in [33] that measurements on just a single component of velocity is sufficient to obtain synchronization. Charney’s question in [20, 43, 44] asks whether temperature observations are enough to determine the entire dynamical state of the system. In [44], an analytical argument suggested that Charney’s conjecture is correct, in particular, for a shallow water model. Further numerical testing in [43] affirmed that it is not certain whether assimilation with temperature data alone will yield initial states of arbitrary accuracy. The authors in [6] concluded that assimilation using coarse temperature measurements only will not always recover the true state of the full system. It was observed that the convergence to the true state using temperature measurements only is actually sensitive to the amount of noise in the measured data as well as to the spacing (the sparsity of the collected data) and the time frequency of such measured temperature data. Rigorous justification for Charney’s conjecture was provided in [32] in the case of the 3D planetary geostrophic model. Earlier, for the specific setting of 3D convection in a porous medium, where inertial effects can be ignored in the fluid velocity, it was shown in [31] that temperature measurements alone suffice to determine the velocity field. By comparison, the thrust of the analysis performed here is to establish the conclusions analogous to [31] while accounting for these inertial effects within the regime of a finite but large Prandtl number.

We consider the nudging approach both analytically and through numerical experiments to explore the range of applicability of the technique in this geophysically interesting context of large Prandtl convective systems. Ultimately, we accomplish the following:

1. We develop a nudging data assimilation scheme for both large and infinite Prandtl number RB convection in the traditional simplified 3D box geometry (see (2.4), (2.5), and (2.6) below) with observations in the temperature field only. In section 2, we formally introduce the governing equations for the 3D RB system. In section 3, we provide the mathematical framework within which our analysis is performed, as well as the relevant well-posedness results for the 3D RB system. We then establish rigorous estimates on the convergence rates for the simpler case of Prandtl number $(Pr) = \infty$ in section 4. The case of large, but finite Pr is addressed in section 5.
2. We perform high-resolution direct numerical simulations on the 2D version of this problem for moderately turbulent flows, in an effort to shed some light on the practical applicability of the rigorous estimates. In particular, we probe the values of the relaxation parameter and the number of required modes for the nudging scheme to converge. This is done in sections 4.2 and 5.2, immediately after their respective mathematical analysis.
3. We consider a practical scenario of “model error,” in which the assimilated variables are nudged “incorrectly.” Specifically, we assume that the modeling system corresponds to the infinite Prandtl system nudged by *data corresponding to a finite Prandtl system* (2.4), (2.5). This situation is studied both analytically and numerically in section 6.

We note that the choice of finitely many Fourier modes as the manifestation of our observables is made for ease of both exposition and numerical implementation. We reserve establishing estimates on the convergence rates for more general observables to a subsequent study.

2. The RB system and nudging equation. This initial section recalls the RB system and its nondimensional formulation. We then present the precise form of the nudging algorithm which we will study in sections 4 and 5.

The RB system for convection originates from the Boussinesq equations for an incompressible fluid with appropriate boundary conditions. The Boussinesq system over a d -dimensional domain, where $d = 2, 3$, $\Omega = [0, \tilde{L}]^{d-1} \times [0, h]$, is given by

$$(2.1) \quad \partial_t \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u} = \nu \Delta \mathbf{u} - \nabla p + \alpha g \mathbf{e}_d T, \quad \nabla \cdot \mathbf{u} = 0, \quad \partial_t T + (\mathbf{u} \cdot \nabla) T - \kappa \Delta T = 0,$$

where $\mathbf{u} = (u_1, \dots, u_d)$ is the velocity vector field, p is the scalar pressure field, and T denotes the temperature of a buoyancy driven fluid. The parameter $\nu > 0$ denotes the kinematic viscosity of the fluid, $\kappa > 0$ its thermal diffusivity, $\alpha > 0$ the thermal expansion coefficient, $g > 0$ denotes the constant gravitational force, and $\mathbf{e}_d$ is a constant vector antiparallel to the gravitational force. To model convection, (2.1) is then supplemented by the boundary conditions

$$(2.2) \quad \mathbf{u}|_{x_d=0} = \mathbf{u}|_{x_d=h} = \mathbf{0}, \quad T|_{x_d=0} = \delta T, \quad T|_{x_d=h} = 0, \quad \mathbf{u}, T \text{ are } L\text{-periodic in } x_1, x_{d-1},$$

where δT is a fixed constant determined by the (relative) strength of the bottom heating. The relations (2.1), (2.2) together constitute the RB system for convection in two and three dimensions. We note that variations on these boundary conditions are applicable to the earth's mantle, but as our focus is on the convergence of the data assimilation scheme, we assume that such variations have secondary effects.

Nondimensionalized variables. As is customary, we work with nondimensionalized variables. The system (2.1) is rescaled using h as a length scale, δT as the temperature scale, and the diffusive scale $\frac{\kappa}{h^2}$ as the time scale. The relevant nondimensional physical parameters for the system are the Prandtl number, Pr , and Rayleigh number, Ra , which are defined as

$$(2.3) \quad \text{Pr} := \frac{\nu}{\kappa}, \quad \text{Ra} := \frac{\alpha g (\delta T) h^3}{\nu \kappa}.$$

This leads to nondimensionalized variables over the rescaled domain $\Omega' = [0, L]^{d-1} \times [0, 1]$, $d = 2, 3$, which satisfy

$$(2.4) \quad \frac{1}{\text{Pr}} [\partial_{t'} \mathbf{u}' + (\mathbf{u}' \cdot \nabla') \mathbf{u}'] - \Delta' \mathbf{u}' = -\nabla' p' + \text{Ra} \mathbf{e}_d T', \quad \nabla' \cdot \mathbf{u}' = 0, \quad \mathbf{u}'(\mathbf{x}', 0) = \mathbf{u}'_0(\mathbf{x}'),$$

$$\partial_{t'} T' + \mathbf{u}' \cdot \nabla' T' - \Delta' T' = 0, \quad T'(\mathbf{x}', 0) = T'_0(\mathbf{x}')$$

with the boundary conditions

$$(2.5) \quad \mathbf{u}'|_{x'_d=0} = \mathbf{u}'|_{x'_d=1} = \mathbf{0}, \quad T'|_{x'_d=0} = 1, \quad T'|_{x'_d=1} = 0, \quad \mathbf{u}', T' \text{ are } L\text{-periodic in } x'_1, x'_{d-1}.$$

For notational simplicity, we will drop the $'$ in all that follows.

As previously alluded to, the physical setting of interest in this article is the earth's mantle where the Prandtl number Pr is *large*, namely, on the order of 10^{25} , while the Rayleigh number Ra that are of interest for mantle convection are typically between 10^7 and 10^8 [68]. Upon formally setting $\text{Pr} = \infty$ in the system (2.4)–(2.5), one arrives at

$$(2.6) \quad \begin{aligned} -\Delta \mathbf{u} &= -\nabla p + \text{Ra} \mathbf{e}_d T, \quad \nabla \cdot \mathbf{u} = 0, \\ \partial_t T + \mathbf{u} \cdot \nabla T - \Delta T &= 0, \quad T(\mathbf{x}, 0) = T_0(\mathbf{x}), \\ \mathbf{u}|_{x_d=0} &= \mathbf{u}|_{x_d=1} = \mathbf{0}, \quad T|_{x_d=0} = 1, \quad T|_{x_d=1} = 0, \quad \mathbf{u}, T \text{ are } L\text{-periodic in } x_1, x_{d-1}. \end{aligned}$$

The initial velocity $\mathbf{u}(\mathbf{x}, 0)$ is determined by T_0 and the corresponding momentum equations. Although there are several additional physical effects relevant to mantle convection that are omitted from the RB model considered here, we consider (2.6) to be an appropriate “zeroth-order” representation of mantle convection; it provides the starting point and test model for mantle convection simulations (see [13]). Although we anticipate eventually extending the results of the current investigation to more realistic models of mantle convection, we believe a more in-depth understanding of the problem at hand is a necessary first step.

Nudging setup. The idea, following [15, 8] and sketched in the introduction, is to nudge the assimilated system (1.2) with a projection of the “truth” that represents the exactly realizable observations of the original system. More precisely, the nudging is accomplished by introducing an affine feedback control term to the original “forecast” model (1.1), whose purpose is to enforce the asymptotic convergence of the solution of the assimilated system (1.2) towards that of the original system (1.1), *but only on the scales at which the observations are made*; it is this “relaxed” imposition that ensures the practicality of the nudging scheme. For this article, our “truth” is assumed to be represented by (2.4), (2.5), or (2.6).

Let $(\mathbf{u}, T)$ satisfy (2.4), (2.5) over $\Omega = [0, L]^2 \times [0, 1]$, from which we have obtained partial observations in the form of finitely many Fourier coefficients corresponding to wave-numbers $|k| \leq N$ for some integer $N > 0$. Let $(\tilde{\mathbf{u}}, \tilde{T})$ denote the assimilated or modeled system variables, which satisfy

$$(2.7) \quad \begin{aligned} \frac{1}{\text{Pr}} [\partial_t \tilde{\mathbf{u}} + (\tilde{\mathbf{u}} \cdot \nabla) \tilde{\mathbf{u}}] - \Delta \tilde{\mathbf{u}} &= -\nabla \tilde{p} + \text{Ra} \mathbf{e}_d \tilde{T}, \quad \nabla \cdot \tilde{\mathbf{u}} = 0, \quad \tilde{\mathbf{u}}(\mathbf{x}, 0) = \tilde{\mathbf{u}}_0(\mathbf{x}), \\ \partial_t \tilde{T} + \tilde{\mathbf{u}} \cdot \nabla \tilde{T} - \Delta \tilde{T} &= -\mu P_N(\tilde{T} - T), \quad \tilde{T}(\mathbf{x}, 0) = \tilde{T}_0(\mathbf{x}), \\ \tilde{\mathbf{u}}|_{x_d=0} &= \tilde{\mathbf{u}}|_{x_d=1} = \mathbf{0}, \quad \tilde{T}|_{x_d=0} = 1, \quad \tilde{T}|_{x_d=1} = 0, \quad \tilde{\mathbf{u}}, \tilde{T} \text{ are } L\text{-periodic in } x_1, x_{d-1}, \end{aligned}$$

where P_N denotes the projection onto Fourier wave-numbers $|k| \leq N$ (see (3.8) below). Its infinite Prandtl counterpart is given by

$$(2.8) \quad \begin{aligned} -\Delta \tilde{\mathbf{u}} &= -\nabla \tilde{p} + \text{Ra} \mathbf{e}_d \tilde{T}, \quad \nabla \cdot \tilde{\mathbf{u}} = 0, \\ \partial_t \tilde{T} + \tilde{\mathbf{u}} \cdot \nabla \tilde{T} - \Delta \tilde{T} &= -\mu P_N(\tilde{T} - T), \quad \tilde{T}(\mathbf{x}, 0) = \tilde{T}_0(\mathbf{x}), \\ \tilde{\mathbf{u}}|_{x_d=0} &= \tilde{\mathbf{u}}|_{x_d=1} = \mathbf{0}, \quad \tilde{T}|_{x_d=0} = 1, \quad \tilde{T}|_{x_d=1} = 0, \quad \tilde{\mathbf{u}}, \tilde{T} \text{ are } L\text{-periodic in } x_1, x_{d-1}, \end{aligned}$$

where $(\mathbf{u}, T)$ comes from either (2.4), (2.5), or (2.6).

One of the basic goals of this paper is to show that $\tilde{T} - T$, $\tilde{p} - p \rightarrow 0$, and $\tilde{\mathbf{u}} - \mathbf{u} \rightarrow \mathbf{0}$ as $t \rightarrow \infty$ in the appropriate space for specific conditions on μ and the number of projected modes

N relative to Ra and Pr . This indicates that for specified Rayleigh and Prandtl numbers, one can determine a sufficiently large number of modes and a sufficiently large nudging parameter μ to ensure that the assimilated system $(\tilde{\mathbf{u}}, \tilde{T})$ will asymptotically match the true system.

3. Mathematical background. For the sake of completeness, this section presents some preliminary material and notation commonly used in the mathematical study of hydrodynamic systems, in particular in the study of the NSEs and the Euler equations for incompressible fluids. For more detailed discussion on these topics, we refer the reader to, e.g., [22, 67].

Let $\Omega = [0, L]^{d-1} \times [0, 1]$, where $d = 2, 3$ and we denote the spatial variable $\mathbf{x} = (x_1, \dots, x_d)$. We consider the function spaces

$$(3.1) \quad \begin{aligned} \mathcal{F} &:= \{v \in C^\infty(\Omega) : v \text{ is } L\text{-periodic in } x_j, j = 1, d-1, \text{ and compactly supported in } x_d\}, \\ \mathcal{F}_\sigma^d &:= \{\mathbf{v} \in \mathcal{F}^d : \nabla \cdot \mathbf{v} = 0\}, \end{aligned}$$

$$(3.2) \quad H := \overline{\mathcal{F}}^{L^2}, \quad \mathcal{H} := \overline{\mathcal{F}_\sigma^d}^{L^2},$$

$$(3.3) \quad V := \overline{\mathcal{F}}^{H^1}, \quad \mathcal{V} := \overline{\mathcal{F}_\sigma^d}^{H^1},$$

$$(3.4) \quad W := \overline{\mathcal{F}}^{H^2}, \quad \mathcal{W} := \overline{\mathcal{F}_\sigma^d}^{H^2},$$

where $H^1(\Omega)$ and $H^2(\Omega)$ denote the classical Sobolev class of first-order and second-order weakly differentiable functions over Ω , respectively. We use the notation $X^{\times d}$ to denote the d -fold product of a set X , and $\langle \cdot, \cdot \rangle$ to denote the usual L^2 inner product over Ω ,

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^d \int_{\Omega} u_i(\mathbf{x}) v_i(\mathbf{x}) d\mathbf{x}, \quad \langle f, g \rangle = \int_{\Omega} f(\mathbf{x}) g(\mathbf{x}) d\mathbf{x}$$

for $\mathbf{u} = (u_1, \dots, u_d)$, $\mathbf{v} = (v_1, \dots, v_d) \in \mathcal{H}$, and $f, g \in L^2(\Omega)$. The inner product on $H^k(\Omega)$, $k = 1, 2, \dots$, is given by

$$\langle f, g \rangle_{H^k} := \sum_{|\gamma|=0}^k \langle D^\gamma f, D^\gamma g \rangle,$$

where $\gamma = (\gamma_1, \dots, \gamma_d)$ is a multi-index and $D^\gamma = (\partial_{x_1}^{\gamma_1}, \dots, \partial_{x_d}^{\gamma_d})$. The spaces H, V, W are then endowed with a Hilbert space structure, whose respective inner products are given by

$$(3.5) \quad (f, g)_H := \langle f, g \rangle, \quad (f, g)_V := \langle \nabla f, \nabla g \rangle, \quad (f, g)_W := (f, b)_V + \sum_{|\gamma|=2} \langle D^\gamma f, D^\gamma g \rangle.$$

The spaces $\mathcal{H}, \mathcal{V}, \mathcal{W}$ have analogous Hilbert space structures. We denote by H', V', W' and $\mathcal{H}', \mathcal{V}', \mathcal{W}'$ the dual spaces of H, V, W , and $\mathcal{H}, \mathcal{V}, \mathcal{W}$, respectively. We then have the following continuous injections:

$$(3.6) \quad \begin{aligned} W &\hookrightarrow V \hookrightarrow H \hookrightarrow H' \hookrightarrow V' \hookrightarrow W', \\ \mathcal{W} &\hookrightarrow \mathcal{V} \hookrightarrow \mathcal{H} \hookrightarrow \mathcal{H}' \hookrightarrow \mathcal{V}' \hookrightarrow \mathcal{W}'. \end{aligned}$$

In what follows we will denote the $L^2(\Omega)$ norm by $\|\cdot\|$. For all other Banach spaces X , e.g., $L^p(\Omega)$ for $p \neq 2$, $H^k(\Omega)$, etc., we denote the associated norms explicitly as $\|\cdot\|_X$.

Let

$$(3.7) \quad \{(\lambda_n, \phi_n(\mathbf{x}))\}_{n=1}^{\infty}$$

denote the orthonormal eigenpairs corresponding to the Laplace operator $-\Delta$ on the domain Ω , ordered so that λ_n is a nondecreasing function of n , supplemented with the mixed horizontally periodic-vertically Dirichlet boundary condition as in (3.1). Then each $f \in W$ can be expressed in terms of the eigenfunctions as

$$f(\mathbf{x}, t) = \sum_{n=1}^{\infty} f_n(t) \phi_n(\mathbf{x}), \quad f_n(t) = \langle f(t), \phi_n \rangle = \int_{\Omega} f(\mathbf{x}, t) \phi_n(\mathbf{x}) \, d\mathbf{x},$$

and the eigenfunctions satisfy the orthogonality relation

$$\langle \phi_i, \phi_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

For each $N \geq 0$, define the projections P_N, Q_N by

$$(3.8) \quad P_N(f) = \sum_{n=1}^N f_n \phi_n, \quad Q_N(f) = [I - P_N](f) = \sum_{n=N+1}^{\infty} f_n \phi_n,$$

where I is the identity operator. In other words, P_N is a truncation of the eigenfunction expansion, and Q_N is its orthogonal complement. The orthogonality of the eigenbasis yields the identities

$$(3.9) \quad \langle P_N(f), Q_N(f) \rangle = 0,$$

$$(3.10) \quad \|P_N(f)\|^2 + \|Q_N(f)\|^2 = \|f\|^2$$

for any $f \in H^2(\Omega)$.

We next recall the following well-known a priori estimates for Stokes' equations:

$$(3.11) \quad -\Delta \mathbf{u} + \nabla p = \mathbf{f},$$

$$(3.12) \quad \nabla \cdot \mathbf{u} = 0,$$

$$(3.13) \quad \mathbf{u}|_{x_3=0} = \mathbf{u}|_{x_3=1} = \mathbf{0}, \quad \mathbf{u} \text{ is } L\text{-periodic in } x_1 \text{ and } x_2;$$

see [22] and the drift-diffusion equation (3.15), where the advecting velocity field is divergence free as in, e.g., [67, 73]. We adapt these results here to our notation and setting as follows.

Lemma 3.1. *Let $d = 2, 3$ and $\mathbf{f} \in H^{\times d}$. There exists a unique $\mathbf{u} \in \mathcal{W}$ and (up to constants) $p \in V$ such that (3.11) is satisfied. Moreover, there exists a constant $C = C(\Omega) > 0$ such that*

$$\|\mathbf{u}\|_{H^2} + \|p\|_{H^1} \leq C \|\mathbf{f}\|.$$

Remark 3.2. We will use C below to stand for an absolute constant that may change line to line.

The next lemma follows from the theory of linear transport equations in [26] and a variant of the maximum principle proved in [36]. We will refer to the following notation for the “positive parts” and “negative parts” of a function:

$$(3.14) \quad \psi^+ := \max\{\psi, 0\}, \quad \psi^- := \max\{-\psi, 0\}.$$

Lemma 3.3. *Let $d = 2, 3$ and $\tau > 0$. Let $\mathbf{u} \in L^1(0, \tau; \mathcal{V})$ and $T_0 \in L^\infty(\Omega)$ be a.e. L -periodic in x_1, x_{d-1} and $T_0|_{x_d=0} = 0, T_0|_{x_d=1} = 1$ (in the sense of trace). Suppose that $T \in L^\infty(0, \tau; L^\infty(\Omega)) \cap L^2(0, \tau; H^1(\Omega))$ satisfies*

$$(3.15) \quad \begin{aligned} \partial_t T + \mathbf{u} \cdot \nabla T - \Delta T &= 0, \quad T(\mathbf{x}, 0) = T_0(\mathbf{x}), \\ T|_{x_d=0} &= 1, \quad T|_{x_d=1} = 0, \quad T \text{ is a.e. } L\text{-periodic in } x_1, x_{d-1}, \end{aligned}$$

where the boundary values on $\{x_d = 0\} \cup \{x_d = 1\}$ are interpreted in the sense of trace. Then there exists a constant $C_0 = C_0(\Omega, \|T_0\|) > 0$ and functions $\bar{T}, \eta$ such that $T = \bar{T} + \eta$ and

$$0 \leq \bar{T}(t) \leq 1, \quad \eta = (T - 1)^+ - T^-, \quad \|\eta(t)\| \leq C_0 e^{-t}$$

for all $t > 0$.

For the system (2.6), when $\text{Pr} = \infty$, the velocity field $\mathbf{u}$ is determined by the evolution of T . The well-posedness of this system then follows in a standard way for either dimension $d = 2, 3$, and its solution satisfies the estimates stated in Lemmas 3.1 and 3.3. We formally state this result as the following theorem.

Theorem 3.4. *Let $d = 2, 3$. Suppose that $T_0 \in L^\infty(\Omega)$. Suppose that $T_0 \in L^\infty(\Omega)$ is a.e. L -periodic in x_1, x_{d-1} , and $T_0|_{x_d=0} = 0, T_0|_{x_d=1} = 1$ (in the sense of trace). Then there exists a unique $(\mathbf{u}, T)$ satisfying (2.6) such that*

$$\mathbf{u} \in L^\infty(0, \tau; \mathcal{W}), \quad T \in L^\infty(0, \tau; L^2(\Omega)) \cap L^2(0, \tau; H^1(\Omega)) \cap C_w([0, \tau]; L^2(\Omega))$$

for all $\tau > 0$. Moreover, there exist positive constants $\gamma_0 = \gamma_0(\Omega, \|T_0\|)$ and $C_0 = C_0(\Omega, \|T_0\|)$ such that

$$\text{Ra}^{-1} \|\mathbf{u}(t)\|_{H^2} \leq \gamma_0, \quad \|T(t)\| \leq \gamma'_0$$

for all $t \geq 0$, where $\gamma'_0 := |\Omega|^{1/2} + C_0 e^{-t}$.

We consider a change of variable, denoted by $\theta(\mathbf{x}, t)$, that represents the fluctuation of the temperature around the steady state background temperature profile $1 - x_d$:

$$(3.16) \quad \theta(\mathbf{x}, t) = T(\mathbf{x}, t) - (1 - x_d).$$

The functional setting determined by (3.2)–(3.4) accommodates a rigorous mathematical analysis for the perturbed variable θ . The results derived for θ are then transferred naturally to the desired results for the original variable T . We appeal to [73, 74] for the global

existence and eventual regularity of suitable weak solutions for (2.4)–(2.5), as well as the existence of the “global attractor” for the dynamics, although we will not make explicit use of this fact in this article. We will also say that a solution $(\mathbf{u}, T)$ of (2.4)–(2.5) is *regular on* $[0, \tau]$ if $(\mathbf{u}, \theta) \in L^\infty(0, \tau; \mathcal{V}) \times L^\infty(0, \tau; V)$. If $\tau = \infty$, we say that the solution is a *global regular solution*.

Theorem 3.5 (see [73, 74]). *Let $d = 2, 3$. Recalling the notation (3.2)–(3.4), let $(\mathbf{u}_0, \theta_0) \in \mathcal{H} \times H$ and $T_0 := \theta_0 + (1 - x_d)$.*

(i) *(Global existence of weak solutions.) For any $\tau > 0$, there exists $(\mathbf{u}, T)$ such that*

$$\begin{aligned} \mathbf{u} &\in L^\infty(0, \tau; \mathcal{H}) \cap L^2(0, \tau; \mathcal{V}) \cap C_w([0, \tau]; \mathcal{H}), & \frac{d\mathbf{u}}{dt} &\in L^{4/3}(0, \tau; \mathcal{V}'), \\ \theta &\in L^\infty(0, \tau; H) \cap L^2(0, \tau; V) \cap C_w([0, \tau]; H), & \frac{d\theta}{dt} &\in L^{4/3}(0, \tau; V'), \end{aligned}$$

where θ is related to T by the relation (3.16). This $(\mathbf{u}, T)$ satisfies (2.4)–(2.5) in the usual weak sense and maintains the initial condition $(\mathbf{u}_0, T_0)$. Moreover, the following energy inequality and maximum principle are satisfied for all $t \leq \tau$:

$$\begin{aligned} (3.17) \quad & \frac{1}{\text{Pr}} \|\mathbf{u}(t)\|^2 + 2 \int_0^t \|\nabla \mathbf{u}(s)\|^2 ds \leq \frac{1}{\text{Pr}} \|\mathbf{u}_0\|^2 + 2 \text{Ra} \int_0^t \langle \theta(s), u_3(s) \rangle ds, \\ & \|(T - 1)^+(t)\|^2 + 2 \int_0^t \|\nabla (T - 1)^+(s)\|^2 ds \leq \|(T_0 - 1)^+\|^2, \\ & \|T^-(t)\|^2 + 2 \int_0^t \|\nabla T^-(s)\|^2 ds \leq \|(T_0)^-\|^2, \end{aligned}$$

where $(T - 1)^+, T^-$ are defined as in (3.14).

(ii) *(Eventual regularity of weak solutions.) There exists a constant $K_0 > 0$ such that if $\text{Pr Ra}^{-1} \geq K_0$, then there exists a time $\tau^* > 0$ for which all suitable weak solutions corresponding to $(\mathbf{u}_0, \theta_0) \in \mathcal{H} \times H$ become regular solutions on $[\tau^*, \infty)$. In particular, when $d = 3$, there exist constants $\kappa_1, \kappa_2, \kappa_3 > 0$ and $\kappa_0'', \kappa_1'', \kappa_2'' > 0$, depending on Ω , such that*

$$(3.18) \quad \begin{aligned} \|\mathbf{u}(t)\|_{H^1} &\leq \kappa_1 \text{Ra}, & \|\mathbf{u}(t)\|_{H^2} &\leq \kappa_2 \text{Ra}^{5/2}, & \|\partial_t \mathbf{u}(t)\| &\leq \kappa_3 \text{Ra}^{7/2}, \\ \|\theta(t)\| &\leq \kappa_0'', & \|\theta(t)\|_{H^1} &\leq \kappa_1'' \text{Ra}^{5/2}, & \|\theta(t)\|_{H^2} &\leq \kappa_2'' \text{Ra}^8 \end{aligned}$$

for all $t \geq \tau^*$.

For the remainder of the article, we will therefore assume that $(\mathbf{u}, T), (\tilde{\mathbf{u}}, \tilde{T})$ have evolved for a sufficiently long time so that $(\mathbf{u}, T)$ is a regular solution to either (2.4)–(2.5) or (2.6). Physically speaking, the setup of our study assumes that reality is represented exactly by a solution to (2.4)–(2.5) or (2.6) and that we have been observing the system after the point in time at which it has become globally regular. Thus, for the purposes of our analysis, we will henceforth make the following standing hypotheses for the remainder of the paper.

Standing hypotheses. Let $d = 2, 3$ and $\text{Ra} \geq 1$. Let $\gamma_0 > 0$ be the constants from Theorem 3.4. When $\text{Pr} = \infty$, we assume the following:

- (I1) T_0 is a.e. periodic in x_1, x_{d-1} and $T_0|_{x_d=0} = 0$, $T_0|_{x_d=1} = 1$ (in the sense of trace);
- (I2) $T_0 \in L^\infty(\Omega)$;
- (I3) $(\mathbf{u}, T)$ is the unique global solution to (2.6) corresponding to T_0 and guaranteed by Theorem 3.4. In particular, $(\mathbf{u}, T)$ satisfies

$$\text{Ra}^{-1} \|\mathbf{u}(t)\|_{H^2} \leq \gamma_0, \quad \|T(t)\| \leq \gamma'_0$$

for all $t > 0$.

On the other hand, let $K_0 > 0$, $\kappa_1, \kappa_2, \kappa_3 > 0$, and $\kappa''_0, \kappa''_1, \kappa''_2 > 0$ be the constants in Theorem 3.5(ii).

When $\text{Pr} < \infty$, we assume the following:

- (F1) $\text{Pr Ra}^{-1} \geq K_0$;
- (F2) $(\mathbf{u}, T)$ is the unique regular solution to (2.4), (2.5);
- (F3) when $d = 3$, $\mathbf{u}(t)$ satisfies

$$\|\mathbf{u}(t)\|_{H^1} \leq \kappa_1 \text{Ra}, \quad \|\mathbf{u}(t)\|_{H^2} \leq \kappa_2 \text{Ra}^{5/2}, \quad \|\partial_t \mathbf{u}(t)\| \leq \kappa_3 \text{Ra}^{7/2}$$

for all $t \geq 0$;

- (F4) when $d = 3$, T satisfies

$$\|T(t)\| \leq \kappa'_0, \quad \|T(t)\|_{H^1} \leq \kappa'_1 \text{Ra}^{5/2}, \quad \|T(t)\|_{H^2} \leq \kappa'_2 \text{Ra}^8$$

for all $t \geq 0$, where $\kappa'_j = \kappa''_j + 2|\Omega|^{1/2} + 1$, $j = 0, 1, 2$.

Remark 3.6. Since we have nondimensionalized our variables, we point out that although the bounds in (F3), (F4) are derived for the case $d = 3$, they are also valid for the case $d = 2$, up to constants, provided that $\text{Ra} \geq 1$, which have have assumed as one of our standing hypotheses.

4. Infinite Prandtl assimilation. We will first treat the data assimilation problem for the infinite Prandtl number system (2.6), (2.8). Hence, throughout this section we assume that $\text{Pr} = \infty$, i.e., (I1)–(I3) holds.

A rigorous mathematical analysis is performed in dimensions $d = 2, 3$, while the numerical component of our studies are carried out for $d = 2$. Due to the structure of the nudged system (2.8), we do not have a maximum principle. Instead, we require only that (2.8) have a well-defined solution in the weak sense, which one can do by establishing that the differences $\mathbf{w} = \tilde{\mathbf{u}} - \mathbf{u}$, $S = \tilde{T} - T$ satisfy their respective evolution equations in the weak sense. Since one of the relevant a priori estimates to this end is performed below for the proof of Theorem 4.2, we simply state this result as the following theorem.

Theorem 4.1. *Let $\mu > 0$ and $\{\lambda_n\}_{n=1}^\infty$ be as in (3.7). Let $\tilde{T}_0 \in L^2(\Omega)$ a.e. in Ω such that $\tilde{T}_0$ is a.e. L -periodic in x_1, x_{d-1} with $\tilde{T}_0|_{x_d=0} = 0$, $\tilde{T}_0|_{x_d=1} = 1$ (in the sense of trace). Suppose $N > 0$ satisfies $\frac{1}{4}\lambda_N \geq \mu$. Then there exists a unique $(\tilde{\mathbf{u}}, \tilde{T})$ satisfying (2.8) in the weak sense such that*

$$\tilde{\mathbf{u}} \in L^\infty(0, \tau; \mathcal{W}), \quad \tilde{T} \in L^\infty(0, \tau; L^2(\Omega)) \cap L^2(0, \tau; H^1(\Omega)) \cap C_w([0, \tau], L^2(\Omega))$$

for all $\tau > 0$.

4.1. Synchronization. We are now ready to state and prove the main theorem of this section. We show that (2.6) and (2.8) synchronize under certain conditions detailed below.

Theorem 4.2. *Let $\mu > 0$ and $\{\lambda_n\}_{n=1}^\infty$ be as in (3.7). Let $N > 0$ satisfy $\frac{1}{4}\lambda_N \geq \mu$ and $\tilde{T}_0$ be given as in Theorem 4.1. Let $(\tilde{\mathbf{u}}, \tilde{T})$ be the corresponding unique solution to (2.8) guaranteed by Theorem 4.1. There exists a constant $C_0 = C_0(\Omega, \|T_0\|) > 0$ such that if*

$$(4.1) \quad \mu \geq C_0 \text{Ra}^2,$$

then for all $t \geq 0$,

$$(4.2) \quad \|(\tilde{T} - T)(t)\|^2 + \text{Ra}^{-2} \|(\tilde{\mathbf{u}} - \mathbf{u})(t)\|_{H^2}^2 \leq C_1 e^{-\mu t}$$

for $C_1 = \|\tilde{T}_0 - T_0\|$.

Proof. Let $S = \tilde{T} - T$, $\mathbf{w} = \tilde{\mathbf{u}} - \mathbf{u}$, and $q = \tilde{p} - p$. Subtracting (2.6) from (2.8) yields the system

$$(4.3) \quad \begin{aligned} -\Delta \mathbf{w} &= -\nabla q + \text{Ra} \mathbf{e}_3 S, \quad \nabla \cdot \mathbf{w} = 0, \\ \partial_t S + \tilde{\mathbf{u}} \cdot \nabla S + \mathbf{w} \cdot \nabla T - \Delta S &= -\mu P_N S, \\ \mathbf{w}|_{x_3=0} &= \mathbf{w}|_{x_3=1} = \mathbf{0}, \quad S|_{x_3=0} = S|_{x_3=1} = 0, \quad \mathbf{w}, S \text{ are periodic in } x_1 \text{ and } x_2 \end{aligned}$$

with the initial condition $S(\mathbf{x}, 0) = \tilde{T}_0(\mathbf{x}) - T_0(\mathbf{x})$. The momentum equation in (4.3) satisfies Lemma 3.1, so there exists a constant $C > 0$ such that

$$(4.4) \quad \|\mathbf{w}\|_{H^2} + \|q\|_{H^1} \leq C \text{Ra} \|S\|.$$

Therefore, to establish (4.2), it is sufficient to show that $S \rightarrow 0$ with an exponential rate in $L^2(\Omega)$.

Upon multiplying the S equation in (4.3) by S and integrating over Ω , we obtain

$$(4.5) \quad \frac{1}{2} \frac{d}{dt} \|S\|^2 + \|\nabla S\|^2 + \mu \|P_N(S)\|^2 = - \int_{\Omega} (\mathbf{w} \cdot \nabla T) S \, d\mathbf{x}.$$

Assume that μ is chosen sufficiently large so that (4.1) holds, where $C_0 > 0$ is, as of yet, unspecified. After integrating by parts, an application of the Sobolev embedding $H^2(\Omega) \hookrightarrow L^\infty(\Omega)$, (4.4), and then (I3) of the standing hypotheses implies

$$(4.6) \quad \begin{aligned} \left| \int_{\Omega} (\mathbf{w} \cdot \nabla T) S \, d\mathbf{x} \right| &\leq \|\mathbf{w}\|_{L^\infty} \|T\| \|\nabla S\| \leq C \text{Ra} \|S\| \|T\| \|\nabla S\| \\ &\leq \frac{1}{2} \|\nabla S\|^2 + C \text{Ra}^2 \|T\|^2 \|S\|^2 \leq \frac{1}{2} \|\nabla S\|^2 + C \text{Ra}^2 \gamma'_0 \|S\|^2. \end{aligned}$$

On the other hand, due to (3.10) and the inverse Poincaré inequality, and since $N > 0$ satisfies $\frac{1}{4}\lambda_N \geq \mu$, it follows that

$$(4.7) \quad \begin{aligned} \frac{1}{2} \|\nabla S\|^2 + \mu \|P_N S\|^2 &= \frac{1}{2} \|\nabla S\|^2 - \mu \|Q_N S\|^2 + \mu \|S\|^2 \\ &\geq \left(\frac{1}{2} - \frac{\mu}{\lambda_N} \right) \|\nabla S\|^2 + \mu \|S\|^2 \geq \mu \|S\|^2. \end{aligned}$$

Joining (4.5) with (4.6) and (4.7) and combining like terms, we arrive at

$$\frac{d}{dt}\|S\|^2 + 2(\mu - C\text{Ra}^2\gamma'_0 t)\|S\|^2 \leq 0.$$

Finally, by Grönwall's inequality and the second condition in (4.1), we deduce that

$$(4.8) \quad \|S(t)\|^2 \leq \|S(0)\|^2 \exp(-2\mu t + C\text{Ra}^2\gamma'_0 t) \leq \|S(0)\|^2 \exp(-\mu t),$$

which completes the proof upon choosing $C_0 = C\gamma'_0$. ■

Theorem 4.2 shows that given Ra and reasonable boundary conditions T_0 and $\tilde{T}_0$, we can always choose μ and N large enough so that $(\mathbf{u}, T)$ and $(\tilde{\mathbf{u}}, \tilde{T})$ eventually match in the infinite-time limit. In the next subsection, we use numerical simulations to verify that this is indeed the case.

4.2. Numerical results at infinite Prandtl number. Rather than focusing on a handful of highly turbulent 3D simulations, for the same computational cost we consider more detailed and extensive simulations in the 2D setting where $\Omega = [0, L] \times [0, 1]$ with coordinates $\mathbf{x} = (x_1, x_3)$ in order to search through the relevant parameters. We simulate (2.6) and (2.8) using a stream function formulation with *Dedalus* [17], a Python package that uses pseudospectral methods to solve partial differential equations on spectrally representable domains. All of the simulations in this section and all that follow are completed with a 4-stage 3rd order Runge–Kutta implicit-explicit time stepping scheme that treats the linear terms implicitly and the nonlinear terms explicitly (see [17] for details on the selected time-stepping mechanism). The selected time-stepping algorithm employs an adaptive time step that is based on a default Courant–Friedrichs–Lewy condition taken within the stability region of the 4-stage, 3rd order Runge–Kutta scheme, and using the horizontal and vertical velocities as the “winds.” We note that as pointed out in [63], Runge–Kutta methods are not completely valid for this type of data assimilation as there is no information available between time steps. Thus, the simulations presented here are only a proof of concept to demonstrate the usability of the nudging algorithm. Each simulation is run with $L = 4$ and 256 Fourier grid points in the horizontal and 128 Chebyshev points in the vertical with a standard 3/2 dealiasing. All of the reported simulations were also performed at 384×192 resolution with little to no effect on the results. All simulations were run until the time-averaged Nusselt number and other pertinent statistics were temporally well-converged (meaning that the average over the last half of the total integration period was within 5% of the same average over the last 1/4 of the integration period) for all cases considered here. For the sake of reproducibility, the Python code that produces the data and figures reported here is publicly available at <https://github.com/shanemcq18/DAiLPRBCfTM-Paper>.

Choosing the initial conditions T_0 and $\tilde{T}_0$ in our numerical experiments is nontrivial. For T_0 , we have two options.

- Set $T_0(x_1, x_3) = 1 - x_3 + \varepsilon(x_1, x_3)$ for some small perturbation function $\varepsilon : \Psi \rightarrow \mathbb{R}$. This begins the simulation close to the conductive state $1 - x_3$, which—though not physically relevant—is a suitable starting point for initial experiments.
- Load T_0 from the final state of a previous simulation with similar parameters. If this previous simulation has run long enough, T will thus begin in a reasonable state for the new set of parameters.

Finding a suitable choice for $\tilde{T}_0$ is less obvious. Setting $\tilde{T}_0(x_1, x_3) = 1 - x_3$ assumes no prior knowledge of T and results in an overly stiff initialization due to the strength of the initial temperature difference, and is therefore numerically impractical. In addition, initiating T at this conductive state ignores observations taken at the initial time $t = 0$, which would not be advantageous in practice. On the other hand, setting $\tilde{T}_0 = P_N(T_0)$ for large N results in a very weak temperature difference so that (2.6) and (2.8) are nearly synchronized at $t = 0$. As a compromise, we set $\tilde{T}_0$ as a low-mode projection of T_0 such as $\tilde{T}_0 = P_4(T_0)$. This permits some initial knowledge of the true state without driving it directly to the “truth” immediately. Preliminary investigation indicates that our results are independent of the choice $N = 4$, as $N = 6, 8$, etc. yield similar results with smaller initial error for $N > 4$ but qualitatively the same convergence rate.

Values of the Rayleigh number Ra that are of interest for mantle convection are typically between 10^7 and 10^8 [68, 13]. However, the greater the value of Ra , the more computationally expensive the simulation, and it is numerically unstable to start T in a state such as $T_0(x_1, x_3) \approx 1 - x_3$ at large Ra . We therefore select logarithmically spaced Ra values from 10^4 to 10^9 and run a simulation for each Ra , one after another. For the very first simulation ($Ra = 10^4$), we set $T_0(x_1, x_3) = 1 - x_3 + \varepsilon(x_1, x_3)$ where $\varepsilon(x_1, x_3) = 10^{-4} \sin(kx_1) \sin(2\pi x_3)$ and $k = 3.117$ (near the first unstable mode that bifurcates away from conduction at onset), and evolve the systems forward until reaching a statistical equilibrium state as described above (defined by the convergence of temporal averages of the Nusselt number). Thereafter, we set T_0 as the final T from the previous simulation (at a lower value of Ra) and incrementally increase Ra . This yields realistic initial conditions for any given Ra in the range considered. These “initial conditions” are saved and used for the finite Prandtl setting considered in the following sections as well. Figure 1 shows an example of T_0 and $\tilde{T}_0$, together with subsequent end states of T and $\tilde{T}$ for $Ra \approx 5.22 \times 10^7$.

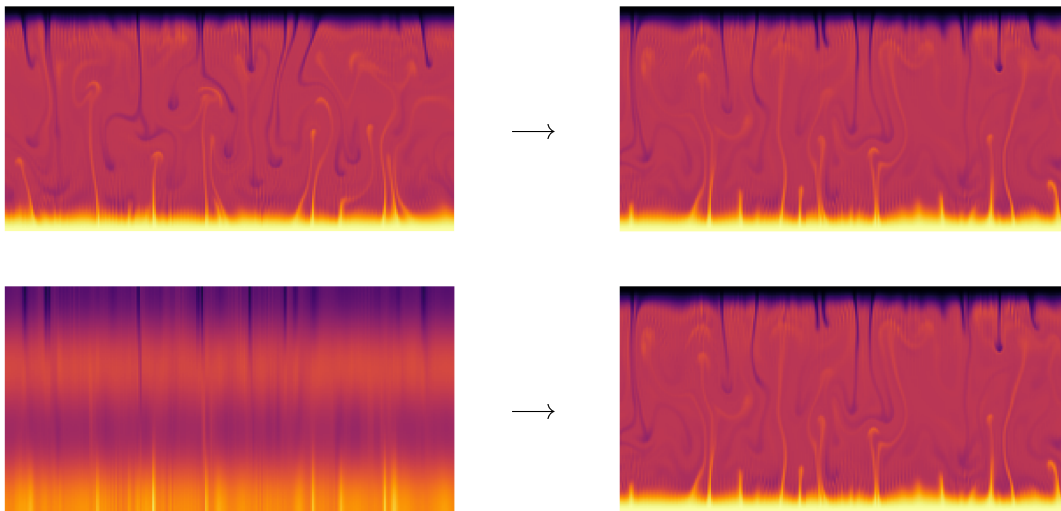


Figure 1. Typical “before” and “after” snapshots of a successfully assimilated setup. On the top, the initial condition T_0 compared with T at the end of a simulation. On the bottom, the initial projected temperature $\tilde{T}_0 = P_4(T_0)$ and the eventual $\tilde{T}$ at the end of the simulation. Note that the postsimulation T and $\tilde{T}$ appear identical to the naked eye. This particular simulation uses $Ra \approx 5.22 \times 10^7$, $\mu = 12,700$, and $N = 32$.

To measure the synchronization of (2.6) and (2.8), we keep track of the $L^2(\Omega)$, $H^1(\Omega)$, and $H^2(\Omega)$ norms of $\tilde{T} - T$ and $\tilde{\mathbf{u}} - \mathbf{u}$ throughout the simulation. Each error is normalized by dividing by the norms of the truth system's variables. For example, to measure the temperature difference in $L^2(\Omega)$, we compute $\frac{\|(\tilde{T}-T)(t)\|}{\|T(0)\|}$ at each simulation time t . For simplicity, we write these errors without the denominators in all that follows.

Figure 2 shows how the tracked error norms decrease in the simulation from Figure 1. Though the analysis indicates that these norms should converge to $\varepsilon_{\text{machine}} \approx 10^{-16}$, with the current setup, we are not able to see such a convergence. It appears that the cause of this discrepancy at these moderately large Rayleigh numbers is due to the Runge–Kutta (RK) time-stepping: as described earlier, a RK time-stepping algorithm steps forward by first evaluating the evolution equation at times in between the final steps wherein the “true” system is not available. For $\text{Ra} < 10^6$, this effect isn't noticeable, but for $\text{Ra} > 10^6$, the convergence using an RK method is less than desirable, while we found that near $\varepsilon_{\text{machine}}$ convergence occurs when using an Adams–Bashforth (AB) time-stepper. For the sake of computational time (high order AB methods have much smaller stability regions than RK), we have continued to use the RK method here despite this drawback but plan to address this issue in a later study. Changing the time-stepping algorithm clearly has an effect on the synchronization (for $\text{Ra} \sim 5 \times 10^6$ a 3rd order AB method yields synchronization within 10^{-15} , whereas the RK method gives at best 10^{-14}), but it does not appear to affect the *rate* of synchronization, which is more of the focus here.

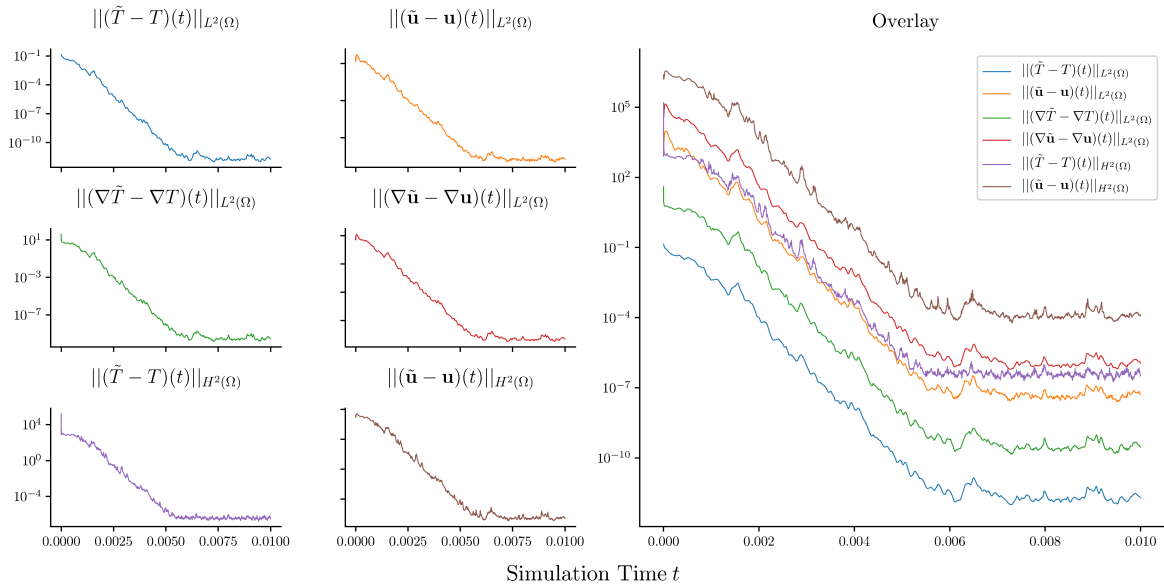


Figure 2. Synchronization in various norms for $\text{Ra} \approx 5.22 \times 10^7$, $\mu = 12,700$, and $N = 32$ (the same conditions used in Figure 1). The temperature and velocity differences decrease exponentially until flattening out at sufficiently small values. The temperature difference is the smallest, which seems to be a consequence of using temperature-only observations for the assimilation.

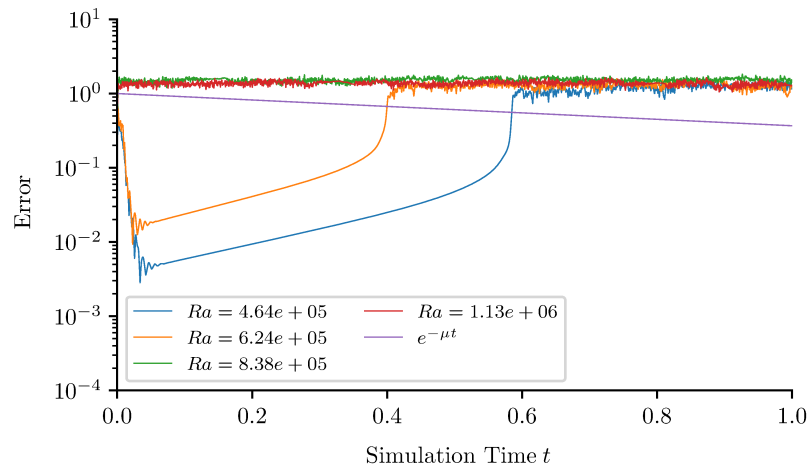


Figure 3. Convergence (or lack thereof) of $\|(\tilde{\mathbf{u}} - \mathbf{u})(t)\|_{H^2} + \|(\tilde{T} - T)(t)\|_{H^2}$ with $\mu = 1$ and $N = 32$ for various values of Ra . Note that even though a transient period of near-synchronization appears for lower values of Ra , actual synchronization doesn't occur in any case for $\mu = 1$. To demonstrate the lack of synchronization, a comparison of $e^{-\mu t}$ is included which would provide an estimate for the convergence were it to occur. The results presented below for much larger values of μ will have a far shorter expected convergence time (as dictated by $e^{-\mu t}$ again), so the time interval displayed is significantly shorter.

4.2.1. Dependencies of nudging parameter. Previous studies, [6, 28], of data assimilation for RB convection fixed the value of the nudging parameter at $\mu = 1$ in their experiments and instead focused on the effects of other parameters, such as the number of projected modes. However, a sufficient condition in Theorem 4.2 is that μ be proportional to Ra^2 , while it is not a necessary condition; synchronization with $\mu = 1$ is not expected as Ra increases. Indeed, Figure 3 shows that for Ra in the range of 4×10^7 to 2×10^8 , using $\mu = 1$ results in either extremely slow convergence or in no convergence at all.

To explore the relationship between Ra , μ , and N needed for synchronization to occur, we fix two of the parameters at a time and run several simulations with various values for the remaining parameter. First, we fix Ra , set $N = 32$, and vary μ until synchronization can be observed within 0.005 units of simulation time (a simulation of this length usually requires several thousand iterations). An a posteriori choice of this cutoff time is motivated by the inherent time scale dictated by the choice of μ , i.e., assuming a unitary prefactor on the decay estimate ($e^{-\mu t}$), and hence setting $e^{-\mu T^*} \approx \epsilon_{\text{machine}} = 10^{-16}$ for the minimal value $\mu = 8000$ shown in Figure 4 will yield $T^* \approx 0.0046$. Table 1 records the smallest value of μ where synchronization was observed in this time frame; Figure 4 shows the convergence rates for a few different μ values when $Ra \approx 3.89 \times 10^7$ and $Ra \approx 7.02 \times 10^7$.

Given the requirement (4.1) from Theorem 4.2, it is surprising to discover that—based on Table 1—the relationship between μ and Ra is more linear than quadratic. Indeed, a least squares fit of the data to a general quadratic of the form $f(x) = a + bx + cx^2$ yields coefficients of $a \approx 4.03 \times 10^3$, $b \approx 1.77 \times 10^{-4}$, and $c \approx -1.37 \times 10^{-13}$ (essentially $c = 0$). See Figure 5.

4.2.2. Relation between relaxation and number of observables. The relaxation parameter μ is a system parameter without a clear physical interpretation. The number of projected

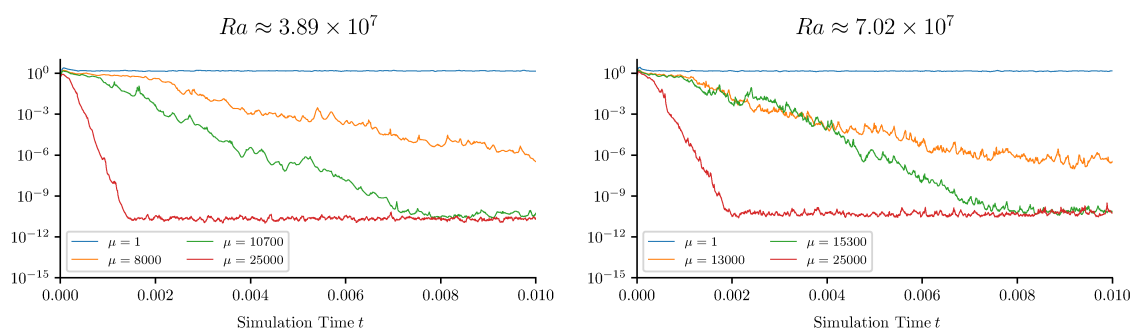


Figure 4. $\|(\tilde{\mathbf{u}} - \mathbf{u})(t)\|_{H^2} + \|(\tilde{T} - T)(t)\|_{H^2}$ with $N = 32$ and various μ for two different values of Ra . We begin to see satisfactory convergence around $\mu = 10,700$ and $\mu = 15,300$ for $Ra \approx 3.89 \times 10^7$ and $Ra \approx 7.02 \times 10^7$, respectively.

Table 1

Minimal values of μ that result in synchronization within about 0.005 units of simulation time for the given Ra with $N = 32$. These values represent the edge of what works when $N = 32$: a lower μ may not stimulate convergence, but a larger μ will. See Figure 5 for a quadratic least-squares fit of this data.

Ra	μ
1.1937766×10^7	6,300
1.6037187×10^7	6,500
2.1544346×10^7	7,900
2.8942661×10^7	9,400
3.8881551×10^7	10,700
5.2233450×10^7	12,700
7.0170382×10^7	15,300
9.4266845×10^7	19,400
1.26638017×10^8	24,900
1.70125427×10^8	29,900

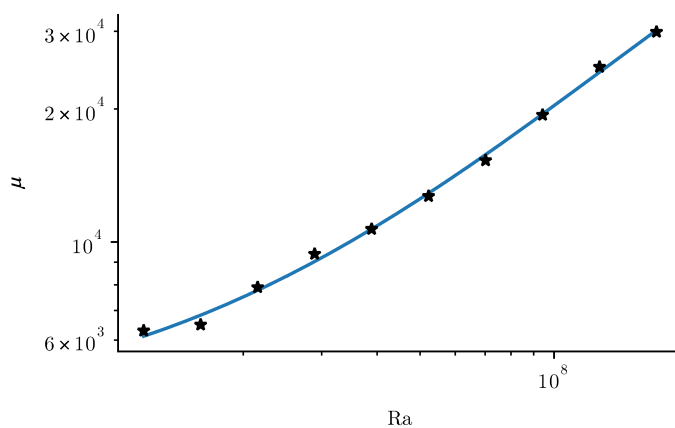


Figure 5. Quadratic (but nearly linear) least-squares fit for the data in Table 1.

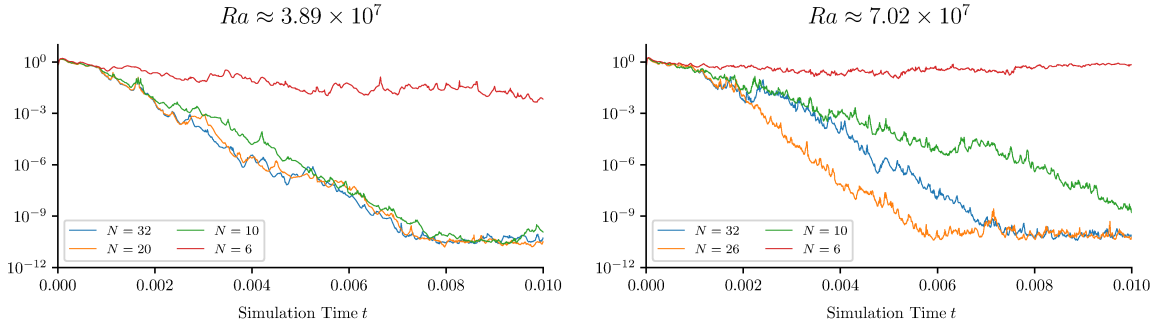


Figure 6. $\|(\tilde{\mathbf{u}} - \mathbf{u})(t)\|_{H^2} + \|(\tilde{T} - T)(t)\|_{H^2}$ for two different values of Ra , with μ as listed in Tables 1 and 2 and for various values of N . For both $Ra \approx 3.89 \times 10^7$ and $Ra \approx 7.02 \times 10^7$, $N = 10$ appears to be the fewest number of modes that results in synchronization.

Table 2

Minimal values of N that result in synchronization within about 0.01 units of simulation time for the given Ra and μ . The constant value Δt is the average time step (averaged over the last half of the integration interval) for each simulation. Note that $\mu \Delta t$ is nearly constant between the different simulations presented here, indicating that the same relative nudging is applied at each time step despite the differences between each simulation. In general, as long as an appropriate μ is chosen, $N = 16$ is sufficient for physically relevant values of Ra .

Ra	μ	$\mu \Delta t$	N
1.1937766×10^7	6,300	6.05×10^{-3}	6
1.6037187×10^7	6,500	5.35×10^{-3}	8
2.1544346×10^7	7,900	5.32×10^{-3}	8
2.8942661×10^7	9,400	5.31×10^{-3}	8
3.8881551×10^7	10,700	3.35×10^{-3}	10
5.2233450×10^7	12,700	4.96×10^{-3}	10
7.0170382×10^7	15,300	3.33×10^{-3}	10
9.4266845×10^7	19,400	5.22×10^{-3}	10
1.26638017×10^8	24,900	5.54×10^{-3}	12
1.70125427×10^8	29,900	5.49×10^{-3}	14

Fourier modes N , on the other hand, indicates the amount of data that is “visible” to the assimilating system. Therefore, a lower bound on N represents how much data is required in order to maintain an accurate model. Fixing μ as given by Table 1, we decrease N to see how it affects synchronization. See Figure 6. As expected, the fewer modes retained in the projection, the slower synchronization is achieved, and if a sufficiently low number of modes ($N \leq 6$) is observed, then synchronization never occurs. In Table 2 we summarize our findings concerning the number of modes N needed as a function of Ra (given a suitably large choice for μ).

4.2.3. Summary. The numerical simulations verify that over physically relevant values of Ra , we can pick μ large enough to nudge (2.6) toward (2.8); in particular, the experiments confirm Theorem 4.2 and show that the conditions stated therein are stronger than necessary. Furthermore, synchronization can be obtained with a relatively low number of modes even for large Ra , as long as μ is chosen large enough. We took the approach of varying μ first and then N , but it is just as feasible to vary μ for a fixed Ra and N .

5. Finite Prandtl assimilation. We now turn to the case when Pr is finite but large. Thus, throughout this section, we will assume that $\text{Pr} < \infty$ and that (F1)–(F4) hold.

As in section 4, we first state the relevant well-posedness result for the nudged equation. We then rigorously establish synchronization of the nudged signal with the true signal, then proceed with a numerical study for the 2D setting.

It is not immediately clear how to properly adapt the argument for (2.6)–(2.8) in [67, 73] to establish a maximum principle (see Lemma 3.3) for the assimilating variable $\tilde{T}$ due to the presence of the additional term $-\mu P_N(\tilde{T} - T)$ in (2.8) (see [50] for a case where it is necessary to develop a maximum principle in spite of this term); the maximum principle would then, in turn, be used to establish global well-posedness of (2.6)–(2.8). However, since we are assuming $(\mathbf{u}, T)$ is a regular solution to (2.4)–(2.5), it is straightforward to verify global existence and uniqueness in the case $d = 2$ for the associated nudged system by instead considering the corresponding system for the difference $(\mathbf{w}, S)$, where $\mathbf{w} = \tilde{\mathbf{u}} - \mathbf{u}$ and $S = \tilde{T} - T$. In this setting, we need only appeal to a maximum principle for T , rather than for $\tilde{T}$. The well-posedness of the nudged system then follows in a standard fashion, so that we refer the reader to [29] for the appropriate details. For $d = 3$, on the other hand, global existence of strong solutions or uniqueness of weak solutions is an outstanding open problem, whether the Prandtl number is large or not. Nevertheless, for our purposes, it suffices to have only *global existence of weak solutions* for the nudged equation, as we will show that any solution of the nudged equation will eventually converge to the true solution corresponding to the observed data. For this reason, we state the global existence of weak solutions and local existence and uniqueness of regular solutions for $d = 2, 3$ together, although, as we have just discussed, global well-posedness holds in $d = 2$; we refer the reader to [73] for relevant details. Note that for the same reason as the $d = 2$ case, it is important to consider the corresponding systems for the differences $\mathbf{w}$ and S . Note also that we refer below to a *Leray–Hopf weak solution* as a weak solution that satisfies the energy inequality in (3.17) for the corresponding velocity field.

Theorem 5.1. *Let $\mu > 0$ and $\{\lambda_n\}_{n=1}^\infty$ be as in (3.7). Let $(\tilde{\mathbf{u}}_0, \tilde{T}_0 - (1 - x_d)) \in \mathcal{H} \times H$. Suppose $N > 0$ satisfies $\frac{1}{4}\lambda_N \geq \mu$. Then there exists a Leray–Hopf weak solution $(\tilde{\mathbf{u}}, \tilde{T})$ satisfying (2.7) such that*

$$(5.1) \quad \begin{aligned} \tilde{\mathbf{u}} &\in L^\infty(0, \tau; \mathcal{H}) \cap L^2(0, \tau; \mathcal{V}) \cap C_w([0, \tau]; \mathcal{H}), & \frac{d\mathbf{u}}{dt} &\in L^{4/3}(0, \tau; \mathcal{V}') \\ \tilde{T} - (1 - x_d) &\in L^\infty(0, \tau; H) \cap L^2(0, \tau; V) \cap C_w([0, \tau]; H), & \frac{dT}{dt} &\in L^{4/3}(0, \tau; V') \end{aligned}$$

for all $\tau \geq 0$.

Moreover, if $(\tilde{\mathbf{u}}_0, \tilde{T}_0 - (1 - x_d)) \in \mathcal{V} \times V$, then there exists $\tau_0 > 0$ and a unique solution $(\tilde{\mathbf{u}}, \tilde{T})$ satisfying (2.7), such that

$$(\tilde{\mathbf{u}}, \tilde{T} - (1 - x_d)) \in L^\infty(0, \tau; \mathcal{V} \times L^\infty(0, \tau), V)$$

for all $\tau < \tau_0$.

5.1. Synchronization. We will establish the finite Prandtl analog to Theorem 4.2. For this, it will be convenient to establish the following stability estimate first.

Lemma 5.2. *Assuming the conditions of Theorem 5.1, with $(\tilde{\mathbf{u}}_0, \tilde{T}_0 - (1 - x_d)) \in \mathcal{V} \times V$, there exists a constant $C_0 = C_0(\Omega) > 0$ such that if*

$$(5.2) \quad \mu \geq \left(\frac{1}{2} + \text{Ra}^2 \right) \text{Pr},$$

then

$$(5.3) \quad \begin{aligned} & \|(\tilde{\mathbf{u}} - \mathbf{u})(t)\|^2 + \|(\tilde{T} - T)(t)\|^2 \\ & \leq \exp \left(-t \text{Pr} + C_0 \int_0^t \left(\frac{1}{\text{Pr}^3} \|\nabla \mathbf{u}(s)\|^4 + \frac{1}{\text{Pr}} \|T(s)\|_{L^6}^4 \right) ds \right) (\|\tilde{\mathbf{u}}_0 - \mathbf{u}_0\|^2 + \|\tilde{T}_0 - T_0\|^2) \end{aligned}$$

holds for $t \geq 0$.

Proof. Similar to the analysis performed for Theorem 4.2, we consider the solution $(\mathbf{w}, S)$ to the difference system

$$(5.4) \quad \begin{aligned} & \frac{1}{\text{Pr}} [\partial_t \mathbf{w} + (\tilde{\mathbf{u}} \cdot \nabla) \mathbf{w} + (\mathbf{w} \cdot \nabla) \tilde{\mathbf{u}}] - \Delta \mathbf{w} = -\nabla q + \text{Ra} \mathbf{e}_3 S, \quad \nabla \cdot \mathbf{w} = 0, \\ & \partial_t S + \tilde{\mathbf{u}} \cdot \nabla S + \mathbf{w} \cdot \nabla T - \Delta S = -\mu P_N S, \\ & \mathbf{w}|_{x_3=0} = \mathbf{w}|_{x_3=1} = \mathbf{0}, \quad S|_{x_3=0} = S|_{x_3=1} = 0, \quad \mathbf{w}, S \text{ are } L\text{-periodic in } x_1 \text{ and } x_2. \end{aligned}$$

Multiplying the second equation by S integrating and arguing as in (4.7), we obtain

$$(5.5) \quad \frac{d}{dt} \|S\|^2 + 2 \|\nabla S\|^2 + 2\mu \|S\|^2 \leq 2 \left| \int \mathbf{w} \cdot \nabla T S dx \right| =: I.$$

On the other hand, an analogous energy calculation for $\mathbf{w}$ together with the Poincaré inequality yield

$$(5.6) \quad \begin{aligned} & \frac{d}{dt} \|\mathbf{w}\|^2 + \text{Pr} \|\nabla \mathbf{w}\|^2 + \text{Pr} \|\mathbf{w}\|^2 \leq \frac{d}{dt} \|\mathbf{w}\|^2 + 2 \text{Pr} \|\nabla \mathbf{w}\|^2 \\ & \leq 2 \left| \int \mathbf{w} \cdot \nabla \tilde{\mathbf{u}} \cdot \mathbf{w} dx \right| + 2 \text{Ra} \text{Pr} \left| \int \mathbf{e}_3 S \cdot \mathbf{w} dx \right| =: II + III. \end{aligned}$$

We estimate the right hand side of (5.5) using Hölder's inequality, the Gagliardo–Nirenberg interpolation inequality, and Young's inequality to obtain

$$\begin{aligned} I & \leq 2 \|\mathbf{w}\|_{L^3} \|\nabla S\| \|T\|_{L^6} \leq C \|\mathbf{w}\|^{1/2} \|\nabla \mathbf{w}\|^{1/2} \|\nabla S\| \|T\|_{L^6} \\ & \leq C \|T\|_{L^6}^2 \|\mathbf{w}\| \|\nabla \mathbf{w}\| + \|\nabla S\|^2 \leq \frac{C}{\text{Pr}} \|T\|_{L^6}^4 \|\mathbf{w}\|^2 + \frac{\text{Pr}}{2} \|\nabla \mathbf{w}\|^2 + \|\nabla S\|^2. \end{aligned}$$

Next, for II in (5.6), we have from Hölder's inequality, the Sobolev embedding $L^6(\Omega) \hookrightarrow H^1$, Gagliardo–Nirenberg interpolation inequality, and Young's inequality that

$$\begin{aligned} II & \leq 2 \|\mathbf{w}\|_{L^6} \|\nabla \tilde{\mathbf{u}}\| \|\mathbf{w}\|_{L^3} \leq C \|\nabla \mathbf{w}\|^{3/2} \|\mathbf{w}\|^{1/2} \|\nabla \tilde{\mathbf{u}}\| \\ & \leq \frac{\text{Pr}}{2} \|\nabla \mathbf{w}\|^2 + \frac{C}{\text{Pr}^3} \|\nabla \tilde{\mathbf{u}}\|^4 \|\mathbf{w}\|^2. \end{aligned}$$

For *III*, we use the Cauchy–Schwarz inequality to obtain

$$(5.7) \quad III \leq 2 \text{Ra}^2 \text{Pr} \|S\|^2 + \frac{\text{Pr}}{2} \|\mathbf{w}\|^2.$$

Combining the bounds for *I–III*, we have

$$\frac{d}{dt}(\|\mathbf{w}\|^2 + \|S\|^2) + \text{Pr} \|\mathbf{w}\|^2 + 2(\mu - \text{Ra}^2 \text{Pr}) \|S\|^2 \leq \left(\frac{C}{\text{Pr}^3} \|\nabla \mathbf{u}\|^4 + \frac{C}{\text{Pr}} \|T\|_{L^6}^4 \right) \|\mathbf{w}\|^2.$$

Thus, by the second condition in (5.2), it follows that

$$\frac{d}{dt}(\|\mathbf{w}\|^2 + \|S\|^2) + \left(\text{Pr} - \frac{C}{\text{Pr}^3} \|\nabla \mathbf{u}\|^4 - \frac{C}{\text{Pr}} \|T\|_{L^6}^4 \right) (\|\mathbf{w}\|^2 + \|S\|^2) \leq 0.$$

Therefore, by Grönwall’s inequality, the desired bound (5.3) now follows. ■

Theorem 5.3. *Assume the conditions of Lemma 5.2. There exists a constant $C_1 = C_1(\Omega) > 0$ such that if*

$$(5.8) \quad \text{Pr} \geq C_1 \text{Ra}^5,$$

then

$$(5.9) \quad \|(\tilde{\mathbf{u}} - \mathbf{u})(t)\| + \|(\tilde{T} - T)(t)\| \leq C_2 e^{-(\text{Pr}/4)t}$$

holds for all $t \geq 0$, where $C_2 = \|\tilde{\mathbf{u}}_0 - \mathbf{u}_0\| + \|\tilde{T}_0 - T_0\|$.

Proof. By (F3) of the standing hypotheses, it follows that

$$(5.10) \quad \frac{C}{\text{Pr}^3} \int_0^t \|\nabla \mathbf{u}(s)\|^4 ds \leq \frac{C \kappa_1^4}{\text{Pr}^3} \text{Ra}^4 t.$$

By (F4) and the Sobolev embedding, $H^1(\Omega) \hookrightarrow L^6(\Omega)$, it follows that

$$(5.11) \quad \frac{C}{\text{Pr}} \int_0^t \|T(s)\|_{L^6}^4 ds \leq \frac{C(\kappa'_1)^6}{\text{Pr}} \text{Ra}^{10} t.$$

Now upon combining (5.10), (5.11), it follows that there exists $C = C(\Omega) > 0$ such that

$$\exp \left(C \int_0^t \left(\frac{1}{\text{Pr}^3} \|\nabla \mathbf{u}(s)\|^4 + \frac{1}{\text{Pr}} \|T(s)\|_{L^6}^4 \right) ds \right) \leq \exp \left(C \frac{\text{Ra}^4}{\text{Pr}} \left(\text{Ra}^6 + \frac{1}{\text{Pr}^2} \right) t \right).$$

Then, having chosen Pr according to (5.8), upon taking square roots, we deduce (5.9) from and Lemma 5.2, as desired. ■

Remark 5.4. Observe that (5.8) requires the Prandtl number to be very large to ensure that synchronization occurs. Indeed, for $\text{Ra} \sim 10^7$ this condition indicates that $\text{Pr} \gtrsim 10^{35}$ which would place the Prandtl number in a regime beyond the one that occurs for the earth’s mantle. In addition, the lower bound on μ stated here would imply that $\mu \gtrsim \text{Pr} \text{Ra}^2$, which would yield a very stiff problem numerically, particularly if $\text{Pr} \sim \text{Ra}^5$ as indicated. Gratefully, as seen in the next subsection, these rigorous estimates are pessimistic, and synchronization is achieved for much lower values of Pr and μ than are indicated here.

5.2. Numerical results. The numerical results are similar to those presented in section 4.2, i.e., the same spatial discretization and time-stepping algorithm are employed here, but now we also consider variations in Pr . In addition, for finite Pr the velocity field is no longer slaved directly to the temperature field, requiring us to specify an initial condition for the velocity as well as the temperature. This is done as in the infinite Pr case by setting the initial state of the assimilated system to a low-order projection of the “true” system. Simulations at higher Rayleigh numbers are initiated with flow fields from previous simulations at incrementally lower Ra . To ensure that the algorithm is working, we first check that synchronization still occurs for reasonable choices of Ra , μ , and N given a finite Pr , say, $\text{Pr} = 100$. See Figure 7.

To simplify our exploration, for each value of Ra listed in Tables 1 and 2, we pick μ so that the systems synchronize quickly. Then, with $N = 32$, we run simulations for logarithmically spaced values of Pr from 1 to 100. Even with these larger-than-necessary choices of μ , convergence is lost for small enough Pr . See Figure 8 for a few additional examples and Table 3 for the chosen μ and the lowest Pr where synchronization still occurs.

There are three main points to take away from these results.

- Finite-but-large Prandtl data assimilation through only temperature measurements is possible, though the convergence is slower than in the infinite Prandtl setting explored in section 4.
- The relationship between Ra , μ , N , and Pr remains unclear and would require further measurements to precisely quantify. However, the fact that the assimilation works at all with reasonably small Pr indicates that the inequality constraints in Theorem 5.3 are strongly overstated or that the constants in (5.8) are extremely small.

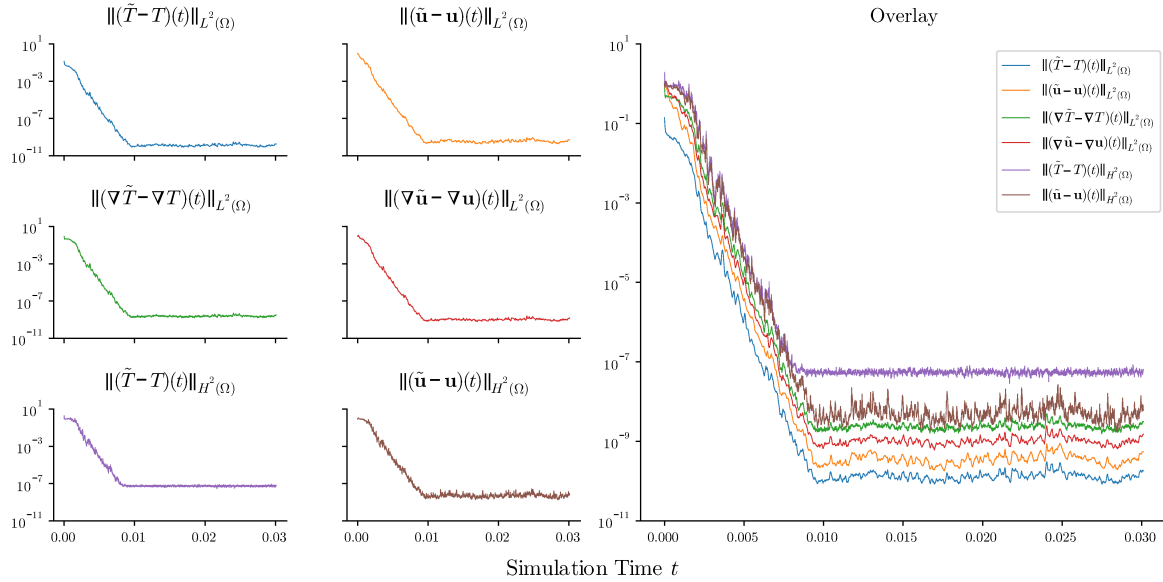


Figure 7. Synchronization in various norms for $\text{Ra} \approx 5.22 \times 10^7$, $\mu = 18,000$, $N = 32$, and $\text{Pr} = 100$. The temperature and velocity differences still decrease exponentially to zero, despite the fact that $\text{Pr} < \infty$. However, the convergence is slow compared to the $\text{Pr} = \infty$ case (see Figure 8).

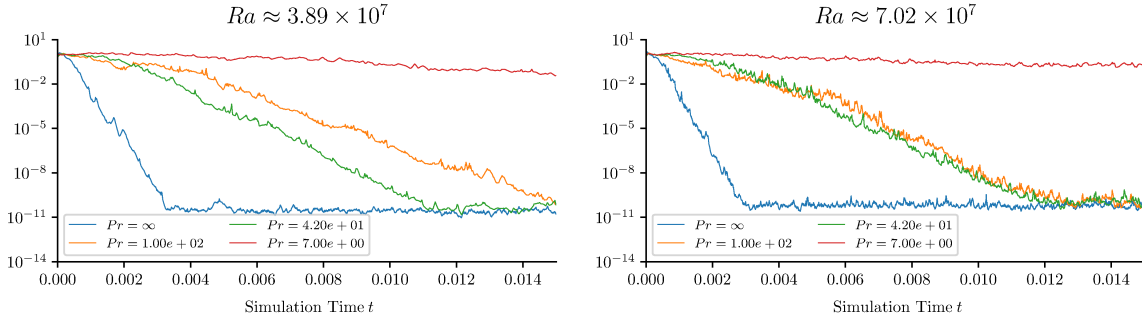


Figure 8. $\|(\tilde{\mathbf{u}} - \mathbf{u})(t)\|_{H^2} + \|(\tilde{T} - T)(t)\|_{H^2}$ with $N = 32$ and various Pr for two different values of Ra , with μ as given in Table 3. The convergence is generally slower as Pr decreases, until eventually there is no synchronization if Pr is too small.

Table 3

Minimal values of Pr that result in synchronization within about 0.01 units of simulation time for the given Ra and μ with $N = 32$. The constant value Δt is the average time step (averaged over the last half of the integration interval) for each simulation. Note that $\mu\Delta t$ is nearly constant between the different simulations presented here, indicating that the same relative nudging is applied at each time step despite the differences between each simulation. The relationship here is much less precise than those described in Tables 1 and 2 (for example, $Ra \approx 2.89 \times 10^7$ is a bit of an outlier), but it is interesting to note that lower Ra seem to require larger Pr . Whether or not this is an effect of Ra increasing or μ increasing, however, is unclear, particularly because the strength of the nudging per time step remains relatively constant.

Ra	μ	$\mu\Delta t$	Pr
1.1937766×10^7	10,000	9.65×10^{-3}	75
1.6037187×10^7	11,000	5.93×10^{-3}	56
2.1544346×10^7	12,000	8.10×10^{-3}	56
2.8942661×10^7	13,000	7.42×10^{-3}	100
3.8881551×10^7	14,000	6.60×10^{-3}	42
5.2233450×10^7	18,000	7.14×10^{-3}	42
7.0170382×10^7	20,000	6.72×10^{-3}	42
9.4266845×10^7	25,000	7.08×10^{-3}	31
1.26638017×10^8	32,000	7.57×10^{-3}	31
1.70125427×10^8	40,000	7.99×10^{-3}	31

- While the estimates obtained in Theorem 5.3 are clearly pessimistic (or perhaps the particular numerical setup—choice of initial conditions, etc.—explored here is uniquely advantageous for synchronization), it is interesting to note that the qualitative behavior is as expected. That is, the finite Prandtl system will synchronize so long as Pr is sufficiently large. This is not unexpected as the limit of $Pr \rightarrow 0$ will be dominated by inertial effects in which the temperature and velocity field are nearly decoupled, so we would anticipate that temperature observations alone will not suffice to recover the full flow field.

6. A scenario of model error. Finally, we address a realistic scenario of assimilating observables that correspond to solutions of the finite-but-large Prandtl number system (2.4), (2.5) into the “incorrect,” albeit more computationally tractable, infinite Prandtl data

assimilation system (2.8). Thus, we suppose $(\mathbf{u}, T)$ satisfy (F1)–(F4) and simply choose a suitable $\tilde{T}(\mathbf{x}, 0) = \tilde{T}_0(\mathbf{x})$ for (2.8) since the corresponding initial velocity is enslaved by the temperature evolution for the later nudging equation. Before we perform the analysis for the error estimates, we state the well-posedness result corresponding to the nudged equation, whose proof follows along similar lines to that of Theorem 4.1.

Theorem 6.1. *Let $\mu > 0$, and let $\{\lambda_n\}_{n=1}^\infty$ be as in (3.7). Let $(\mathbf{u}, T)$ of (2.4)–(2.5) satisfying (F1)–(F4). Let $\tilde{T}_0 \in L^2(\Omega)$ a.e. in Ω such that $\tilde{T}_0$ is a.e. L -periodic in x_1, x_{d-1} with $\tilde{T}_0|_{x_d=0} = 0, \tilde{T}_0|_{x_d=1} = 1$ (in the sense of trace). Suppose $N > 0$ satisfies $\frac{1}{4}\lambda_N \geq \mu$. Then there exists a unique $(\tilde{\mathbf{u}}, \tilde{T})$ satisfying (2.8) in the weak sense such that*

$$\tilde{\mathbf{u}} \in L^\infty(0, \tau; \mathcal{W}), \quad \tilde{T} \in L^\infty(0, \tau; L^2(\Omega)) \cap L^2(0, \tau; H^1(\Omega)) \cap C_w([0, \tau], L^2(\Omega))$$

for all $\tau > 0$.

6.1. Error estimates. Since our data assimilation equation (2.8) does not correspond to the true evolution of the observables (2.4), (2.5), we do not expect to obtain an exact synchronization. Instead, we derive estimates that quantify the maximal error possible and which will vanish as the Prandtl number is taken increasingly large.

Theorem 6.2. *Let $N, \mu > 0$ satisfy $\frac{1}{4}\lambda_N \geq \mu$ and $\tilde{T}_0$ be given as in Theorem 6.1. Let $(\mathbf{u}, T)$ of (2.4)–(2.5) satisfy (F1)–(F4), and let $(\tilde{\mathbf{u}}, \tilde{T})$ be the corresponding unique solution to (2.8) guaranteed by Theorem 6.1. There exists a constant $C_0 = C_0(\Omega) > 0$ such that if*

$$(6.1) \quad \mu \geq 4C_0 \left(\text{Ra}^{21/2} + \text{Ra}^2 \right),$$

then there exist positive constants $C_1 = \|\tilde{T}_0 - T_0\|, C_2 = C_2(\Omega), C_3 = C_3(\Omega)$ such that

$$(6.2) \quad \text{Ra}^{-1} \|\tilde{\mathbf{u}}(t) - \mathbf{u}(t)\|_{H^2} + \|\tilde{T}(t) - T(t)\| \leq C_1 e^{-(\mu/2)t} + \frac{C_2}{\text{Pr}} \frac{(\text{Ra}^{7/2} + \text{Ra})}{\mu^{1/2}} + \frac{C_3}{\text{Pr}} (\text{Ra}^{5/2} + \text{Ra}^{7/4})$$

for all $t \geq 0$.

Proof. Let $S = \tilde{T} - T, \mathbf{w} = \tilde{\mathbf{u}} - \mathbf{u}, q = \tilde{p} - p$. Then

$$(6.3) \quad -\Delta \mathbf{w} + \nabla q = \text{Ra} \mathbf{e}_3 S + \frac{1}{\text{Pr}} [\partial_t \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u}],$$

$$(6.4) \quad \partial_t S + \mathbf{w} \cdot \nabla S + \mathbf{u} \cdot \nabla S = -\Delta S - \mu P_N S - \mathbf{w} \cdot \nabla T.$$

Upon taking the L^2 -inner product of $\mathbf{w}, S$ with (6.3), (6.4), respectively, then adding the consequent relations, we obtain

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} \|S\|^2 + \|\nabla S\|^2 + \mu \|S\|^2 + \|\nabla \mathbf{w}\|^2 \\ &= \mu \|Q_N S\|^2 - \langle \mathbf{w} \cdot \nabla T, S \rangle + \text{Ra} \langle \mathbf{e}_3 S, \mathbf{w} \rangle + \frac{1}{\text{Pr}} \langle \partial_t \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u}, \mathbf{w} \rangle \\ &= I + II + III + IV. \end{aligned}$$

Next, applying Hölder's inequality, the Poincaré inequality, the Sobolev embedding theorem, Gagliardo–Nirenberg interpolation inequality, Young's inequality, and Theorem 3.5(ii), we derive

$$\begin{aligned}
|I| &\leq \frac{\mu}{\lambda_N} \|\nabla S\|^2 \\
|II| &\leq \|\mathbf{w}\|_{L^6} \|\nabla T\|_{L^3} \|S\| \\
&\leq C \|\nabla \mathbf{w}\| \|\mathbf{w}\|_{H^2}^{1/2} \|\mathbf{w}\|_{H^1}^{1/2} \|S\| \\
&\leq C \text{Ra}^{21/2} \|S\|^2 + \frac{1}{8} \|\nabla \mathbf{w}\|^2 \\
|III| &\leq \text{Ra} \|S\| \|\mathbf{w}\| \\
&\leq C \text{Ra}^2 \|S\|^2 + \frac{1}{8} \|\nabla \mathbf{w}\|^2 \\
|IV| &\leq \frac{1}{\text{Pr}} (\|\partial_t \mathbf{u}\| \|\mathbf{w}\| + \|\mathbf{u}\|_{L^4}^2 \|\nabla \mathbf{w}\|) \\
&\leq \frac{C}{\text{Pr}} (\text{Ra}^{7/2} \|\mathbf{w}\| + \text{Ra}^2 \|\nabla \mathbf{w}\|) \\
&\leq \frac{C}{\text{Pr}^2} (\text{Ra}^7 + \text{Ra}^2) + \frac{1}{8} \|\nabla \mathbf{w}\|^2.
\end{aligned}$$

Upon combining estimates for $I - IV$ with the conditions in (6.1) and the fact that $\frac{1}{4}\lambda_N \geq \mu$, it follows that

$$\frac{d}{dt} \|S\|^2 + \mu \|S\|^2 + \|\nabla \mathbf{w}\|^2 \leq \frac{C}{\text{Pr}^2} (\text{Ra}^7 + \text{Ra}^2).$$

Hence, by Grönwall's inequality, we arrive at

$$(6.5) \quad \|S(t)\|^2 \leq e^{-\mu t} \|S_0\|^2 + \mu^{-1} \frac{C}{\text{Pr}^2} (\text{Ra}^7 + \text{Ra}^2) (1 - e^{-\mu t}).$$

Lastly, by Lemma 3.1 and (F3) of the standing hypotheses, we have

$$(6.6) \quad \text{Ra}^{-1} \|\mathbf{w}(t)\|_{H^2} \leq \|S(t)\| + \frac{C}{\text{Pr}} (\text{Ra}^{5/2} + \text{Ra}^{7/4}).$$

We take the square root of (6.5) and add the result to (6.6) to complete the proof. ■

6.2. Numerical results. To numerically verify Theorem 6.2 in a way that is consistent with the numerical simulations corresponding to Theorems 4.2 and 5.3, we compare $(\tilde{\mathbf{u}}, \tilde{T})$ to $(\mathbf{u}, T)$. To begin, consider a simulation with $\text{Ra} \approx 5.22 \times 10^7$, $\mu = 18,000$, $N = 32$, and $\text{Pr} = 100$. For the finite Prandtl model in section 5, this set of parameters results in synchronization (see Figure 7). In this situation, however, the synchronization appears to be limited by the $O(\text{Pr}^{-1})$ error from the $\text{Pr} = \infty$ model to the $\text{Pr} < \infty$ reality.

This apparent lack of convergence illustrated in Figure 9 is expected, however, since Theorem 6.2 only guarantees that the error between $(\tilde{\mathbf{u}}, \tilde{T})$ and $(\mathbf{u}, T)$ decreases to $O(\text{Pr}^{-1})$ as time increases. Using the same set of parameters, but with larger and larger Pr , results in tighter and tighter synchronization. To more carefully match the results to the statement of Theorem 6.2, we calculate $\text{Ra}^{-1} \|(\tilde{\mathbf{u}} - \mathbf{u})(t)\|_{H^2} + \|(\tilde{T} - T)(t)\|$ at each simulation time t . See Figure 10.

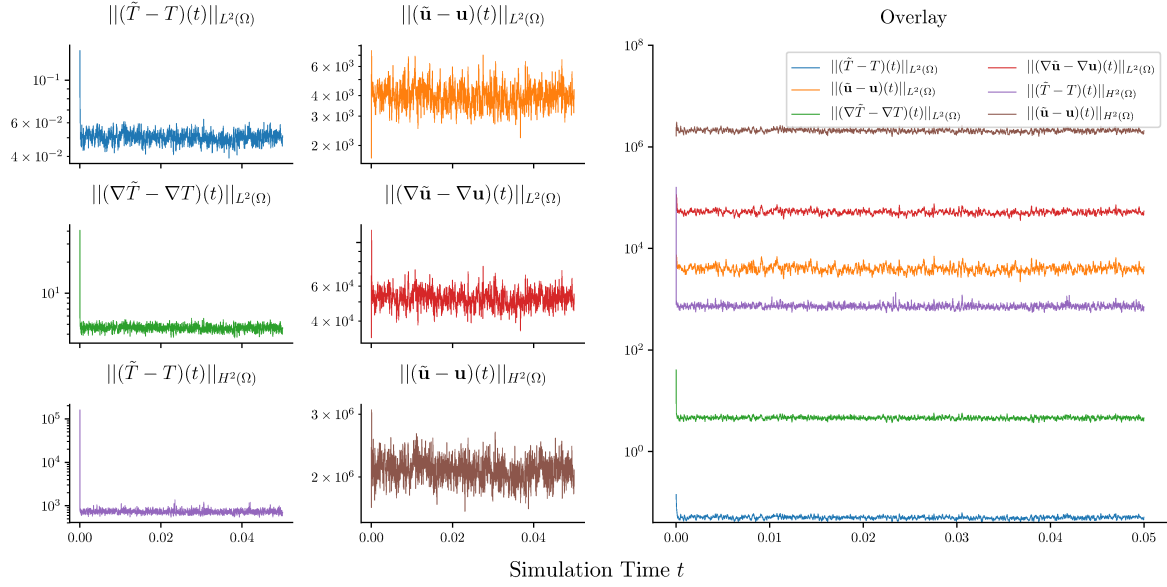


Figure 9. Hybrid assimilation with $Ra \approx 5.22 \times 10^7$, $\mu = 18,000$, $N = 32$, and $Pr = 100$. The temperature and velocity differences remain almost constant, with no hint of convergence. Note the difference in vertical axes relative to figures exhibited previously in sections 4.2, 5.2, above.

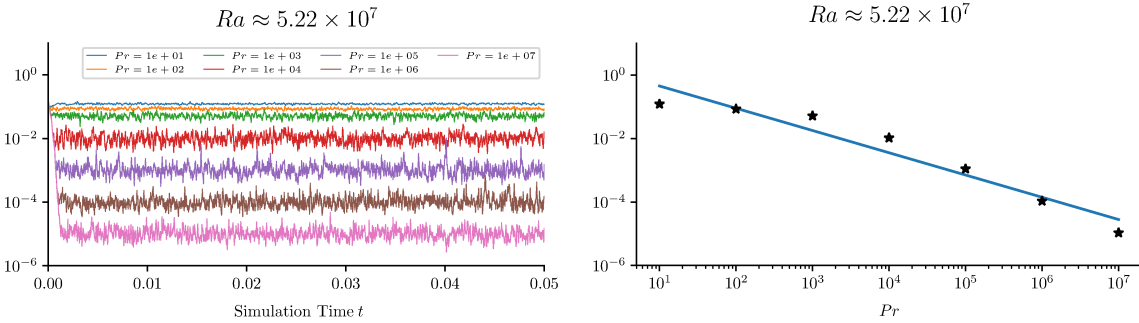


Figure 10. $Ra^{-1} \|\tilde{\mathbf{u}} - \mathbf{u}\|(t)_{H^2} + \|\tilde{T} - T\|(t)$ with $Ra \approx 5.22 \times 10^7$, $\mu = 18,000$, $N = 32$, and various Pr . As Pr increases, the minimal error between $(\tilde{\mathbf{u}}, \tilde{T})$ and $(\mathbf{u}, T)$ decreases. The plot on the right describes the relationship between Pr and the error with a logarithmic linear least-squares fit. The slope of this curve indicates that the error scales with Pr^γ with $\gamma \approx -0.7015$. Note that this fitted curve relies on very few data points, sampled from a very roughly defined synchronization criteria.

While Figure 10 is encouraging, it also highlights a weakness in the data assimilation scheme unique to this hybrid setting. In both of the previous settings considered here, the rigorous estimates were pessimistic (at least relative to the numerical experiments performed here), particularly for the large-but-finite Prandtl case. It appears that the estimates provided here for this hybrid setting are closer to the reported numerical observations, i.e., the dependence of the error on Pr as seen in Figure 10 is clearly monotonic, and although a fitted exponent with so few data points should be considered with some caution, it is interesting

to note that the fitted exponent is near the rigorous bound of $\gamma = -1$. Although the precise value of $\gamma = -0.7015$ should not be taken too seriously in light of the qualitative and sparse approach to collecting the data points used in Figure 10, this dependence is clearly closer to the rigorously estimated value than the parameter dependence of the previous cases.

These results suggest that the practical success of these types of data assimilation schemes is highly dependent on the data coming from the same model as the simulated system, an unrealistically stringent restriction if the application of interest is as complicated a process as weather and/or climate. Recall that the Boussinesq approximation is effectively a “zeroth-order” approximation for the mantle, meaning the true physical system has several complicated secondary effects (some of which are unknown) that are not included in the model. The lack of numerical synchronization in such a simple setting suggests that data assimilation may not be as adaptable to settings where the exact model is not known. On the other hand, it may be that these results are indicating that the $\text{Pr} = \infty$ model does not adequately reflect the large Pr setting; in other words, does the lack of synchronization imply a fundamental difference between the finite and infinite Pr settings, or do the $O(\text{Pr}^{-1})$ differences (which are accurately estimated from the analysis) dominate the potential synchronization?

7. Conclusions and outlook. In section 4, we examined a data assimilation scheme for the RB system with $\text{Pr} = \infty$ and showed rigorously that synchronization occurs between the data and assimilating equations under certain conditions on the relaxation parameter μ and the number of projected modes N relative to the Rayleigh number Ra when measurements of the temperature only are observed. That is, as long as there is enough data (i.e., N is not too small), μ can be chosen large enough to guarantee synchronization. Though this is a satisfying theoretical result, the numerical experiments in section 4.2 show that synchronization often occurs under much weaker conditions on μ and N than Theorem 4.2 requires. In particular, the inequality conditions on μ is shown to be at least an order of magnitude away from being sharp. This gap between the rigorous estimates and the numerical observations may be due to the particular numerical experiment considered here, or to the lack of strictness in the relevant estimates (see [18] for some improved bounds on the RB problem for certain boundary conditions). In addition, the numerical results also demonstrate situations in which synchronization fails, namely, when μ and/or N are not large enough.

Section 5 shows that synchronization in the temperature measurements only and at finite Pr is also possible, although the rate of convergence is slower than with infinite Pr and the relationship between Ra , Pr , μ , and N needed to achieve synchronization remains somewhat ambiguous. As in the infinite Pr case, the conditions imposed on μ appear quite pessimistic when compared to numerical experiments, where it is not clear what the source of this difference between the analysis and the numerical experiments may be.

Finally, when the true values are taken from simulations with $\text{Pr} < \infty$, but the assimilating equations use $\text{Pr} = \infty$, the synchronization is highly dependent on Pr , as predicted by the rigorous bounds. This hybrid setting illustrates that the difference between the two systems is dominating the error, rather than the dynamical error in the synchronization process. Although the numerical simulations agree well with the rigorous predictions in this setting, they do indicate a pessimistic outlook for additional settings wherein the exact evolution of the dynamics for a data assimilation system of this type is unknown. In particular, as noted

above, we have omitted several details in our model of mantle convection that play a vital role in the evolution and may have an effect similar to the difference between finite and infinite Pr. To investigate this further, data assimilation applied to the internally heated convective setting (see [45, 77, 78], for example) and possibly the anelastic or compressible convective systems [55] will be explored.

The current consideration of the difference between the infinite and near-infinite Prandtl number convective systems lends itself to further investigations wherein the assimilating model is different from the physical system from which the observations are obtained. For example, one might consider the effects of imprecisely defined boundary conditions, i.e., what if the observations were obtained from a convective simulation in which the velocity satisfied a Navier-slip condition, but the nudged system was modeled with a no-slip condition? Other variations in the model itself might include slight variations in the geometry between the two systems, and additional terms in the equations themselves such as internal heating mentioned above. The rub of the matter is that data assimilation techniques, if they are meant to apply to physical settings such as weather, climate, and investigations of the earth's mantle, must consider the fallibility of the model they are relying on; that is, do variations in the underlying model itself allow for synchronization of the model with the observed truth? On the other hand, do the results obtained here indicate that the infinite Prandtl model is *not* an adequate model of large Pr convection, i.e., is the lack of synchronization indicative of fundamental differences between the models that are otherwise not accounted for? Further consideration of these questions for the convective problem, and for the nudging approach to data assimilation for more complicated systems, is imperative not only to understand the potential application of these methods but also to quantify the impact that model assumptions can have on the original system.

Acknowledgments. The authors wish to thank two anonymous reviewers whose comments greatly enhanced the presentation and discussion of these results. This paper was initiated by conversations between AF and NEGH (and eventually JPW) which took place during a workshop at the Institute for Pure and Applied Mathematics (IPAM), for which all the authors are indebted to IPAM, the National Science Foundation (NSF), and organizers of the program on Mathematics of Turbulence in fall 2014. SM and JPW acknowledge computational resources and support from the Fulton Supercomputing Laboratory at Brigham Young University where all the simulations were performed, as well as extensive feedback from the Dedalus developers (see <http://dedalus-project.org/>). JPW also acknowledges the hospitality and generosity of the Mathematics Department at Tulane University where some of the finishing touches on this work were carried out.

REFERENCES

- [1] G. AHLERS, *Low-temperature studies of the Rayleigh-Bénard instability and turbulence*, Phys. Rev. Lett., 33 (1974), 1185.
- [2] G. AHLERS AND R. BEHRINGER, *Evolution of turbulence from the Rayleigh-Bénard instability*, Phys. Rev. Lett., 40 (1978), 712.
- [3] G. AHLERS, S. GROSSMANN, AND D. LOHSE, *Heat transfer and large scale dynamics in turbulent Rayleigh-Bénard convection*, Rev. Modern Phys., 81 (2009), 503.

- [4] D. ALBANEZ AND M. BENVENUTTI, *Continuous data assimilation algorithm for simplified Bardina model*, *Evol. Equ. Control Theory*, 7 (2018), pp. 33–52.
- [5] D. ALBANEZ, H. NUSSENZVEIG LOPES, AND E. S. TITI, *Continuous data assimilation for the three-dimensional Navier-Stokes- α model*, *Asymptotic Anal.*, 97 (2016), pp. 165–174.
- [6] M. ALTAF, E. S. TITI, T. GEBRAEL, O. KNIO, L. ZHAO, AND M. MCCABE, *Downscaling the 2D Bénard convection equations using continuous data assimilation*, *Comput. Geosci.*, 21 (2017), pp. 393–410.
- [7] D. AUROUX AND J. BLUM, *A nudging-based data assimilation method for oceanographic problems: The back and forth nudging (BFN) algorithm*, *Nonlinear Process. Geophys.*, 15 (2008), pp. 305–219.
- [8] A. AZOUANI, E. OLSON, AND E. S. TITI, *Continuous data assimilation using general interpolant observables*, *J. Nonlinear Sci.*, 24 (2014), pp. 277–304.
- [9] H. BESSAIH, E. OLSON, AND E. S. TITI, *Continuous data assimilation with stochastically noisy data*, *J. Nonlinear Sci.*, 28 (2015), pp. 729–753.
- [10] A. BISWAS, C. FOIAS, C. MONDAINI, AND E. S. TITI, *Downscaling data assimilation algorithm with applications to statistical solutions of the Navier-Stokes equations*, *Ann. Inst. Henri Poincaré Anal. Non Linéaire*, 36 (2019), pp. 295–326.
- [11] A. BISWAS, J. HUDSON, A. LARIOS, AND Y. PEI, *Continuous data assimilation for the magnetohydrodynamics equations in 2D using one component of the velocity and magnetic fields*, *Asymptotic Anal.*, 108 (2018), pp. 1–43.
- [12] A. BISWAS AND V. MARTINEZ, *Higher-order synchronization for a data assimilation algorithm for the 2D Navier-Stokes equations*, *Nonlinear Anal. Real World Appl.*, 35 (2017), pp. 132–157.
- [13] B. BLANKENBACH, F. BUSSE, U. CHRISTENSEN, L. CSEREPES, D. GUNKEL, U. HANSEN, H. HARDER, G. JARVIS, M. KOCH, G. MARQUART, D. MOORE, P. OLSON, H. SCHMELING, AND T. SCHNAUBELT, *A benchmark comparison for mantle convection codes*, *Geophys. J. Int.* 98 (1989), pp. 23–28.
- [14] J. BLOCHER, V. MARTINEZ, AND E. OLSON, *Data assimilation using noisy time-averaged measurements*, *Phys. D*, 376 (2018), pp. 49–59.
- [15] D. BLOMKER, K. LAW, A. STUART, AND K. ZYGALAKIS, *Accuracy and stability of the continuous-time 3DVAR filter for the Navier-Stokes equation*, *Nonlinearity*, 26 (2013), pp. 2193–2219.
- [16] E. BODENSCHATZ, J. R. DE BRUYN, G. AHLERS, AND D. S. CANNELL, *Transitions between patterns in thermal convection*, *Phys. Rev. Lett.*, 67 (1991), 3078.
- [17] K. J. BURNS, G. M. VASIL, J. S. OISHI, D. LECOANET, AND B. P. BROWN, *Dedalus: A Flexible Framework for Numerical Simulations with Spectral Methods*, [arXiv:1905.10388](https://arxiv.org/abs/1905.10388), 2019.
- [18] Y. CAO, M. S. JOLLY, E. S. TITI, AND J. P. WHITEHEAD, *Algebraic Bounds on the Rayleigh-Bénard Attractor*, [arXiv:1905.01399](https://arxiv.org/abs/1905.01399), 2019.
- [19] E. CARLSON, J. HUDSON, AND A. LARIOS, *Sensitivity Analysis and Parameter Recovery for the 2D Navier-Stokes Equations via Continuous Data Assimilation*, [arXiv:1812.07646](https://arxiv.org/abs/1812.07646), 2018.
- [20] J. CHARNEY, M. HALEM, AND R. JASTROW, *Use of incomplete historical data to infer the present state of the atmosphere*, *J. Atmos. Sci.*, 26 (1969), pp. 1160–1163.
- [21] B. COCKBURN, D. JONES, AND E. S. TITI, *Estimating the number of asymptotic degrees of freedom for nonlinear dissipative systems*, *Math. Comp.*, 66 (1997), pp. 1073–1087.
- [22] P. CONSTANTIN AND C. FOIAS, *Navier-Stokes Equations*, *Chicago Lectures in Mathematics*, University of Chicago Press, Chicago, IL, 1988.
- [23] R. DALEY, *Atmospheric Data Analysis*, *Cambridge Atmospheric and Space Science Series*, Cambridge University Press, Cambridge, UK, 1991.
- [24] S. DESAMSETTI, H. DASARI, S. LANGODAN, E. S. TITI, O. KNIO, AND I. HOTEIT, *Efficient dynamical downscaling of general circulation models using continuous data assimilation*, *Quart. J. Roy. Meteorol. Soc.*, 145 (2019), pp. 3175–3194.
- [25] P. DI LEONI, A. MAZZINO, AND L. BIFERALE, *Inferring flow parameters and turbulent configuration with physics-informed data assimilation and spectral nudging*, *Phys. Rev. Fluids*, 3 (2018), 104604.
- [26] R. DIPERNA AND P.-L. LIONS, *Ordinary differential equations, transport theory and Sobolev spaces*, *Invent. Math.*, 98 (1989), pp. 511–547.
- [27] R. ERRICO AND D. BAUMHEFNER, *Predictability experiments using a high-resolution limited-area model*, *Monthly Weather Rev.*, 115 (1987), pp. 488–504.
- [28] A. FARHAT, H. JOHNSTON, M. S. JOLLY, AND E. S. TITI, *Assimilation of nearly turbulent Rayleigh-Bénard flow through vorticity or local circulation measurements: A computational study*, *J. Sci. Comput.*, 77 (2018), pp. 1519–1533.

- [29] A. FARHAT, M. S. JOLLY, AND E. S. TITI, *Continuous data assimilation for the 2D Bénard convection through velocity measurements alone*, Phys. D, 303 (2015), pp. 59–66.
- [30] A. FARHAT, E. LUNASIN, AND E. S. TITI, *Abridged continuous data assimilation for the 2D Navier-Stokes equations utilizing measurements of only one component of the velocity field*, J. Math. Fluid Mech., 18 (2016), pp. 1–23.
- [31] A. FARHAT, E. LUNASIN, AND E. S. TITI, *Data assimilation algorithm for 3D Bénard convection in porous media employing only temperature measurements*, J. Math. Anal. Appl., 438 (2016), pp. 492–506.
- [32] A. FARHAT, E. LUNASIN, AND E. S. TITI, *On the Charney conjecture of data assimilation employing temperature measurements alone: The paradigm of 3D planetary geostrophic model*, Math. Climate Weather Forecast., 2 (2016), pp. 61–74.
- [33] A. FARHAT, E. LUNASIN, AND E. S. TITI, *Continuous data assimilation for a 2D Bénard convection system through horizontal velocity measurements alone*, J. Nonlinear Sci., 27 (2017), pp. 1065–1087.
- [34] A. FARHAT, E. LUNASIN, AND E. S. TITI, *A data assimilation algorithm: The paradigm of the 3D Leray- α model of turbulence*, in Partial Differential Equations Arising from Physics and Geometry: A volume in memory of Abbas Bahri, M. Ben Ayed, M. A. Jendoubi, Y. Rébaï, H. Riahi, and H. Zaag, eds., Cambridge University Press, Cambridge, UK, 2019, pp. 253–273. London Mathematical Society. Lecture Notes Series 450, Cambridge University Press, 2019.
- [35] C. FOIAS, O. MANLEY, R. ROSA, AND R. TEMAM, *Navier-Stokes Equations and Turbulence*, Cambridge University Press, Cambridge, UK, 2001.
- [36] C. FOIAS, O. MANLEY, AND R. TEMAM, *Attractors for the Bénard problem: Existence and physical bounds on their fractal dimension*, Nonlinear Anal., 11 (1987), pp. 939–967.
- [37] C. FOIAS, C. MONDAINI, AND E. S. TITI, *A discrete data assimilation scheme for the solutions of the 2D Navier-Stokes equations and their statistics*, SIAM J. Appl. Dyn. Syst., 15 (2016), pp. 2019–2142.
- [38] C. FOIAS AND G. PRODI, *Sur le comportement global des solutions non-stationnaires des équations de Navier-Stokes en dimension 2*, Rend. Semin. Mat. Univ. Padova, 39 (1967), pp. 1–34.
- [39] C. FOIAS AND R. TEMAM, *Determination of the solutions of the Navier-Stokes equations by a set of nodal values*, Math. Comp., 43 (1984), pp. 117–133.
- [40] C. FOIAS AND E. S. TITI, *Determining nodes, finite difference schemes and inertial manifolds*, Nonlinearity, 4 (1991), p. 135.
- [41] J. FOLDES, N. GLATT-HOLTZ, AND G. RICHARDS, *Large Prandtl number asymptotics in randomly forced turbulent convection*, NoDEA Nonlinear Differential Equations Appl., 26 (2019), 43.
- [42] M. GESHO, E. OLSON, AND E. S. TITI, *A computational study of a data assimilation algorithm for the two-dimensional Navier-Stokes equations*, Commun. Comput. Phys., 19 (2016), pp. 1094–1110.
- [43] M. GHIL, M. HALEM, AND R. ATLAS, *Time-continuous assimilation of remote-sounding data and its effect on weather forecasting*, Monthly Weather Rev., 107 (1978), pp. 140–171.
- [44] M. GHIL, B. SHKOLLER, AND V. YANGARBER, *A balanced diagnostic system compatible with a barotropic prognostic model*, Monthly Weather Rev., 105 (1977), pp. 1223–1238.
- [45] D. GOLUSKIN, *Internally Heated Convection and Rayleigh-Bénard Convection*, Springer, New York, NY, 2016.
- [46] J. HOKE AND R. ANTHES, *The initialization of numerical models by a dynamic initialization technique*, Monthly Weather Rev., 104 (1976), pp. 1551–1556.
- [47] M. HOLST AND E. S. TITI, *Determining projections and functionals for weak solutions of the Navier-Stokes equations*, Contemp. Math., 204 (1997), pp. 125–138.
- [48] H. IBDIAH, C. MONDAINI, AND E. S. TITI, *Fully discrete numerical schemes of a data assimilation algorithm uniform-in-time error estimates*, IMA J. Numer. Anal., in Press .
- [49] M. S. JOLLY, V. MARTINEZ, E. OLSON, AND E. S. TITI, *Continuous data assimilation with blurred-in-time measurements of the surface quasi-geostrophic equation*, Chin. Ann. Math. Ser. B, 40 (2019), pp. 721–764.
- [50] M. S. JOLLY, V. MARTINEZ, S. SADIGOV, AND E. S. TITI, *A determining form for the subcritical surface quasi-geostrophic equation*, J. Dynam. Differential Equations, (2018), pp. 1–38.
- [51] M. S. JOLLY, V. MARTINEZ, AND E. S. TITI, *A data assimilation algorithm for the 2D subcritical surface quasi-geostrophic equation*, Adv. Nonlinear Stud., 35 (2017), pp. 167–192.
- [52] D. JONES AND E. S. TITI, *Determining finite volume elements for the 2D Navier-Stokes equations*, Phys. D, 60 (1992), pp. 165–174.

- [53] D. JONES AND E. S. TITI, *On the number of determining nodes for the 2D Navier-Stokes equations*, J. Math. Anal., 168 (1992), pp. 72–88.
- [54] E. KALNAY, *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge University Press, Cambridge, UK, 2003.
- [55] S. KING, C. LEE, P. VAN KEKEN, W. LENG, S. ZHONG, E. TAN, N. TOSI, AND M. KAMEYAMA, *A community benchmark for 2-D Cartesian compressible convection in the Earth's mantle*, Geophys. J. Int., 180 (2010), pp. 73–87.
- [56] A. LARIOS, L. G. REBHOLZ, AND C. ZERFAS, *Global in time stability and accuracy of imex-fem data assimilation schemes for Navier–Stokes equations*, Comput. Methods Appl. Mech. Engrg., 345 (2019), pp. 1077–1093.
- [57] D. LOHSE AND K. Q. XIA, *Small-scale properties of turbulent Rayleigh–Bénard convection*, Ann. Rev. Fluid Mech., 42 (2010).
- [58] A. LORENC, R. BELL, AND B. MACPHERSON, *The Meteorological Office analysis correction data assimilation scheme*, Quart. J. Roy. Meteorol. Soc., 117 (1991), pp. 59–89.
- [59] E. LORENZ, *Deterministic nonperiodic flow*, J. Atmos. Sci., 20 (1963), pp. 130–141.
- [60] W. LYNE, R. SWINBANK, AND B. N.T., *A data assimilation experiment and the global circulation during the fgge special observing periods*, Quart. J. Roy. Meteorol. Soc., 108 (1982), pp. 575–594.
- [61] P. MARKOWICH, E. S. TITI, AND S. TRABELSI, *Continuous data assimilation for the 3D Brinkman–Forchheimer-extended Darcy model*, Nonlinearity, 24 (2016), pp. 1292–1328.
- [62] C. MONDAINI AND E. S. TITI, *Uniform-in-time error estimates for the postprocessing Galerkin method applied to a data assimilation algorithm*, SIAM J. Numer. Anal., 56 (2018), pp. 78–110.
- [63] E. OLSON AND E. S. TITI, *Determining modes and Grashof number in 2D turbulence: A numerical case study*, Theoret. Comput. Fluid Dynam., 22 (2008), pp. 327–339.
- [64] L. RAYLEIGH, LIX, *On convection currents in a horizontal layer of fluid, when the higher temperature is on the under side*, Philos. Mag. Ser. 6 32 (1916), pp. 529–546.
- [65] D. STAUFFER AND J. BAO, *Optimal determination of nudging coefficients using the adjoint equations*, Tellus, 45A (1993), pp. 358–369.
- [66] D. STAUFFER AND N. SEAMAN, *Use of four-dimensional data assimilation in a limited area mesoscale model - Part 1: Experiments with synoptic-scale data*, Quart. J. Roy. Meteorol. Soc., 118 (1990), pp. 1551–1556.
- [67] R. TEMAM, *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, Appl. Math. Sci. 68, Springer-Verlag, New York, 1997.
- [68] D. TURCOTTE AND G. SCHUBERT, *Geodynamics*, Cambridge University Press, Cambridge, UK, 2014.
- [69] J. VERRON AND W. HOLLAND, *Impact de données d'altimétrie satellitaire sur les simulations numériques des circulations générales océaniques aux latitudes moyennes*, Ann. Geophys., 7 (1989), pp. 31–46.
- [70] P. VIDARD, F.-X. LE DIMET, AND A. PIACENTINI, *Determination of optimal nudging coefficients*, Tellus, 55A (2003), pp. 1–15.
- [71] X. WANG, *Infinite Prandtl number limit of Rayleigh–Bénard convection*, Comm. Pure Appl. Math., 57 (2004), pp. 1265–1282.
- [72] X. WANG, *Large Prandtl number behavior of the Boussinesq system of Rayleigh–Bénard convection*, Appl. Math. Lett., 17 (2004), pp. 821–825.
- [73] X. WANG, *A note on long time behavior of solutions to the Boussinesq system at large Prandtl number*, Contemp. Math., 371 (2005), pp. 315–323.
- [74] X. WANG, *Asymptotic behavior of the global attractors to the Boussinesq system for Rayleigh–Bénard convection at large Prandtl number*, Comm. Pure Appl. Math., 60 (2007), pp. 1293–1318.
- [75] X. WANG, *Bound on vertical heat transport at large Prandtl number*, Phys. D, 237 (2008), pp. 854–858.
- [76] X. WANG, *Stationary statistical properties of Rayleigh–Bénard convection at large Prandtl number*, Comm. Pure Appl. Math., 61 (2008), pp. 789–815.
- [77] J. P. WHITEHEAD AND C. R. DOERING, *Internal heating driven convection at infinite Prandtl number*, J. Math. Phys., 52 (2011), p. 11 pages.
- [78] J. P. WHITEHEAD AND C. R. DOERING, *Rigid rigorous bounds on heat transport in a slippery container*, J. Fluid Mech., 707 (2012), pp. 241–259.
- [79] X. ZOU, I. NAVON, AND F.-X. LE DIMET, *An optimal nudging data assimilation scheme using parameter estimation*, Quart. J. Roy. Meteorol. Soc., 118 (1992), pp. 1163–1186.