

On nonlinear transitions, minimal seeds and exact solutions for the geodynamo

Calum S. Skene^{1,2} , Florence Marcotte³ and Steven M. Tobias^{1,2}

¹Department of Applied Mathematics, University of Leeds, Leeds LS2 9JT, UK

²School of Physics and Astronomy, The University of Edinburgh, Edinburgh EH9 3FD, UK

³Inria, CNRS, LJAD, Université Côte d'Azur, France

Corresponding author: Calum S. Skene, cskene3@ed.ac.uk

(Received 7 November 2024; revised 7 April 2025; accepted 16 August 2025)

Nearly fifty years ago, Roberts (1978) postulated that the Earth's magnetic field, which is generated by turbulent motions of liquid metal in its outer core, likely results from a subcritical dynamo instability characterised by a dominant balance between Coriolis, pressure and Lorentz forces (requiring a finite-amplitude magnetic field). Here, we numerically explore subcritical convective dynamo action in a spherical shell, using techniques from optimal control and dynamical systems theory to uncover the nonlinear dynamics of magnetic field generation. Through nonlinear optimisation, via direct-adjoint looping, we identify the minimal seed – the smallest magnetic field that attracts to a nonlinear dynamo solution. Additionally, using the Newton-hookstep algorithm, we converge stable and unstable travelling wave solutions to the governing equations. By combining these two techniques, complex nonlinear pathways between attracting states are revealed, providing insight into a potential subcritical origin of the geodynamo. This paper showcases these methods on the widely studied benchmark of Christensen *et al.* (2001, *Phys. Earth Planet. Inter.*, vol. 128, pp. 25–34), laying the foundations for future studies in more extreme and realistic parameter regimes. We show that the minimal seed reaches a nonlinear dynamo solution by first approaching an unstable travelling wave solution, which acts as an edge state separating a hydrodynamic solution from a magnetohydrodynamic one. Furthermore, by carefully examining the choice of cost functional, we establish a robust optimisation procedure that can systematically locate dynamo solutions on short time horizons with no prior knowledge of its structure.

Key words: geodynamo, magneto convection, variational methods

1. Introduction

1.1. *Balances in the geodynamo: weak and strong field branches*

The problem of generation of the Earth's magnetic field is an important and challenging one for geophysics; the magnetic field plays a key role in deflecting the solar wind and therefore is important in helping to protect us from the harmful effects of the charged particles and the depletion of ozone leading to increased exposure to harmful ultraviolet radiation (see e.g. Arsenović *et al.* 2024). The geomagnetic field is predominantly dipolar, thus large scale, with a dipole axis that is offset from the axis of rotation. It has variability on a wide range of time scales, from years to centuries, and includes such dynamical features as excursions and reversals (see e.g. Constable & Korte 2006; Roberts 2008).

The Earth's magnetic field is believed to be generated by dynamo action in its fluid outer core where convective motions (either compositionally or thermally driven) in the electrically conducting medium can drive currents and hence magnetic fields via induction (see e.g. Moffatt & Dormy 2019). The geodynamo problem involves the simultaneous solution of the partial differential equations governing the evolution of the fluid (the momentum equation), the magnetic field (the induction equation) and the evolution equation for the co-density. The theoretical challenge arises owing to the vast range of temporal scales (and to a lesser extent spatial scales) that need to be resolved in the momentum equation (with the induction and co-density – a variable that combines the buoyancy effects of temperature and chemical species concentration (Braginsky & Roberts 1995) – equations causing fewer problems); the large range of temporal scales leads to the requirement that the non-dimensional equations be solved at extreme parameter values, and hence the impossibility of direct solution via computational methods.

Nonetheless, progress has been made on a number of fronts. Heroic computations (following on from the pioneering work of Glatzmaier & Roberts 1995) are starting to achieve small enough parameters (though still extremely far from their true values) that interesting dynamics is achieved (see e.g. Aubert, Gastine & Fournier 2017; Schaeffer *et al.* 2017; Schwaiger, Gastine & Aubert 2019). This computational endeavour has been predicated on the understanding that it is important to approach the correct regime by keeping the balances in the momentum equation and co-density equations realistic as one increases the separation of time scales. This is achieved by considering a distinguished limit, as suggested by Dormy (2016), so that all the non-dimensional parameters scale together to achieve the correct balance.

What is the correct balance for the geodynamo? It is widely thought that the magnetic field in the Earth's core is strong enough to play a significant role in the dynamics of the flow, at least at some length scale. For a detailed discussion of how, and at what scale, the magnetic Lorentz force may enter in the dynamics, see Aurnou & King (2017) or Tobias (2021). Theoretical support for the importance of the magnetic field arises from theoretical and computational studies of magnetoconvection (see e.g. Proctor 1994; Fearn 1998). Roberts (1978) recognised that there may be two branches of the Earth's dynamo: a weak field branch and a strong field branch. The weak field branch arises in a traditional bifurcation scenario: as the thermal driving (as measured by the non-dimensional Rayleigh number) is increased, hydrodynamic thermal convection sets in. Because the Earth is a rapid rotator, this convection sets in at a very small length scale and at a very high Rayleigh number. Eventually, the convection becomes strong enough so that the magnetic Reynolds number of the flow is large enough for the onset of dynamo action. This leads to the formation of weak multipolar magnetic fields (Petitdemange 2018); here, these fields are weak in the sense that their magnetic energy is equivalent to, or less than, the kinetic energy of the flow.

A more efficient situation occurs when the magnetic field is strong. The strong magnetic field breaks the restriction imposed on the flow by the fast rotation, and leads to the onset of convection on larger length scales at more moderate values of the Rayleigh number measuring thermal driving. Thus one might expect to find a more efficient subcritical strong field branch for values of the Rayleigh number well below onset of weak field dynamo action (and potentially even below the onset of hydrodynamic convection). On the strong field branch, the magnetic energy should be much larger than the kinetic energy of the flow, and the Lorentz force may even be strong enough to enter into the leading-order balance, competing with the Coriolis force and the pressure gradient to give a magnetostrophic balance (at least at certain scales; Dormy 2016). The magnetic field in this situation acts so as to make the dynamo behave in a more efficient manner – an example of an essentially nonlinear dynamo (Tobias, Cattaneo & Brummell 2011). There is therefore a strong theoretical drive to understand subcritical dynamo action in the context of rapidly rotating spherical shells.

1.2. Computing essentially nonlinear dynamos

It is, however, non-trivial to compute nonlinear dynamos in the subcritical regime. Although time-stepping methods are able sometimes to locate these, the results are often haphazard, and sometimes owe their success to a slice of luck. A more systematic approach relies on the computation of fixed points. The latter builds on quasi-Newton iterative solvers to compute equilibrium dynamo solutions by continuation while progressing through the parameter space. Such solvers, however, require fairly good first guesses for the procedure to converge towards a solution. This usually implies finding a starting point in the linearly unstable regime (whenever possible) and then following the saturated solution down as the control parameter is gradually decreased toward the linearly stable regime. Alternatively, one may turn the subcritical problem into a supercritical one, e.g. through the addition of a suitable physical forcing, which is then gradually removed (Waleffe 2003) – in a dynamo context, this method was successfully used by Rincon, Ogilvie & Proctor (2007) and Deguchi (2020, 2022) to identify subcritical equilibrium solutions in magnetohydrodynamic (MHD) plane shear flows. Moreover, while this continuation approach has the considerable advantage of computing unstable as well as stable equilibria, it can only compute exact solutions that are travelling waves or at least recurring states, which limits its applicability.

In the context of wall-bounded turbulence, the objective of such computations of equilibria is not to find the attracting, statistically steady turbulent state. This state is fairly easy to reproduce numerically (or indeed experimentally). The theoretical challenge rather lies with the identification of unstable coherent states and their dynamical role in the transition to turbulence (Toh & Itano 2003; Schneider, Eckhardt & Yorke 2007; Wang, Gibson & Waleffe 2007; Khapko *et al.* 2016), with the control of the transition as a long-term objective. In the context of subcritical dynamo problems, however, identifying the attracting dynamo state in itself can be very uncertain. It is imperative that the computed dynamo states be allowed to be highly fluctuating. Furthermore, it is desirable that as little knowledge of the dynamo mechanism as possible be required for the computation. To overcome these limitations, a robust way to numerically identify an elusive subcritical state without prior guesses of the equilibrium is through a combination of short-time optimal control and long-time direct numerical simulations, as demonstrated by Mannix, Ponty & Marcotte (2022) with nonlinear dynamos.

This alternative approach builds on the mathematical framework of transient growth and non-modal stability analysis (see e.g. Schmid (2007) or Kerswell (2018) for reviews

of both linear and nonlinear non-modal stability analysis, respectively). It has been widely used in the last decade to compute ‘minimal’ (i.e. least-energy) perturbations that nonlinearly trigger transition to hydrodynamic turbulence in shear flows (Cherubini *et al.* 2010; Duguet, Brandt & Larsson 2010; Pringle & Kerswell 2010; Vavaliaris, Beneitez & Henningson 2020), which no alternative methods can currently compute. In particular, fixed point methods cannot identify minimal seeds since these are not steady or recurring solutions of the underlying dynamical system. The review of Kerswell, Pringle & Willis (2014) illustrates this approach for identifying minimal seeds in dynamical systems in general, and identifies three key considerations. The first outlines the importance of the optimisation time horizon, which must be large enough to overcome initial transients, but not so large as to render convergence difficult due to the increasing sensitivity of the optimum to the initial condition. Second, including nonlinearities is imperative since the most-amplified perturbations in the linear problem are (generally) not the minimal seeds for subcritical transition. Finally, the functional to be optimised to compute least-energy perturbations need not be the energy. Further to these points, Mannix *et al.* (2022) found that the spatial structure of minimal dynamo seeds bears no resemblance to that of the subcritical, nonlinear state – a finding that is consistent with studies of minimally triggered turbulence in shear flows – and that the sequence of events that they trigger is revealing of the mechanisms of subcritical dynamo action in the considered systems.

1.3. *Triggering nonlinear transition to subcritical, convective dynamos*

In this paper, we investigate the geodynamo benchmark of Christensen *et al.* (2001) as an ideal starting point for developing the methodology needed to identify minimal seeds that lead to subcritical dynamo action in a geodynamo set-up. This benchmark study contains three cases, of which we consider the first two: a purely hydrodynamic benchmark, and a subcritical dynamo benchmark. In the hydrodynamic benchmark, convective instability gives rise to Taylor columns with $m = 4$ symmetry that ultimately persist as a travelling wave. Although at the benchmark parameters the conducting (steady base state) is unstable to convection, the resulting convecting state is linearly stable to dynamo action. In this manner, the form of the magnetic field initial condition is crucial, and needs to have a large enough amplitude as well as the correct structure in order to reach a nonlinear dynamo.

Although originally intended as a benchmark for Boussinesq MHD spherical shell codes, the ‘modest’ benchmark parameters provide rich dynamical behaviour that has subsequently been studied. The hydrodynamic bifurcation behaviour has been explored by Feudel *et al.* (2015), who show that there are multiple stable nonlinear hydrodynamic travelling waves present at the benchmark parameters. Subsequently, Feudel *et al.* (2017) studied the bifurcation nature of nonlinear dynamo solutions possible in the system, showing that at the benchmark parameters, there are stable and unstable MHD states that arise via a saddle node bifurcation, and that the MHD solution branch is smoothly connected to a purely hydrodynamic branch where a secondary Hopf bifurcation occurs. In these studies, the nonlinear travelling waves were converged using a matrix-free Newton solver using implicit integration of the linear terms as a preconditioner (Mamun & Tuckerman 1995). More recently, Skene & Tobias (2024) used a weakly nonlinear analysis to study the saturation of the convective instability in the purely hydrodynamic case near the onset of convection.

While these studies illuminate the dynamical states possible for the system, they do not provide information on the complex nonlinear pathways between these states or their basins of attraction. A more systematic programme for mapping out the landscape of subcritical dynamos is needed, with tools developed for uncovering dynamics in a

spherical shell. It is the first steps of just such a programme that we will describe in this paper, which combines nonlinear optimisation with the identification of the unstable and stable travelling waves present in the system. Optimal control, on the one hand, can identify the magnetic field with the smallest energy that attracts to a dynamo solution. Variational optimisation in dynamo theory dates back to Backus (1958), who derived a bound on the magnetic Reynolds number required for dynamo action in a sphere. More recently, optimisation of steady velocity fields has been performed in various geometries to maximise linear dynamo growth rates (Willis 2012; Chen *et al.* 2015, 2018; Herreman 2018; Holdenried-Chernoff, Chen & Jackson 2019). In the present paper, we use optimal control on the full set of MHD equations to provide a limit on the basins of attraction for reaching a nonlinear dynamo. By considering the cost functional to extremise, we furnish a robust procedure that can identify dynamos using short time horizon optimisations, enabling this procedure to be scaled up in future studies. On the other hand, we also use a Newton-hookstep method (Dennis & Schnabel 1996) to find stable and unstable travelling waves in the system, and examine them in light of the dynamical pathways between important states. We demonstrate that the Newton-hookstep procedure is able to converge unstable magnetic states without preconditioning, and therefore provide the foundation for future studies where these exact solutions will lie within a turbulent and chaotic flow field. The structure of the paper is as follows: the mathematical and numerical set-up is outlined in § 2, the results are detailed in § 3, and conclusions are offered in § 4.

2. Mathematical and numerical set-up

2.1. Governing equations

We model the geodynamo by considering the set-up illustrated in figure 1. Spherical coordinates (θ, ϕ, r) are used, where θ , ϕ and r are the co-latitudinal, azimuthal and radial coordinates, respectively. Our domain $\mathcal{V} = \mathbb{R}^3$ is split into three parts, i.e. $\mathcal{V} = \mathcal{V}_i \cup \mathcal{V}_s \cup \mathcal{V}_o$. The spherical shell region $\mathcal{V}_s$, where $r_i < r < r_o$, is filled with a conducting fluid. The inner region $\mathcal{V}_i$ ($r < r_i$) and outer region $\mathcal{V}_o$ ($r > r_o$) do not consist of a fluid, but can support a magnetic field and therefore are needed to complete our geodynamo description: the inner region $\mathcal{V}_i$ models the Earth's solid inner core, while the unbounded outer region $\mathcal{V}_o$ provides a simple model for the screening of internal geomagnetic fields by an insulating, rocky mantle. The whole system is taken to rotate around the z -axis with angular velocity $\boldsymbol{\Omega} = \Omega \mathbf{e}_z$.

The motion of the conducting fluid in $\mathcal{V}_s$ is governed by the Navier–Stokes equations under the Boussinesq approximation (which is an approximation often used in geodynamo simulations):

$$\begin{aligned} \frac{\partial \hat{\mathbf{U}}}{\partial \hat{t}} + \hat{\mathbf{U}} \cdot \hat{\nabla} \hat{\mathbf{U}} &= -\frac{1}{\rho_0} \hat{\nabla} \hat{P} - 2\Omega \mathbf{e}_z \times \hat{\mathbf{U}} + \frac{g_o \alpha \hat{r}}{r_o} \hat{T} \mathbf{e}_r + \nu \hat{\nabla}^2 \hat{\mathbf{U}} + \frac{1}{\mu} (\hat{\nabla} \times \hat{\mathbf{B}}) \times \hat{\mathbf{B}}, \\ \nabla \cdot \hat{\mathbf{U}} &= 0, \\ \frac{\partial \hat{\mathbf{B}}}{\partial \hat{t}} &= \hat{\nabla} \times (\hat{\mathbf{U}} \times \hat{\mathbf{B}}) + \eta \hat{\nabla}^2 \hat{\mathbf{B}}, \\ \hat{\nabla} \cdot \hat{\mathbf{B}} &= 0, \\ \frac{\partial \hat{T}}{\partial \hat{t}} + \hat{\mathbf{U}} \cdot \hat{\nabla} \hat{T} &= \kappa \hat{\nabla}^2 \hat{T}. \end{aligned} \quad (2.1)$$

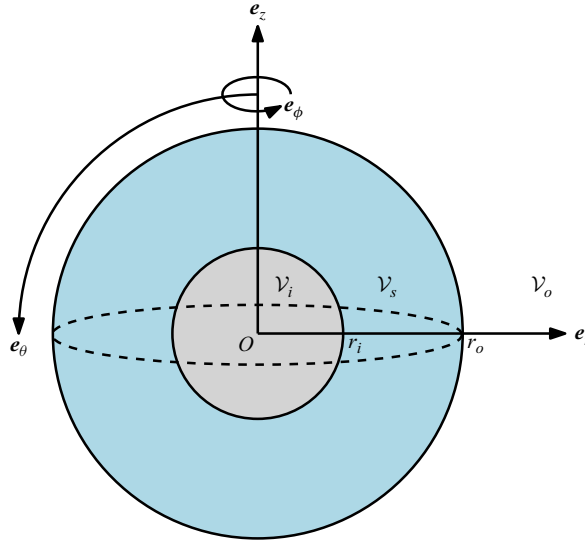


Figure 1. Sketch of the numerical domain. The fluid domain $\mathcal{V}_s$ (in blue) is surrounded by two insulating, solid domains: the inner core $\mathcal{V}_i$ (in grey) and an outer mantle $\mathcal{V}_o$.

These dimensional equations describe the evolution of the velocity field $\hat{\mathbf{U}}$, temperature field $\hat{T}$, magnetic field $\hat{\mathbf{B}}$ and modified pressure $\hat{P}$ (which also accounts for the centrifugal acceleration). Although we use the temperature, we note that one could also formulate (2.1) in terms of the co-density. In writing the equations in dimensional form, denoted with ‘hats’, we have used the kinematic viscosity ν , the magnetic permeability μ , the magnetic diffusivity η , the thermal expansion coefficient α , the gravity at the outer radius g_o , and the thermal diffusivity κ . The outer region $\mathcal{V}_o$ is taken to be an insulator, so the electrical current vanishes there ($\hat{\mathbf{J}} = \hat{\nabla} \times \hat{\mathbf{B}} = 0$), so that the magnetic field satisfies a Laplace equation in $\mathcal{V}_o$. Whilst different approaches have been used to model the inner region $\mathcal{V}_i$, such as taking it to be finitely conducting (Glatzmaier & Roberts 1995), we will take it to be insulating.

These equations are non-dimensionalised as follows: length is scaled with the shell width $d = \hat{r}_o - \hat{r}_i$, time with the viscous diffusion time d^2/ν , temperature with the temperature difference $\Delta\hat{T}$ across the shell, and the magnetic field using $\sqrt{\rho\mu\eta\Omega}$. This results in the non-dimensional equations

$$\begin{aligned} Ek \left(\frac{\partial \mathbf{U}}{\partial t} + \mathbf{U} \cdot \nabla \mathbf{U} - \nabla^2 \mathbf{U} \right) + 2\mathbf{e}_z \times \mathbf{U} + \nabla P &= \tilde{Ra} \frac{\mathbf{r}}{r_0} T + \frac{1}{Pm} (\nabla \times \mathbf{B}) \times \mathbf{B}, \\ \nabla \cdot \mathbf{U} &= 0, \\ \frac{\partial \mathbf{B}}{\partial t} &= \nabla \times (\mathbf{U} \times \mathbf{B}) + \frac{1}{Pm} \nabla^2 \mathbf{B}, \\ \nabla \cdot \mathbf{B} &= 0, \\ \frac{\partial T}{\partial t} + \mathbf{U} \cdot \nabla T &= \frac{1}{Pr} \nabla^2 T, \end{aligned} \quad (2.2)$$

where we have introduced the notation $\mathbf{r} = r\mathbf{e}_r$. These equations are governed by the Ekman number Ek , the modified Rayleigh number $\tilde{Ra}$ (which accounts for the stabilising effect of rotation on convection, and is related to the classical Rayleigh number $Ra = \alpha g_o \Delta T d^3/(\nu\kappa)$ via $\tilde{Ra} = Ra Ek/Pr$), the magnetic Prandtl number Pm , and the Prandtl

number Pr , respectively defined as

$$Ek = \frac{\nu}{\Omega d^2}, \quad \widetilde{Ra} = \frac{\alpha g_o \Delta \hat{T} d}{\nu \Omega}, \quad Pm = \frac{\nu}{\eta}, \quad Pr = \frac{\nu}{\kappa}. \quad (2.3)$$

The boundary conditions for the temperature and velocity fields are

$$T(r_i) = 1, \quad (2.4)$$

$$T(r_o) = 0, \quad (2.5)$$

$$\mathbf{U}(r_i) = \mathbf{0}, \quad (2.6)$$

$$\mathbf{U}(r_o) = \mathbf{0}. \quad (2.7)$$

We enforce the continuity of the magnetic field through the boundaries with the insulating regions by matching with a potential field at the inner and outer radii, as detailed in § 2.4. Under this non-dimensionalisation, the non-dimensional kinetic and magnetic energies take the form

$$K = \frac{1}{2} \iiint \mathbf{U} \cdot \mathbf{U} \, dV, \quad M = \frac{1}{2 Ek Pm} \iiint \mathbf{B} \cdot \mathbf{B} \, dV, \quad (2.8)$$

respectively.

2.2. The adjoint-based optimal control procedure

We briefly outline the optimisation procedure for determining the initial condition $\mathbf{q}_0$ that maximises a chosen cost functional under the strong constraint that the system, described by the state variable $\mathbf{q}(\mathbf{x}, t)$, evolves according to the governing equations, written here in general form:

$$\mathbf{M} \frac{\partial \mathbf{q}}{\partial t} = \mathcal{N}(\mathbf{q}), \quad \mathbf{q}(\mathbf{x}, 0) = \mathbf{q}_0. \quad (2.9)$$

This constrained optimisation problem can be solved by determining the critical points of an augmented Lagrangian

$$\mathcal{L}(\mathbf{q}, \mathbf{q}_0, \mathbf{q}^\dagger, \mathbf{q}_0^\dagger) = \mathcal{J}(\mathbf{q}) - [\mathbf{q}^\dagger, \mathbf{M} \frac{\partial \mathbf{q}}{\partial t} - \mathcal{N}(\mathbf{q})] - \langle \mathbf{q}_0^\dagger, \mathbf{q}(\mathbf{x}, 0) - \mathbf{q}_0 \rangle, \quad (2.10)$$

where $\mathcal{J}$ is the cost functional that we need to maximise, $\mathbf{q}^\dagger$ is a Lagrange multiplier (denoted as an adjoint variable), and the inner products $[\cdot, \cdot]$ and $\langle \cdot, \cdot \rangle$ are defined as

$$[\mathbf{y}^\dagger, \mathbf{y}] = \int_0^{t_{opt}} \iiint (\mathbf{y}^\dagger)^T \mathbf{y} \, dV \, dt = \int_0^{t_{opt}} \langle \mathbf{y}^\dagger, \mathbf{y} \rangle \, dt. \quad (2.11)$$

In what follows, the cost functional that we will consider is expressed as

$$\mathcal{J}(\mathbf{q}) = \int_0^{t_{opt}} \iiint J_I \, dV \, dt, \quad (2.12)$$

i.e. as the time and volume integral of some quantity of interest J_I , with t_{opt} a chosen time horizon (fixed throughout the optimisation procedure). Quantity J_I will be defined later, and depends here on the instantaneous state variable $\mathbf{q}$. Setting all the first variations of the Lagrangian to zero means that (i) $\mathbf{q}$ must evolve according to (2.9) (also called the ‘direct problem’), (ii) $\mathbf{q}^\dagger$ must solve an adjoint problem, which we will specify below for our particular system, and (iii) the initial condition for the adjoint variable $\mathbf{q}^\dagger(\mathbf{x}, 0)$ must

vanish here:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{q}_0} = \mathbf{q}_0^\dagger = \mathbf{M}^T \mathbf{q}^\dagger(\mathbf{x}, 0) = 0. \quad (2.13)$$

In practice, an optimum is iteratively computed by evolving the direct problem forwards in time, starting from some initial (typically random) guess $\mathbf{q}_0$, then evolving the adjoint problem backwards in time, down to $t = 0$. At this stage, (2.13) provides the required gradient information to update $\mathbf{q}_0$ in order to increase the value of the cost functional $\mathcal{J}$, and hence move towards a local maximum.

With our direct problem defined by the governing MHD equations (2.2), and using the initial magnetic field $\mathbf{B}_0 = \mathbf{B}(\mathbf{x}, 0)$ as the optimisation variable, we obtain the following adjoint problem (the full details of the derivation of the adjoint equations are given in Appendix A):

$$\begin{aligned} Ek \left(-\frac{\partial \mathbf{u}^\dagger}{\partial t} + \nabla \times (\mathbf{U} \times \mathbf{u}^\dagger) + \mathbf{u}^\dagger \times \boldsymbol{\omega} - \nabla^2 \mathbf{u}^\dagger \right) - 2\mathbf{e}_z \times \mathbf{u}^\dagger + \nabla p^\dagger \\ = -T^\dagger \nabla T + \mathbf{B} \times (\nabla \times \mathbf{b}^\dagger) + \frac{\partial J_I}{\partial \mathbf{U}}, \\ \nabla \cdot \mathbf{u}^\dagger = 0, \\ -\frac{\partial \mathbf{b}^\dagger}{\partial t} = -\nabla \Pi^\dagger + \frac{1}{Pm} \nabla \times (\mathbf{B} \times \mathbf{u}^\dagger) - \frac{1}{Pm} (\nabla \times \mathbf{B}) \times \mathbf{u}^\dagger - \mathbf{U} \times (\nabla \times \mathbf{b}^\dagger) \\ - \frac{1}{Pm} \nabla \times (\nabla \times \mathbf{b}^\dagger) + \frac{\partial J_I}{\partial \mathbf{B}}, \\ \nabla \cdot \mathbf{b}^\dagger = 0, \\ -\frac{\partial T^\dagger}{\partial t} - \nabla \cdot (\mathbf{U} T^\dagger) = \widetilde{Ra} \frac{\mathbf{r}}{r_0} \cdot \mathbf{u}^\dagger + \frac{1}{Pr} \nabla^2 T^\dagger + \frac{\partial J_I}{\partial T}, \end{aligned} \quad (2.14)$$

where $\mathbf{q}^\dagger = (\mathbf{u}^\dagger, p^\dagger, \mathbf{b}^\dagger, \Pi^\dagger, T^\dagger)^T$ is the adjoint state, and where we have introduced the vorticity $\boldsymbol{\omega} = \nabla \times \mathbf{U}$. These adjoint equations, solved backwards in time, are ‘initialised’ with the final time condition (found by cancelling the first variations of the Lagrangian with respect to $\mathbf{q}$) that $\mathbf{u}^\dagger(t_{opt}) = \mathbf{b}^\dagger(t_{opt}) = \mathbf{0}$ and $T^\dagger(t_{opt}) = 0$. The boundary conditions for the adjoint variables are homogeneous Dirichlet boundary conditions for $T^\dagger$ and all components of $\mathbf{u}^\dagger$ on both $r = r_o$ and $r = r_i$, along with vacuum boundary conditions on $\mathbf{b}^\dagger$, and homogeneous Dirichlet boundary conditions on $\Pi^\dagger$ (see § A.2). Note that the integration of the adjoint equations requires that the whole direct solution be stored in memory, which can lead to high memory costs. While we did not encounter this issue within the present study, the memory requirements in future studies might become large enough to prevent the storage of the entire forward solution. This difficulty is classically overcome by the use of checkpointing schemes, which alleviate the large memory requirement at the cost of recomputing the forward solution (Griewank 1992). Let us mention in particular the `checkpoint_schedules` Python library (Dolci *et al.* 2024), which can easily be used with our developed code and has implementations of many popular optimal checkpointing schemes – including Revolve (Stumm & Walther 2009), disk-revolve (Aupy *et al.* 2016), periodic-disk-revolve (Aupy & Herrmann 2017), two-level (Pringle *et al.* 2016), H-Revolve (Herrmann & Aupy 2020), and mixed storage checkpointing (Maddison 2024).

Throughout the optimisation, we will require that the initial magnetic field is constrained such that $\|\mathbf{B}_0\|^2 = M_0$, where M_0 is a specified magnetic energy budget. Whilst this

constraint can be handled by introducing an extra Lagrange multiplier (Pringle, Willis & Kerswell 2012), we will constrain it using Riemannian optimisation methods (Absil, Mahony & Sepulchre 2007), of which rotation-based projections (Douglas, Amari & Kung 1998; Fournes, Caulfield & Schmid 2013) are a special case (see the discussion of e.g. Skene *et al.* 2022). In this manner, the Euclidean gradient returned by the optimisation procedure,

$$\frac{\partial \mathcal{L}}{\partial \mathbf{B}_0} = \mathbf{b}^\dagger(0), \quad (2.15)$$

is projected into the Riemannian gradient on the hypersphere $\|\mathbf{B}_0\|^2 = M_0$ where optimisation is taking place.

The minimal seed that triggers dynamo action is found by performing a series of optimisations with decreasing values of M_0 . The initial condition for the hydrodynamic variables is defined as an equilibrium solution of the purely hydrodynamic equations. As our time horizon is smaller than the time needed for a dynamo to be established, we perform one long forward run after each optimisation has converged (following Mannix *et al.* 2022), in order to determine whether the system evolves towards a self-sustained magnetic state. The minimal dynamo seed corresponds to the optimal initial condition with smallest M_0 that still triggers a dynamo.

2.3. Nonlinear travelling wave solutions

Many nonlinear states of interest in this system are found to take the form of travelling waves, which rotate around the z -axis with the non-dimensional drift frequency ω relative to our already-rotating reference frame. In order to systematically identify these states, whether they be stable or unstable, we will converge them using the Newton-hookstep algorithm (Dennis & Schnabel 1996). To that end, we need to adopt the rotating reference frame in which the travelling waves correspond to stationary states; we thus introduce the flow map $\Phi_t(\mathbf{q}_0, \omega)$, which returns the state at time t starting from initial condition $\mathbf{q}_0$, in a rotating reference frame with angular frequency $1 + \omega$. A travelling wave solution is then an initial condition $\mathbf{q}_0$ and a drift frequency ω such that $\Phi_t(\mathbf{q}_0, \omega) = \mathbf{q}_0$ for all times t , which the Newton-hookstep algorithm identifies by computing the zeros of $\mathbf{f}(\mathbf{q}_0, \omega) = \Phi_{t_f}(\mathbf{q}_0, \omega) - \mathbf{q}_0$, where the time is now fixed at a small value t_f . Although the value of t_f is arbitrary for a travelling wave (as it is stationary in the identified reference frame, so $\Phi_{t_f}(\mathbf{q}_0, \omega) = \mathbf{q}_0$ for all t_f), a small value is used following the advice of Willis (2019b), which prevents the algorithm converging to a time-periodic solution.

The classic Newton method finds a root of a vector function $\mathbf{g}(\mathbf{x})$ through the iterative procedure

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \delta \mathbf{x}_i, \quad (2.16)$$

$$\frac{\partial \mathbf{g}}{\partial \mathbf{x}} \delta \mathbf{x}_i = -\mathbf{g}(\mathbf{x}_i). \quad (2.17)$$

The Newton-hookstep variant aims to alleviate issues that arise when the initial guess $\mathbf{x}_0$ is far from the true solution, which can cause the algorithm to not converge. This is particularly pertinent for unstable travelling wave solutions, where finding a good initial guess is difficult. In the Newton-hookstep algorithm, (2.17) is replaced with

$$\delta \mathbf{x}_{i+1} = \arg \min_{\|\delta \mathbf{x}\| < \delta} \left\| \frac{\partial \mathbf{g}}{\partial \mathbf{x}} \delta \mathbf{x} + \mathbf{g}(\mathbf{x}_i) \right\|, \quad (2.18)$$

where δ is an (adjustable) trust region size. In this manner, the exact Newton step given by (2.17) is relaxed in order to constrain the size of the update to lie within the trust region. The result is a more stable algorithm that has become the primary algorithm of choice for converging unstable equilibria, travelling waves and relative periodic orbits (Viswanath 2007, 2009; Duguet, Pringle & Kerswell 2008; Chandler & Kerswell 2013; Budanur *et al.* 2017).

For our system, we have $\mathbf{x} = (\mathbf{q}_0, \omega)^T$, which means that we have one more degree of freedom, given by ω , than in equations *f*. Therefore, in order to obtain a square system to solve in (2.18), we introduce the additional constraint that the update $\delta\mathbf{x}_i$ cannot simply rotate $\mathbf{x}_i$ around the z -axis. Rotating $\mathbf{x}_i(\theta, \phi, r)$ by a small amount around the z -axis can be achieved by setting $\mathbf{x}_i(\theta, \phi + \delta\phi, r)$, i.e. we increment ϕ by the small amount $\delta\phi$. Using Taylor expansions, one can write

$$\mathbf{x}_i(\theta, \phi + \delta\phi, r) = \mathbf{x}_i(\theta, \phi, r) + \nabla\mathbf{x}_i \cdot \mathbf{e}_\phi \delta\phi + \mathcal{O}((\delta\phi)^2), \quad (2.19)$$

where $\mathbf{e}_\phi$ is the unit vector in the ϕ direction. This shows that small rotations are given by the term $\nabla\mathbf{x}_i \cdot \mathbf{e}_\phi$, and our constraint is that $\delta\mathbf{x}_i$ should be orthogonal to this direction, which provides the extra equation required for solvability, addressing the underlying rotational symmetry of the underlying system.

2.4. Numerical implementation

All equations are solved using the open-source partial differential equation solver Dedalus (Burns *et al.* 2020). Dedalus is a pseudo-spectral solver that is able to solve equations in spherical geometries (Lecoanet *et al.* 2019; Vasil *et al.* 2019). It parses the governing equations that have been input by the user in plain text, and provides access to the spectral discretisation and routines for time-stepping them. For time-stepping, we use the second-order semi-implicit backward differentiation formula implicit–explicit multistep scheme (Wang & Ruuth 2008). Diffusion and pressure-like terms, which allow for divergence-free conditions to be enforced, are time-stepped implicitly. In this manner, divergence-free conditions are solved for along with all problem variables, and do not necessitate any operator splitting methods or poloidal–toroidal decompositions. Spin-weighted spherical harmonics are used for the bases in the θ and ϕ directions, and Jacobi polynomials are used to discretise the radial direction. The typical resolution used in our study consists of all spherical harmonics up to a maximum degree $\ell_{max} = 63$, Jacobi polynomial degree $N_{max} = 63$, and a fixed time step $\Delta t = 0.5 \times 10^{-4}$. We have verified that our code with this resolution satisfies the Boussinesq hydrodynamic and dynamo benchmark cases of Christensen *et al.* (2001).

When solving the direct equations (2.2), we must consider how to handle the divergence-free constraint as well as the potential boundary conditions on the magnetic field. In order to enforce the divergence-free condition on $\mathbf{B}$, we solve for the vector potential, i.e. we solve for $\mathbf{A}$ such that $\nabla \times \mathbf{A} = \mathbf{B}$. The vector potential satisfies the equation

$$\frac{\partial \mathbf{A}}{\partial t} = \mathbf{U} \times \mathbf{B} + \nabla\phi + \frac{1}{Pm} \nabla^2 \mathbf{A}, \quad (2.20)$$

where ϕ is a scalar that is determined here by enforcing the Coulomb gauge condition $\nabla \cdot \mathbf{A} = 0$, although other gauges may be possible (Cattaneo, Bodo & Tobias 2020). Then the potential boundary conditions for the magnetic field on the inner and outer spheres can also be enforced directly on the vector potential (Lecoanet *et al.* 2019),

taking the form

$$\left(\frac{\partial \mathbf{A}_{\ell,\sigma}}{\partial r} - \frac{\ell + \sigma}{r} \mathbf{A}_{\ell,\sigma} \right) \Big|_{r=r_i} = 0, \quad (2.21)$$

$$\left(\frac{\partial \mathbf{A}_{\ell,\sigma}}{\partial r} + \frac{\ell + 1 + \sigma}{r} \mathbf{A}_{\ell,\sigma} \right) \Big|_{r=r_o} = 0, \quad (2.22)$$

where $\mathbf{A}_{\ell,\sigma}$ is the coefficient of the ℓ th spherical harmonic for regularity component $\sigma \in \{-1, 1, 0\}$ (Vasil *et al.* 2019).

The vector potential formulation (2.20) of the direct induction equations is possible since the direct equations naturally preserve the divergence of $\mathbf{B}$. However, the adjoint equations (2.14) do not naturally satisfy this property, with the divergence of the adjoint magnetic field $\mathbf{b}^\dagger$ being a gauge choice that is enforced with the Lagrange multiplier $\Pi^\dagger$ (see Appendix A for more details). Hence we solve directly for $\mathbf{b}^\dagger$. Note that $\mathbf{b}^\dagger$ must also satisfy a divergence-free constraint and continuous matching with a potential field on the spherical boundaries. A commonly used, alternative approach (e.g. Dormy 1997; Wicht 2002; Schaeffer *et al.* 2017) to enforce both conditions at the same time is to decompose the considered field $\mathbf{X}$ into its poloidal–toroidal form as

$$\mathbf{X} = \nabla \times (\mathcal{T}\mathbf{r}) + \nabla \times (\nabla \times (\mathcal{P}\mathbf{r})). \quad (2.23)$$

Under this decomposition, the potential boundary conditions take a very simple form in terms of the spherical harmonic decomposition of the poloidal and toroidal parts (see Hollerbach (2000) for more details). However, as we do not solve the adjoint equations for the poloidal and toroidal parts since it is more natural to solve the full vector form of the equations in Dedalus, care is needed in order to properly enforce the potential boundary conditions on $\mathbf{b}^\dagger$. To that end, we make use of the relations $\mathbf{r} \cdot \mathbf{b}^\dagger = -\nabla_\parallel^2 \mathcal{P}^\dagger$ and $\mathbf{r} \cdot \mathbf{J}^\dagger = -\nabla_\parallel^2 \mathcal{T}^\dagger$, where $\nabla_\parallel$ is the surface Laplacian, $\mathbf{J}^\dagger = \nabla \times \mathbf{b}^\dagger$, and $\mathcal{P}^\dagger$ and $\mathcal{T}^\dagger$ are the poloidal and toroidal parts of the adjoint magnetic field, respectively. By also using the fact that the action of the surface Laplacian on spherical harmonic Y_ℓ^m is $\nabla_\parallel^2 Y_\ell^m = -\ell(\ell + 1)Y_\ell^m$, we obtain the following boundary conditions for the adjoint magnetic field:

$$\mathbf{e}_r \cdot \left(\frac{\partial \mathbf{b}_\ell^\dagger}{\partial r} - \frac{\ell - 1}{r} \mathbf{b}_\ell^\dagger \right) \Big|_{r=r_i} = 0, \quad (2.24)$$

$$\mathbf{e}_r \cdot \left(\frac{\partial \mathbf{b}_\ell^\dagger}{\partial r} + \frac{\ell + 2}{r} \mathbf{b}_\ell^\dagger \right) \Big|_{r=r_o} = 0, \quad (2.25)$$

$$\mathbf{e}_r \cdot \mathbf{J}^\dagger \Big|_{r=r_o, r_i} = 0, \quad (2.26)$$

which indirectly give the correct conditions on the poloidal and toroidal parts of $\mathbf{b}^\dagger$. This gives two conditions at each boundary, which, together with the boundary conditions imposed on $\Pi^\dagger$, gives the total number of boundary conditions needed for the adjoint equations.

The numerical implementation of the optimisation procedure outlined in § 2.2 requires a few considerations. Solving the direct equations followed by the adjoint equations provides both the value of the cost functional, and its gradient with respect to the initial magnetic field. This update should not change the specified energy of the initial magnetic field – which, as previously indicated, is handled by directly optimising on a Riemannian manifold. We use for this the SphereManOpt library (Mannix *et al.* 2024), building on Pymanopt (Townsend, Koep & Weichwald 2016) and modified to take place on

spherical manifolds, with a conjugate gradient algorithm from Sato (2022) and Armijo line search (Wright & Nocedal 1999). We note that while this study uses the continuous adjoint method, Dedalus now supports a discrete adjoint approach via its automatic differentiation capabilities (Skene & Burns 2025) – a feature that was not available at the time of this study.

For implementing the Newton-hookstep algorithm, we use the routines available in the JFNK-Hookstep Github repository (Willis 2019a,b). Specifically, we use the FORTRAN implementation of the code developed for Openpipeflow (Willis 2017), which is compiled into a Python module using f2py (Peterson 2009). This enables the algorithm to be easily used together with Dedalus. The flow map $\Phi_{t_f}(\mathbf{q}_0, \omega)$ is provided to the Newton-hookstep algorithm by using Dedalus to solve (2.2) with the Coriolis term changed to $(2 + 2\omega)\mathbf{e}_z \times \mathbf{U}$, and the boundary conditions for $\mathbf{U}$ now taking the form $\mathbf{U}(r = r_i) = -Ek^{-1}\omega r_i \sin\theta \mathbf{e}_\phi$ and $\mathbf{U}(r = r_o) = -Ek^{-1}\omega r_o \sin\theta \mathbf{e}_\phi$. When the Newton-hookstep algorithm converges, this corresponds to solving the equations in a reference frame that rotates with the rotation rate of the travelling wave, rather than the Earth. Hence the fluid will appear stationary, with the boundary conditions giving the rotation rate of the Earth relative to the travelling wave. Note that as with the direct equations (2.2), the centrifugal term is absorbed in the pressure gradient. While there are sophisticated ways for generating initial guesses for the Newton-hookstep algorithm, including those based on near recurrences (Viswanath 2007; Chandler & Kerswell 2013), dynamic mode decomposition (Page & Kerswell 2020), variational methods (Lan & Cvitanović 2004; Parker & Schneider 2022) and convolutional autoencoders (Page *et al.* 2021, 2024), at the parameters used for this study, we can find initial guesses directly from snapshots obtained from the evolution of the governing equations. The travelling waves are converged to a tolerance of 10^{-6} , with error measured using the L_2 norm of the spectral discretisation of the fields. In order to calculate the stability of each of the travelling wave solutions, we take a matrix-free approach similar to that described by Skene & Tobias (2023). The stability problem is solved in a reference frame co-rotating with the travelling wave, reducing a Floquet stability problem to that of a fixed point.

3. Results

3.1. Benchmark set-up

For the purpose of the present study, and to test the numerical scheme, we adopt the parameters of the dynamo benchmark introduced by Christensen *et al.* (2001): $Ek = 10^{-3}$, $Ra = 100$, $Pm = 5$, $Pr = 1$ and $r_i/r_o = 0.35$. This benchmark is a widely studied subcritical dynamo solution, and therefore provides an ideal case on which to test our optimisation procedure. The initial temperature field for the benchmark is specified as

$$T = \frac{r_i r_o}{r} - r_i + (a(1 - 3x^2 + 3x^4 - x^6) Y_4^4(\phi, \theta) + \text{c.c.}), \quad (3.1)$$

with $x = 2r - r_i - r_o$ and $a = 0.1$, driving an $m = 4$ convective instability (‘case 0’ of Christensen *et al.* (2001): non-magnetic convection). While the resulting hydrodynamic state is linearly stable to magnetic perturbations, it is (nonlinearly) unstable to some finite-amplitude magnetic perturbations that trigger dynamo action. Christensen *et al.* (2001) suggest that the basin of attraction of this dynamo solution may be small, and therefore recommend prescribing a mainly dipolar, initial magnetic field of the form

$$\mathcal{T} = \frac{4\sqrt{5}\pi}{3} \sin(\pi[r - r_i]) Y_2^0(\phi, \theta), \quad (3.2)$$

Label	K	M	ω	Stable	Growth rate
TW0	58.35	0	0.000182	Yes	−2.06 (−0.80)
TW1	30.77	626.5	−0.00310	Yes	−0.56
TW2	29.72	459.9	−0.00250	No	1.13

Table 1. Travelling wave solutions of the governing equations with $m = 4$ symmetry. Here, ω is the frequency at which they drift relative to the reference frame, and K and M give their dimensionless kinetic and magnetic energies, respectively. Also displayed are the (dimensionless) least stable growth rates of the solutions to perturbations. For TW0, the growth rate is displayed for a purely hydrodynamic perturbation, with the growth rate for a purely magnetic perturbation also shown in parentheses. With our definition of ω , the numerical values differ from the drift frequencies of Christensen *et al.* (2001) by a factor Ek .

$$\mathcal{P} = -\frac{5}{4r^2} \sqrt{\frac{\pi}{3}} (r^3[3r - 4r_o] + r_i^4) Y_1^0(\phi, \theta), \quad (3.3)$$

with (dimensionless) magnetic energy $M_0 = 1215$, which kickstarts an $m = 4$ dynamo solution. We note that as at the benchmark parameters the system is hydrodynamically unstable to multiple modes of convection, other dynamo solutions are possible using different thermal forcing wavenumbers – see Feudel *et al.* (2017), for instance. In fact, we have checked that our procedure works without modification for identifying solutions with $m = 5$ symmetries by replacing the Y_4^4 spherical harmonic in (3.1) with Y_5^5 . However, as the overall dynamical picture is identical to the $m = 4$ case, we do not present it here.

Using the Newton-hookstep procedure outlined in § 2.3, we find the system to possess three $m = 4$ travelling wave (TW) solutions at our working parameters, characterised by the energies and drift frequency (relative to the already-rotating reference frame) displayed in table 1. The table shows one purely hydrodynamic solution (TW0) which corresponds to the ‘case 0’ benchmark of Christensen *et al.* (2001). This hydrodynamic solution arises via a supercritical Hopf bifurcation at $\tilde{Ra}_c = 55.9$ (see the weakly nonlinear analysis of Skene & Tobias (2024), for instance), and is linearly stable to both hydrodynamic and magnetic perturbations at our working parameters (with growth rates displayed in table 1). The two other TW solutions are MHD solutions that form at a saddle node bifurcation at $\tilde{Ra} \sim 98$ (Feudel *et al.* 2017) connected to the purely convecting solution: one stable (TW1) and corresponding to the dynamo (‘case 1’ of Christensen *et al.* 2001); the other unstable (TW2). Figure 2 shows slices of TW1, showing that the nonlinear state is mainly dipolar, with a quadrupolar toroidal part (evident from the radial component of the current). We note here that TW2 has a very similar structure to TW1, which is to be expected since we are close to the saddle node bifurcation, although it shows slightly more spatial localisation than TW1. This localisation becomes more apparent at higher Rayleigh numbers, where the two states start to deviate in structure.

We now let the system evolve to the saturated, purely hydrodynamic, $m = 4$ convective solution (TW0). This state is then seeded with a magnetic field of the form given by Christensen *et al.* (2001), i.e. (3.2)–(3.3); however, the field amplitude is rescaled to have a given initial magnetic energy M_0 , which is gradually decreased (through a series of repeated simulations) down to the point where the initial field fails to trigger a dynamo. The energy time series for two slightly different energy budgets close to the threshold are shown in figure 3. Initially, the evolution is very similar for both cases and is characterised by a transient amplification of the magnetic field, which extracts energy from the convection. This transient growth occurs on a dynamical time scale and is followed by a plateau until $t \approx 2.5$ (i.e. half an ohmic time scale). At this point, the evolutions differ, with the $M_0 = 345$ case achieving further growth in magnetic energy to reach a new stable

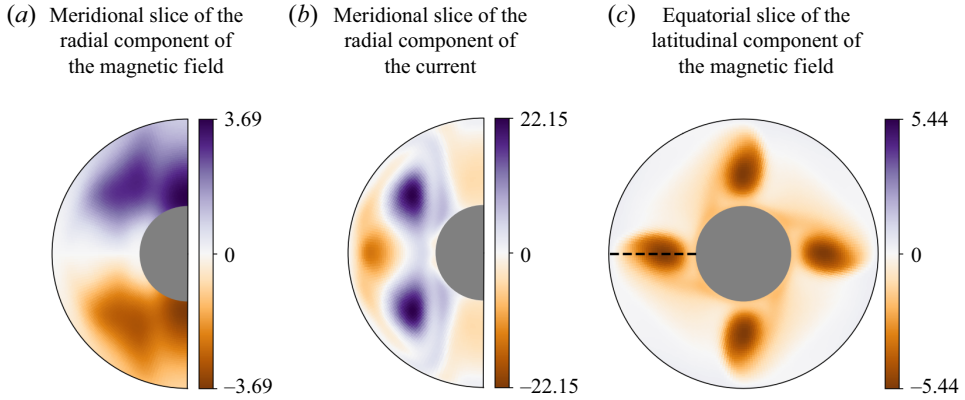


Figure 2. Slices of the $m = 4$ nonlinear dynamo state (TW1). The dashed line on the equatorial slice indicates where the meridional slices were taken.

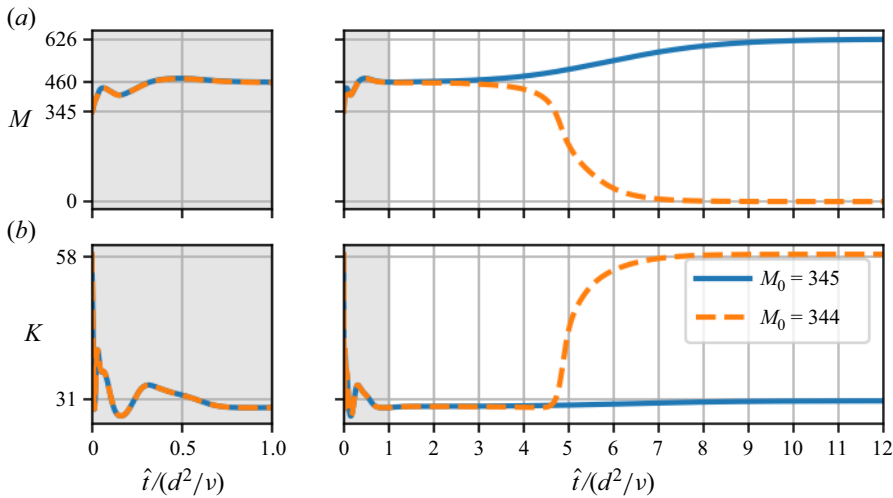


Figure 3. (a) Magnetic energy and (b) kinetic energy evolution, with the rescaled benchmark with two different initial energies.

MHD state (corresponding to the TW1 dynamo) over a few ohmic time scales, and the $M_0 = 344$ case magnetic field decaying back to the purely hydrodynamic state TW0 (for future reference we label the rescaled benchmark initial condition with $M_0 = 344$ as RB). Remarkably, the flow state near $t \approx 2.5$ corresponds to the unstable MHD state TW2. By performing a stability calculation of TW2, we confirm that it has a single unstable direction (see table 1). Thus TW2 is found to be a saddle point, with both solutions initially approaching TW2 along its stable directions. The flow can then either approach TW0 or TW1 along this unstable direction, with which travelling wave the flow approaches depending on which side of the saddle the state lies.

3.2. Optimised seeds

With the baseline behaviour for our system established, we now seek to identify the spatial structure of the lowest-energy magnetic perturbation required to reach the dynamo solution, using the optimisation procedure outlined in § 2.2. In other words, what is the

smallest magnetic field initial condition, with the hydrodynamic initial condition fixed as TW0, that leads to TW1? Following Mannix *et al.* (2022), we initially consider to that end the cost functional with cost integrand

$$J_I = \frac{1}{2EkPm} \mathbf{B} \cdot \mathbf{B}, \quad (3.4)$$

which corresponds to maximising the time-integrated magnetic energy. As our expected magnetic field is large-scale, we initialise the optimisation procedure with a random guess with spherical harmonic degrees up to $\ell = m = 2$, and a random radial dependence with Jacobi polynomial degree up to four. We note here that we have tested random initial guesses up to $\ell = m = 4$. The behaviour of the optimisation is initially similar, with the seed magnetic field converging to a large-scale solution. However, the less sparse spectra of the $\ell = m = 4$ guess causes slower convergence, with the optimisation spending a long time removing the increased amount of fine-scale components that build up quicker when using this guess. Even though this guess is still large-scale, the initial strongly nonlinear transient period, in which the magnetic field takes energy from the convection, leads to updates that initially increase the higher-scale components of $\mathbf{B}_0$. We further describe the convergence behaviour of the optimisation routine in Appendix C. Based on the quick evolution of transients in the previous section, we start by considering a short optimisation window $t_{opt} = 0.2$. The initial magnetic energy budget is specified to be $M_0 = 344$ as this is below the threshold at which a dynamo can be reached from the baseline solution.

Because the optimisation procedure only identifies local extrema, it was repeated multiple times starting from different realisations of the noisy initial guess. This resulted in identifying several optimised seeds (all well converged), corresponding to distinct local extrema. While all the optimised seeds did trigger strong transient growth of the magnetic energy when used as initial conditions for direct simulations, it was nevertheless found upon long-time integration ($t \gg t_{opt}$) that most of these seeds did not successfully kickstart a dynamo. Figure 4 shows (solid blue line) the magnetic and kinetic energy evolution starting from an (ultimately non-dynamo) total-energy-based optimal seed. The figure shows that despite success of the optimisation leading to the transient growth of magnetic energy being considerably larger than that of the baseline solution (dashed orange line), the magnetic energy eventually decays and no dynamo solution is reached. We also show (dotted red line) the magnetic and kinetic energy evolution of an optimised seed that reaches a dynamo solution. This solution shows increased transient growth in the optimisation window, and eventually reaches a dynamo solution, showcasing the coexistence of multiple local optima.

Increasing the time horizon of the optimisation procedure to $t_{opt} = 0.4$ was not found to add robustness to the computation of optimal seeds, as multiple local extrema characterised by strong transient growth (but not necessarily a long-time dynamo) continue to coexist at such target times. While it is tempting to further increase t_{opt} – indeed, we expect that optimising over (very) long times would ultimately discriminate between dynamo seeds and magnetic perturbations yielding only transient growth and ultimately decay – for long target times convergence becomes increasingly difficult (numerical cost notwithstanding). It is therefore more realistic to define another cost functional that could possibly reduce the number of local extrema, yielding more robust seed identification. This is consistent with the observation of Kerswell *et al.* (2014) that ‘the measure used as the objective functional and the norm constraining the initial condition do not need to be the same’ for computing minimal seeds.

To that end, we now consider an alternative cost functional, and optimise the energy in the $m = 0$, or axisymmetric, part of $\mathbf{B}$. This can be achieved by considering the cost

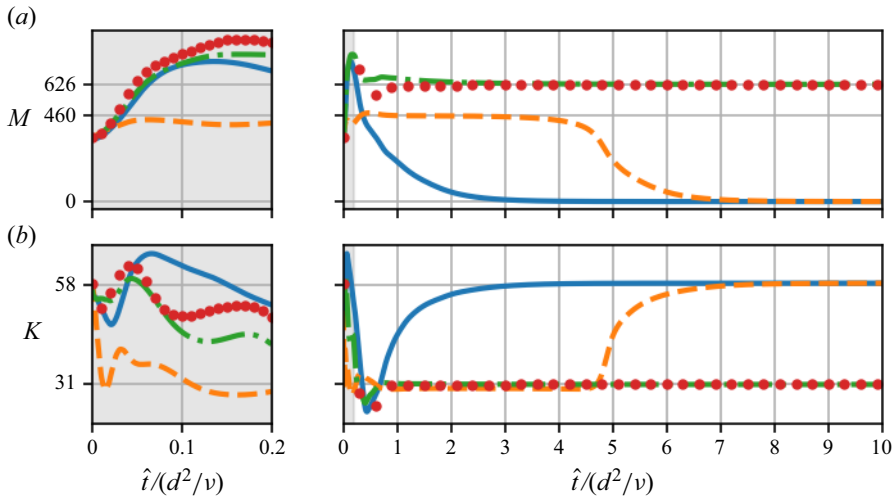


Figure 4. Comparison of the (a) magnetic energy and (b) kinetic energy evolution with four different initial conditions with $M_0 = 344$. The initial conditions are the optimised seeds obtained by optimising the total magnetic energy with a random initial guess (solid blue line), optimising the total magnetic energy with an initial guess of the rescaled benchmark initial condition (RB) (dotted red line), and optimised seed obtained by optimising the energy in the $m = 0$ part of the magnetic field starting from a random guess (dash-dotted green line). We also show the time series obtained without optimisation, starting directly from the rescaled benchmark initial condition RB (dashed orange line). The plots on the left show the energy evolution in the optimisation window $t_{opt} = 0.2$, and the plots on the right show the long-time evolution.

integrand

$$J_I = \frac{1}{2 Ek Pm} \bar{\mathbf{B}} \cdot \bar{\mathbf{B}}, \quad (3.5)$$

with

$$\bar{\mathbf{B}} = \frac{1}{2\pi} \int_0^{2\pi} \mathbf{B} d\phi. \quad (3.6)$$

The resulting cost functional has a physical motivation. It is known that dynamos with a strong magnetic field have a tendency to be more dipolar and large-scale. Optimising the zonal mean magnetic energy promotes flows that are large-scale and on which diffusion operates slowly; this will promote the possibility of dynamo action with a strong field. It is hoped that this physically motivated cost functional will therefore give more robust results in finding the subcritical dynamo.

The result of optimising this cost functional is also shown in figure 4 (dash-dotted green line), and we label this optimal seed OS1. In contrast to optimising the total magnetic energy, optimising this new cost functional is robust in the sense that it was found to always identify the same optimal seed (regardless of the random initial guess used to initialise the algorithm). Furthermore, the computed optimal seed does indeed lead to a dynamo. Incidentally, it is also found to outperform some ‘optimal’ seeds that (locally) maximise the total magnetic energy (as exemplified in figure 4). Based on these observations, one could perhaps speculate that using this new cost functional might be a way to ‘smooth out’ the optimisation landscape, possibly suppressing the multiple local extrema and thus providing a good proxy for the global extremum of the first (total energy) optimisation problem – although this is highly uncertain.

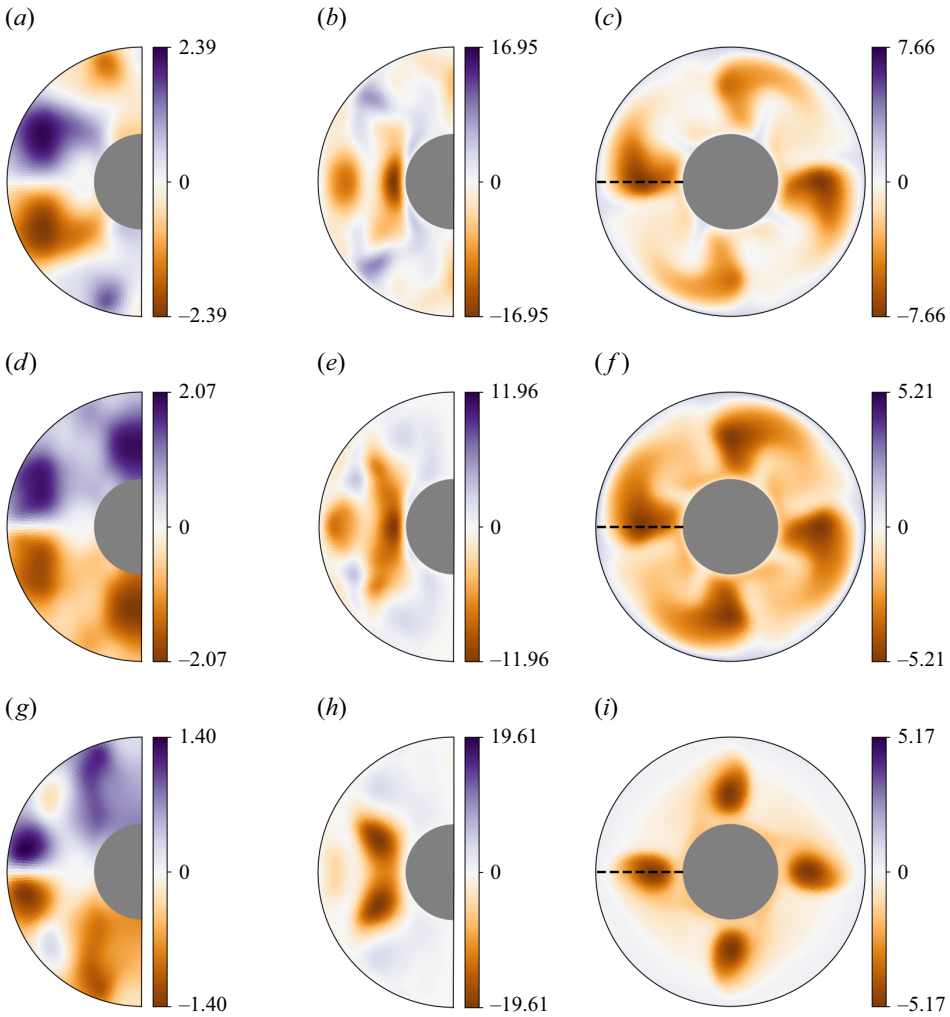


Figure 5. (a–c) Slices of an example optimal seed (non-dynamo) obtained with the total magnetic energy cost functional, with $t_{opt} = 0.2$ and $M_0 = 344$. (d–f) Slices of the optimal seed (dynamo) obtained with the axisymmetric magnetic energy cost functional with $t_{opt} = 0.2$ and $M_0 = 344$. (g–i) Slices of the optimal seed (dynamo) obtained with the axisymmetric magnetic energy cost functional with $t_{opt} = 0.4$ and $M_0 = 162$. Here, (a,d,g) show meridional slices of the radial component of the magnetic field, (b,e,h) show meridional slices of the radial component of the current, and (c,f,i) show equatorial slices of the latitudinal component of the magnetic field. The dashed lines on the equatorial slices indicate where the meridional slices were taken.

The optimal ‘non-dynamo’ seed obtained with the total energy cost functional, and optimal dynamo seed obtained by maximising the axisymmetric energy, are shown in figure 5. Clearly, there are significant differences between the two seeds, with the non-dynamo and dynamo seeds having a dominant quadrupolar and dipolar poloidal part, respectively. From the equatorial slice, it is also evident that the non-dynamo seed has strong $m = 2$ and $m = 4$ components, whereas the dynamo seed is predominantly $m = 4$. This visual comparison is also confirmed by computing the spectra of each seed, which also reveals the presence of a strong quadrupolar toroidal field for the dynamo seed (similarly to the benchmark field). Although not shown, the optimal dynamo seed obtained with the total energy cost functional has a structure similar to that obtained using the

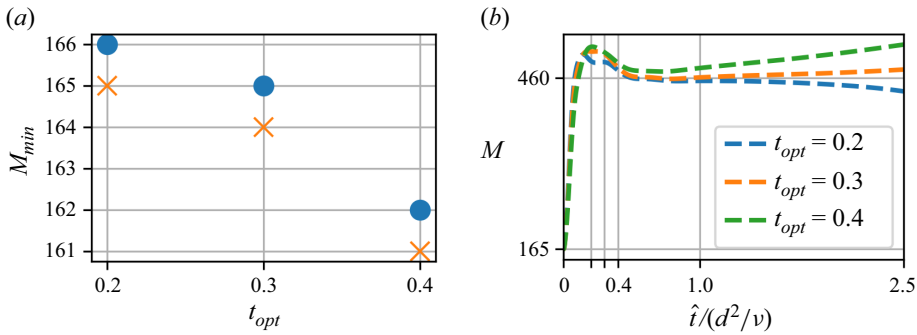


Figure 6. Robustness of the optimisation results with respect to the time horizon t_{opt} . (a) The minimum initial magnetic energy budget to find a dynamo solution for a given time horizon t_{opt} ($m = 0$ cost functional). Circles show runs that succeed in finding a dynamo, and crosses indicate that no dynamo was found. (b) Magnetic energy time series for $M_0 = 165$ starting from optimised seeds ($m = 0$ cost functional) computed with different time horizons.

axisymmetric energy cost functional. From figure 4, we see that the non-dynamo seed does not attract to the edge state, and quickly decays to TW0, which can now be attributed to it having symmetries different to those of TW1 and TW2. Although the dynamo seeds found with the total energy and axisymmetric energy cost functionals are similar, we also see from figure 4 that the seed obtained by optimising the total energy shows more transient growth in the optimisation window. This is to be expected, as it is found by directly optimising the total energy, whereas the seed obtained with the axisymmetric energy only indirectly maximises this quantity. However, we also see that after the optimisation window, the $m = 0$ seed more directly approaches the dynamo. This indicates that the slightly differing structure of the total-energy optimising dynamo seed, whilst leading to more transient growth, comes at the expense of not directly targeting a dynamo solution. The result is that the $m = 0$ energy based cost functional is the more appropriate choice for finding seeds that lead to dynamo action with short optimisation time horizons.

3.3. ‘Minimal’ seeds and dynamical landscape

With a robust cost functional identified, we now turn our attention to lowering the energy of the initial seed. To this end, we perform a series of optimisations maximising the total axisymmetric energy, for gradually decreasing initial energy budgets at fixed time horizon. Each optimal seed then serves as initial condition for a long-time direct numerical simulation: if a dynamo state is found at long times, then M_0 is decreased further, and a new optimisation is performed. We thus determine the initial energy M_0^* below which the computed optimal seed no longer triggers a subcritical dynamo, and repeat the whole procedure for increasing time horizon. Indeed, Pringle & Kerswell (2010) and Kerswell *et al.* (2014) already noted that the choice of time horizon is crucial in accurately determining the amplitude of the minimal seed for transition to turbulence. As shown in figure 6(a), M_0^* was found not to vary much as the target time was increased to $t_{\text{opt}} = 0.4$. Moreover, the effect of lengthening the optimisation time horizon is shown in figure 6(b) for an initial energy budget $M_0 = 165$; at this energy, $t_{\text{opt}} = 0.2$ is not long enough to optimise an initial condition that yields a dynamo. We see that lengthening the time horizon to $t_{\text{opt}} = 0.3$ and $t_{\text{opt}} = 0.4$ enables the optimisation procedure to converge to a dynamo solution. Although the early time series in figure 6(b) show that there are differences in the magnetic energy evolution as t_{opt} is changed, we see that these changes become significantly fewer as t_{opt} is increased. This indicates that the computed extremum

eventually becomes insensitive to t_{opt} , which we can attribute to the fact that for $t_{opt} \approx 0.4$, the transient growth regime has ended, and the flow has approached the unstable MHD state TW2. For longer time horizons, the optimisation procedure struggles to converge. The minimal seed obtained with $M_0 = 162$ and $t_{opt} = 0.4$ is labelled OS2. Additionally, we also label the optimal seed obtained with $M_0 = 161$ and $t_{opt} = 0.4$, which does not reach a dynamo solution, OS3.

Based on these findings, we conclude that $t_{opt} = 0.4$ is the most practical choice of time horizon, enabling the optimisation procedure to converge, while at the same time being long enough to ensure robustness with respect to changes in t_{opt} . Recall that this is non-dimensionalised with respect to the viscous diffusion time, and this is a dynamical time scale. In other words, we only need to optimise up to when the state approaches the unstable edge state, at which point we require the instability of the edge state to lead to the dynamo solution. This is similar to previous studies finding minimal seeds for flows with edge states (e.g. see the work of Duguet *et al.* 2010; Juniper 2011). The spatial structure of the ‘minimal’ seed (i.e. the lowest-energy dynamo seed that we could find) is shown in figure 5(g–i). It is similar to the optimal seed found at higher energy (OS1) $M_0 = 344$ (figure 5d–f); however, it shows more localisation (indicating a more concentrated magnetic energy density profile in all spatial directions). We note that neither of these magnetic fields closely resembles either the nonlinear dynamo state (figure 2) or the unstable state (similar to figure 2). We see that although they have some similar features, such as a dominant dipolar poloidal part, and quadrupolar toroidal part, the radial structure and energy spectra are significantly different between the two solutions – and are more complex in the case of the optimal seed. This showcases the need for a systematic procedure that can work from random initial guesses (Mannix *et al.* 2022).

4. Conclusions

Motivated by the search for strong branch solutions of the geodynamo, we have developed code to perform adjoint-based optimal control on a convection-driven, magnetohydrodynamic (MHD) flow in a rotating spherical shell, using the Dedalus environment. We have revisited the well-studied dynamo benchmark of Christensen *et al.* (2001) as an example case to systematically identify minimal magnetic perturbations (in the sense of least magnetic energy) that nonlinearly trigger a convective dynamo, relying on transient growth optimisation.

Whereas in the systems considered by Mannix *et al.* (2022), maximising the time-integrated magnetic energy over a fraction of the ohmic time scale was found suitable to robustly identify dynamo solutions and their minimal seeds, in the present system this approach appears less reliable. Indeed, the optimisation procedure identified multiple magnetic conditions that (locally) maximise magnetic energy growth but do not trigger a sustained dynamo at longer times. We are facing here two practical difficulties: first, the optimisation problem is highly non-convex; second, we are restricted to relatively short time horizons for the optimisation. Indeed, we found optimisation to become highly sensitive to small changes in the initial conditions as the target time approached $t_{opt} = 0.4$, rendering convergence difficult; this observation is, however, readily explained by the fact that the system was approaching the (unstable) edge state by this time. Moreover, our use of a differentiate-then-discretise approach to solve the adjoint equations (Mannix *et al.* 2024) causes our gradient estimate to deteriorate as the time horizon is increased. Yet the optimisation question in which we are truly interested is best investigated over (very) long time horizons; indeed, the very definition of dynamo action is the existence of an initial condition such that magnetic energy does not vanish as time goes to infinity! Were t_{opt}

allowed to become long enough while maximising the magnetic energy, then all initial conditions yielding failed dynamos would essentially yield negligible objective functions, while any local maximum would have to be a dynamo seed, whereas for short target times, maximising the total magnetic energy does not penalise local extrema corresponding to initial conditions that trigger strong transient growth ultimately followed by decay. In other words, we need for practical reasons to devise a cost functional that when extremised over a relatively short time horizon, acts as a robust proxy for discriminating long-time evolution.

This situation is reminiscent of a similar issue encountered in optimal mixing problems, where multiscale cost functions can act as shorter target time proxies for long-time flow evolution. Specifically, minimising Sobolev norms of negative indexes called the mix-norms (Mathew, Mezić & Petzold 2005; Doering & Thiffeault 2006) by a short target time has been shown to effectively yield the same results (in terms of both the optimal perturbations and the achieved objective functions) as minimising the flow variance over long target times, whether in unstratified (Foures, Caulfield & Schmid 2014; Heffernan & Caulfield 2022) or stratified (Marcotte & Caulfield 2018) fluids. Mix-norms thus provide a computationally efficient way of identifying the long-term, physical mixing processes while working with affordable target times.

In the present case, this difficulty was overcome by considering an alternative ‘more physics-based’ cost functional such that the magnetic energy of the axisymmetric component of the magnetic field is maximised. Choosing this cost functional not only appears to bias the optimisation procedure towards local extrema that do systematically trigger a sustained subcritical dynamo at long times, but is also found more robust in the sense that repeated optimisations consistently identified the same dynamo seed, which is also found not to significantly vary as the target time is increased.

Moreover, the identification of travelling wave solutions using a Newton-hookstep algorithm (Willis 2019a) confirms that the evolution initiated by these optimal dynamo seeds brings the system close to an unstable MHD solution within the time horizon that we used for the computation of the minimal seed. From that unstable state, the fate of the system is determined by the energy budget allowed for the initial seeds, travelling away towards either the dynamo attractor (which also corresponds to the linearly stable, magnetised travelling wave solution identified by the Newton-hookstep algorithm) or a non-dynamo one (the linearly stable, purely hydrodynamic travelling wave solution).

Therefore, we can now form a picture of the dynamical landscape of the various basins of attraction at our working parameters, which summarises the findings of the present paper. Figure 7 shows the trajectory initiated by different initial conditions in the kinetic–magnetic energy phase space, along with the travelling wave solutions (hydrodynamic TW0, stable MHD TW1, unstable MHD TW2). Whereas the Christensen *et al.* (2001) benchmark initial condition (rescaled here to the smallest energy at which it can trigger a dynamo) approaches the unstable edge state TW2 in a (somewhat) inefficient way, undergoing repeated energy transfers between the kinetic, magnetic and thermal energy reservoirs, the less energetic, minimal dynamo seed yields an important amplification of the magnetic energy by maintaining a stronger convection at early times, while energy transfers at later times are largely suppressed. This strategy rapidly propels the system to the vicinity of the edge state, from which it is repelled away towards the dynamo attractor or (when the initial energy budget is too small) the purely hydrodynamic one.

We conclude by returning to our motivation. What is the relevance of our approach for studies of the geodynamo? Finding minimal seeds and exact nonlinear solutions will be important in promoting our understanding in a field where it is impossible to perform numerical calculations in the correct parameter regime. We intend to build on this work by investigating how the minimal seeds and exact nonlinear solutions change

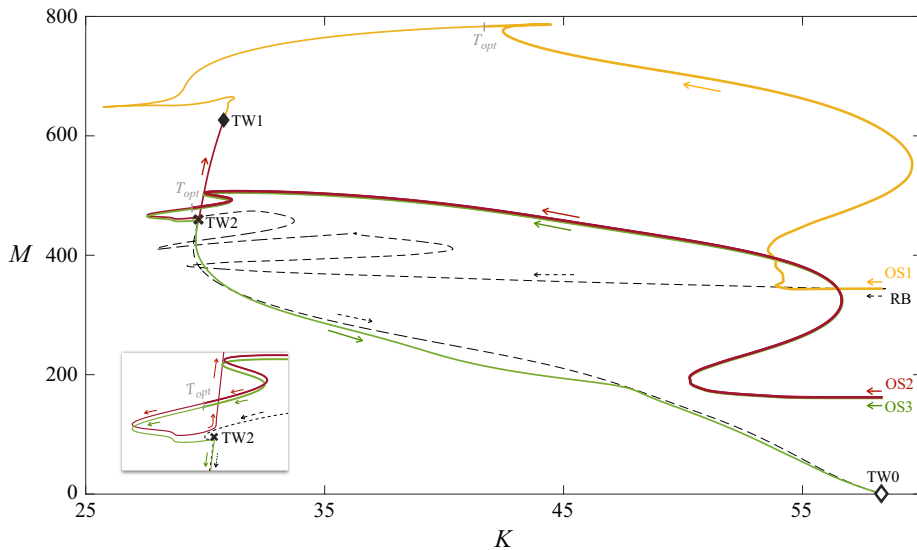


Figure 7. Dynamical landscape, projected in the kinetic energy (K) versus magnetic energy (M) phase space. The three travelling wave states are indicated by TW0 (empty diamond, linearly stable hydrodynamic state), TW1 (full diamond, linearly stable dynamo state) and TW2 (black cross, linearly unstable MHD state, edge state). Four trajectories are shown, corresponding to simulations initiated with, respectively, a rescaled benchmark magnetic field $M_0 = 344$, dashed black line), and three optimal seeds identified with the $m = 0$ energy cost: OS1 (solid yellow line, $M_0 = 344$, $t_{opt} = 0.2$), the minimal dynamo seed OS2 (solid red line, $M_0 = 162$, $t_{opt} = 0.4$), and OS3 (solid green line, $M_0 = 161$, $t_{opt} = 0.4$). Thicker lines mark the span of the time horizons for the optimisation procedures.

as the underlying parameters are varied. For example, how does the minimal energy of the magnetic field perturbation change as one varies the parameters along the dynamo path (Aubert *et al.* 2017) taking one to the distinguished limit of Dormy (2016). Does the energy required to take one to a fully nonlinear strong field solution decrease as the Ekman number is decreased? This behaviour would be reminiscent of the behaviour with Reynolds number of the energy of the minimal seed in wall-bounded shear flows (see e.g. Duguet *et al.* 2013). If this were the case, then even a tiny magnetic energy perturbation would be enough to recover the dynamo if it entered a non-dynamo state – even if the hydrodynamic solution were formally linearly stable. Furthermore, tracing the behaviour of exact solutions as parameters are varied is a very efficient way of examining whether (admittedly specialised) solutions could maintain balance as parameters are varied. The bifurcation structure of these solutions (particularly the presence or absence of global bifurcations) may also give hints about mechanisms for achieving reversal of solutions whilst maintaining the correct balance.

Acknowledgements. C.S.S. would like to acknowledge J. Page for helpful discussions on the Newton-hookstep algorithm. The authors thank Y. Duguet for providing valuable feedback on the manuscript. This work was undertaken on ARC4, part of the High Performance Computing facilities at the University of Leeds, UK.

Funding. C.S.S. and S.M.T. acknowledge partial support from a grant from the Simons Foundation (grant no. 662962, GF). They would also like to acknowledge support of funding from the European Union Horizon 2020 research and innovation programme (grant agreement no. D5S-DLV-786780). F.M. acknowledges support from the European Research Council under grant agreement CIRCE 101117412.

Data availability statement. Data that supports the findings of this study is openly available at <https://doi.org/10.5281/zenodo.17313967>.

Term	Adjoint	Boundary term
$\langle \mathbf{y}^\dagger, \nabla^2 \mathbf{y} \rangle$	$\langle \nabla^2 \mathbf{y}^\dagger, \mathbf{y} \rangle$	$\mathbf{y}^\dagger \cdot \nabla \cdot \mathbf{y} - \mathbf{y} \cdot \nabla \cdot \mathbf{y}^\dagger + \mathbf{y}^\dagger \times (\nabla \times \mathbf{y}) + (\nabla \times \mathbf{y}^\dagger) \times \mathbf{y}$
$\langle \mathbf{y}^\dagger, \nabla^2 \mathbf{y} \rangle$	$\langle \nabla^2 \mathbf{y}^\dagger, \mathbf{y} \rangle$	$\mathbf{y}^\dagger \nabla \mathbf{y} - \mathbf{y} \nabla \mathbf{y}^\dagger$
$\langle \mathbf{y}^\dagger, \nabla \mathbf{y} \rangle$	$\langle -\nabla \cdot \mathbf{y}^\dagger, \mathbf{y} \rangle$	$\mathbf{y}^\dagger \mathbf{y}$
$\langle \mathbf{y}^\dagger, \nabla \times \mathbf{y} \rangle$	$\langle \nabla \times \mathbf{y}^\dagger, \mathbf{y} \rangle$	$-\mathbf{y}^\dagger \times \mathbf{y}$
$\langle \mathbf{y}^\dagger, \mathbf{a} \times \mathbf{y} \rangle$	$\langle -\mathbf{a} \times \mathbf{y}^\dagger, \mathbf{y} \rangle$	0

Table 2. Table of terms and their corresponding adjoints, as well as boundary contributions. Effectively, each row of this table can be read as the result of integration by parts. Adjoint for combinations of these operations can be obtained by applying these rules sequentially.

Declaration of interests. The authors report no conflict of interest.

Appendix A. Derivation of the adjoint equations

A.1. Adjoints of vector calculus operations

In order to find the adjoint of our equations, we first describe how to take the adjoints of common vector calculus operations. Vector calculus gives $\mathbf{y}^\dagger \cdot \nabla \mathbf{y} = \nabla \cdot (\mathbf{y}^\dagger \mathbf{y}) - \mathbf{y} \nabla \cdot \mathbf{y}^\dagger$, hence

$$\iiint \mathbf{y}^\dagger \cdot \nabla \mathbf{y} \, dV = \oint (\mathbf{y}^\dagger \mathbf{y}) \cdot d\mathbf{S} - \iiint \mathbf{y} \nabla \cdot \mathbf{y}^\dagger \, dV, \quad (\text{A1})$$

$$\langle \mathbf{y}^\dagger, \nabla \mathbf{y} \rangle = \langle -\nabla \cdot \mathbf{y}^\dagger, \mathbf{y} \rangle + \{\mathbf{y}^\dagger \mathbf{y}\}, \quad (\text{A2})$$

where we have used the divergence theorem. Note that we have also introduced the notation $\{\cdot\}$ to represent the boundary terms. This identity gives the adjoint for both the gradient and divergence operators. For the curl, we can use the identity $\nabla \cdot (\mathbf{y}^\dagger \times \mathbf{y}) = \mathbf{y} \cdot \nabla \times \mathbf{y}^\dagger - \mathbf{y}^\dagger \cdot \nabla \times \mathbf{y}$ to obtain

$$\langle \mathbf{y}^\dagger, \nabla \times \mathbf{y} \rangle = \langle \mathbf{y}, \nabla \times \mathbf{y}^\dagger \rangle - \{\mathbf{y}^\dagger \times \mathbf{y}\}. \quad (\text{A3})$$

To find the adjoint of the scalar Laplacian, we can use our results for the divergence and gradient to obtain

$$\langle \mathbf{y}^\dagger, \nabla^2 \mathbf{y} \rangle = \langle \nabla^2 \mathbf{y}^\dagger, \mathbf{y} \rangle + \{\mathbf{y}^\dagger \nabla \mathbf{y}\} - \{\mathbf{y} \nabla \mathbf{y}^\dagger\}. \quad (\text{A4})$$

The adjoint of the vector Laplacian can also be found from the rules that we have derived so far, by writing $\nabla^2 \mathbf{y} = \nabla(\nabla \cdot \mathbf{y}) - \nabla \times (\nabla \times \mathbf{y})$, yielding

$$\langle \mathbf{y}^\dagger, \nabla^2 \mathbf{y} \rangle = \langle \nabla^2 \mathbf{y}^\dagger, \mathbf{y} \rangle + \{\mathbf{y}^\dagger \nabla \cdot \mathbf{y}\} - \{\mathbf{y} \nabla \cdot \mathbf{y}^\dagger\} + \{\mathbf{y}^\dagger \times (\nabla \times \mathbf{y})\} + \{(\nabla \times \mathbf{y}^\dagger) \times \mathbf{y}\}. \quad (\text{A5})$$

These results are summarised in [table 2](#).

A.2. The adjoint equations

With the adjoints of general vector calculus operations now derived, we can proceed to find the adjoint of our equations by cancelling all first variations of the Lagrangian (2.10). Taking the first variation with respect to the adjoint state ensures that the governing equations are satisfied. The first variation with respect to the state $\mathbf{q}$ gives

$$\nabla_{\mathbf{q}} \mathcal{L} \cdot \delta \mathbf{q} = \left[\frac{\partial \mathcal{I}}{\partial \mathbf{q}}, \delta \mathbf{q} \right] + \left[\mathbf{M}^T \frac{\partial \mathbf{q}^\dagger}{\partial t}, \delta \mathbf{q} \right] + \left[\mathbf{q}^\dagger, \frac{\partial \mathcal{N}(\mathbf{q})}{\partial \mathbf{q}} \delta \mathbf{q} \right], \quad (\text{A6})$$

where integration by parts has been used in time. By finding the adjoint operator $\mathcal{A}^\dagger$ such that

$$\left\langle \mathbf{y}^\dagger, \frac{\partial \mathcal{N}(\mathbf{q})}{\partial \mathbf{q}} \mathbf{y} \right\rangle = \langle \mathcal{A}^\dagger \mathbf{y}^\dagger, \mathbf{y} \rangle, \quad (\text{A7})$$

we can then rearrange (A6) to

$$\nabla_{\mathbf{q}} \mathcal{L} \cdot \delta \mathbf{q} = \left[\mathbf{M}^\top \frac{\partial \mathbf{q}^\dagger}{\partial t} + \mathcal{A}^\dagger \mathbf{q}^\dagger + \frac{\partial \mathcal{J}_I}{\partial \mathbf{q}}, \delta \mathbf{q} \right]. \quad (\text{A8})$$

Requiring this term to be zero is achieved by setting the adjoint to satisfy

$$-\mathbf{M}^\top \frac{\partial \mathbf{q}^\dagger}{\partial t} = \mathcal{A}^\dagger \mathbf{q}^\dagger + \frac{\partial \mathcal{J}_I}{\partial \mathbf{q}}. \quad (\text{A9})$$

In order to find the adjoint operator $\mathcal{A}^\dagger$, (A7) shows us that we must first linearise the governing equations. Integration by parts using the rules in table 2 then yields the adjoint operator. For our equations, this is done using the steps outlined by Chen *et al.* (2018), where the governing equations are rewritten in a form in which Ohm's law (where Ampère's law has been used to relate the magnetic field to the current)

$$\sigma \mathbf{E} = -\mathbf{U} \times \mathbf{B} + \frac{1}{Pm} \nabla \times \mathbf{B}, \quad (\text{A10})$$

and Faraday's law

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E}, \quad (\text{A11})$$

appear separately. In other words, equations for the electric field $\mathbf{E}$ and magnetic field are constrained separately using adjoint variables $\mathbf{E}^\dagger$ and $\mathbf{b}^\dagger$, respectively. As in Chen *et al.* (2018), σ is a relative electrical conductivity that takes the value 1 for $\mathbf{x} \in \mathcal{V}_s$, and 0 for $\mathbf{x} \in \mathbb{R}^3 \setminus \mathcal{V}_s$. This allows the adjoint formulation to more easily incorporate electrically insulating inner and outer regions, where the insulating boundary condition is simply continuity of the radial part of the magnetic field and tangential part of the electric field at the boundaries between $\mathcal{V}_s$ and the insulating regions. Additionally, we introduce a Lagrange multiplier $\Pi^\dagger$, which constrains the solenoidal condition $\nabla \cdot \mathbf{B} = 0$.

Following this procedure, we obtain the adjoint equations

$$\begin{aligned} Ek \left(-\frac{\partial \mathbf{u}^\dagger}{\partial t} + \nabla \times (\mathbf{U} \times \mathbf{u}^\dagger) + \mathbf{u}^\dagger \times \boldsymbol{\omega} - \nabla^2 \mathbf{u}^\dagger \right) - 2\mathbf{e}_z \times \mathbf{u}^\dagger + \nabla p^\dagger \\ = -T^\dagger \nabla T - \mathbf{B} \times \mathbf{E}^\dagger + \frac{\partial \mathcal{J}_I}{\partial \mathbf{U}}, \\ \nabla \cdot \mathbf{u}^\dagger = 0, \\ -\frac{\partial T^\dagger}{\partial t} - \nabla \cdot (\mathbf{U} T^\dagger) = \widetilde{Ra} \frac{\mathbf{r}}{r_0} \cdot \mathbf{u}^\dagger + \frac{1}{Pr} \nabla^2 T^\dagger + \frac{\partial \mathcal{J}_I}{\partial T}, \\ -\frac{\partial \mathbf{b}^\dagger}{\partial t} = -\nabla \Pi^\dagger + \frac{1}{Pm} \nabla \times (\mathbf{B} \times \mathbf{u}^\dagger) - \frac{1}{Pm} (\nabla \times \mathbf{B}) \times \mathbf{u}^\dagger + \mathbf{U} \times \mathbf{E}^\dagger \\ + \frac{1}{Pm} \nabla \times \mathbf{E}^\dagger + \frac{\partial \mathcal{J}_I}{\partial \mathbf{B}}, \\ \mathbf{E}^\dagger = -\nabla \times \mathbf{b}^\dagger, \end{aligned} \quad (\text{A12})$$

for $\mathbf{x} \in \mathcal{V}_s$, and

$$\nabla \times \mathbf{b}^\dagger = 0, \quad (\text{A13})$$

for $\mathbf{x} \in \mathcal{V}_i \cup \mathcal{V}_o$. Alternatively, using the fact that $\mathbf{E}^\dagger = -\nabla \times \mathbf{b}^\dagger$ for $\mathbf{x} \in \mathcal{V}_s$ gives the form used in (2.14).

These equations are similar to the form derived by Li *et al.* (2011, 2014) for data assimilation applications; in data assimilation, the optimisation procedure is designed to align the three-dimensional dynamics with the observations, and is used not only for prediction but for making inferences about the three-dimensional evolution of the system. We see that the adjoint equations currently have more unknowns than equations due to the inclusion of the $\nabla \Pi^\dagger$ term in the adjoint magnetic field equation. As in the study of Chen *et al.* (2018), we see that the adjoint equations imply a gauge freedom for the adjoint magnetic field $\mathbf{b}^\dagger$. As it is only the curl of the adjoint magnetic field that occurs, setting $\mathbf{b}^\dagger \rightarrow \mathbf{b}^\dagger + \nabla \chi$ does not change the equations. This gauge freedom is a consequence of the direct equations being overdetermined for $\mathbf{B}$, with the induction equation naturally preserving the divergence of $\mathbf{B}$, which is physically always initialised with a divergence-free field. It turns out to be convenient to then fix this gauge freedom by choosing $\mathbf{b}^\dagger$ to be solenoidal, i.e. we solve (2.14) together with $\nabla \cdot \mathbf{b}^\dagger = 0$, which determines the value of $\Pi^\dagger$. This gauge choice has two main benefits: first, that the update condition (2.13) is now guaranteed to give an updated magnetic field that is divergence-free; and second, the boundary conditions for the adjoint magnetic field will be the same as for the direct magnetic field (see e.g. Chen *et al.* 2018). Cancelling the boundary terms that arise from integration by parts of terms involving the velocity and temperature fields gives the adjoint boundary conditions.

Appendix B. Cost functionals

For each cost functional considered, we need to be able to find its contribution to the adjoint equation. For optimising the total magnetic energy (3.4), we obtain

$$\frac{\partial \mathcal{J}_I}{\partial \mathbf{B}} = \frac{\mathbf{B}}{Ek Pm}. \quad (\text{B1})$$

Similarly, for optimising the axisymmetric energy (3.5), we find

$$\frac{\partial \mathcal{J}_I}{\partial \mathbf{B}} = \frac{\bar{\mathbf{B}}}{Ek Pm}. \quad (\text{B2})$$

Appendix C. Convergence

The typical convergence behaviour of the optimisation procedure is illustrated in figure 8, which shows the evolution of both the cost functional and the relative residual over the iterations. We have defined the residual as

$$r = \frac{\iint \mathbf{g}_\perp \cdot \mathbf{g}_\perp dV}{\mathcal{J}}, \quad (\text{C1})$$

where $\mathbf{g}_\perp$ is the Riemannian gradient that does not change the norm of $\mathbf{B}_0$. The cost functional rapidly increases at first, as the large-scale structure of the initial magnetic field undergoes rapid changes from its random initial guess. We perform the optimisation procedure until either the residual is less than 10^{-3} or 120 iterations have occurred. After approximately 20 iterations, the cost functional tends towards a constant value as minor, mostly small-scale adjustments occur, while the relative residual steadily decreases.

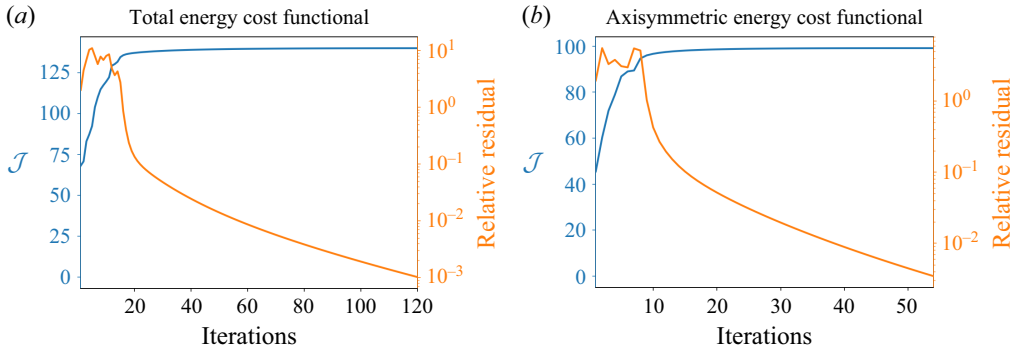


Figure 8. Convergence behaviour for the optimisation procedure described in § 2.2 with two different cost functionals, with $t_{opt} = 0.2$ and $M_0 = 344$ – and with $l = m = 2$ initially.

REFERENCES

- ABSIL, P.-A., MAHONY, R. & SEPULCHRE, R. 2007 *Optimization Algorithms on Matrix Manifolds*. Princeton University Press.
- ARSENOVIĆ, P., *et al.* 2024 Global impacts of an extreme solar particle event under different geomagnetic field strengths. *Proc. Natl Acad. Sci. USA* **121** (28), e2321770121.
- AUBERT, J., GASTINE, T. & FOURNIER, A. 2017 Spherical convective dynamos in the rapidly rotating asymptotic regime. *J. Fluid Mech.* **813**, 558–593.
- AUPY, G. & HERRMANN, J. 2017 Periodicity in optimal hierarchical checkpointing schemes for adjoint computations. *Optim. Meth. Softw.* **32** (3), 594–624.
- AUPY, G., HERRMANN, J., HOVLAND, P. & ROBERT, Y. 2016 Optimal multistage algorithm for adjoint computation. *SIAM J. Sci. Comput.* **38** (3), C232–C255.
- AURNOU, J.M. & KING, E.M. 2017 The cross-over to magnetostrophic convection in planetary dynamo systems. *Proc. R. Soc. A: Math. Phys. Engng Sci.* **473**, 20160731.
- BACKUS, G. 1958 A class of self-sustaining dissipative spherical dynamos. *Ann. Phys.* **4** (4), 372–447.
- BRAGINSKY, S.I. & ROBERTS, P.H. 1995 Equations governing convection in Earth’s core and the geodynamo. *Geophys. Astrophys. Fluid Dyn.* **79** (1–4), 1–97.
- BUDANUR, N.B., SHORT, K.Y., FARAZMAND, M., WILLIS, A.P. & CVITANOVIĆ, P. 2017 Relative periodic orbits form the backbone of turbulent pipe flow. *J. Fluid Mech.* **833**, 274–301.
- BURNS, K.J., VASIL, G.M., OISHI, J.S., LECOANET, D. & BROWN, B.P. 2020 Dedalus: a flexible framework for numerical simulations with spectral methods. *Phys. Rev. Res.* **2**, 023068.
- CATTANEO, F., BODO, G. & TOBIAS, S.M. 2020 On magnetic helicity generation and transport in a nonlinear dynamo driven by a helical flow. *J. Plasma Phys.* **86** (4), 905860408.
- CHANDLER, G.J. & KERSWELL, R.R. 2013 Invariant recurrent solutions embedded in a turbulent two-dimensional Kolmogorov flow. *J. Fluid Mech.* **722**, 554–595.
- CHEN, L., HERREMAN, W. & JACKSON, A. 2015 Optimal dynamo action by steady flows confined to a cube. *J. Fluid Mech.* **783**, 23–45.
- CHEN, L., HERREMAN, W., LI, K., LIVERMORE, P.W., LUO, J.W. & JACKSON, A. 2018 The optimal kinematic dynamo driven by steady flows in a sphere. *J. Fluid Mech.* **839**, 1–32.
- CHERUBINI, S., DE PALMA, P., ROBINET, J.-C. & BOTTARO, A. 2010 Rapid path to transition via nonlinear localized optimal perturbations in a boundary-layer flow. *Phys. Rev. E* **82** (6), 066302.
- CHRISTENSEN, U., *et al.* 2001 A numerical dynamo benchmark. *Phys. Earth Planet. Inter.* **128**, 25.
- CONSTABLE, C. & KORTE, M. 2006 Is Earth’s magnetic field reversing? *Earth Planet. Sci. Lett.* **246** (1), 1–16.
- DEGUCHI, K. 2020 Streaky dynamo equilibria persisting at infinite Reynolds numbers. *J. Fluid Mech.* **884**, R3.
- DEGUCHI, K. 2022 Shear-driven Hall-magnetohydrodynamic dynamos. *J. Fluid Mech.* **932**, A46.
- DENNIS, J.E. & SCHNABEL, R.B. 1996 *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Society for Industrial and Applied Mathematics.
- DOERING, C.R. & THIFFEAULT, J.-L. 2006 Multiscale mixing efficiencies for steady sources. *Phys. Rev. E* **74**, 025301.

- DOLCI, D.I., MADDISON, J.R., HAM, D.A., PALLEZ, G. & HERRMANN, J. 2024 checkpoint_schedules: schedules for incremental checkpointing of adjoint simulations. *J. Open Source Softw.* **9** (95), 6148.
- DORMY, E. 1997 Modélisation numérique de la dynamo terrestre. PhD thesis, Institut de physique du globe, Paris.
- DORMY, E. 2016 Strong-field spherical dynamos. *J. Fluid Mech.* **789**, 500–513.
- DOUGLAS, S.C., AMARI, S. & KUNG, S.-Y. 1998 Gradient adaptation under unit-norm constraints. In *Ninth IEEE Signal Processing Workshop on Statistical Signal and Array Processing (Cat. No.98TH8381)*, pp. 144–147. IEEE.
- DUGUET, Y., BRANDT, L. & LARSSON, B.R.J. 2010 Towards minimal perturbations in transitional plane Couette flow. *Phys. Rev. E* **82**, 026316.
- DUGUET, Y., MONOKROUSOS, A., BRANDT, L. & HENNINGSON, D.S. 2013 Minimal transition thresholds in plane Couette flow. *Phys. Fluids* **25** (8), 084103.
- DUGUET, Y., PRINGLE, C.C.T. & KERSWELL, R.R. 2008 Relative periodic orbits in transitional pipe flow. *Phys. Fluids* **20** (11), 114102.
- FEARN, D.R. 1998 Hydromagnetic flow in planetary cores. *Rep. Prog. Phys.* **61** (3), 175.
- FEUDEL, F., TUCKERMAN, L.S., GELLERT, M. & SEEHAFFER, N. 2015 Bifurcations of rotating waves in rotating spherical shell convection. *Phys. Rev. E* **92**, 053015.
- FEUDEL, F., TUCKERMAN, L.S., ZAKS, M. & HOLLERBACH, R. 2017 Hysteresis of dynamos in rotating spherical shell convection. *Phys. Rev. Fluids* **2**, 053902.
- FOURES, D.P.G., CAULFIELD, C.P. & SCHMID, P.J. 2013 Localization of flow structures using ∞ -norm optimization. *J. Fluid Mech.* **729**, 672–701.
- FOURES, D.P.G., CAULFIELD, C.P. & SCHMID, P.J. 2014 Optimal mixing in two-dimensional plane Poiseuille flow at finite Péclet number. *J. Fluid Mech.* **748**, 241–277.
- GLATZMAIER, G.A. & ROBERTS, P.H. 1995 A three-dimensional convective dynamo solution with rotating and finitely conducting inner core and mantle. *Phys. Earth Planet. Inter.* **91** (1), 63–75.
- GRIEWANK, A. 1992 Achieving logarithmic growth of temporal and spatial complexity in reverse automatic differentiation. *Optim. Meth. Softw.* **1** (1), 35–54.
- HEFFERNAN, C. & CAULFIELD, C.P. 2022 Robust and efficient identification of optimal mixing perturbations using proxy multiscale measure. *Phil. Trans. R. Soc. A* **380**, 38020210026.
- HERREMAN, W. 2018 Minimal perturbation flows that trigger mean field dynamos in shear flows. *J. Plasma Phys.* **84** (3), 735840305.
- HERRMANN, J. & AUPY, G.P. 2020 H-Revolve: a framework for adjoint computation on synchronous hierarchical platforms. *ACM Trans. Math. Softw.* **46** (2), 1–25.
- HOLDENRIED-CHERNOFF, D., CHEN, L. & JACKSON, A. 2019 A trio of simple optimized axisymmetric kinematic dynamos in a sphere. *Proc. R. Soc. A: Math. Phys. Engng Sci.* **475** (2229), 20190308.
- HOLLERBACH, R. 2000 Magnetohydrodynamic flows in spherical shells. In *Physics of Rotating Fluids* (ed. C. Egber & G. Pfister), pp. 295–316. Springer.
- JUNIPER, M.P. 2011 Triggering in the horizontal Rijke tube: non-normality, transient growth and bypass transition. *J. Fluid Mech.* **667**, 272–308.
- KERSWELL, R.R. 2018 Nonlinear nonmodal stability theory. *Annu. Rev. Fluid Mech.* **50**, 319–345.
- KERSWELL, R.R., PRINGLE, C.C.T. & WILLIS, A.P. 2014 An optimization approach for analysing nonlinear stability with transition to turbulence in fluids as an exemplar. *Rep. Prog. Phys.* **77** (8), 085901.
- KHAPKO, T., KREILOS, T., SCHLATTER, P., DUGUET, Y., ECKHARDT, B. & HENNINGSON, D.S. 2016 Edge states as mediators of bypass transition in boundary-layer flows. *J. Fluid Mech.* **801**, R2.
- LAN, Y. & CVITANOVIĆ, P. 2004 Variational method for finding periodic orbits in a general flow. *Phys. Rev. E* **69**, 016217.
- LECOANET, D., VASIL, G.M., BURNS, K.J., BROWN, B.P. & OISHI, J.S. 2019 Tensor calculus in spherical coordinates using Jacobi polynomials. Part II: Implementation and examples. *J. Comput. Phys.* **X 3**, 100012.
- LI, K., JACKSON, A. & LIVERMORE, P.W. 2011 Variational data assimilation for the initial-value dynamo problem. *Phys. Rev. E* **84**, 056321.
- LI, K., JACKSON, A. & LIVERMORE, P.W. 2014 Variational data assimilation for a forced, inertia-free magnetohydrodynamic dynamo model. *Geophys. J. Intl* **199** (3), 1662–1676.
- MADDISON, J.R. 2024 Step-based checkpointing with high-level algorithmic differentiation. *J. Comput. Sci.* **82**, 102405.
- MAMUN, C.K. & TUCKERMAN, L.S. 1995 Asymmetry and Hopf bifurcation in spherical Couette flow. *Phys. Fluids* **7** (1), 80–91.
- MANNIX, P.M., PONTY, Y. & MARCOTTE, F. 2022 Systematic route to subcritical dynamo branches. *Phys. Rev. Lett.* **129**, 024502.

- MANNIX, P.M., SKENE, C.S., AUROUX, D. & MARCOTTE, F. 2024 A robust, discrete-gradient descent procedure for optimisation with time-dependent PDE and norm constraints. *SMAI J. Comput. Maths* **10**, 1–28.
- MARCOTTE, F. & CAULFIELD, C.P. 2018 Optimal mixing in two-dimensional stratified plane Poiseuille flow at finite Péclet and Richardson numbers. *J. Fluid Mech.* **853**, 359–385.
- MATHEW, G., MEZIĆ, I. & PETZOLD, L. 2005 A multiscale measure for mixing. *Phys. D: Nonlinear Phenom.* **211** (1), 23–46.
- MOFFATT, K. & DORMY, E. 2019 *Self-Exciting Fluid Dynamos. Cambridge Texts in Applied Mathematics*, vol. 59. Cambridge University Press.
- PAGE, J., BRENNER, M.P. & KERSWELL, R.R. 2021 Revealing the state space of turbulence using machine learning. *Phys. Rev. Fluids* **6**, 034402.
- PAGE, J., HOLEY, J., BRENNER, M.P. & KERSWELL, R.R. 2024 Exact coherent structures in two-dimensional turbulence identified with convolutional autoencoders. *J. Fluid Mech.* **991**, A10.
- PAGE, J. & KERSWELL, R.R. 2020 Searching turbulence for periodic orbits with dynamic mode decomposition. *J. Fluid Mech.* **886**, A28.
- PARKER, J.P. & SCHNEIDER, T.M. 2022 Variational methods for finding periodic orbits in the incompressible Navier–Stokes equations. *J. Fluid Mech.* **941**, A17.
- PETERSON, P. 2009 F2PY: a tool for connecting Fortran and Python programs. *Intl J. Comput. Sci. Engng* **4** (4), 296–305.
- PETITDEMANGE, L. 2018 Systematic parameter study of dynamo bifurcations in geodynamo simulations. *Phys. Earth Planet. Inter.* **277**, 113–132.
- PRINGLE, C.C.T. & KERSWELL, R.R. 2010 Using nonlinear transient growth to construct the minimal seed for shear flow turbulence. *Phys. Rev. Lett.* **105**, 154502.
- PRINGLE, C.C.T., WILLIS, A.P. & KERSWELL, R.R. 2012 Minimal seeds for shear flow turbulence: using nonlinear transient growth to touch the edge of chaos. *J. Fluid Mech.* **702**, 415–443.
- PRINGLE, G.C., JONES, D.C., GOSWAMI, S., NARAYANAN, S.H.K. & GOLDBERG, D. 2016 Providing the ARCHER community with adjoint modelling tools for high-performance oceanographic and cryospheric computation. *Tech. Rep.* EPCC.
- PROCTOR, M.R.E. 1994 Convection and magnetoconvection in a rapidly rotating sphere. In *Lectures on Solar and Planetary Dynamos* (ed. M.R.E. Proctor & A.D. Gilbert), Publications of the Newton Institute 2, pp. 97–116. Cambridge University Press.
- RINCON, F., OGILVIE, G.I. & PROCTOR, M.R.E. 2007 Self-sustaining nonlinear dynamo process in Keplerian shear flows. *Phys. Rev. Lett.* **98**, 254502.
- ROBERTS, A.P. 2008 Geomagnetic excursions: knowns and unknowns. *Geophys. Res. Lett.* **35** (17), L17307.
- ROBERTS, P.H. 1978 Magneto-convection in a rapidly rotating fluid. In *Rotating Fluids in Geophysics* (ed. P.H. Roberts & A.M. Soward), pp. 421–435. Academic Press.
- SATO, H. 2022 Riemannian conjugate gradient methods: general framework and specific algorithms with convergence analyses. *SIAM J. Optim.* **32** (4), 2690–2717.
- SCHAEFFER, N., JAULT, D., NATAF, H.-C. & FOURNIER, A. 2017 Turbulent geodynamo simulations: a leap towards Earth’s core. *Geophys. J. Intl* **211** (1), 1–29.
- SCHMID, P.J. 2007 Nonmodal stability theory. *Annu. Rev. Fluid Mech.* **39** (2007), 129–162.
- SCHNEIDER, T.M., ECKHARDT, B. & YORKE, J.A. 2007 Turbulence transition and the edge of chaos in pipe flow. *Phys. Rev. Lett.* **99**, 034502.
- SCHWAIGER, T., GASTINE, T. & AUBERT, J. 2019 Force balance in numerical geodynamo simulations: a systematic study. *Geophys. J. Intl* **219**, S101–S114.
- SKENE, C.S. & TOBIAS, S.M. 2024 Weakly nonlinear analysis of the onset of convection in rotating spherical shells. *Geophys. Astrophys. Fluid Dyn.* 1–22.
- SKENE, C.S. & BURNS, K.J. 2025 Fast automated adjoints for spectral PDE solvers. [arXiv:2506.14792](https://arxiv.org/abs/2506.14792).
- SKENE, C.S. & TOBIAS, S.M. 2023 Floquet stability and Lagrangian statistics of a nonlinear time-dependent ABC dynamo. *Phys. Rev. Fluids* **8**, 083701.
- SKENE, C.S., YEH, C.-A., SCHMID, P.J. & TAIRA, K. 2022 Sparsifying the resolvent forcing mode via gradient-based optimisation. *J. Fluid Mech.* **944**, A52.
- STUMM, P. & WALTHER, A. 2009 Multistage approaches for optimal offline checkpointing. *SIAM J. Sci. Comput.* **31** (3), 1946–1967.
- TOBIAS, S.M. 2021 The turbulent dynamo. *J. Fluid Mech.* **912**, P1.
- TOBIAS, S.M., CATTANEO, F. & BRUMMELL, N.H. 2011 On the generation of organized magnetic fields. *Astrophys. J.* **728**, 153.
- TOH, S. & ITANO, T. 2003 A periodic-like solution in channel flow. *J. Fluid Mech.* **481**, 67–76.

- TOWNSEND, J., KOEP, N. & WEICHWALD, S. 2016 Pymanopt: a Python toolbox for optimization on manifolds using automatic differentiation. *J. Machine Learning Res.* **17** (137), 1–5.
- VASIL, G.M., LECOANET, D., BURNS, K.J., OISHI, J.S. & BROWN, B.P. 2019 Tensor calculus in spherical coordinates using Jacobi polynomials. Part I: Mathematical analysis and derivations. *J. Comput. Phys.* **X** **3**, 100013.
- VAVALIARIS, C., BENEITEZ, M. & HENNINGSON, D.S. 2020 Optimal perturbations and transition energy thresholds in boundary layer shear flows. *Phys. Rev. Fluids* **5**, 062401.
- VISWANATH, D. 2007 Recurrent motions within plane Couette turbulence. *J. Fluid Mech.* **580**, 339–358.
- VISWANATH, D. 2009 The critical layer in pipe flow at high Reynolds number. *Phil. Trans. R. Soc. A: Math. Phys. Engng Sci.* **367** (1888), 561–576.
- WALEFFE, F. 2003 Homotopy of exact coherent structures in plane shear flows. *Phys. Fluids* **15** (6), 1517–1534.
- WANG, D. & RUUTH, S.J. 2008 Variable step-size implicit–explicit linear multistep methods for time-dependent partial differential equations. *J. Comput. Maths* **26** (6), 838–855.
- WANG, J., GIBSON, J. & WALEFFE, F. 2007 Lower branch coherent states in shear flows: transition and control. *Phys. Rev. Lett.* **98** (20), 204501.
- WICHT, J. 2002 Inner-core conductivity in numerical dynamo simulations. *Phys. Earth Planet. Inter.* **132** (4), 281–302.
- WILLIS, A.P. 2012 Optimization of the magnetic dynamo. *Phys. Rev. Lett.* **109**, 251101.
- WILLIS, A.P. 2017 The Openpipeflow Navier–Stokes solver. *SoftwareX* **6** (4), 124–127.
- WILLIS, A.P. 2019*a* Equilibria, periodic orbits and computing them, [arXiv:1908.06730](https://arxiv.org/abs/1908.06730).
- WILLIS, A.P. 2019*b* JFNK-hookstep. Available at: <https://github.com/apwillis1/JFNK-Hookstep>.
- WRIGHT, S. & NOCEDAL, J. 1999 *Numerical Optimization*. Springer Science.