Automatic adjoint-based inversion schemes for geodynamics: Reconstructing the evolution of Earth's mantle in space and time

Sia Ghelichkhan[1, 2], Angus Gibson[1], D. Rhodri Davies[1], Stephan C. Kramer[3], and David A. Ham[4]

[1]Research School of Earth Sciences, The Australian National University, Canberra, ACT, Australia.
[2]Institute for Water Futures, The Australian National University, Canberra, ACT, Australia.
[3]Department of Earth Science and Engineering, Imperial College London, London, UK.
[4]Department of Mathematics, Imperial College London, London, UK.

Correspondence: Sia Ghelichkhan (siavash.ghelichkhan@anu.edu.au)

Abstract. Reconstructing the thermo-chemical evolution of Earth's mantle and its diverse surface manifestations is a widely-recognised grand challenge for the geosciences. It requires the creation of a digital twin: a digital representation of Earth's mantle across space and time that is compatible with available observational constraints on the mantle's structure, dynamics and evolution. This has led geodynamicists to explore adjoint-based approaches that reformulate mantle convection modelling as an inverse problem, in which unknown model parameters can be optimised to fit available observational data. Whilst there has been a notable increase in the use of adjoint-based methods in geodynamics, the theoretical and practical challenges of deriving, implementing and validating adjoint systems for large-scale, non-linear, time-dependent problems, such as global mantle flow, has hindered their broader use. Here, we present the Geoscientific Adjoint Optimisation Platform (G-ADOPT), an advanced computational modelling framework that overcomes these challenges for coupled, non-linear, time-dependent systems. By integrating three main components: (i) Firedrake, an automated system for the solution of partial differential equations using the finite element method; (ii) Dolfin-Adjoint, which automatically generates discrete adjoint models in a form compatible with Firedrake; and (iii) the Rapid Optimisation Library, ROL, an efficient large-scale optimisation toolkit; G-ADOPT enables the application of adjoint methods across geophysical continua, showcased herein for geodynamics. Through two sets of synthetic experiments, we demonstrate application of this framework to the initial condition problem of mantle convection, in both square and annular geometries, for both isoviscous and non-linear rheologies. We confirm the validity of the gradient computations underpinning the adjoint approach, for all cases, through second-order Taylor remainder convergence tests, and subsequently demonstrate excellent recovery of the unknown initial conditions. Moreover, we show that the framework achieves theoretical computational efficiency. Taken together, this confirms the suitability of G-ADOPT for reconstructing the evolution of Earth's mantle in space and time. The framework overcomes the significant theoretical and practical challenges of generating adjoint models, and will allow the community to move from idealised forward models to data-driven simulations that rigorously account for observational constraints and their uncertainties using an inverse approach.

1 Introduction

Mantle convection is the 'engine' driving our dynamic planet. It is the principal control on Earth's thermal and chemical evolution, and underpins tectonic and geological activity at Earth's surface (e.g., Davies and Richards, 1992; Coltice et al., 2017).
Through interactions with Earth's crust, it introduces heat and fluids that contribute to the formation and concentration of ore deposits (e.g., Hoggard et al., 2020). Mantle flow also induces vertical movements of Earth's surface (so-called dynamic topography, see Davies et al., 2023, for a review), leading to regional and global changes in sea level and climate (e.g., Poore et al., 2006; Moucha et al., 2008; Cloetingh and Haq, 2015). The lithosphere, considered here to be the mantle's upper thermal boundary layer, serves as a window into the form and time-dependence of mantle convection, recorded through tectonic plate motions (e.g., Iaffaldano and Bunge, 2015; Müller et al., 2016; Stotz et al., 2017, 2018; Wang et al., 2023). Although substantial progress has been made in reconstructing the history of plate tectonic motions (e.g., Seton et al., 2012; Gurnis et al., 2012; Müller et al., 2019; Merdith et al., 2021), the quest for a dynamic reference, revealing the force equilibria within the underlying mantle, remains ongoing. In other words, the veracity of plate motion reconstructions is not matched by an equivalent knowledge of the thermochemical structure and flow history of the underlying mantle. This is a major limitation, as it inhibits our ability to understand fundamental processes that depend on time-dependent interactions between Earth's surface and its deep interior.

The primary challenge in reconstructing past mantle flow stems from the lack of knowledge surrounding the initial state, compounded by uncertain physical and chemical parameters. Mantle convection is an *initial condition problem*, uniquely determined by an initial condition some time in the past: starting from some point in time, it can be uniquely modelled by solving conservation equations for mass, momentum and energy (e.g., Ricard, 2007; Zhong et al., 2007). However, the lack of knowledge on this initial condition — specifically the thermochemical state of Earth's mantle at some time in the past — renders reconstructions of mantle flow through conventional forward calculations intractable (Bunge et al., 2003) (Figure 1-a). Moreover, current global mantle convection models employ billions of degrees of freedom and require multiple time steps to resolve the multi-scale dynamics of Earth's mantle (e.g., Davies and Davies, 2009; Wolstencroft et al., 2009; Stadler et al., 2010; Weismüller et al., 2015; Dannberg and Gassmöller, 2018; Bauer et al., 2019, 2020). Owing to the resulting computational expense, the use of conventional geophysical inverse methods, including Monte Carlo techniques (e.g. Sambridge and Mosegaard, 2002), are considered impractical for determining the mantle's past structure, dynamics and evolution.

The initial condition problem can be partially addressed through *sequential data assimilation* techniques. In essence, the objective of sequential data assimilation is to leverage all accessible information to improve predictions of mantle flow in space and time. Data assimilation is commonly achieved through sequential filtering (e.g., Wunsch, 1996), in which the model is advanced in time over the period in which observations exist. Whenever observations become available, the model is adjusted or 'corrected' (e.g., Bunge and Grand, 2000; Bocher et al., 2016). The magnitude of the correction can be optimally determined using methods such as the Kalman filter (e.g., Bocher et al., 2018), with the model subsequently restarted from the updated state, and this process is repeated until all available information has been utilised (Figure 1-b). In a geodynamic context, the most commonly exploited dataset consists of paleo-surface velocities from plate tectonic reconstruction models, primarily
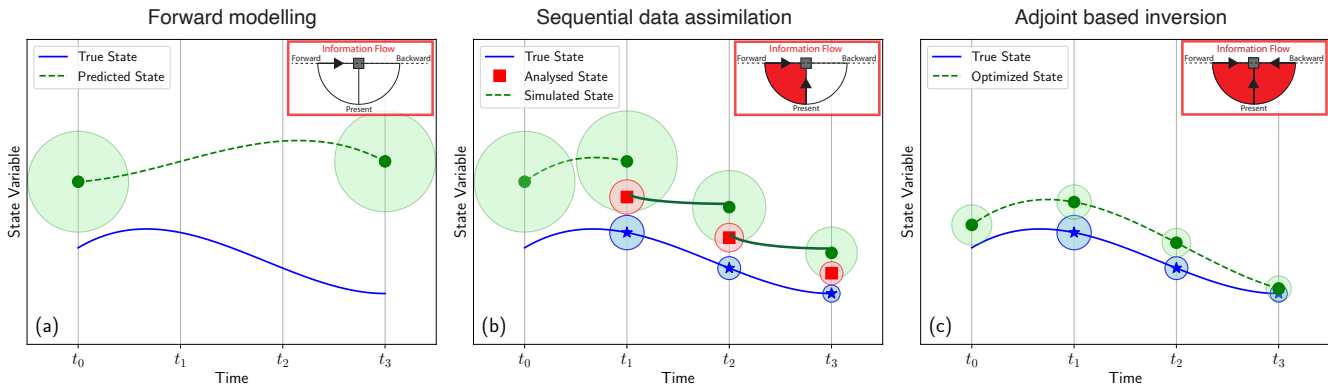
Figure 1. Illustration of different procedures available for estimating past mantle structure: (a) forward modelling prediction, where an unknown initial condition is estimated at t_0, with modelling error, measured as the difference between the modelled and true states. This difference is represented by the distance along the y-axis, which typically grows in time. Variations of forward modelling includes "nudging" of the initial condition to better match present-day structure; (b) sequential data assimilation, or "Kalman Filtering" – starting from an initial condition at t_0, the forward model is run until t_1. An analysis is subsequently undertaken from the resulting model and the available observation. The corrected model is then subsequently integrated in time until t_2, when the next observation is available. This process is repeated until t_3 or the last time step with available observational data. The information flow diagram depicts how information is carried from both the past and present, using current data; (c) adjoint-based or four dimensional variational data-assimilation, which is capable of carrying information on present-day structure (e.g., images from seismic tomography) explicitly backward in time. In (c) observational data that constrains present-day mantle structure, alongside data from different points in space and time, are used to optimise the unknown initial condition. Here, all available observations between t_3 and t_0 contribute to the analysis. The true (unknown) signal is represented by the solid line. Observations (stars), predictions (dark circles) and analyses (squares) are surrounded by ellipsoids of a size proportional to the estimated model uncertainty. Modified from Carrassi et al. (2018) and Davies et al. (2023).

through a kinematic surface boundary condition in time-dependent models. In recent years, there has been a notable increase in the use of this approach (e.g., Bunge et al., 2002; Davies et al., 2012; Bower et al., 2013; Zhong and Rudolph, 2015; Nerlich et al., 2016; Young et al., 2022; Panton et al., 2023). This can be attributed to two main factors: (i) improved confidence in the validity of plate tectonic reconstruction models, and their extension further back in time (e.g., Merdith et al., 2021; Young et al., 2022; Müller et al., 2022); and (ii) the enhanced accessibility of such models, facilitated via open-source community frameworks like the GPlates project (e.g., Gurnis et al., 2012; Müller et al., 2018).

Sequential data-assimilation methods, however, come with an inherent limitation: due to the sequential nature of the assimilation process, each observation is incorporated to influence the model only at later times. Consequently, information propagates from the past into the future but cannot be transmitted back into the past. This drawback poses a significant limitation, as our knowledge of the mantle at the present day is substantially more detailed than at any other time. Thus, it becomes imperative to explore approaches that explicitly carry information backwards in time, or more precisely, enable the estimation of a time-dependent model that best fits all available observational constraints.

Inverse geodynamics is a rapidly evolving field that embarks on this very idea. The foundation of this field is an optimisation approach, in which mantle convection modelling is reformulated as an inverse problem. Using inverse theory, unknown model parameters can be optimised to fit available observational data, via the so-called adjoint method (e.g., Bunge et al., 2003; Ismail-Zadeh et al., 2004), through which the sensitivities of a performance measure (the so-called 'objective functional'), with respect to model parameters (e.g., the choice of initial condition) can be computed. The resulting sensitivity information can be used to adjust model parameters, generating a model flow trajectory that matches observational constraints (e.g., present-day mantle thermochemical structure) (Figure 1-c). The geodynamic adjoint equations for reconstructing the initial condition have been derived for isochemical, incompressible (e.g., Bunge et al., 2003; Horbach et al., 2014), compressible (Ghelichkhan and Bunge, 2016), and thermochemical mantle flow (Ghelichkhan and Bunge, 2018). Moreover, the method has been enhanced for simultaneous recovery of initial temperature conditions and rheological parameters (Li et al., 2017), despite the inherent ill-posedness introduced in such a problem. Growing adoption of adjoint-based methods within a broader geodynamic context is evidenced by their application in Reuber et al. (2020) to decipher subsurface structures and rheological parameters via inversion of principal stress directions, and by Crawford et al. (2018) to quantify the sensitivity of post-glacial sea level changes to lateral variations in mantle viscosity. Recent growth in computational power has led to multiple adjoint-based mantle reconstructions, with a particular focus on regional geological events, as recorded in the geological record of the Americas (e.g., Spasojevic et al., 2009; Liu and Gurnis, 2010; Shephard et al., 2010) and the Atlantic realm (e.g., Colli et al., 2018). To supplement these findings, Ghelichkhan et al. (2021) undertook a systematic global-scale comparison between adjoint model predictions and independent geological constraints, with implications for the expected rates of gravitational and dynamic ellipticity changes resulting from convection within Earth's mantle (Ghelichkhan et al., 2018, 2020).

While substantial strides have been made in the application of adjoint-based methods in geodynamics, there remain widespread obstacles to the broader use of these techniques within the geodynamic modelling community. One key challenge is the complexity involved with deriving, implementing, and validating adjoint models for large-scale, non-linear, time-dependent problems, core features of global mantle convection models. Owing to these difficulties, studies that have used the adjoint method to explore the thermochemical evolution of Earth's mantle often resort to a simplified strategy. This typically involves either a highly idealised treatment of mantle rheology (e.g., Colli et al., 2018; Ghelichkhan et al., 2021), omitting certain coupling terms in the adjoint equations, or doing both (e.g., Liu et al., 2008; Spasojevic et al., 2009; Liu and Gurnis, 2010; Shephard et al., 2010). Such simplifications limit the applicability of these results. The work of Li et al. (2017) is a notable exception to this trend, although its focus on a 2-D rectangular domain, specifically aimed at reconstructing the dynamics of subduction, inherently limits its applicability to global mantle convection simulations. Furthermore, previous global applications of the adjoint approach have been hampered by their reliance on legacy community codes (e.g., Bunge et al., 2003; Liu et al., 2010; Shephard et al., 2010; Colli et al., 2018; Ghelichkhan et al., 2021). These codes are not easily extensible to different geometries or approximations of the underlying physics, employ solver strategies that have since been superseded, and often have limits on model resolution due to the fully structured discretisations encoded. Moreover, the complex and extensive low-level code implementation of coupled adjoint and forward calculations remains obscured, making it difficult to extend and validate, thus restricting the types of observational datasets that can be incorporated. An example of these datasets are the (uncertain)

constraints provided by plate tectonic reconstruction models that are prescribed kinematically (e.g., Spasojevic et al., 2009; Shephard et al., 2010; Zhou and Liu, 2017; Colli et al., 2018; Ghelichkhan et al., 2021), as opposed to being formally incorpo-

105   rated through the objective functional. The kinematic prescription of the surface velocities, however, can only improve mantle reconstructions forward in time, and therefore prohibits their influence on previous system states. In light of these limitations, there is a need for a general framework that can robustly handle the rheological complexities of Earth's mantle and is easily extensible and transferable to other problems in mantle and lithosphere dynamics. Such a framework must also be capable of utilising a variety of observational constraints to comprehensively unravel the historical evolution of Earth's interior's and its

110   diverse impacts at Earth's surface surface (Davies et al., 2023).

In this paper, we introduce the Geoscientific Adjoint Optimisation Platform (G-ADOPT), research software infrastructure that allows us to overcome these limitations. G-ADOPT is built around three state-of-the-art software libraries — (i) Firedrake, an automated system for solving a range of partial differential equations using the finite element method (Rathgeber et al., 2016), recently validated for geodynamics (Davies et al., 2022); (ii) Dolfin-Adjoint, which automatically derives the discrete adjoint

115   equations in a form compatible with Firedrake (Farrell et al., 2013a; Mitusch et al., 2019); and (iii) the Rapid Optimisation Library (ROL), a Trillions package for performing highly-efficient large-scale optimisation (The ROL Project Team, 2022). When combined, they provide a geodynamic inversion framework that is highly efficient, with fully-consistent forward and adjoint calculations that achieve theoretical computational efficiency.

We structure our paper as follows: in Section 2.1 we describe the geodynamic forward problem and our solution strategy

120   using G-ADOPT. Section 2.2 describes the inverse problem considered herein, where we focus on finding the (unknown) initial condition using an objective functional that accounts for observations of surface velocity over time and the final state temperature field. This is broken into three subsections: (i) Section 2.2.1 describes discrete and continuous approaches for obtaining the adjoint systems followed by a detailed derivation of the discrete adjoint systems; (ii) Section 2.2.2 introduces Dolfin-Adjoint, the underlying approach utilised in G-ADOPT to compute the discrete adjoint and derivative fields; and (iii) Section 2.2.6

125   provides an overview of the gradient-based optimisation approach utilised here, facilitated by ROL. We demonstrate the applicability of our approach in Section 3, using two sets of twin experiments with increasing complexity, where we reconstruct the spatial and temporal evolution of a reference simulation. The first set of experiments (Section 3.1) involves a simple isoviscous simulation within an enclosed square domain, while the second set (Section 3.2) examines convection with a non-linear (temperature, depth and strain-rate dependent) rheology, at Earth-like convective vigour, within an annular domain, which has

130   direct applicability to convection within Earth's mantle. We finalise by discussing our results and conclusions in Sections 4 and 5, respectively.

2 Method

2.1 Forward Problem

2.1.1 Governing Equations and Boundary Conditions

135 Mantle flow is described by conservation laws for mass, momentum and energy. We solve these equations in their simplest form, assuming incompressibility and the Boussinesq approximation. The three non-dimensional conservation equations are

$$-\nabla \cdot (2\eta \dot{\boldsymbol{\epsilon}}(\boldsymbol{u})) + \nabla p + \mathrm{Ra} T \hat{\boldsymbol{k}} = 0, \tag{1a}$$

$$\nabla \cdot \boldsymbol{u} = 0, \tag{1b}$$

$$\frac{\partial T}{\partial t} + \boldsymbol{u} \cdot \nabla T - \kappa \nabla^2 T - H = 0, \tag{1c}$$

140 with the vector field $\boldsymbol{u}$, and scalar fields p and T as the principal unknowns of velocity, pressure and temperature, respectively. Table 1 summarises other symbols used in Eqs. 1 and elsewhere. In Eq. 1a, the strain-rate tensor $\dot{\boldsymbol{\epsilon}}(\boldsymbol{u})$ is given by

$$\dot{\boldsymbol{\epsilon}}(\boldsymbol{u}) = \frac{1}{2} \left(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^T \right), \tag{2}$$

and the Rayleigh number is defined by

$$\mathrm{Ra} = \frac{\rho_0 \alpha \Delta T g d^3}{\mu_0 \kappa}. \tag{3}$$

145 For this problem, we define the time interval of interest as $I = [t_I, t_F]$, with the computational domain V bounded by ∂V, and S and C denoting the top and bottom boundaries, respectively. For all simulations, free-slip and impermeable velocity boundary conditions are specified on all boundaries, whilst temperature boundary conditions are set to constant values of T_S and T_C, at top and bottom boundaries, respectively. For the simulations considered in an enclosed square domain, natural temperature boundary conditions (zero heat-flux) are specified on side walls. The set of boundary conditions, for top and bottom boundaries,

150 are listed in Eqs. 4:

$$\boldsymbol{u}(x,t) \cdot \boldsymbol{n} = 0, \qquad\qquad x \in \partial V,\ t \in I, \tag{4a}$$

$$[\eta \dot{\boldsymbol{\epsilon}}(\boldsymbol{u}(x,t)) \cdot \boldsymbol{n}] \cdot \boldsymbol{s} = 0, \qquad\qquad x \in \partial V,\ t \in I, \tag{4b}$$

$$T(x,t) = T_S, \qquad\qquad x \in S,\ t \in I, \tag{4c}$$

$$T(x,t) = T_C, \qquad\qquad x \in C,\ t \in I. \tag{4d}$$

155 In Eqs. 4, $\boldsymbol{n}$ denotes the outer normal vector, and $\boldsymbol{s}$ any tangential vector. Finally, as mantle convection is an initial value problem we require a prescribed temperature field at initial time t_I:

$$T(x, t_I) = T_{IC}(x), \qquad x \in V. \tag{5}$$

Symbol	Description	Symbol	Description
η	Dynamic viscosity	ρ_0	Reference density
$\hat{\boldsymbol{k}}$	Unit vector in opposite gravity direction	α	Thermal expansion coefficient
H	Internal heating rate	ΔT	Temperature difference between domain top and bottom
$\dot{\epsilon}$	Strain-rate	g	Gravitational acceleration
Ra	Rayleigh number	d	Depth of domain
κ	Thermal diffusivity	μ_0	Reference dynamic viscosity
$\boldsymbol{u}$	Velocity	p	Dynamic pressure
T	Temperature	$\delta(x)$	Dirac-delta function
$\beta_{u,d,s}$	Weightings in objective functional	$\alpha_{u,d,s}$	User-defined scaling for functional terms
$\boldsymbol{u}_{obs}$	"Observed" surface velocity	T_{obs}	"Observed" present-day temperature
$G(r)$	Radial Profile for damping		

Table 1. Symbols used in this study.

2.1.2 Solution Strategy: Leveraging Firedrake through G-ADOPT

160 We use the Finite Element method to solve the coupled system of partial differential equations presented in Eqs. 1. For the Stokes system, we use a $Q_2 - Q_1$ (piecewise biquadratic, resp. bilinear) finite element pair for velocity and pressure, with Q_2 elements used for temperature. We strongly impose Dirichlet boundary conditions for temperature at top and bottom boundaries. Free-slip velocity boundary conditions are imposed in two ways: (i) in our square domain cases, we impose strong Dirichlet boundary conditions for $\boldsymbol{u}$; (ii) in our annular domain cases, where boundaries do not align with Cartesian

165 directions, we employ a symmetric Nitsche penalty method (Nitsche, 1971), which weakly enforces boundary conditions via a modification of the variational formulation. An implicit mid-point scheme is used for time integration in the energy equation.

To solve the coupled system of forward (and adjoint) equations, we employ Firedrake, which provides an automated system designed for the solution of partial differential equations using the Finite Element method (Ham et al., 2023). It incorporates principles and some aspects of the code-base from the FEniCS project (Logg et al., 2012), including its use of the Unified Form

170 Language (UFL) (Alnæs et al., 2014) for the representation of variational problems. UFL is a high-level language utilised for the symbolic description of the governing equations in a form that closely mimics their mathematical formulation. The key advantage of UFL in this work lies in its high-level abstraction, which allows for an innovative automatic derivation approach when deriving adjoint systems. We will address the significance of UFL in automatic derivation of adjoint systems in Section 2.2.2.

175 Firedrake provides an array of features that are particularly conducive to tackling problems in geophysical fluid dynamics. Key among these features is support for a variety of Finite Element discretisations, including a highly efficient implementation of discretisations based on extruded meshes, programmable non-linear solvers and operator-aware solver preconditioners that can be combined in a flexible manner to create linear or non-linear systems, which are solved by PETSc (e.g., Balay et al., 1997; Dalcin et al., 2011; Balay et al., 2023). The suitability of Firedrake for geodynamics has been demonstrated via comparison

with a comprehensive set of analytical solutions and community benchmarks in Davies et al. (2022). We refer the reader to this study, alongside Rathgeber et al. (2016), for a more in-depth discussion on Firedrake and its dependencies, alongside an outline of the solution strategy employed herein.

Building upon the foundations provided by Firedrake, G-ADOPT introduces an array of features tailored for cutting-edge geodynamical simulations. It integrates smoothly with Firedrake's advanced data-loading capabilities, enabling finite-element consistent point-data interpolation (Nixon-Hill et al., 2023) and facilitating the integration of diverse observational and physical datasets. Various non-linear rheological laws can be effortlessly incorporated using the symbolic representation provided by UFL. G-ADOPT provides a selection of time discretisation schemes, including second-order Runge-Kutta and backward/forward Euler methods. Visualisation capabilities are provided through integration with the *Visualization Toolkit* (VTK), allowing comprehensive analysis in ParaView, among other visualisation packages. Compatibility with meshes generated by *gmsh* and seamless integration with *NETGEN* provide straightforward support for a range of mesh configurations and adaptive mesh refinement. G-ADOPT offers a diverse array of boundary conditions, encompassing free- and no-slip conditions, a free-surface (Kramer et al., 2012), and kinematic forcing, for example, via seamless integration with pyGplates (Gurnis et al., 2012; Müller et al., 2018). The collection of features, elaborately detailed in the Firedrake and G-ADOPT online documentation and referenced in Davies et al. (2022), highlights the platform's robustness and adaptability for geodynamical research.

2.2 Inverse problem: An Optimisation Approach

Representative reconstructions of the spatial and temporal evolution of mantle flow require knowledge of the initial condition (i.e. the thermochemical state of the mantle at some point in the past), the determination of which is an inverse problem. We therefore seek the best fitting initial condition (T_{IC}), that results in the minimum of an objective functional that measures the difference between predictions and observations of mantle states and irregularity in the solutions. We use the following mathematical description of the objective functional:

$$
\begin{aligned}
J = \quad & \frac{1}{2} \int_I \int_V (T - T_{obs})^2 \, \delta(t - t_F) \, \mathrm{d}x \, \mathrm{d}t \\
& + \frac{\beta_u}{2} \int_I \int_S (\boldsymbol{u} - \boldsymbol{u}_{obs}) \cdot (\boldsymbol{u} - \boldsymbol{u}_{obs}) \, \mathrm{d}s \, \mathrm{d}t \\
& + \frac{\beta_d}{2} \int_I \int_V G(r) \left(T - \bar{T}\right)^2 \delta(t - t_I) \, \mathrm{d}x \, \mathrm{d}t \\
& + \frac{\beta_s}{2} \int_I \int_V \left[\nabla \left(T - \bar{T}\right)\right]^2 \delta(t - t_I) \, \mathrm{d}x \, \mathrm{d}t
\end{aligned}
\tag{6}
$$

The first term in Eq. 6 accounts for the misfit between model predictions and temperature recorded at the final instance, whilst the second term measures the misfit in surface velocities through time. The other terms are Tikhonov regularisation components (Tikhonov, 1963; Hansen, 1992), the first of which penalises deviations from an a-priori depth-averaged profile, $\bar{T}$ (i.e. the damping term), and the second penalises the gradient of these deviations to produce less complex solutions (i.e. the smoothing term). For the damping term, we employ a depth-dependent pre-factor, $G(r)$, that is zero within the mid mantle but

transitions to one in the thermal boundary layers. In these areas, lateral material transport and diffusive processes dominate, leading to diminished sensitivity between the choice of initial condition and the final temperature field. $G(r)$, therefore, helps to minimise the amplitude of the solution within the top and bottom thermal boundary layers, removing noise that would otherwise diffuse over the simulation. The three weighting terms, β_u, β_s and β_d in Eq. 6 are defined as,

$$\beta_u = \alpha_u \frac{\int_V T_{obs}^2 \, \mathrm{d}x}{\Delta t \int_S \left(\boldsymbol{u}_{obs}^{t=t_F} \cdot \boldsymbol{u}_{obs}^{t=t_F} \right) \mathrm{d}x}, \tag{7a}$$

$$\beta_d = \alpha_d \frac{\int_V T_{obs}^2 \, \mathrm{d}x}{\int_V \left(\bar{T} \right)^2 \mathrm{d}x}, \tag{7b}$$

$$\beta_s = \alpha_s \frac{\int_V T_{obs}^2 \, \mathrm{d}x}{\int_V \left(\nabla \bar{T} \right)^2 \mathrm{d}x}. \tag{7c}$$

Note that Δt is the total duration of the simulation. The integrals in Eqs. 7 are employed to ensure normalised objective terms relative to the final temperature misfit (first term on the right hand side of Eq. 6). The three scaling parameters, α_u, α_s and α_d, can be set to adjust the importance of these terms relative to the final temperature misfit. We perform a parameter search to find the best performing combinations of α values.

For our inverse problem we are interested in optimising J described by Eq. 6. J depends on some 'control' parameters, which in this case is the initial temperature field, and some 'state' variables (i.e. surface velocity and the final temperature), which are solutions of the forward problem in Eqs. 1, with the forward system itself depending again on the control. To solve this problem, we define a 'reduced' functional, which is a function of the initial condition, T_{IC}, alone. This reduced functional is typically defined by first solving the forward PDEs (Eqs. 1) for a given value of the control, and then substituting the solutions into the expression for the functional (Eq. 6). The result is a functional that depends only on the control parameters, not the state variables directly (hence the name reduced).

Non-linear optimisation methods provide the means to find the optimal T_{IC} by minimising a reduced functional defined by J. These methods are iterative. They begin with an initial guess of T_{IC} to generate a sequence of improved estimates (called *iterates*) until certain conditions, e.g., residual tolerance, in the objective functional are achieved. Crucial for the efficiency of these iterations are derivatives of Eq. 6 with respect to T_{IC}. Owing to the large number of unknowns in 3-D spherical mantle convection models with Earth-like parameters, obtaining the derivative by means of classical finite differencing techniques is impractical. The *adjoint* method serves as a mathematically elegant and computationally efficient way to obtain the derivatives (e.g., Talagrand, 1997; Giles and Pierce, 2000; Plessix, 2006; Hinze et al., 2008).

2.2.1 Continuous versus Discrete Adjoints

Approaches to deriving, implementing, and obtaining derivatives using the adjoint method primarily fall into two categories: (i) continuous, or the differentiate-then-discretise approach; and (ii) discrete, or the discretise-then-differentiate approach (Gunzburger, 2000).

The continuous approach commences with derivation of the adjoint equations. The resemblance between the forward and resulting adjoint equations allows the adjoint PDEs to be discretised in a consistent manner and, subsequently, implemented

within a numerical framework. By deriving the continuous adjoint PDEs, one can develop an understanding of the key characteristics of adjoint sensitivities, the physical implications of individual terms in the PDEs, and their boundary conditions. Moreover, the continuous method affords complete autonomy in the discretisation and implementation of the adjoint system, often leading to simplified, but more cost-effective approximations of the solutions, as demonstrated by Ismail-Zadeh et al.

245 (2004).

By contrast, the discrete approach relies on already discretised forward equations then differentiates and transposes them to obtain the adjoint equations. This method's primary advantage lies in maintaining consistency in spatial and temporal discretisations, allowing for the automatic determination of the exact gradient of the (discrete) objective functional (Giering and Kaminski, 1998; Gunzburger, 2002). Such consistency ensures full convergence of second-order Newtonian optimisation

250 methods and simplifies the debugging of adjoint programs (Giles and Pierce, 2000). For example, with the discrete approach, even minor inconsistencies in the derivative can highlight numerical or programming errors that must be rectified. It also permits the automatic creation of the adjoint program, stemming from the property that a transposed (adjoint) matrix shares the same eigenvalues with the original linear matrix, ensuring convergence for the adjoint problem's iterative solution methods (Giles and Pierce, 2000). This advantage has facilitated the development of various Automatic Differentiation (AD) tools in

255 recent decades, including those used in Tensor-Flow (Abadi et al., 2015), PyTorch (Paszke et al., 2019) and Enzyne (Moses et al., 2022). It is essential to note that the continuous and discrete methods are equivalent in the limit of infinite spatial and temporal resolution. However, in practical terms, it is the discrete method that typically provides more accurate gradient information. For more details on both approaches we refer the reader to Giles and Pierce (2000) and Gunzburger (2002).

2.2.2 Dolfin-Adjoint

260 Robust and efficient derivative calculations for large-scale simulations using automatic differentiation is challenging and often too slow for the purpose of large-scale optimisation problems (Naumann, 2011). This inefficiency is often attributed to the usually employed approach of treating a numerical model as a sequence of elementary instructions such as addition, multiplication or exponentiation, known as *blocks*. Once the AD tool establishes a sequence of blocks with their dependencies (a process often called *taping*), each block is individually differentiated, and one arrives at the derivative of the entire model using the

265 chain rule.

Dolfin-Adjoint uses an innovative approach to achieve theoretical efficiency by using so-called operator overloading differentiation (Tijskens et al., 2002). By leveraging the high-level mathematical language used by Firedrake and FEniCS (UFL), Dolfin-Adjoint performs the taping process at the highest abstraction level. This can result in blocks that symbolise whole PDE system solves, for which the adjoint derivation is performed at the same level of abstraction. The derived adjoint operation to

270 a block is itself a Firedrake operation. This facilitates generation of the low-level adjoint code using the same finite element form compiler as the forward model. That is, while the taping operation is, in essence, similar to the fundamental abstraction in automatic differentiation techniques, Dolfin-Adjoint operates at a higher level of abstraction and, accordingly, can achieve maximum efficiency and robustness. Consequently, platforms such as Firedrake and FEniCS have incorporated Dolfin-Adjoint for precise derivative computations, which are essential for solving various optimisation and inverse problems. We note the

pursuit of similar automatic differentiation techniques, which have been central to the development of numerical tools like PETSc's TSAdjoint (Zhang et al., 2022) and JuliaAdFEM (Xu and Darve, 2022).

To efficiently compute derivatives, Dolfin-Adjoint tracks and manipulates the computational operations involved in solving the forward and adjoint problem. Central to Dolfin-Adjoint's design is the `Tape` class, responsible for recording a series of high-level finite-element operations — such as (non-)linear solves, functions, and interpolations — that map the initial temperature field T_{IC} into an objective functional J through a series of operations. Operations executed during the forward problem are encapsulated within instances of `Block` subclasses, each representing a distinct computational step. These `Block` instances, by keeping track of their inputs and outputs through `BlockVariable` instances, establish a computational graph that is used by Dolfin-Adjoint to navigate to compute adjoints. To seamlessly integrate this functionality, Dolfin-Adjoint employs *operator overloading*, allowing the creation of overloaded functions and data types that, to the user, behave identically to their original counterparts but are augmented to support automatic differentiation. This approach enables the overloading of data types through inheritance from both the original data type and the `OverloadedType` class in Dolfin-Adjoint, ensuring that each data object carries a `BlockVariable` for tracking its computational lineage. The elegance of Dolfin-Adjoint's design lies in its ability to abstract the computational details, allowing the scientific end-user to focus on designing the forward problem, and simultaneously leverage the computational efficiency of automatic differentiation. Mimicking Firedrake's strategy in utilising Dolfin-Adjoint, the forward operations in G-ADOPT are overloaded by importing the inverse module of G-ADOPT (`gadopt.inverse`), which will be used to populate a tape and establish a `ReducedFunctional`, providing the necessary functionalities for computing a functional and its derivative with respect to controls.

2.2.3 Discrete forward model

As explained in more detail in Davies et al. (2022), the forward geodynamical model can be described as a series of linear and nonlinear solves. Although we solve for temperature using Q_2 elements, we choose the control T_{IC} to be in the Q_1 function space as a means to regularise the inversion problem. This means we need to project T_{IC} to the discrete function space $Q = Q_2$, to obtain a temperature T^0 used in the first timestep. This can be formulated as solving the following system for T^0:

$$F_{\text{project}}(q; T^0, T_{IC}) := \int_V q\left(T^0 - T_{IC}\right) \mathrm{d}x = 0 \quad \text{for all } q \in Q, \tag{8}$$

where q are test functions in Q. Subsequently we solve in each timestep $n = 0, \dots N-1$, the following two systems for $\boldsymbol{u}^n, p^n$, and T^{n+1}, respectively:

$$F_{\text{Stokes}}(\boldsymbol{v}, w; \boldsymbol{u}^n, p^n, T^n) := \int_V (\nabla \boldsymbol{v}) : \eta(\boldsymbol{u}^n, T^n)\left[\nabla \boldsymbol{u}^n + (\nabla \boldsymbol{u}^n)^T\right] \mathrm{d}x - \int_V (\nabla \cdot \boldsymbol{v}) p^n \mathrm{d}x - \int_V \mathrm{Ra} T^n \boldsymbol{v} \cdot \hat{\boldsymbol{k}} \mathrm{d}x$$
$$- \int_V w \nabla \cdot \boldsymbol{u}^n \mathrm{d}x = 0 \quad \text{for all } \boldsymbol{v} \in V, w \in W, \tag{9}$$

$$F_{\text{energy}}(q; T^{n+1}, T^n, \boldsymbol{u}^n) := \int_V q \frac{T^{n+1} - T^n}{\Delta t} \mathrm{d}x + \int_V q \boldsymbol{u}^n \cdot \nabla T^{n+\theta} \mathrm{d}x + \int_V (\nabla q) \cdot \left(\kappa \nabla T^{n+\theta}\right) \mathrm{d}x = 0 \quad \text{for all } q \in Q, \tag{10}$$

where V, W are the discrete function spaces for velocity and pressure (here $V = [Q_2]^{\dim}, W = Q_1$) with test functions v, and w. $T^{n+\theta}$ is the weighted average $\theta T^{n+1} + (1-\theta)T^n$. Note that we assume a strain rate and temperature dependent rheology, and thus write $\eta = \eta(u, T)$. This makes Eq. (9) a non-linear system, which we solve through Newton's method. The discrete functional is calculated as

$$J(T^N, u^0, \ldots, u^{N-1}, T_{IC}) = \frac{1}{2} \int_V \left(T^N - T_{obs}\right)^2 dx + \frac{\beta_u}{2} \sum_{n=0}^{N-1} \int_S \left(u^n - u^n_{obs}\right)^2 ds$$
$$+ \frac{\beta_d}{2} \int_V G(r)\left(T_{IC} - \bar{T}\right)^2 dx + \frac{\beta_s}{2} \int_V \left[\nabla(T_{IC} - \bar{T})\right]^2 dx \tag{11}$$

2.2.4 Calculating Gradients Using the Adjoint Method

We denote the entire forward solution as $z = (T^0, \ldots T^N, u^0, \ldots u^{N-1}, p^0, \ldots p^{N-1})$, so that the functional can be written as a function $J(z, T_{IC})$, of z and the control T_{IC}. The forward solution itself is also dependent on T_{IC}, as for each choice of T_{IC} we can solve the forward model to obtain $z(T_{IC})$. We define the reduced functional $\hat{J}$ as

$$\hat{J}(T_{IC}) = J(z(T_{IC}), T_{IC}). \tag{12}$$

Thus we can reformulate the PDE-constrained minimisation problem:

minimise $J(z, T_{IC})$ under the constraints (8–10),

as an unconstrained minimisation problem for $\hat{J}(T_{IC})$. To use efficient gradient-based optimisation algorithms we do, however, need a means of computing its gradient for which we will employ the adjoint method. In addition to the forward solution z, we also define an adjoint solution $\lambda = (\Psi^0, \ldots \Psi^N, \phi^0, \ldots \phi^{N-1}, \xi^0 \ldots \xi^{N-1})$, where each component is associated with one of the constraints: $\Psi^0 \in Q$ with (8), $\phi^n \in V, \xi^n \in W$ with (9), and $\Psi^{n+1} \in Q$ with (10) for $n = 0, \ldots N-1$. Using these we define the following sum of the constraints, with each constraint weighted by the corresponding adjoint solution λ:

$$F(\lambda; z, T_{IC}) = F_{\text{project}}(\Psi^0; T^0, T_{IC}) + \sum_{n=0}^{N-1} F_{\text{Stokes}}(\phi^n, \xi^n; u^n, p^n, T^n) + \sum_{n=0}^{N-1} F_{\text{energy}}(\Psi^{n+1}; T^{n+1}, T^n, u^n). \tag{13}$$

Since, by definition, for any choice of T_{IC} the associated forward solution $z(T_{IC})$ satisfies all constraints, we have:

$$F(\lambda; z(T_{IC}), T_{IC}) = 0,$$

and thus, for any choice of λ we have:

$$\frac{\partial F(\lambda; z(T_{IC}), T_{IC})}{\partial T_{IC}} = \left.\frac{\partial F(\lambda; z, T_{IC})}{\partial z}\right|_{z=z(T_{IC})} \frac{\partial z(T_{IC})}{\partial T_{IC}} + \left.\frac{\partial F(\lambda; z, T_{IC})}{\partial T_{IC}}\right|_{z=z(T_{IC})} = 0. \tag{14}$$

If we choose λ to be the solution to the following, so called, *adjoint equation*

$$\left.\frac{\partial F(\lambda; z, T_{IC})}{\partial z}\right|_{z=z(T_{IC})} = \left.\frac{\partial J(z, T_{IC})}{\partial z}\right|_{z=z(T_{IC})}, \tag{15}$$

we can work out the gradient of the reduced functional:

$$\frac{\partial \hat{J}(T_{IC})}{\partial T_{IC}} = \left.\frac{\partial J(z,T_{IC})}{\partial z}\right|_{z=z(T_{IC})} \frac{\partial z(T_{IC})}{\partial T_{IC}} + \left.\frac{\partial J(z,T_{IC})}{\partial T_{IC}}\right|_{z=T_{IC}} \tag{16}$$

$$= \left.\frac{\partial F(\lambda;z,T_{IC})}{\partial z}\right|_{z=z(T_{IC})} \frac{\partial z(T_{IC})}{\partial T_{IC}} + \left.\frac{\partial J(z,T_{IC})}{\partial T_{IC}}\right|_{z=T_{IC}} \tag{17}$$

$$= -\left.\frac{\partial F(\lambda;z,T_{IC})}{\partial T_{IC}}\right|_{z=z(T_{IC})} + \left.\frac{\partial J(z,T_{IC})}{\partial T_{IC}}\right|_{z=T_{IC}} \tag{18}$$

2.2.5 Discrete Backward Model

Although the adjoint equation (15) is derived symbolically from the forward model (8–10), and solved for fully automatically by Dolfin-Adjoint, we here briefly work out the discrete adjoint equations to show that these equations can still be interpreted as the solution process of a backwards-in-time PDE, similar to the continuous adjoint approach, but that a specific time-discretisation is derived, which is necessary to obtain a gradient that is consistent with the discrete forward model. Split out by component, the adjoint equations read

$$\frac{\partial F}{\partial T^0} = \frac{\partial J}{\partial T^0}, \tag{19}$$

$$\frac{\partial F}{\partial \boldsymbol{u}^n} = \frac{\partial J}{\partial \boldsymbol{u}^n}, \quad \text{for } n=0,\dots N-1, \tag{20}$$

$$\frac{\partial F}{\partial p^n} = \frac{\partial J}{\partial p^n}, \quad \text{for } n=0,\dots N-1, \tag{21}$$

$$\frac{\partial F}{\partial T^{n+1}} = \frac{\partial J}{\partial T^{n+1}}, \quad \text{for } n=0,\dots N-1, \tag{22}$$

which can be solved for Ψ^n, ϕ^n, ξ^n.

These equations can be solved by going backwards through the timesteps $n = N-1\dots 0$. In each, we first solve (22) for Ψ^{n+1}. Starting at the last timestep $n = N-1$, we get:

$$\frac{\partial F_{\text{energy}}(\Psi^N; T^N, T^{N-1}, \boldsymbol{u}^{N-1})}{\partial T^N} \delta T = \frac{\partial J}{\partial T^N} \delta T, \tag{23}$$

where we have applied the gradient of F with respect to T^N to an arbitrary perturbation δT. This allows us to interpret this equation as a weak form, tested with δT, that we can solve for Ψ^N:

$$\int_V \Psi^N \frac{\delta T}{\Delta t} \mathrm{d}x + \theta \int_V \Psi^N \boldsymbol{u}^{N-1} \cdot \nabla \delta T \mathrm{d}x + \theta \int_V \left(\nabla \Psi^N\right) \cdot \left(\kappa \nabla \delta T\right) \mathrm{d}x = \int_V \left(T^N - T_{obs}\right) \delta T \mathrm{d}x \quad \text{for all } \delta T \in Q. \tag{24}$$

Defining an auxiliary $\Psi^{N+1} = \Delta t(T^N - T_{obs})$, and integrating the advection term by parts, we get:

$$\int_V \left[\frac{\Psi^N - \Psi^{N+1}}{\Delta t} \delta T - \theta \nabla \cdot \left(\Psi^N \boldsymbol{u}^{N-1}\right) \delta T + \theta \left(\nabla \Psi^N\right) \cdot \kappa \nabla \delta T\right] \mathrm{d}x = 0 \quad \text{for all } \delta T \in Q, \tag{25}$$

which for $\theta = 1$, we may recognise an advection-diffusion equation run backwards in time.

For $n < N - 1$, equation (22) contains more terms:

$$\frac{\partial F_{\text{energy}}(\Psi^{n+1}; T^{n+1}, T^n, \boldsymbol{u}^n)}{\partial T^{n+1}} \delta T = -\frac{\partial F_{\text{energy}}(\Psi^{n+2}; T^{n+2}, T^{n+1}, \boldsymbol{u}^{n+1})}{\partial T^{n+1}} \delta T$$

$$-\frac{\partial F_{\text{Stokes}}(\boldsymbol{\phi}^{n+1}, \xi^{n+1}; \boldsymbol{u}^{n+1}, p^{n+1}, T^{n+1})}{\partial T^{n+1}} + \frac{\partial J}{\partial T^{n+1}} \delta T \quad \text{for all } \delta T \in Q. \quad (26)$$

as the energy equation in both forward time steps n and $n+1$ depends on T^{n+1}, as does the Stokes system in time step $n+1$. Going backwards through the equations however, we can still solve for Ψ^{n+1}, associated with time step n, as we have already solved for $\boldsymbol{\phi}^{n+1}, \xi^{n+1}$, and Ψ^{n+2} associated with time step $n+1$. Note that the $\partial J / \partial T^{n+1}$–term vanishes in this case as J does not explicitly depend on intermediate temperature solutions. Similar to (25), we may rewrite to

$$\int_V \left[\frac{\Psi^{n+1} - \Psi^{n+2}}{\Delta t} \delta T - \nabla \cdot \left(\theta \Psi^{n+1} \boldsymbol{u}^n + (1 - \theta) \Psi^{n+2} \boldsymbol{u}^{n+1} \right) \delta T + \nabla \left(\theta \Psi^{n+1} + (1 - \theta) \Psi^{n+2} \right) \cdot \kappa \nabla \delta T \right] \mathrm{d}x =$$

$$-\int_V \nabla \boldsymbol{\phi}^{n+1} : \frac{\partial \eta(\boldsymbol{u}^{n+1}, T^{n+1})}{\partial T^{n+1}} \left[\nabla \boldsymbol{u}^{n+1} + \left(\nabla \boldsymbol{u}^{n+1} \right)^T \right] \delta T \mathrm{d}x + \int_V \mathrm{Ra} \boldsymbol{\phi}^{n+1} \cdot \hat{\boldsymbol{k}} \, \delta T \mathrm{d}x \quad \text{for all } \delta T \in Q, \quad (27)$$

which we can interpret as a backward-in-time theta-weighted advection diffusion step for Ψ, with source terms associated with sensitivity of the rheology and buoyancy to temperature. Note, however, that here in the advection term we weight both Ψ and $\boldsymbol{u}$ with θ, in contrast to the forward timestep (10).

After solving for Ψ^{n+1}, we can solve (20) together with (21) for $\boldsymbol{\phi}^n$ and ξ^n:

$$\frac{\partial F_{\text{Stokes}}(\boldsymbol{\phi}^n, \xi^n; \boldsymbol{u}^n, p^n, T^n)}{\partial \boldsymbol{u}^n} \delta \boldsymbol{u} + \frac{\partial F_{\text{Stokes}}(\boldsymbol{\phi}^n, \xi^n; \boldsymbol{u}^n, p^n, T^n)}{\partial p^n} \delta p = -\frac{\partial F_{\text{energy}}(\Psi^{n+1}; T^{n+1}, T^n, \boldsymbol{u}^n)}{\partial \boldsymbol{u}^n} \delta \boldsymbol{u} + \frac{\partial J}{\partial \boldsymbol{u}^n} \delta \boldsymbol{u} \quad (28)$$

for all velocity and pressure perturbations $\delta \boldsymbol{u} \in V, \delta p \in W$. This leads to the following weak-form equation for $\boldsymbol{\phi}^n$ and ξ^n:

$$\int_V [\nabla \boldsymbol{\phi}^n + (\nabla \boldsymbol{\phi}^n)] : \eta(\boldsymbol{u}^n, T^n) \nabla \delta \boldsymbol{u} \mathrm{d}x + \int_V [\nabla \boldsymbol{\phi}^n + (\nabla \boldsymbol{\phi}^n)] : \frac{\partial \eta(\boldsymbol{u}^n, T^n)}{\partial \boldsymbol{u}^n} \delta \boldsymbol{u} \nabla \boldsymbol{u} \mathrm{d}x - \int_V \xi^n \nabla \cdot \delta \boldsymbol{u}^n \mathrm{d}x - \int_V (\nabla \cdot \boldsymbol{\phi}^n) \delta p \mathrm{d}x$$

$$= -\int_V \Psi^{n+1} \delta \boldsymbol{u} \cdot \nabla T^{n+\theta} + \beta_u \int_S (\boldsymbol{u}^n - \boldsymbol{u}_{obs}) \delta \boldsymbol{u} \mathrm{d}x \quad \text{for all } \delta \boldsymbol{u} \in V, \delta p \in W \quad (29)$$

The left-hand side is similar to the Stokes system in (9), except for an additional $\partial \eta / \partial \boldsymbol{u}$ term, alongside the fact that this is now just a linear system as the rheology only depends on the forward variables $\boldsymbol{u}^n$ and T^n. Instead of a buoyancy term, we now have forcing terms associated with sensitivity of temperature advection and the mismatch with observed velocities at the surface.

Finally, after having solved either (25) (n=N-1) or (27) (n<N-1) for Ψ^{n+1}, and (29) for $\boldsymbol{\phi}^n, \xi^n$ going backward through the timesteps $n = N - 1 \to n = 0$, we can solve (19) for Ψ^0:

$$\frac{\partial F_{\text{project}}(\Psi^0; T^0, T_{IC})}{\partial T^0} \delta T = -\frac{\partial F_{\text{energy}}(\Psi^1; T^1, T^0, \boldsymbol{u}^0)}{\partial T^0} \delta T - \frac{\partial F_{\text{Stokes}}(\boldsymbol{\phi}^0, \xi^0; \boldsymbol{u}^0, p^0, T^0)}{\partial T^0} \delta T \quad \text{for all } \delta T \in Q \quad (30)$$

which can be worked out as a projection

$$\int_V \Psi^0 \delta T \mathrm{d}x = \int_V \frac{\Psi^1}{\Delta t} \delta T \mathrm{d}x + (1-\theta) \int_V \left(\nabla \cdot \Psi^1 \boldsymbol{u}^1 \right) \delta T \mathrm{d}x - (1-\theta) \int_V \nabla \psi^{n+1} \cdot \kappa \nabla \delta T \mathrm{d}x$$

$$- \int_V \nabla \boldsymbol{\phi}^1 : \frac{\partial \eta(\boldsymbol{u}^1, T^1)}{\partial T^1} \left[\nabla \boldsymbol{u}^1 + \left(\nabla \boldsymbol{u}^1 \right)^T \right] \delta T \mathrm{d}x + \int_V \mathrm{Ra} \boldsymbol{\phi}^1 \cdot \hat{\boldsymbol{k}} \delta T \mathrm{d}x \quad \text{for all } \delta T \in Q. \quad (31)$$

Finally, the gradient of the reduced functional with respect to the control is obtained by:

$$\frac{\partial \hat{J}(T_{IC})}{\partial T_{IC}} \delta T = - \int_V \Psi^0 \, \delta T \, \mathrm{d}x + \frac{\beta_d}{2} \int_V \left(T_{IC} - \bar{T} \right) \delta T \, \mathrm{d}x + \frac{\beta_s}{2} \int_V \nabla \left(T_{IC} - \bar{T} \right) \cdot \nabla \delta T \, \mathrm{d}x \quad (32)$$

2.2.6 Gradient-Based Non-linear Optimisation

To find the solution to the inverse problem, the gradient fields from Firedrake and Dolfin-Adjoint can be redirected to an optimisation package with a Python interface (e.g., scipy.optimize: Virtanen et al., 2020). However, the majority of well-established optimisation packages are hard-coded to apply the Euclidean (l^2) inner product for optimisation-specific operations. l^2, typically referred to as a sequence space, is often used in signal processing or discrete mathematics where functions are treated as sequences, i.e. discrete data, and do not represent continuous functions. The Euclidean inner product is therefore not suitable for finite element function-based optimisation and, unlike L^2 or Sobolev spaces, it cannot produce mesh and/or basis function independent convergence (Schwedes et al., 2017). The Rapid Optimisation Library (ROL, The ROL Project Team, 2022), a Trilinos package for large-scale optimisation, resolves this issue by introducing a generic interface for data structures that can be overloaded to perform inner-product aware operations (e.g., in L^2 space) and achieve mesh-independent convergence results (Schwedes et al., 2016). ROL has been primarily used for the solution of optimal design, optimal control and inverse problems in large-scale engineering applications (Iglesias et al., 2018; Kouri et al., 2021a, b, 2023) and has a comprehensive catalogue of gradient-based optimisation algorithms. For the results presented herein, we use the python interface for ROL in Firedrake, which we have supplemented with additional checkpointing functionality. This allows us to checkpoint intermediate optimisation states and variables, including those related to step lengths or Hessian estimates, and subsequently restart the optimisation procedure without loss of performance. This, in particular, is relevant on modern High Performance Computing facilities with strict wall time limits.

The distinguishing factor between different optimisation algorithms (in ROL or elsewhere) is the strategy used to move from one iterate to the next. Broadly speaking, there are two strategies for moving from the current iterate, x_k, to a new one, x_{k+1}: line-search and trust-region strategies (Nocedal and Wright, 1999). A fundamental distinction between line-search and trust-region methods lies in the sequence of selecting the direction and distance (Nocedal and Wright, 2006). In line search methods, the direction is initially fixed, and an appropriate step length is subsequently determined. Conversely, trust-region methods commence by establishing an initial size for a trusted region (hence the name), and then simultaneously constrain the direction and step to achieve sufficient amount of improvement within this trusted region. The size of this trusted region around the current iterate is determined according to a model, m, that approximates the region around the current iterate, x_k, according to a quadratic approximation. At each iteration, the accuracy of this model then is assessed based on its agreement with the actual changes in the function, f. If the new value, $f(x_k + p_k)$ is greater than the current value of $f(x_k)$, m is not

a good approximation of the objective functional around x_k and the size of the trusted-region measured by the trust-region *radius* is therefore reduced to improve the applicability of the model. By bounding the calculations to a trusted region where model m is applicable, trust-region methods prohibit overly aggressive steps, which make them suitable for handling negative curvature situations (non-convexity optimisation problems) more gracefully compared to line-search methods. Nonetheless, faster convergence can be achieved by expanding trust region in case of a predictive model, ensuring robust minimisation of the objective functional.

In this study, we employ the trust region method of Lin and Moré (1999) implemented in ROL. Lin-More is a truncated Newton method, consisting of repeated application of an iterative algorithm to approximately solve Newton's equations (Dembo and Steihaug, 1983). Lin-More can effectively handle provided bound constraints by ensuring that variables remain within their specified bounds: At each iteration, variables are classified into "active" and "inactive" sets. Variables at their bounds and not allowing descent are considered active and are fixed during the iteration. The remaining variables, which can change without violating the bounds, are inactive. The described properties renders the algorithm as a robust and efficient method for solving bound-constrained optimisation problems.

3 Numerical Experiments and Results

Twin experiments serve as a means to illustrate the feasibility of geophysical inverse methods. In our experimental setup, we generate a synthetic reference simulation that advances forward in time, starting from a user-defined initial condition. We use this reference simulation to emulate a real-Earth reconstruction scenario, where the resulting temperature field at the final time, $T(x, t = t_F)$, and the corresponding surface velocities at all times, $u(x = x_S, t)$, are stored for subsequent use as 'observations' in reconstructing the initial state of the mantle and its evolution through space and time. These fields are used to mimic fundamental datasets in mantle reconstruction models, drawing parallels to 3-D models of mantle temperature inferred from seismic tomography images, and surface velocities derived from plate tectonic reconstructions. We examine two sets of 2-D twin experiments: (i) simulations of a single upwelling in an enclosed square domain, with an isoviscous rheology; and (ii) convection within an annular domain, incorporating a non-linear visco-plastic rheology, as has previously been employed to generate plate-like behaviour in mantle convection models (see Coltice et al., 2017, for a review).

3.1 A Single Upwelling in an Enclosed Square Domain

3.1.1 Forward Problem

We start our experiments by reconstructing the evolution of a single upwelling plume within an enclosed square computational domain (free-slip boundary conditions on all boundaries). We assume an isoviscous rheology, incompressible flow under the Bousinessq approximation assumption, and a Rayleigh number of $Ra = 10^6$. The model is heated from below ($T_C = 1$) and cooled from above ($T_S = 0$). The initial condition is generated by a Gaussian anomaly of amplitude 0.1, centred at $x_0 = (0.5, 0.2)$, superimposed on top of an average temperature profile generated by two error functions, representing top and

```
1: from gadopt import *
2:
3: # load mesh
4: mesh1d = IntervalMesh(150, length_or_left=0.0, right=1.0)  # Unit interval mesh
5: mesh_temp = ExtrudedMesh(
6:          mesh1d, layers=150, layer_height=1.0/150, extrusion_type="uniform")
7:
8: # Set up function spaces for the Q2Q1 pair for velocity+pressure and Q2 for temperature
9: V = VectorFunctionSpace(mesh, "CG", 2)  # Velocity function space (vector)
10: W = FunctionSpace(mesh, "CG", 1)  # Pressure function space (scalar)
11: Z = MixedFunctionSpace([V, W])  # mixed space for Stokes
12: Q = FunctionSpace(mesh, "CG", 2)  # Temperature function space (scalar)
13:
14: z = Function(Z)  # A field over the mixed function space Z
15: u, p = z.subfunctions  # Symbolic UFL expressions for u and p
16:
17: T = Function(Q, name="Temperature")  # Temperature Field
18: X = SpatialCoordinate(mesh)  # Spatial Coordinates
19: # Initial condition
20: T.interpolate(
21:      0.5 * (erf((1 - X[1]) * 3.0) + erf(-X[1] * 3.0) + 1)
22:      + 0.1 * exp(-0.5 * ((X - as_vector((0.5, 0.2))) / Constant(0.1)) ** 2))
23:
24: Z_nullspace = create_stokes_nullspace(Z, closed=True, rotational=False)  # Pressure nullspace
25:
26: checkpoint_file = CheckpointFile("Checkpoint_State.h5", "w")
27: checkpoint_file.save_mesh(mesh)  # saving the mesh
28: Ra = Constant(1e6)  # Rayleigh number
29: approximation = BoussinesqApproximation(Ra)  # Incompressible (Boussinesq) approximation
30: delta_t = Constant(4e-6)  # Constant time step
31: max_timesteps = 80  # Number of time steps
32:
33: # Boundary conditions
34: stokes_bcs = {"bottom": {"uy": 0}, "top": {"uy": 0}, 1: {"ux": 0}, 2: {"ux": 0}}  # Velocity BCs
35: temp_bcs = {"bottom": {"T": 1.0}, "top": {"T": 0.0}}  # Temperature BCs
36:
37: energy_solver = EnergySolver(T, u, approximation, delta_t, ImplicitMidpoint, bcs=temp_bcs)
38: stokes_solver = StokesSolver(z, T, approximation, bcs=stokes_bcs,
39:                              nullspace=Z_nullspace, transpose_nullspace=Z_nullspace)
40:
41: # time loop:
42: for timestep in range(0, max_timesteps):
43:      stokes_solver.solve()  # solving Stokes
44:      energy_solver.solve()  # solving energy
45:      checkpoint_file.save_function(u, name="Velocity", idx=timestep)  # Store velocity (uobs)
46:
47: checkpoint_file.save_function(T, name="Temperature", idx=max_timesteps - 1)  # Store final temperature (Tobs)
48: checkpoint_file.close()
```

Listing 1. Selected lines from G-ADOPT code, demonstrating generation of our reference isoviscous simulation in an enclosed square domain.

bottom thermal boundary layers. Starting from this initial condition, the model is run for 80 time steps of $\delta t = 4 \times 10^{-6}$: the

440   time required for the temperature anomaly to form a plume that reaches the domain's top boundary.

Listing 1 shows selected lines of a G-ADOPT script used to generate this reference synthetic experiment. The first step, illustrated on line 1, is to import the G-ADOPT module (which provides access to Firedrake and associated functionality). We next need a mesh, for which we use a built-in Firedrake meshing function. The computational domain is a unit square with 150×150 elements, loaded on lines 4-6. The function spaces, within which our solutions are defined, are specified as follows:

445   1. A vector function space, v, is specified for the velocity field (line 9), employing a $Q2$ discretisation.

2. A scalar function space, w, is specified for pressure (line 10), utilising a $Q1$ discretisation.

3. These are combined on line 11 to create the mixed function space, z, for the Stokes (velocity and pressure) system.

4. A function space Q, is specified for the temperature field (line 12), using a $Q2$ discretisation.

We specify functions to hold our solutions on lines 14-18. The temperature field, T, is initialised on line 20 using a symbolic

450   expression for the coordinates from line 18. The initialisation includes a 1-D profile along the y-axis (line 21) and a Gaussian anomaly, as specified above (line 22). This problem has a constant pressure nullspace, defined as z_nullspace on line 24, which
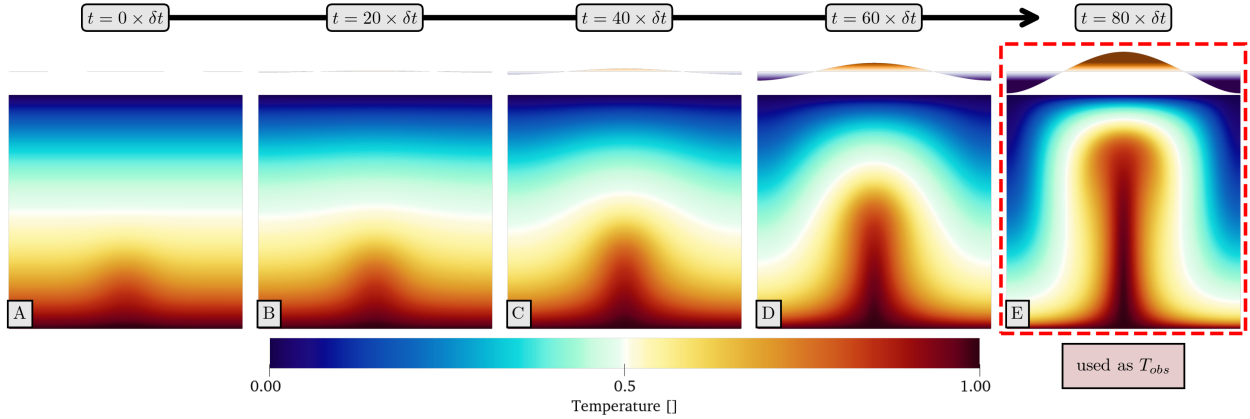
Figure 2. Reference forward simulation: The initial condition is generated by superimposing a gaussian anomaly of maximum amplitude 0.1 centred at $x_0 = (0.5, 0.2)$ on top of an adiabatic profile. The simulation runs forward in time from this initial condition, until the plume-like feature approaches the top boundary after 80 times-steps. Also visualised at top of each snapshot are the normal stresses acting on the top boundary, which are proportional to dynamic topography. Boxed in dashed red is the final state of the simulation that is subsequently used as T_{obs} in our synthetic adjoint simulation.

will subsequently be passed to the solver, and PETSc will seek a solution in the space orthogonal to the provided nullspace. A checkpoint file is initiated on lines 26-27 to retain the necessary fields for the subsequent adjoint inversion. Important constants in this problem (Rayleigh Number, Ra, time-stepping parameters, `delta_t` and `max_timesteps`) are defined on lines 28-31, with

455   the Boussinesq approximation specified on line 29 (later passed on to the Stokes and energy systems to determine which terms are to be assembled). Boundary conditions for velocity and temperature are specified on lines 34 and 35, respectively. The latter uses integer mesh markers to tag entities of meshes, with boundaries tagged as follows: tag 1 corresponds to the plane $x = 0$ (left); 2 to $x = 1$ (right); `"bottom"` to $y = 0$; and `"top"` to $y = 1$.

     We now solve the variational problem, with solver objects for the energy, `energy_solver`, and Stokes, `stokes_solver`, systems

460   created on lines 37 and 38. For the energy system we pass in the solution field T, velocity $\boldsymbol{u}$, the physical approximation, time step, temporal discretisation approach (i.e. implicit middle point) and boundary conditions. For the Stokes system, we pass in the solution fields `z`, Temperature, the physical approximation, boundary conditions and the nullspace object. Solution of the two variational problems is undertaken by PETSc. The time-loop is defined on lines 42-45, with the Stokes system solved on line 43, the energy equation on line 44, and velocities (for later use as time-dependent surface constraints in our adjoint

465   inversions) checkpointed on line 45. Figs. 2-A to E show temporal snapshots of the reference forward simulation. We note that the final temperature field (Fig. 2-E), subsequently utilised in the adjoint inversion, is also checkpointed on line 47.

3.1.2 Inverse Problem

Only trivial changes are required to convert the forward problem outlined in Listing 1 into its corresponding adjoint, which are outlined in Listing 2. We first augment our imports with `gadopt.inverse` (line 2): this provides access to crucial Dolfin-Adjoint functionalities harnessed by Firedrake, enabling overloading differentiation and taping of finite element operations. On line 4, we activate disk checkpointing of intermediate forward solutions, ensuring that these fields — otherwise retained in memory — are available as inputs for solving the adjoint equations. For this problem, we select the initial temperature as the control, symbolised as `Tic` on line 7, using a $Q1$ function space defined on line 6. On line 8, we define the average temperature field for regularisation terms using the same $Q1$ function space. Despite utilising $Q2$ elements for the forward temperature computation, opting for a basis function with a lower polynomial degree for the control is advantageous, as it curtails computational expense during the internal optimisation algorithm's operations by reducing the number of degrees of freedom in the solution. Furthermore, using a lower polynomial degree reduces the complexity and regularisation requirements of the optimisation problem, helping to avoid over-fitting of the solution (e.g., Hastie et al., 2009).

On line 10, a checkpoint file is opened. The reference final temperature field is subsequently loaded on line 11, with our guess at the initial temperature condition specified on line 12 (noting that it corresponds to the terminal temperature field of the reference simulation). We specify the control for Dolfin-Adjoint on line 15, followed by projection of the initial temperature condition onto `T` on line 18. During execution of the time loop (lines 20-24), we solve Stokes and energy systems, after which we load velocities from the reference simulation on line 24, used when accumulating contributions to the surface velocity misfit, `u_misfit`, on line 24. After the time loop, we define several components of the objective functional. Specifically, we establish the damping term and its associated normalisation factor (lines 29-31), the smoothing term and associated normalisation (lines 32-33), two normalisation terms associated with final state temperature and surface velocity misfits (lines 34-35), and the misfit associated with the final temperature field (line 36). These terms are combined on lines 38-42 as our `objective`, using weights (α_u, α_d, and α_s) specified on line 37, and is later utilised on line 44 to define the reduced functional. We note that values for α_u, α_d, and α_s are systematically tested herein.

3.1.3 Investigating the Derivative

Our *initial guess* for T_{IC} is set to the final *'observed'* temperature field T_{obs} (i.e., the terminal temperature field of the forward model). This choice is grounded in the findings of Horbach et al. (2014), and our own tests, which demonstrate that the minimisation problem possesses a strong minimum, rendering it insensitive to the initial guess. The first optimisation iteration starts from this initial guess (Fig. 3-A) and runs forward in time to arrive at the first modelled terminal temperature field $T_{t=t_F}$ (Fig. 3-B). Compared to the reference T_{obs} (Fig. 3-C), the model is further advanced in time, with the plume tail in the lower mantle narrower, hot buoyant anomalies spread throughout the upper part of the domain, generating a cold return flow and resulting in thinning of the thermal boundary layers. Previously, we noted that for our reconstruction simulations, the objective functional, denoted as J, encompasses four distinct terms: (i) final temperature field; (ii) surface velocity misfit; (iii) smoothing terms; and (iv) damping terms. The relevance of the three latter terms is gauged by their respective weighting

```
1: ...
2: from gadopt.inverse import *
3: ...
4: enable_disk_checkpointing()  # Enable checkpointing to disk for adjoint
5: ...
6: Q1 = FunctionSpace(mesh, "CG", 1)  # Control (Tic) function space
7: Tic = Function(Q1, name="Initial Temperature")   # Control: Initial temperature
8: Taverage = Function(Q1, name="Average Temperature")  # Average temperature
9:
10: checkpoint_file = CheckpointFile("Checkpoint_State.h5", "r")   # File for loading reference constraints
11: Tobs = checkpoint_file.load_function(mesh, "Temperature", idx=max_timesteps - 1)   # Load reference final state
12: # Initialise control to reference final T
13: Tic.project(checkpoint_file.load_function(mesh, "Temperature", idx=max_timesteps - 1))
14: ...
15: control = Control(Tic)   # definition of control
16: T.project(Tic, bcs=energy_solver.strong_bcs)   # projection from Q1 to Q2,
17:                                                 # and imposing boundary conds
18: # Populate the tape by running the forward simulation
19: # and calculate surface velocity misfit
20: for timestep in range(0, max_timesteps):
21:     stokes_solver.solve()
22:     energy_solver.solve()
23:     uobs = checkpoint_file.load_function(mesh, name="Velocity", idx=timestep)   # load uobs
24:     u_misfit += assemble(dot(u - uobs, u - uobs) * ds_t)   # accumulate misfit
25:
26: checkpoint_file.close()
27:
28: # Define the objective functional
29: damping_mask = gaussian(X[1], 1.0, 0.1) + gaussian(X[1], 0.0, 0.1)
30: damping = assemble(damping_mask * (Tic - Taverage) ** 2 * dx)
31: norm_damping = assemble(damping_mask * Taverage**2 * dx)
32: smoothing = assemble(dot(grad(Tic - Taverage), grad(Tic - Taverage)) * dx)   # smoothing term
33: norm_smoothing = assemble(dot(grad(Tobs), grad(Tobs)) * dx)   # normalisation for smoothing
34: norm_obs = assemble(Tobs**2 * dx)   # normalisation for Tobs
35: norm_u_surface = assemble(dot(uobs, uobs) * ds_t)   # normalisation for u surface
36: t_misfit = assemble((T - Tobs) ** 2 * dx)   # Tobs misfit
37: alpha_u, alpha_d, alpha_s = 0.1, 0.1, 0.1
38: objective = (
39:     t_misfit
40:     + alpha_u * (norm_obs * u_misfit / max_timesteps / norm_u_surface)
41:     + alpha_d * (norm_obs * damping / norm_damping)
42:     + alpha_s * (norm_obs * smoothing / norm_smoothing))
43: # Define the reduced functional
44: reduced_functional = ReducedFunctional(objective, control)
```

Listing 2. Changes compared to the forward script in Listing 1 required to define the reduced functional used in the adjoint problem. The three `alpha` parameters on line 36 denote the three weighting parameters for surface, damping and smoothing terms, respectively.

500   factors, namely, α_u, α_s, and α_d. When combining terms linearly in the objective functional, the superposition principle dictates that the gradient of the entire objective functional is essentially the cumulative sum of the gradients of its constituent terms in J. Consequently, it becomes instructive to visualise and validate the gradient of each individual term within the objective functional.

Fig. 3-D to G display the gradient fields corresponding to the four terms in the objective functional. The gradient for the
505   final temperature misfit is presented in Fig. 3-D: it implies that a better match to the final state temperature field is achievable through changes to the initial condition that include a major reduction in temperature in the domain's centre, complemented by an increase in temperature moving towards the domain's edges, particularly towards the domain's upper regions. In Fig. 3-E, we illustrate the gradient of the cumulative surface velocity misfit with respect to T_{IC}: it reveals sensitivities extending to the domain's base, indicating that to better align with the 'observed' surface velocities, an optimal T_{IC} should be decreased in
510   the middle, but increased towards the domain's boundaries. For the smoothing term (Fig. 3-F), the highest values emerge in areas with the most abrupt changes in T_{IC}. Given our optimisation strategy works by moving towards corrections opposite to the gradient's direction, this subsequently results in iterative refinement and smoothing of the solution. For the damping term (Fig. 3-G), the gradient aims to minimise fluctuations of the field in relation to the ambient temperature profile ($\bar{T}$), adjacent to the top and bottom boundaries, implying that this can be achieved by reducing boundary layer temperatures in the centre but
515   increasing boundary layer temperatures towards the domain's sides.
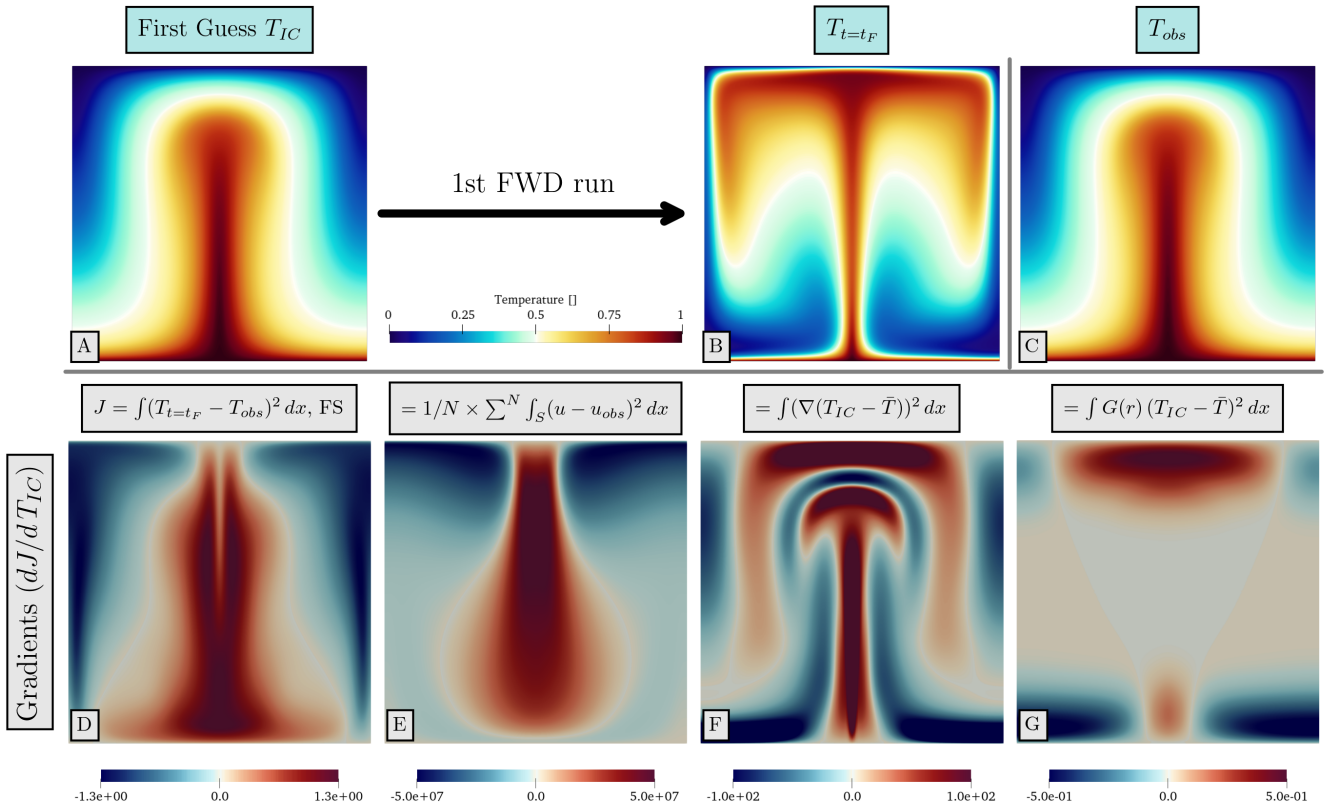
Figure 3. First forward run and associated gradients: **(A)** is the first guess for T_{IC}. For all experiments in this study we choose the *'observed'* temperature field (i.e., the final state) as the initial guess for our optimisations. **(B)** is the final temperature field after integrating forward in time from (A) for 80 time steps. **(C)** is the reference final temperature field, T_{obs}, which is used in the definition of the misfit functional. **(D)**-**(G)** illustrate the gradient fields. **(D)** is gradient of the final temperature misfit. **(E)** is the gradient of the total surface velocity misfit. **(F)** and **(G)** are for the regularising smoothing and damping terms, respectively.

3.1.4 Verification of Gradients: Taylor Remainder Convergence Test

A fundamental tool used in verification of gradients is the Taylor remainder convergence test (Farrell et al., 2013b). For the reduced functional, $J(T_{IC})$ defined in Eq. 6 and its derivative $\frac{\mathrm{d}J}{\mathrm{d}T_{IC}}$, it can be proven that,

$$\left| J(T_{IC} + h\,\delta T_{IC}) - J(T_{IC}) - h\frac{\mathrm{d}J}{\mathrm{d}T_{IC}} \cdot \delta T_{IC} \right| \longrightarrow 0 \text{ at } O(h^2). \tag{33}$$

520   The expression on the left hand side of Eq. 33 is termed the second-order Taylor remainder. This term's convergence rate of $O(h^2)$ serves as a strong foundation for verifying any computational implementation meant for determining $\frac{\mathrm{d}J}{\mathrm{d}T_{IC}}$ (the adjoint code) with respect to a specific functional that computes $J(T_{IC})$ (the forward code). Given any arbitrary selection of h and δT_{IC}, halving the value of h should decrease the magnitude of the second-order Taylor remainder by a factor of 4. Grounded in this theoretical prediction, we employ these so-called *Taylor tests* to confirm the accuracy of the determined gradients.
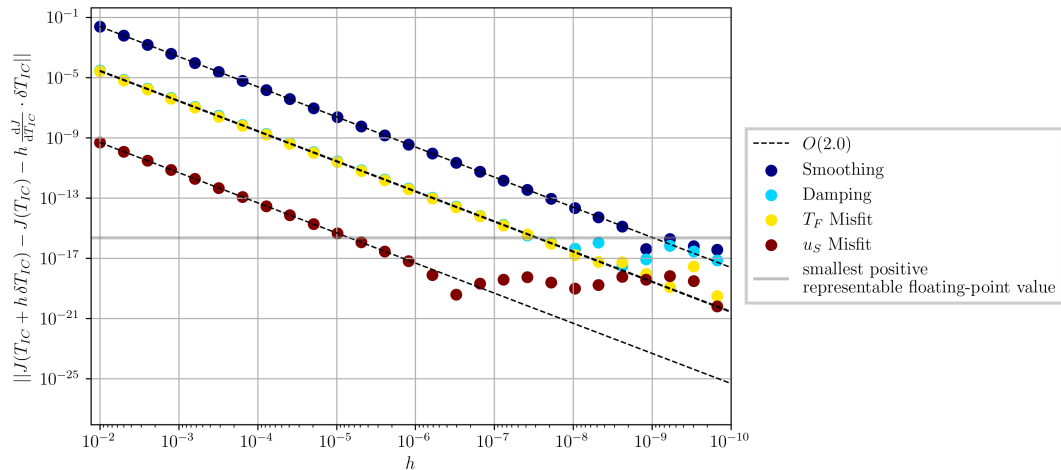
Figure 4. Second-order Taylor remainder test: for each gradient field, the test is performed by computing the functional and the associated gradient when randomly perturbing the initial temperature field T_{IC} and subsequently dividing the perturbations by a factor of two at each level. The dashed line is the theoretical convergence rate of $O(2.0)$. Residuals show accuracy of gradient information down to floating-point precision (defined as the smallest positive ϵ, such that $1.0 + \epsilon$ is distinguishable from 1.0.)

We conduct a second-order Taylor remainder test for each term of the objective functional. The results are displayed in Fig. 4, where the gradient fields are calculated for random perturbations of the initial temperature field T_{IC}, with the amplitude of these perturbations successively halved ($h/2, h/4, h/8, ...$). Notably, all four Taylor remainder tests demonstrate a convergence rate of $O(2.0)$, extending down to floating-point preision, defined as the smallest positive ϵ such that $1.0 + \epsilon$ is distinguishable from 1.0. Our results demonstrate consistency with theoretical expectations, highlighting the robustness of our approach.

3.1.5 Efficiency

Another metric that can be used to assess the suitability of our framework for large-scale mantle convection optimisation problems is the efficiency of derivative calculations. Given the iterative nature of our inverse problem, where derivatives are computed frequently, any efficiency gain, or the lack thereof, can have profound implications for the overall computational cost and feasibility of our automated approach. The computational efficiency can be measured by comparing the computational time of a derivative calculation to that of a forward calculation. Using this reference, a theoretical optimum is defined which measures the ratio of the time that is required to calculate one set of forward and adjoint calculations to one forward calculation. For the Stokes problem we have detailed, in which a linear rheology has been employed, this ratio is considered to be 2.0 (Naumann, 2011; Funke and Farrell, 2013). This is primarily due to the similarity of the forward and adjoint systems. For the simulations presented in this section, we achieve a ratio of 2.01, consistent with theoretical expectations, thus demonstrating the efficiency of our approach.

```
1: ...
2: # Perform a bounded nonlinear optimisation
3: T_lb = Function(Tic.function_space(), name="Lower bound temperature").assign(0.0)
4: T_ub = Function(Tic.function_space(), name="Upper bound temperature").assign(1.0)
5:
6: minimisation_problem = MinimizationProblem(reduced_functional, bounds=(T_lb, T_ub))
7:
8: optimiser = LinMoreOptimiser(minimisation_problem, minimisation_parameters, checkpoint_dir="
       optimisation_checkpoint")
9: optimiser.run()
```

Listing 3. Necessary changes to solve the minimisation problem. `T_lb` and `T_ub` are the lower and upper bounds for the minimisation problem, respectively.

3.1.6 Optimisation

Executing an optimisation task with G-ADOPT is straightforward. Once the reduced functional is set up (see Listing 2), only a few additional lines of Python are required (see Listing 3). As we use a bounded method for our optimisation problem, we specify a set of upper and lower bounds for the algorithm, on lines 3-4. Subsequently, a minimisation problem is outlined (line 6), using both the reduced functional and the designated bounds. This minimisation problem, together with the associated parameters for optimisation, are passed to the Lin-More optimisation algorithm in ROL (line 8), which is executed on line 9.

Using this framework, we perform a suite of 81 different inverse simulations that aim to find the most optimal combination of the three weightings (α_u, α_d, and α_s) that results in the best solution for T_{IC} when compared to the reference initial temperature field. The simulations are obtained by sweeping through values in ranges of $[10^{-1}, 10^{-3}]$, $[10^{-2}, 10^{-4}]$ and $[10^1, 10^{-7}]$ for α_d, α_s and α_u, respectively.

Fig. 5 provides an overview of the outcomes from a subset (16 out of 81) of these optimisation exercises. Minimisation of the objective functional is shown in Fig. 5-A, alongside two additional metrics: (i) the misfit between the reconstructed final temperature field and T_{obs}, termed the *final misfit* (Fig. 5-B); and (ii) the misfit between the reconstructed initial condition, T_{IC}, and the reference initial condition, highlighting the quality of the reconstructed initial condition (Fig. 5-C). The reduction in the metrics in all cases is reported versus the cost, which is the sum of the number of forward and adjoint calculations. A consistent pattern is observed across all reconstruction simulations for these three metrics. Firstly, the solutions are unique, as all converge to a consistent initial condition following roughly 100 iterations (a cost of 200 forward and adjoint calculations). However, the trajectory to this solution varies based on the smoothing weight, denoted by α_s. A significant portion of the simulations with $\alpha_s = 10^{-1}$ exhibit subpar performance in the initial stages, due to over-smoothing, with most of the best performing simulations utilising $\alpha_s = 10^{-2}$. Nevertheless, despite differences in convergence rates, all simulations eventually converge to similar misfits in all three metrics.

Fig. 6 showcases multiple iterations from the best-performing simulation using $\alpha_u = 10^{-2}$, $\alpha_s = 10^{-3}$, and $\alpha_d = 10^{-2}$. As mentioned previously, the reference final condition (Fig. 6-B) is used as our starting guess for the inverse simulation (Fig. 6-C), which yields substantial differences in the modelled final state (Fig. 6-K), as illustrated through the squared difference between reconstructed and reference temperature (Fig. 6-O). Leveraging derivative information, the inverse simulation corrects the initial condition through an iterative approach. The initial conditions obtained after 20, 50, and 100 iterations are shown in Figs. 6-D, E, and F, respectively. These solutions reveals that most corrections occur during the initial iterations, with significant
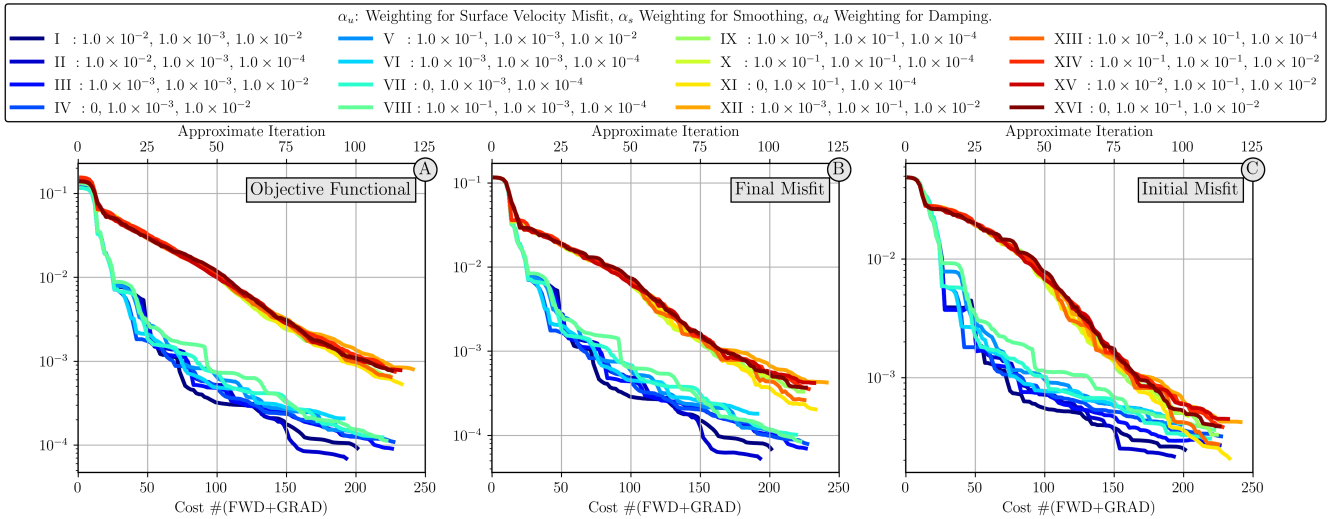
Figure 5. An overview of optimisation results across all experiments: (A) Illustrates the minimisation of the objective functional. (B) Depicts the final misfit, representing the difference between the reconstructed final temperature field and T_{obs}. (C) Highlights the initial misfit, characterising the discrepancy between the reconstructed initial condition T_{IC} and the reference initial condition. The bottom horizontal axis in all figures reports the cost, here defined as the sum of forward and derivative calculations. The top axis shows the approximate equivalent number of iterations, which simply is the cost divided by two. Across all metrics, simulations exhibit a consistent pattern: they converge to the same solution approximately after 100 iterations, though the path to this solution diverges depending on the smoothing weight, α_s.

improvement in the domain's upper part by iteration 20, albeit with some remaining noise in the solution (see misfit in Fig. 6-H). By iteration 100, the solution (visible in Fig. 6-F and N) closely mirrors the reference initial condition, represented by misfit values that have diminished by three orders of magnitude (Fig. 5-B and C).

Fig. 7 compares the evolution of reference forward (A-E) and reconstructed simulations (F-J), with differences highlighted (K-O). Furthermore, surface normal stresses are displayed as indicators of surface dynamic topography. For the reconstructed simulation (Fig. 7-F), the initial condition is well captured as a Gaussian anomaly in the domain's lower section, superimposed on a depth-dependent field, similar to the reference initial condition. This initial temperature anomaly ascends, culminating at the surface after 80 time steps. The precision of the reconstruction is evident from the misfit panels, with negligible differences between reference and reconstructed simulations throughout the model's evolution (less than 0.01). This is further substantiated by a reduction, by over four orders of magnitude, in the objective functional for both initial and end states (Fig. 5-B and C). The accuracy of reconstructed surface normal stresses are evidenced by negligible discrepancies in the associated misfit visualisation.

Parameter	Symbol	Value
Viscosity []	μ	—
Plastic viscosity []	μ_p	—
Linear (temperature-dependent) viscosity []	μ_{lin}	—
Principal strain rate tensor	$\epsilon_i i$	—
Ambient viscosity	$\mu(r)$	—
Yield stress (at Earth's surface)	σ_{yield}	2×10^4
Minimum plastic viscosity	μ^*	0.1
Temperature dependence of viscosity	Δ_T	$ln(80)$

Table 2. Parameters employed for 2-D annulus twin experiments.

3.2 Convection in an Annulus, Incorporating a Visco-plastic Rheology

We next analyse a set of reconstructions which utilise a reference twin that more closely mimics Earth's geometrical and rheological characteristics. We use a 2-D annular domain, generated by extruding 128 times radially from a circular manifold consisting of 512 cells. Inner and outer radii are set to 1.22 and 2.22 respectively, ensuring unit depth and maintaining a comparable ratio between surface and CMB radii as Earth's mantle. We set $Ra = 10^7$ and adopt a composite visco-plastic rheology, with effective viscosity determined via a harmonic mean, represented as:

$$\frac{1}{\mu} = \frac{1}{\mu_p} + \frac{1}{\mu_{lin}}, \qquad \mu_p = \mu^* + \frac{\sigma_{yield}}{\epsilon_{ii}}, \qquad \mu_{lin} = \mu_0(r) \exp\left(\frac{\Delta_{\mu,T}}{T}\right). \qquad (34)$$

Parameters specified in Eq. 34 are listed in Table 2. As demonstrated by Davies et al. (2022), the changes necessary to transform our forward model from a square to an annular domain, and from an isoviscous to a visco-plastic rheology, are only minor, noting the Firedrake has already been validated for simulations of this nature (see Davies et al. (2022) and repository accompanying this paper for a comprehensive script).

To determine the reference simulation's initial condition, we commence from a starting state with small-scale spherical harmonic perturbations and advance forward in time until a quasi-steady state is achieved (i.e., when basal and surface heat-fluxes are roughly in balance). The resulting state (Figs.8-1.A and 2.A), which contains downwellings that descend from the upper thermal boundary layer and upwelling plumes that rise from the lower thermal boundary layer, forms our reference initial condition. We subsequently advance forward for a duration of $t = 200 \times (5 \times 10^{-6})$ to reach the reference final state shown in Figs. 8-1.G. While the non-dimensional time of 10^{-3} translates into 256 Myrs of convection, the domain velocities suggest an equivalent Earth-like time of ≈ 100 Myrs (due to a marginally reduced Ra relative to estimates for Earth's mantle). The sequence presented in Figs. 8-1.A-G and 2.A-G trace the temporal development of reference temperature and viscosity fields, respectively. The viscosity field spans approximately three orders of magnitude, extending from low asthenospheric values of ≈ 1.0, rising to 140 within the cooler segments of the lower mantle, and decreasing to 0.4 in locations of high strain-rates at surface convergent boundaries.

3.2.1 Verification of Gradients

As with the previous Cartesian case examined, reconstruction simulations are conducted with an objective functional encompassing the misfit component related to the terminal temperature field, accumulative surface velocity misfits, and regularisation terms. For consistency with the previous case, we first confirm the accuracy of the calculated gradient fields corresponding to each component in the objective functional by performing second-order Taylor remainder convergence tests. To further analyse the robustness of our results against solver tolerances, specifically concerning the final temperature field and surface velocities, we conduct a set of two additional convergence tests. The first set utilises a Newton (SNES) relative solver tolerance of 10^{-10} and is depicted using solid colors in Fig. 9, while the second set, using a relative tolerance of 10^{-5}, is represented by semi-transparent colors. Our findings indicate that with tighter tolerances, results adhere to an O^{-2} trend consistent with theoretical predictions, confirming accuracy down to the smallest floating-point representation. Conversely, when the tolerance is relaxed to 10^{-5}, this behavior holds only up to perturbations of similar magnitude ($h \approx 10^{-5}$), demonstrating a divergence from the expected $O(h^2)$ trend for smaller perturbations.

3.2.2 Efficiency

In the experiments detailed in this section, an efficiency of 1.45 is achieved. This exceeds the previously outlined theoretical efficiency of 2.0 for the isoviscous Stokes model in Sec. 3.1.5, due to a major difference in the forward and adjoint momentum equations: while the forward momentum equation employs a non-linear visco-plastic rheology and requires multiple linear Newton solves per time-step, the adjoint momentum remains linear. The forward model is therefore more computationally demanding, explaining the improved efficiency ratio.

3.2.3 Optimisation

As highlighted with our previous example, absolute and relative variations in weighting of different objective functional components can generate solutions with distinct properties, some of which provide an improved match to the reference simulation. Accordingly, it is vital to assess the consequences of these distinct weight combinations for the case considered here. To address this, we have undertaken 21 simulations, adjusting the parameters α_u, α_s, and α_d within the intervals $[0.05, 0.1]$, $[0.01, 0.1]$, and $[0.01, 0.1]$ respectively, with values motivated by the results of our previous set of simulations. The collective convergence of all 21 simulations is illustrated in Fig. 10.

Our objective functional (Fig. 10-A), has initial values that range between $\sim 1 \times 10^{-1}$ and 2×10^{-1}. We consistently observe a reduction of an order of magnitude or more in this measure. Notably, the steepest decline is seen within the initial 50 iterations (cost ~ 100). The simulations denoted as X, XI exhibit the largest reduction, with a consistent reduction trajectory even when approaching iteration 200 (cost ~ 400).

The final temperature misfit (Fig. 10-B) exhibits different trends to the objective functional. In the initial iterations, the largest reductions in final misfit are observed for simulations I and II, with simulation XII trailing behind. Despite displaying a lower misfit reduction up until iterations 140 and 180 (cost $\sim 280 - 360$) simulations X and XI eventually display a similar

misfit by iteration 200 (cost ~ 400). When we turn to reduction in the initial misfit (Fig. 10-C), an entirely different trend comes to light: simulations I and II sustain their reduction until around iteration 150 (cost ~ 300), after which they plateau. Conversely, many other simulations plateau at misfit values that are, on average, twice as high.

In our analysis, the reconstruction quality is predominantly governed by three key weighting parameters: surface velocity misfit (α_u), smoothing (α_s), and damping (α_d). These parameters calibrate the significance of different objective terms in relation to the final temperature misfit term. Thus, the primary metric to assess a reconstruction is the simultaneous reduction of misfits for both initial and final states. Incorporating the misfit associated with surface velocities enhances the quality of the reconstructions, noting that higher weightings of the surface velocities require higher values of smoothing. Among the weighting parameters, α_s shows to have the highest effect in convergence outcomes: higher values for α_s lead to over-regularisation, thereby limiting the role of sensitivity information tied to misfit terms in the solution. α_d offers a more varied range of effective values, which aligns with its role confined to thermal boundary layers and, consequently, its lesser impact on the overall numerical domain. Therefore, Case I, characterised by $(\alpha_u, \alpha_s, \alpha_d) = (1 \times 10^{-1}, 1 \times 10^{-1}, 1 \times 10^{-2})$, emerges as the optimal set of weighting parameters for our reconstructions.

In contrast to the three orders of magnitude reduction in the misfit functions from the isoviscous experiment, our non-linear experiment exhibits a modest reduction of $O(1)$. The key factor contributing to this is the extended total simulation time for the non-linear experiment. Representing a far longer reconstruction period, this implies more information loss during the inversion process. Additionally, while the isoviscous experiment focused on a single temperature anomaly, this simulation tackles whole-mantle convection, with numerous, and occasionally complex and highly time-dependent anomalies, reflecting the more intricate visco-plastic rheology. Nonetheless, this order of magnitude reduction in misfit translates into a satisfactory reconstruction of the initial condition, demonstrating the efficacy and robustness of the numerical approaches employed.

This is confirmed by visual inspection of the best reconstruction model, with temperature, viscosity and surface normal stresses presented and compared to the reference case in Fig. 11 (marked case I in Fig. 10, using values of $\alpha_u = 10^{-1}$, $\alpha_s = 10^{-1}$ and $\alpha_d = 10^{-2}$). At $t = 0$, the reconstructed temperature field exhibits upwelling and downwelling features that are reconstructed in the correct locations, although temperature anomalies are generally smoother than those of the reference case (Fig. 11-1.A). The corresponding viscosity field mirrors this smoothness, despite capturing weaker convergence zones at the surface. Despite these variations, the spatial misfit is generally below 10^{-2}, (Fig. 11-5.A) with errors over 0.05 restricted to sharper features that are inevitably smoothed in the reconstruction process. This smoothness is also reflected in recovered surface normal stresses, with highs and lows correctly positioned, albeit at longer wavelengths than the reference case.

Given the application of a free-slip boundary condition at the surface of this simulation, a prominent outcome of this set of experiments is the emergence of sharp subducting slabs and weak zones at the top boundary as the simulation evolves (Figs. 11-3.B and 4.B). The marked decrease in misfit over time (Fig. 11-5.B) confirms the development of more detailed convective patterns. As the simulation evolves, reconstructed plume features become more precise and reconstructed surface normal stresses more closely resembles the reference case. This enhancement progresses up to the final time-step, where the reconstructed thermal field and surface normal stresses are indistinguishable from those in the reference simulation, reflected via an order of magnitude reduction in the spatial misfit field.

4 Discussion

670 Robust reconstructions of the spatial and temporal evolution of Earth's mantle and its diverse surface expressions is critical to scientific progress across the geosciences. It requires the construction of a digital twin: a vital instrument for analysing and revealing the complex interplay between the mantle and Earth's other systems. To this end, the adjoint method provides the necessary means for obtaining and analysing model sensitivities with respect to earlier mantle states. A burgeoning number of studies exploiting this methodology for reconstructions of mantle convection have emerged in recent years (e.g., Bunge et al.,

675 2003; Ismail-Zadeh et al., 2004; Liu et al., 2010; Spasojevic et al., 2009; Li et al., 2017; Price and Davies, 2018; Ghelichkhan et al., 2021). Nevertheless, the derivation, implementation, and validation of adjoint systems for coupled, non-linear, time-dependent systems remains notoriously difficult. It is due to these difficulties that existing applications of the geodynamic adjoint method often include major simplifications, either incorporating an oversimplified treatment of mantle rheology (e.g., Colli et al., 2018; Ghelichkhan et al., 2021), neglecting certain (coupling) terms in the adjoint equations (e.g., Ismail-Zadeh

680 et al., 2004), or both (e.g., Liu and Gurnis, 2010): they are therefore likely limited in their applicability to realistic Earth scenarios. In this study, we leverage the latest advances in scientific computing to overcome these limitations and develop G-ADOPT, an open-source numerical framework for geoscientific adjoint reconstructions, developed in full compliance with FAIR (Findable, Accessible, Interoperable, Reusable) principles (Wilkinson et al., 2016).

G-ADOPT is underpinned by three primary software elements. The first is Firedrake (Ham et al., 2023), an automated

685 system for solving partial differential equations using the finite-element method. In our previous work (Davies et al., 2022) we examined the applicability of Firedrake for geodynamical simulations, confirming its accuracy, efficiency, extensibility and parallel scalability through comprehensive benchmarks and state-of-the-art mantle convection simulations. The second element is Dolfin-Adjoint (Farrell et al., 2013a; Mitusch et al., 2019), a system that automatically generates the discrete adjoint from forward models designed in Firedrake. Dolfin-Adjoint elevates the conventional abstraction of automatic differentiation

690 from individual floating point operations to complete systems of differential equations, leveraging the high-level mathematical abstraction of finite element problems and their symbolic representation in UFL (Alnæs et al., 2014). The adjoint systems derived by Dolfin-Adjoint are UFL expressions and valid Firedrake inputs. Therefore, they inherit the parallel support native to the forward model, which results in optimal computational efficiency. The third element is the Rapid Optimisation Library, ROL, a Trilinos package for large-scale optimisation problems (The ROL Project Team, 2022), enhanced herein with intra-

695 optimisation checkpointing functionality.

We have demonstrated the applicability of G-ADOPT for time-dependent geodynamic reconstructions herein. The objective functional utilised in our reconstructions is composed of two distinct misfit components. The first is a term that quantifies the misfit corresponding to the observed final state temperature field, analogous to the present-day temperature field within Earth's mantle as obtained through a combination of mantle mineralogical models (e.g., Stixrude and Lithgow-Bertelloni, 2011; Chust

700 et al., 2017) and seismic imaging (e.g., Rawlinson et al., 2010; French and Romanowicz, 2014; Simmons et al., 2015; Bozdağ et al., 2016; Koelemeijer et al., 2016; Fichtner et al., 2018). The second term corresponds to observed surface velocities,

accessible through plate tectonic reconstruction models (e.g., Müller et al., 2019). Additionally, smoothing and damping terms have been incorporated to enforce regularity in our solutions.

Our study analysed two sets of reconstructions of systematically increasing complexity. We first examined the evolution of a single ascending hot anomaly in an enclosed isoviscous square domain. By taking advantage of the simplicity of the geometry and rheological properties, we were able to deliver an in-depth examination of the gradients for each term, including a parameter-space search to ascertain optimal weighting parameters. Our results reveal a general convergence of the solutions, notwithstanding substantial variations in convergence rates subject to the weightings. Additionally, although not detailed in this paper, we have explored a number of different optimisation methods and parameters. Through this comprehensive analysis, we are confident that the problem possesses a stable solution that can be found through an appropriate combination of weighting parameters. The second set of reconstruction experiments explored convection with a stress-, depth- and temperature-dependent rheology at the convective vigour of Earth's mantle, demonstrating the feasibility of reconstruction studies for Earth's mantle with a non-linear rheology. The weightings selected for this series of experiments were broadly consistent with the first set. Given the success at reproducing surface velocities and normal stresses, our findings suggest that reconstruction models of Earth's mantle can serve as a powerful means for probing changes in the landscape at Earth's surface induced by mantle dynamics (e.g., Friedrich et al., 2018; Hoggard et al., 2021; Davies et al., 2023).

In both experimental sets, we assessed the numerical efficiency of our framework by evaluating the cost ratio between forward and adjoint calculations. In the first set, our results produced a ratio of 2.01, aligning with the theoretical efficiency of 2.0 (e.g. Naumann, 2011). In the second set, where we solved the nonlinear forward equations, we observed a ratio of 1.45. This efficiency is attributed to the linearised nature of the adjoint method: even when applying nonlinear rheologies in the forward equations, the adjoint equations remain linear. We also conducted second-order Taylor remainder convergence tests for each of the objective functional terms to validate the adjoint calculations. We note our convergence is accurate down to floating-point precision, consistent with results presented by Coltice et al. (2023). Our assessments demonstrate the accuracy of the derivative calculations (Figs.4 and 9). These Taylor remainder convergence tests provide a robust basis for future validations of geodynamic adjoint frameworks.

Our experiments incorporate two significant simplifications relative to realistic-Earth scenarios, which were necessary to facilitate the number of reconstruction simulations analysed: (i) the use of a 2-D computational domain; and (ii) application of the Boussinesq approximation instead of more pertinent approximations such as anelastic-liquid approximations (e.g., Jarvis and McKenzie, 1980). Nevertheless, the composable nature of G-ADOPT should alleviate any concerns regarding the extensibility of our framework to these more realistic problem sets. Our prior work in Davies et al. (2022) demonstrates the flexible nature of our approach: for example, transitioning our 2-D annulus simulations to a 3-D spherical shell domain can be achieved via changes to only a few lines of Python. The application of G-ADOPT for reconstructing Earth's mantle evolution using nonlinear rheologies and compressiblity will be the subject of future investigations, although we note that the forward modelling approach has already been developed (Davies et al., 2022). Moreover, our framework is extensible to various other problems in geodynamics. These include utilising principal stress directions (e.g., Reuber et al., 2020), surface plate velocities (e.g., Ratnaswamy et al., 2015; Bocher et al., 2018), and/or residual depth measurements (e.g., Panasyuk and Hager, 2000; Spasoje-

vic et al., 2009) to explore the mantle's rheological properties, and to study the visco-elastic adjustment of Earth's surface in response to the melting of Earth's polar ice-sheets (Al-Attar and Tromp, 2014; Martinec et al., 2015, e.g.,), and post-seismic deformation following significant subduction earthquakes (e.g., Sabadini and Vermeersen, 1997).

740   Reconstructing past mantle states is fraught with substantial theoretical and practical challenges. In this study, we targeted some of these theoretical and practical hurdles by introducing G-ADOPT. Nevertheless, significant obstacles exist that are beyond the scope of this work. We predicated our work on zero uncertainty in our reference fields (i.e., the present-day temperature field and past surface velocities), thereby committing what is known as the 'inverse crime' (Colton et al., 1998), a term used to describe the situation when the code employed in the inversions is also utilised to generate reference simulations. Estimation

745 of the present-day mantle state from seismic imaging and the assumptions regarding the thermal and compositional interpretation of seismic heterogeneity are both fraught with considerable uncertainty (e.g., Styles et al., 2011; Mosca et al., 2012; Zaroli et al., 2013; Davies et al., 2015). Furthermore, plate tectonic reconstructions can be uncertain, particularly further back in time and within the Pacific region (e.g., Shephard et al., 2012; Williams et al., 2015; Tetley et al., 2019). However, the existence of seafloor spreading isochrons up to approximately 125 Ma for all major plates provides confidence in modelling relative

750 plate movements in more recent geological periods (Seton et al., 2020). An uncertainty impact study carried out by Colli et al. (2020) posits that the presence of uncertainties causes reconstructed and reference flow histories to diverge exponentially back in time, with unrealistic structures materialising within and adjacent to thermal boundary layers. To minimise these impacts, Colli et al. (2020) advocate for terminating the optimisation after a few iterations. Here, however, inclusion of regularisation terms in the objective functional mitigates these impacts, effectively constraining the reconstruction to a smoother solution.

755 This becomes particularly advantageous in real-Earth applications where observational constraints become sparser and more uncertain back in time. Without a smoothing term, the solution to the initial condition can contain high-frequency noise, which would diffuse over the course of the simulation. Smoothing therefore drives the solution towards a longer-wavelength initial state, whilst maintaining sensitivity to shorter wavelength information recorded in seismic tomography images.

     Moreover, by formally introducing past surface velocities into the objective functional, we infuse sensitivity information that

760 propagates further back in time, refining the flow trajectory to improve the accuracy of reconstructions in the upper thermal boundary layer region. This sets our approach apart from the method used in previous adjoint reconstruction simulations (e.g., Vynnytska and Bunge, 2015; Zhou and Liu, 2017; Ghelichkhan et al., 2021), where the sequential-in-time nature of plate velocity assimilation can improve flow trajectories only forward in time. Despite this, it is crucial to recognise that observed surface velocities are the result of a complex force balance, including contributions from within Earth's mantle and at Earth's

765 surface, including those at plate boundaries, such as the genesis and destruction of orogenies, which have been associated with rapid changes in plate velocities (Colli et al., 2014; Iaffaldano and Bunge, 2015). These forces are not captured in current state-of-the-art spherical mantle convection models. Although our inversions utilise absolute plate velocities – thereby accounting for the majority of the forces driving surface movements – it is essential to emphasise that constraints on Earth's present-day thermal field must be prioritised as the primary data source for such inversions.

770   Our study lays the foundations for exploring several unresolved geodynamical questions. Previous research has shown that prescribing plate tectonic reconstruction velocities as a top boundary condition improves the precision of mantle reconstruction

models and diminishes noise (e.g., Colli et al., 2015; Taiwo et al., 2023). Our framework formally incorporates these constraints through misfit terms, and future studies should compare this with other methods to find the most efficient way to integrate this valuable data. Earlier research advocates solving the reduced adjoint system, effectively considering velocities as insensitive to initial conditions (e.g., Ismail-Zadeh et al., 2004; Liu et al., 2008). The second-order Taylor remainder convergence tests, examined herein, provides a robust foundation for evaluating the accuracy of such simplifications. Furthermore, our framework sets the stage for including hitherto unused observations within our inversions, such as geochemical constraints on mantle temperature and pressure (Ball et al., 2021). This stems from the design principle of *composable abstractions* in the software packages used in G-ADOPT, ensuring all components' modularity, interoperability, reusability, scalability, and maintainability. Specifically, Firedrake emphasises a clear separation between using the finite element method and implementing it. Dolfin-Adjoint automates the derivation and computation of the adjoint systems using high-level symbolic language, ensuring the same advanced strategies that are applied for the forward calculation are utilised in the adjoints. Finally, ROL offers large-scale optimisation algorithms that seamlessly integrates with Firedrake and Dolfin-Adjoint. Their integration through G-ADOPT is a groundbreaking development that opens up adjoint problems to a new class of user and developer.

5   Conclusions

Transitioning from idealised forward models to data-driven simulations necessitates an inverse approach that rigorously incorporates observational constraints and their uncertainties. The adjoint method has emerged as a key technique for optimising unknown model parameters against observational data (e.g., Bunge et al., 2003). In this context, we introduce the Geoscientific Adjoint Optimisation Platform (G-ADOPT), which capitalises on cutting-edge developments in computational sciences, particularly through innovations in Firedrake (e.g., Rathgeber et al., 2016; Davies et al., 2022), Dolfin-Adjoint (Farrell et al., 2013a; Mitusch et al., 2019), and the Rapid Optimisation Library (ROL) (The ROL Project Team, 2022).

Our investigation, validated by two distinct sets of twin experiments, demonstrates G-ADOPT's efficacy in deducing the initial conditions for mantle flow and tracing its progression through space and time. These synthetic experiments leverage misfit terms to integrate both contemporary constraints across the computational domain and temporal constraints at Earth's surface. Additionally, we explored regularisation techniques to modulate the amplitude and complexity within the optimal solution. Notwithstanding the simplifications made (e.g., neglecting compressibility and 3-D geometry), G-ADOPT's architecture allows for straightforward adaptation to more complex scenarios with minimal modifications to the Firedrake forward model (as evidenced by Davies et al., 2022), alongside automated adjoint derivation and computation via Dolfin-Adjoint. The employment of the second-order Taylor remainder convergence test further corroborates our methodology, establishing a precedent for the advancement of geodynamic adjoint frameworks and facilitating prompt validation of adjoint models for more complex analyses.

Historically, geoscientific modelling frameworks were tailored to specific equations, which limited their application across disciplines. The current progress in adjoint-based techniques, crucial for data assimilation, sensitivity analysis, and optimisation, has markedly benefited meteorology and oceanography, though similar progress in other fields has been hindered by

derivation and implementation challenges. G-ADOPT's modular design seeks to bridge these divides. Its components are designed for ease of assembly and reusability, promoting a culture of modular, interoperable, scalable, and maintainable methods and frameworks. This philosophy ensures that G-ADOPT can be readily adapted to a broad spectrum of geoscientific research areas.

Code and data availability. For the specific components of G-ADOPT, including the full scripts of the simulations used in this paper see https://doi.org/10.5281/zenodo.10050733 (Gibson et al., 2023). For the specific components of Firedrake project used in this paper see https://zenodo.org/records/10047031 (zenodo/Firedrake-20231027.0).

Author contributions. SG and DRD conceived this study, with all the authors having significant input on the design, development and validation of the examples and cases presented. All authors contributed towards writing the manuscript.

Competing interests. David Ham is chief executive editor of Geoscientific Model Development. The contact author has declared that neither they nor their co-authors have any other competing interest.

Financial support. This research has been supported by the Australian Research Data Commons (ARDC), AuScope, Geosciences Australia and the National Computational Infrastructure (NCI) under G-ADOPT platform grant PL031. It was also supported by the Australian Research Council under grant nos. DP170100058 and DP220100173 and the Engineering and Physical Sciences Research Council [grant no. EP/R029423/1]

Acknowledgements. Numerical simulations were undertaken on the Gadi supercomputer at the National Computational Infrastructure (NCI) in Canberra, Australia, which is supported by the Australian Commonwealth Government. The authors are grateful to the entire Firedrake, Dolfin-Adjoint and ROL development teams for support and advice at various points of this research. We are also grateful to Georg Reuber, Nicolas Coltice and anonymous reviewer: their careful and constructive feedback has significantly improved the quality of this contribution.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Good-
fellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S.,
Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F.,
830   Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X.: TensorFlow: Large-Scale Machine Learning on Heterogeneous
Systems, https://www.tensorflow.org/, software available from tensorflow.org, 2015.

Al-Attar, D. and Tromp, J.: Sensitivity kernels for viscoelastic loading based on adjoint methods, Geophysical Journal International, 196,
34–77, 2014.

Alnæs, M. S., Logg, A., Ølgaard, K. B., Rognes, M. E., and Wells, G. N.: Unified form language: A domain-specific language for weak formu-
835   lations of partial differential equations, ACM Transactions on Mathematical Software (TOMS), 40, 1–37, https://doi.org/10.1145/2566630,
2014.

Balay, S., Gropp, W. D., McInnes, L. C., and Smith, B. F.: Efficient Management of Parallelism in Object Oriented Numerical Software
Libraries, in: Modern Software Tools in Scientific Computing, edited by Arge, E., Bruaset, A. M., and Langtangen, H. P., pp. 163–202,
Birkhäuser Press, 1997.

840   Balay, S., Abhyankar, S., Adams, M. F., Benson, S., Brown, J., Brune, P., Buschelman, K., Constantinescu, E. M., Dalcin, L., Dener, A.,
Eijkhout, V., Faibussowitsch, J., Gropp, W. D., Hapla, V., Isaac, T., Jolivet, P., Karpeev, D., Kaushik, D., Knepley, M. G., Kong, F., Kruger,
S., May, D. A., McInnes, L. C., Mills, R. T., Mitchell, L., Munson, T., Roman, J. E., Rupp, K., Sanan, P., Sarich, J., Smith, B. F., Zampini,
S., Zhang, H., Zhang, H., and Zhang, J.: PETSc Web page, https://petsc.org/, https://petsc.org/, 2023.

Ball, P., Czarnota, K., White, N., Klöcking, M., and Davies, D. R.: Thermal structure of eastern Australia's upper mantle and its
845   relationship to Cenozoic volcanic activity and dynamic topography, Geochemistry, Geophysics, Geosystems, 22, e2021GC009 717,
https://doi.org/10.1029/2021gc009717, 2021.

Bauer, S., Huber, M., Ghelichkhan, S., Mohr, M., Rüde, U., and Wohlmuth, B.: Large-scale simulation of mantle convection based on a new
matrix-free approach, Journal of computational science, 31, 60–76, 2019.

Bauer, S., Bunge, H.-P., Drzisga, D., Ghelichkhan, S., Huber, M., Kohl, N., Mohr, M., Rüde, U., Thönnes, D., and Wohlmuth, B. I.: TerraNeo-
850   Mantle Convection Beyond a Trillion Degrees of Freedom., 2020.

Bocher, M., Coltice, N., Fournier, A., and Tackley, P. J.: A sequential data assimilation approach for the joint reconstruction of mantle
convection and surface tectonics, Geophysical Journal International, 204, 200–214, https://doi.org/10.1093/gji/ggv427, 2016.

Bocher, M., Fournier, A., and Coltice, N.: Ensemble Kalman filter for the reconstruction of the Earth's mantle circulation, Nonlinear Processes
in Geophysics, 25, 99–123, https://doi.org/10.5194/npg-25-99-2018, 2018.

855   Bower, D. J., Gurnis, M., and Seton, M.: Lower mantle structure from paleogeographically constrained dynamic Earth models, Geochemistry,
Geophysics, Geosystems, 14, 44–63, https://doi.org/10.1029/2012GC004267, 2013.

Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N., and Pugmire, D.: Global adjoint tomography:
first-generation model, Geophysical Journal International, 207, 1739–1766, https://doi.org/10.1093/gji/ggw356, 2016.

Bunge, H.-P. and Grand, S. P.: Mesozoic plate-motion history below the northeast Pacific Ocean from seismic images of the subducted
860   Farallon slab, Nature, 405, 337–340, 2000.

Bunge, H.-P., Richards, M., and Baumgardner, J.: Mantle–circulation models with sequential data assimilation: inferring present–day mantle structure from plate–motion histories, Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 360, 2545–2567, https://doi.org/10.1098/rsta.2002.1080, 2002.

Bunge, H.-P., Hagelberg, C. R., and Travis, B. J.: Mantle circulation models with variational data assimilation: inferring past mantle flow and structure from plate motion histories and seismic tomography, Geophysical Journal International, 152, 280–301, https://doi.org/10.1046/j.1365-246X.2003.01823.x, 2003.

Carrassi, A., Bocquet, M., Bertino, L., and Evensen, G.: Data assimilation in the geosciences: An overview of methods, issues, and perspectives, WIREs Climate Change, 9, e535, https://doi.org/https://doi.org/10.1002/wcc.535, 2018.

Chust, T. C., Steinle-Neumann, G., Dolejš, D., Schuberth, B. S. A., and Bunge, H.-P.: MMA-EoS: A Computational Framework for Mineralogical Thermodynamics, Journal of Geophysical Research: Solid Earth, 122, 9881–9920, https://doi.org/10.1002/2017JB014501, 2017.

Cloetingh, S. and Haq, B. U.: Inherited landscapes and sea level change, Science, 347, 1258 375, https://doi.org/10.1126/science.1258375, 2015.

Colli, L., Stotz, I., Bunge, H.-P., Smethurst, M., Clark, S. R., Iaffaldano, G., Tassara, A., Guillocheau, F., and Bianchi, M. C.: Rapid South Atlantic spreading changes and coeval vertical motion in surrounding continents: Evidence for temporal changes of pressure-driven upper mantle flow, Tectonics, 33, 1304–1321, https://doi.org/10.1002/2014TC003612, 2014.

Colli, L., Bunge, H.-P., and Schuberth, B. S. A.: On retrodictions of global mantle flow with assimilated surface velocities, Geophysical Research Letters, 42, 8341–8348, https://doi.org/10.1002/2015GL066001, 2015.

Colli, L., Ghelichkhan, S., Bunge, H.-P., and Oeser, J.: Retrodictions of Mid Paleogene mantle flow and dynamic topography in the Atlantic region from compressible high resolution adjoint mantle convection models: Sensitivity to deep mantle viscosity and tomographic input model, Gondwana Research, 53, 252–272, https://doi.org/10.1016/j.gr.2017.04.027, 2018.

Colli, L., Bunge, H., and Oeser, J.: Impact of model inconsistencies on reconstructions of past mantle flow obtained using the adjoint method, Geophysical Journal International, 221, 617–639, https://doi.org/10.1093/gji/ggaa023, 2020.

Coltice, N., Gérault, M., and Ulvrová, M.: A mantle convection perspective on global tectonics, Earth-Science Reviews, 165, 120–150, https://doi.org/10.1016/j.earscirev.2016.11.006, 2017.

Coltice, N., Blessing, S., Giering, R., and Tackley, P.: Sensitivity Analysis of Global Kinematics on Mantle Structure Using Automatically Generated Adjoint Thermochemical Convection Codes, https://doi.org/10.31223/X5N09Q, 2023.

Colton, D. L., Kress, R., and Kress, R.: Inverse acoustic and electromagnetic scattering theory, vol. 93, Springer, 1998.

Crawford, O., Al-Attar, D., Tromp, J., Mitrovica, J. X., Austermann, J., and Lau, H. C.: Quantifying the sensitivity of post-glacial sea level change to laterally varying viscosity, Geophysical journal international, 214, 1324–1363, https://doi.org/10.1093/gji/ggy184, 2018.

Dalcin, L. D., Paz, R. R., Kler, P. A., and Cosimo, A.: Parallel distributed computing using Python, Advances in Water Resources, 34, 1124 – 1139, https://doi.org/10.1016/j.advwatres.2011.04.013, new Computational Methods and Software Tools, 2011.

Dannberg, J. and Gassmöller, R.: Chemical trends in ocean islands explained by plume–slab interaction, Proceedings of the National Academy of Sciences, 115, 4351–4356, https://doi.org/10.1073/pnas.1714125115, 2018.

Davies, D. R. and Davies, J. H.: Thermally-driven mantle plumes reconcile multiple hot-spot observations, Earth and Planetary Science Letters, 278, 50–54, 2009.

Davies, D. R., Goes, S., Davies, J. H., Schuberth, B. S. A., Bunge, H.-P., and Ritsema, J.: Reconciling dynamic and seismic models of Earth's lower mantle: The dominant role of thermal heterogeneity, Earth and Planetary Science Letters, 353–354, 253–269, https://doi.org/10.1016/j.epsl.2012.08.016, 2012.

Davies, D. R., Goes, S., and Lau, H. C. P.: Thermally Dominated Deep Mantle LLSVPs: A Review, pp. 441–477, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-319-15627-9_14, 2015.

Davies, D. R., Kramer, S. C., Ghelichkhan, S., and Gibson, A.: Towards automatic finite-element methods for geodynamics via Firedrake, Geoscientific Model Development, 15, 5127–5166, https://doi.org/10.5194/gmd-15-5127-2022, 2022.

Davies, D. R., Ghelichkhan, S., Hoggard, M., Valentine, A. P., and Richards, F. D.: Observations and models of dynamic topography: Current status and future directions, Dynamics of Plate Tectonics and Mantle Convection, pp. 223–269, https://doi.org/10.1016/B978-0-323-85733-8.00017-2, 2023.

Davies, G. F. and Richards, M. A.: Mantle Convection, The Journal of Geology, 100, 151–206, https://doi.org/10.1086/629582, 1992.

Dembo, R. S. and Steihaug, T.: Truncated-Newton algorithms for large-scale unconstrained optimization, Mathematical Programming, 26, 190–212, https://doi.org/10.1007/BF02592055, 1983.

Farrell, P. E., Ham, D. A., Funke, S. W., and Rognes, M. E.: Automated derivation of the adjoint of high-level transient finite element programs, SIAM Journal on Scientific Computing, 35, 1–27, https://doi.org/10.1137/120873558, 2013a.

Farrell, P. E., Ham, D. A., Funke, S. W., and Rognes, M. E.: Automated derivation of the adjoint of high-level transient finite element programs, SIAM Journal on Scientific Computing, 35, C369–C393, 2013b.

Fichtner, A., van Herwaarden, D.-P., Afanasiev, M., Simutė, S., Krischer, L., Çubuk-Sabuncu, Y., Taymaz, T., Colli, L., Saygin, E., Villaseñor, A., et al.: The collaborative seismic earth model: generation 1, Geophysical research letters, 45, 4007–4016, https://doi.org/10.1029/2018GL077338, 2018.

French, S. W. and Romanowicz, B. A.: Whole-mantle radially anisotropic shear velocity structure from spectral-element waveform tomography, Geophysical Journal International, 199, 1303–1327, https://doi.org/10.1093/gji/ggu334, 2014.

Friedrich, A. M., Bunge, H.-P., Rieger, S. M., Colli, L., Ghelichkhan, S., and Nerlich, R.: Stratigraphic framework for the plume mode of mantle convection and the analysis of interregional unconformities on geological maps, Gondwana Research, 53, 159–188, https://doi.org/10.1016/j.gr.2017.06.003, 2018.

Funke, S. W. and Farrell, P. E.: A framework for automated PDE-constrained optimisation, arXiv preprint arXiv:1302.3894, 2013.

Ghelichkhan, S. and Bunge, H.-P.: The compressible adjoint equations in geodynamics: Derivation and numerical assessment, GEM-International Journal on Geomathematics, 7, 1–30, https://doi.org/10.1007/s13137-016-0080-5, 2016.

Ghelichkhan, S. and Bunge, H.-P.: The adjoint equations for thermochemical compressible mantle convection: derivation and verification by twin experiments, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 474, 20180329, https://doi.org/10.1098/rspa.2018.0329, 2018.

Ghelichkhan, S., Murböck, M., Colli, L., Pail, R., and Bunge, H.-P.: On the observability of epeirogenic movement in current and future gravity missions, Gondwana Research, 53, 273–284, https://doi.org/10.1016/j.gr.2017.04.016, 2018.

Ghelichkhan, S., Fuentes, J. J., Hoggard, M. J., Richards, F. D., and Mitrovica, J. X.: The precession constant and its long-term variation, Icarus, p. 114172, https://doi.org/10.1016/j.icarus.2020.114172, 2020.

Ghelichkhan, S., Bunge, H.-P., and Oeser, J.: Global mantle flow retrodictions for the early Cenozoic using an adjoint method: evolving dynamic topographies, deep mantle structures, flow trajectories and sublithospheric stresses, Geophysical Journal International, 226, 1432–1460, https://doi.org/10.1093/gji/ggab108, 2021.

Gibson, A., Davies, D. R., Kramer, S., Ghelichkhan, S., Duvernay, T., and Turner, R.: g-adopt/g-adopt: v2.1.0, 2023.

Giering, R. and Kaminski, T.: Recipes for adjoint code construction, ACM Transactions on Mathematical Software (TOMS), 24, 437–474, https://doi.org/10.1145/293686.293695, 1998.

Giles, M. B. and Pierce, N. A.: An introduction to the adjoint approach to design, Flow, turbulence and combustion, 65, 393–415, 2000.

Gunzburger, M.: Adjoint equation-based methods for control problems in incompressible, viscous flows, Flow, Turbulence and Combustion, 65, 249–272, 2000.

940 Gunzburger, M. D.: Perspectives in flow control and optimization, SIAM, 2002.

Gurnis, M., Turner, M., Zahirovic, S., DiCaprio, L., Spasojevic, S., Müller, R. D., Boyden, J., Seton, M., Manea, V. C., and Bower, D. J.: Plate tectonic reconstructions with continuously closing plates, Computers & Geosciences, 38, 35–42, https://doi.org/10.1016/j.cageo.2011.04.014, 2012.

Ham, D. A., Kelly, P. H. J., Mitchell, L., Cotter, C. J., Kirby, R. C., Sagiyama, K., Bouziani, N., Vorderwuelbecke, S., Gregory, T. J.,
945 Betteridge, J., Shapero, D. R., Nixon-Hill, R. W., Ward, C. J., Farrell, P. E., Brubeck, P. D., Marsden, I., Gibson, T. H., Homolya, M., Sun, T., McRae, A. T. T., Luporini, F., Gregory, A., Lange, M., Funke, S. W., Rathgeber, F., Bercea, G.-T., and Markall, G. R.: Firedrake User Manual, Imperial College London and University of Oxford and Baylor University and University of Washington, first edition edn., https://doi.org/10.25561/104839, 2023.

Hansen, P. C.: Analysis of discrete ill-posed problems by means of the L-curve, SIAM review, 34, 561–580, 1992.

950 Hastie, T., Tibshirani, R., Friedman, J. H., and Friedman, J. H.: The elements of statistical learning: data mining, inference, and prediction, vol. 2, Springer, 2009.

Hinze, M., Pinnau, R., Ulbrich, M., and Ulbrich, S.: Optimization with PDE constraints, vol. 23, Springer Science & Business Media, 2008.

Hoggard, M., Austerman, J., Randel, C., and Stephenson, S.: Observing dynamic topography through space and time, AGU Geophysical Monograph 263, Washington, DC, https://doi.org/10.1002/9781119528609.ch15, 2021.

955 Hoggard, M. J., Czarnota, K., Richards, F. D., Huston, D. L., Jaques, A. L., and Ghelichkhan, S.: Global distribution of sediment-hosted metals controlled by craton edge stability, Nature Geoscience, 13, 504–510, https://doi.org/10.1038/s41561-020-0593-2, 2020.

Horbach, A., Bunge, H.-P., and Oeser, J.: The adjoint method in geodynamics: derivation from a general operator formulation and application to the initial condition problem in a high resolution mantle circulation model, GEM: International Journal on Geomathematics, 5, 163–194, https://doi.org/10.1007/s13137-014-0061-5, 2014.

960 Iaffaldano, G. and Bunge, H.-P.: Rapid plate motion variations through geological time: Observations serving geodynamic interpretation, Annual Review of Earth and Planetary Sciences, 43, 571–592, 2015.

Iglesias, J. A., Sturm, K., and Wechsung, F.: Two-dimensional shape optimization with nearly conformal transformations, SIAM Journal on Scientific Computing, 40, A3807–A3830, 2018.

Ismail-Zadeh, A., Schubert, G., Tsepelev, I., and Korotkii, A.: Inverse problem of thermal convection: Numerical approach and application
965 to mantle plume restoration, Physics of the Earth and Planetary Interiors, 145, 99–114, https://doi.org/10.1016/j.pepi.2004.03.006, 2004.

Jarvis, G. T. and McKenzie, D. P.: Convection in a compressible fluid with infinite Prandtl number, Journal of Fluid Mechanics, 96, 515–583, https://doi.org/10.1017/S002211208000225X, 1980.

Koelemeijer, P., Ritsema, J., Deuss, A., and Van Heijst, H.-J.: SP12RTS: a degree-12 model of shear-and compressional-wave velocity for Earth's mantle, Geophysical Journal International, 204, 1024–1039, https://doi.org/10.1093/gji/ggv481, 2016.

970 Kouri, D. P., Jakeman, J. D., Huerta, J. G., Smith, C. B., Walsh, T. F., Udell, M., and Uryasev, S.: Risk-Adaptive Experimental Design for High-Consequence Systems: LDRD Final Report, Tech. rep., Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), 2021a.

Kouri, D. P., Ridzal, D., and Tuminaro, R.: KKT preconditioners for PDE-constrained optimization with the Helmholtz equation, SIAM Journal on Scientific Computing, 43, S225–S248, 2021b.

Kouri, D. P., Staudigl, M., and Surowiec, T. M.: A relaxation-based probabilistic approach for PDE-constrained optimization under uncertainty with pointwise state constraints, Computational Optimization and Applications, 85, 441–478, 2023.

Kramer, S. C., Wilson, C. R., and Davies, D. R.: An implicit free surface algorithm for geodynamical simulations, Physics of the Earth and Planetary Interiors, 194, 25–37, 2012.

Li, D., Gurnis, M., and Stadler, G.: Towards adjoint-based inversion of time-dependent mantle convection with nonlinear viscosity, Geophysical Journal International, 209, 86–105, https://doi.org/10.1093/gji/ggw493, 2017.

Lin, C.-J. and Moré, J. J.: Newton's method for large bound-constrained optimization problems, SIAM Journal on Optimization, 9, 1100–1127, https://doi.org/10.1137/S1052623498345075, 1999.

Liu, L. and Gurnis, M.: Dynamic subsidence and uplift of the Colorado Plateau, Geology, 38, 663–666, https://doi.org/10.1130/g30624.1, 2010.

Liu, L., Spasojević, S., and Gurnis, M.: Reconstructing Farallon plate subduction beneath North America back to the Late Cretaceous, Science, 322, 934–938, https://doi.org/10.1126/science.1162921, 2008.

Liu, L., Gurnis, M., Seton, M., Saleeby, J., Müller, R. D., and Jackson, J. M.: The role of oceanic plateau subduction in the Laramide orogeny, Nature Geoscience, 3, 353–357, https://doi.org/10.1038/ngeo829, 2010.

Logg, A., Mardal, K.-A., and Wells, G.: Automated solution of differential equations by the finite element method: The FEniCS book, vol. 84, Springer Science & Business Media, https://doi.org/10.1007/978-3-642-23099-8, 2012.

Martinec, Z., Sasgen, I., and Velímský, J.: The forward sensitivity and adjoint-state methods of glacial isostatic adjustment, Geophysical Journal International, 200, 77–105, 2015.

Merdith, A. S., Williams, S. E., Collins, A. S., Tetley, M. G., Mulder, J. A., Blades, M. L., Young, A., Armistead, S. E., Cannon, J., Zahirovic, S., and Müller, R. D.: Extending full-plate tectonic models into deep time: Linking the Neoproterozoic and the Phanerozoic, Earth-Science Reviews, 214, 103 477, https://doi.org/https://doi.org/10.1016/j.earscirev.2020.103477, 2021.

Mitusch, S. K., Funke, S. W., and Dokken, J. S.: dolfin-adjoint 2018.1: automated adjoints for FEniCS and Firedrake, Journal of Open Source Software, 4, 1292, https://doi.org/10.21105/joss.01292, 2019.

Mosca, I., Cobden, L., Deuss, A., Ritsema, J., and Trampert, J.: Seismic and mineralogical structures of the lower mantle from probabilistic tomography, Journal of Geophysical Research, 117, https://doi.org/10.1029/2011jb008851, 2012.

Moses, W. S., Narayanan, S. H. K., Paehler, L., Churavy, V., Schanen, M., Hückelheim, J., Doerfert, J., and Hovland, P.: Scalable automatic differentiation of multiple parallel paradigms through compiler augmentation, in: Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis, SC '22, IEEE Press, 2022.

Moucha, R., Forte, A. M., Mitrovica, J. X., Rowley, D. B., Quéré, S., Simmons, N. A., and Grand, S. P.: Dynamic topography and long-term sea-level variations: There is no such thing as a stable continental platform, Earth and Planetary Science Letters, 271, 101–108, 2008.

Müller, R. D., Seton, M., Zahirovic, S., Williams, S. E., Matthews, K. J., Wright, N. M., Shephard, G. E., Maloney, K. T., Barnett-Moore, N., Hosseinpour, M., Bower, D. J., and Cannon, J.: Ocean Basin Evolution and Global-Scale Plate Reorganization Events Since Pangea Breakup, Annual Review of Earth and Planetary Sciences, 44, 107–138, https://doi.org/10.1146/annurev-earth-060115-012211, 2016.

Müller, R. D., Cannon, J., Qin, X., Watson, R. J., Gurnis, M., Williams, S., Pfaffelmoser, T., Seton, M., Russell, S. H. J., and Zahirovic, S.: GPlates: Building a Virtual Earth Through Deep Time, Geochemistry, Geophysics, Geosystems, 19, 2243–2261, https://doi.org/10.1029/2018GC007584, 2018.

1010   Müller, R. D., Zahirovic, S., Williams, S. E., Cannon, J., Seton, M., Bower, D. J., Tetley, M. G., Heine, C., Le Breton, E., Liu, S., et al.: A global plate model including lithospheric deformation along major rifts and orogens since the Triassic, Tectonics, 38, 1884–1907, https://doi.org/10.1029/2018TC005462, 2019.

Müller, R. D., Flament, N., Cannon, J., Tetley, M. G., Williams, S. E., Cao, X., Bodur, O. F., Zahirovic, S., and Merdith, A.: A tectonic-rules-based mantle reference frame since 1 billion years ago – implications for supercontinent cycles and plate–mantle system evolution, Solid

1015   Earth, 13, 1127–1159, https://doi.org/10.5194/se-13-1127-2022, 2022.

Müller, R. D., Cannon, J., Qin, X., Watson, R. J., Gurnis, M., Williams, S., Pfaffelmoser, T., Seton, M., Russell, S. H. J., and Zahirovic, S.: GPlates: Building a Virtual Earth Through Deep Time, Geochemistry, Geophysics, Geosystems, 19, 2243–2261, https://doi.org/https://doi.org/10.1029/2018GC007584, 2018.

Naumann, U.: The Art of Differentiating Computer Programs, Society for Industrial and Applied Mathematics,

1020   https://doi.org/10.1137/1.9781611972078, 2011.

Nerlich, R., Colli, L., Ghelichkhan, S., Schuberth, B., and Bunge, H.-P.: Constraining central Neo-Tethys Ocean reconstructions with mantle convection models, Geophysical Research Letters, 43, 9595–9603, https://doi.org/10.1002/2016gl070524, 2016.

Nitsche, J.: Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind, in: Abhandlungen aus dem mathematischen Seminar der Universität Hamburg, vol. 36, pp. 9–15, Springer, 1971.

1025   Nixon-Hill, R. W., Shapero, D., Cotter, C. J., and Ham, D. A.: Consistent Point Data Assimilation in Firedrake and Icepack, 2023.

Nocedal, J. and Wright, S. J.: Numerical optimization, Springer, 1999.

Nocedal, J. and Wright, S. J.: Quadratic programming, Numerical optimization, pp. 448–492, 2006.

Panasyuk, S. V. and Hager, B. H.: Inversion for mantle viscosity profiles constrained by dynamic topography and the geoid, and their estimated errors, Geophysical Journal International, 143, 821–836, 2000.

1030   Panton, J., Davies, J. H., and Myhill, R.: The Stability of Dense Oceanic Crust Near the Core-Mantle Boundary, Journal of Geophysical Research: Solid Earth, 128, e2022JB025 610, https://doi.org/https://doi.org/10.1029/2022JB025610, 2023.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., and Chintala, S.: PyTorch: An Imperative Style, High-Performance Deep Learning Library, in: Advances in Neural Information Processing Systems 32, pp. 8024–8035, Curran

1035   Associates, Inc., http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf, 2019.

Plessix, R.-E.: A review of the adjoint-state method for computing the gradient of a functional with geophysical applications, Geophysical Journal International, 167, 495–503, 2006.

Poore, H. R., Samworth, R., White, N. J., Jones, S. M., and McCave, I. N.: Neogene overflow of Northern Component Water at the Greenland-Scotland Ridge, Geochemistry, Geophysics, Geosystems, 7, https://doi.org/https://doi.org/10.1029/2005GC001085, 2006.

1040   Price, M. G. and Davies, J. H.: Profiling the robustness, efficiency and limits of the forward-adjoint method for 3D mantle convection modelling, Geophysical Journal International, 212, 1450–1462, https://doi.org/10.1093/gji/ggx489, 2018.

Rathgeber, F., Ham, D. A., Mitchell, L., Lange, M., Luporini, F., Mcrae, A. T. T., Bercea, G.-T., Markall, G. R., and Kelly, P. H. J.: Firedrake: Automating the Finite Element Method by Composing Abstractions, ACT Trans. Math. Softw., 43, 1–24, https://doi.org/10.1145/2998441, 2016.

1045   Ratnaswamy, V., Stadler, G., and Gurnis, M.: Adjoint-based estimation of plate coupling in a non-linear mantle flow model: theory and examples, Geophysical Journal International, 202, 768–786, 2015.

Rawlinson, N., Pozgay, S., and Fishwick, S.: Seismic tomography: a window into deep Earth, Physics of the Earth and Planetary Interiors, 178, 101–135, https://doi.org/10.1016/j.pepi.2009.10.002, 2010.

Reuber, G. S., Holbach, L., Popov, A. A., Hanke, M., and Kaus, B. J.: Inferring rheology and geometry of subsurface structures by adjoint-based inversion of principal stress directions, Geophysical Journal International, 223, 851–861, https://doi.org/10.1093/gji/ggaa344, 2020.

Ricard, Y.: Physics of mantle convection, Treatise on geophysics, 7, 31–87, 2007.

Sabadini, R. and Vermeersen, L.: Influence of lithospheric and mantle stratification on global post-seismic deformation, Geophysical research letters, 24, 2075–2078, 1997.

Sambridge, M. and Mosegaard, K.: Monte Carlo methods in geophysical inverse problems, Reviews of Geophysics, 40, 3–1, https://doi.org/10.1029/2000rg000089, 2002.

Schwedes, T., Funke, S. W., and Ham, D. A.: An iteration count estimate for a mesh-dependent steepest descent method based on finite elements and Riesz inner product representation, arXiv preprint arXiv:1606.08069, 2016.

Schwedes, T., Ham, D. A., Funke, S. W., Piggott, M. D., Schwedes, T., Ham, D. A., Funke, S. W., and Piggott, M. D.: Mesh dependence in PDE-constrained optimisation, Springer, https://doi.org/10.1007/978-3-319-59483-5_2, 2017.

Seton, M., Müller, R. D., Zahirovic, S., Gaina, C., Torsvik, T. H., Shephard, G., Talsma, A., Gurnis, M., Turner, M., Maus, S., and Chandler, M.: Global continental and ocean basin reconstructions since 200 Ma, Earth-Science Reviews, 113, 212–270, https://doi.org/10.1016/j.earscirev.2012.03.002, 2012.

Seton, M., Müller, R. D., Zahirovic, S., Williams, S., Wright, N. M., Cannon, J., Whittaker, J. M., Matthews, K. J., and McGirr, R.: A Global Data Set of Present-Day Oceanic Crustal Age and Seafloor Spreading Parameters, Geochemistry, Geophysics, Geosystems, 21, e2020GC009 214, https://doi.org/https://doi.org/10.1029/2020GC009214, e2020GC009214 2020GC009214, 2020.

Shephard, G., Müller, R., Liu, L., and Gurnis, M.: Miocene drainage reversal of the Amazon River driven by plate–mantle interaction, Nature Geoscience, 3, 870–875, 2010.

Shephard, G., Bunge, H.-P., Schuberth, B., Müller, R., Talsma, A., Moder, C., and Landgrebe, T.: Testing absolute plate reference frames and the implications for the generation of geodynamic mantle heterogeneity structure, Earth and Planetary Science Letters, 317-318, 204–217, https://doi.org/https://doi.org/10.1016/j.epsl.2011.11.027, 2012.

Simmons, N. A., Myers, S. C., Johannesson, G., Matzel, E., and Grand, S. P.: Evidence for long-lived subduction of an ancient tectonic plate beneath the southern Indian Ocean, Geophysical Research Letters, 42, 9270–9278, https://doi.org/10.1002/2015gl066237, 2015.

Spasojevic, S., Liu, L., and Gurnis, M.: Adjoint models of mantle convection with seismic, plate motion, and stratigraphic constraints: North America since the Late Cretaceous, Geochemistry, Geophysics, Geosystems, 10, https://doi.org/10.1029/2008gc002345, 2009.

Stadler, G., Gurnis, M., Burstedde, C., Wilcox, L. C., Alisic, L., and Ghattas, O.: The Dynamics of Plate Tectonics and Mantle Flow: From Local to Global Scales, Science, 329, 1033–1038, https://doi.org/10.1126/science.1191223, 2010.

Stixrude, L. and Lithgow-Bertelloni, C.: Thermodynamics of mantle minerals - II. Phase equilibria, Geophysical Journal International, 184, 1180–1213, https://doi.org/10.1111/j.1365-246X.2010.04890.x, 2011.

Stotz, I., Iaffaldano, G., and Davies, D. R.: Pressure-driven Poiseuille flow: A major component of the torque-balance governing Pacific plate motion, Geophysical Research Letters, 45, 117–125, 2018.

Stotz, I. L., Iaffaldano, G., and Davies, D. R.: Late Miocene Pacific plate kinematic change explained with coupled global models of mantle and lithosphere dynamics, Geophysical Research Letters, 44, 7177–7186, https://doi.org/10.1002/2017GL073920, 2017.

Styles, E., Davies, D. R., and Goes, S.: Mapping spherical seismic into physical structure: biases from 3-D phase-transition and thermal boundary-layer heterogeneity, Geophysical Journal International, 184, 1371–1378, https://doi.org/10.1111/j.1365-246x.2010.04914.x, 2011.

Taiwo, A., Bunge, H.-P., Schuberth, B. S. A., Colli, L., and Vilacis, B.: Robust global mantle flow trajectories and their validation via dynamic topography histories, Geophysical Journal International, 234, 2160–2179, https://doi.org/10.1093/gji/ggad188, 2023.

Talagrand, O.: Assimilation of observations, an introduction (gtspecial issueltdata assimilation in meteology and oceanography: Theory and practice), Journal of the Meteorological Society of Japan. Ser. II, 75, 191–209, https://doi.org/10.2151/jmsj1965.75.1B_191, 1997.

Tetley, M. G., Williams, S. E., Gurnis, M., Flament, N., and Müller, R. D.: Constraining Absolute Plate Motions Since the Triassic, Journal of Geophysical Research: Solid Earth, 124, https://doi.org/10.1029/2019JB017442, 2019.

The ROL Project Team: The ROL Project Website, https://trilinos.github.io/rol.html, 2022.

Tijskens, E., Roose, D., Ramon, H., and De Baerdemaeker, J.: Automatic differentiation for solving nonlinear partial differential equations: an efficient operator overloading approach, Numerical Algorithms, 30, 259–301, 2002.

Tikhonov, A. N.: Solution of incorrectly formulated problems and the regularization method., Sov Dok, 4, 1035–1038, 1963.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., et al.: SciPy 1.0: fundamental algorithms for scientific computing in Python, Nature methods, 17, 261–272, https://doi.org/10.1038/s41592-019-0686-2, 2020.

Vynnytska, L. and Bunge, H.-P.: Restoring past mantle convection structure through fluid dynamic inverse theory: Regularisation through surface velocity boundary conditions, GEM-International Journal on Geomathematics, 6, 83–100, https://doi.org/10.1007/s13137-014-0060-6, 2015.

Wang, Z. R., Stotz, I. L., Bunge, H.-P., Vilacís, B., Hayek, J. N., Ghelichkhan, S., and Lebedev, S.: Cenozoic upper mantle flow history of the Atlantic realm based on Couette/Poiseuille models: Towards paleo-mantle-flowgraphy, Physics of the Earth and Planetary Interiors, 340, 107 045, 2023.

Weismüller, J., Gmeiner, B., Ghelichkhan, S., Huber, M., John, L., Wohlmuth, B., Rüde, U., and Bunge, H.-P.: Fast asthenosphere motion in high-resolution global mantle flow models, Geophysical Research Letters, 42, 7429–7435, https://doi.org/10.1002/2015GL063727, 2015.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., et al.: The FAIR Guiding Principles for scientific data management and stewardship, Scientific data, 3, 1–9, 2016.

Williams, S., Flament, N., Müller, R. D., and Butterworth, N.: Absolute plate motions since 130 Ma constrained by subduction zone kinematics, Earth and Planetary Science Letters, 418, 66–77, 2015.

Wolstencroft, M., Davies, J. H., and Davies, D. R.: Nusselt–Rayleigh number scaling for spherical shell Earth mantle simulation up to a Rayleigh number of 109, Physics of the Earth and Planetary Interiors, 176, 132–141, 2009.

Wunsch, C.: The Ocean Circulation Inverse Problem, Cambridge University Press, Cambridge, https://doi.org/10.1017/CBO9780511629570, 1996.

Xu, K. and Darve, E.: Physics constrained learning for data-driven inverse modeling from sparse observations, Journal of Computational Physics, 453, 110 938, 2022.

Young, A., Flament, N., Williams, S. E., Merdith, A., Cao, X., and Müller, R. D.: Long-term Phanerozoic sea level change from solid Earth processes, Earth and Planetary Science Letters, 584, 117 451, https://doi.org/10.1016/j.epsl.2022.117451, 2022.

Zaroli, C., Sambridge, M., Lévêque, J.-J., Debayle, E., and Nolet, G.: An objective rationale for the choice of regularisation parameter with application to global multiple-frequency S-wave tomography, Solid Earth, 4, 357–371, https://doi.org/10.5194/se-4-357-2013, 2013.

zenodo/Firedrake-20231027.0: Software used in 'Software for g-adopt v2.1.0', https://doi.org/10.5281/zenodo.10047031, 2023.

Zhang, H., Constantinescu, E. M., and Smith, B. F.: PETSc TSAdjoint: a discrete adjoint ODE solver for first-order and second-order sensitivity analysis, SIAM Journal on Scientific Computing, 44, C1–C24, 2022.

Zhong, S. and Rudolph, M. L.: On the temporal evolution of long-wavelength mantle structure of the Earth since the early Paleozoic, Geochemistry, Geophysics, Geosystems, 16, 1599–1615, https://doi.org/https://doi.org/10.1002/2015GC005782, 2015.

Zhong, S., Yuen, D. A., Moresi, L. N., and Knepley, M.: Numerical methods for mantle convection, Treatise on geophysics, 7, 227–252, 2007.

Zhou, Q. and Liu, L.: A hybrid approach to data assimilation for reconstructing the evolution of mantle dynamics, Geochemistry, Geophysics, Geosystems, 18, 3854–3868, 2017.

1125
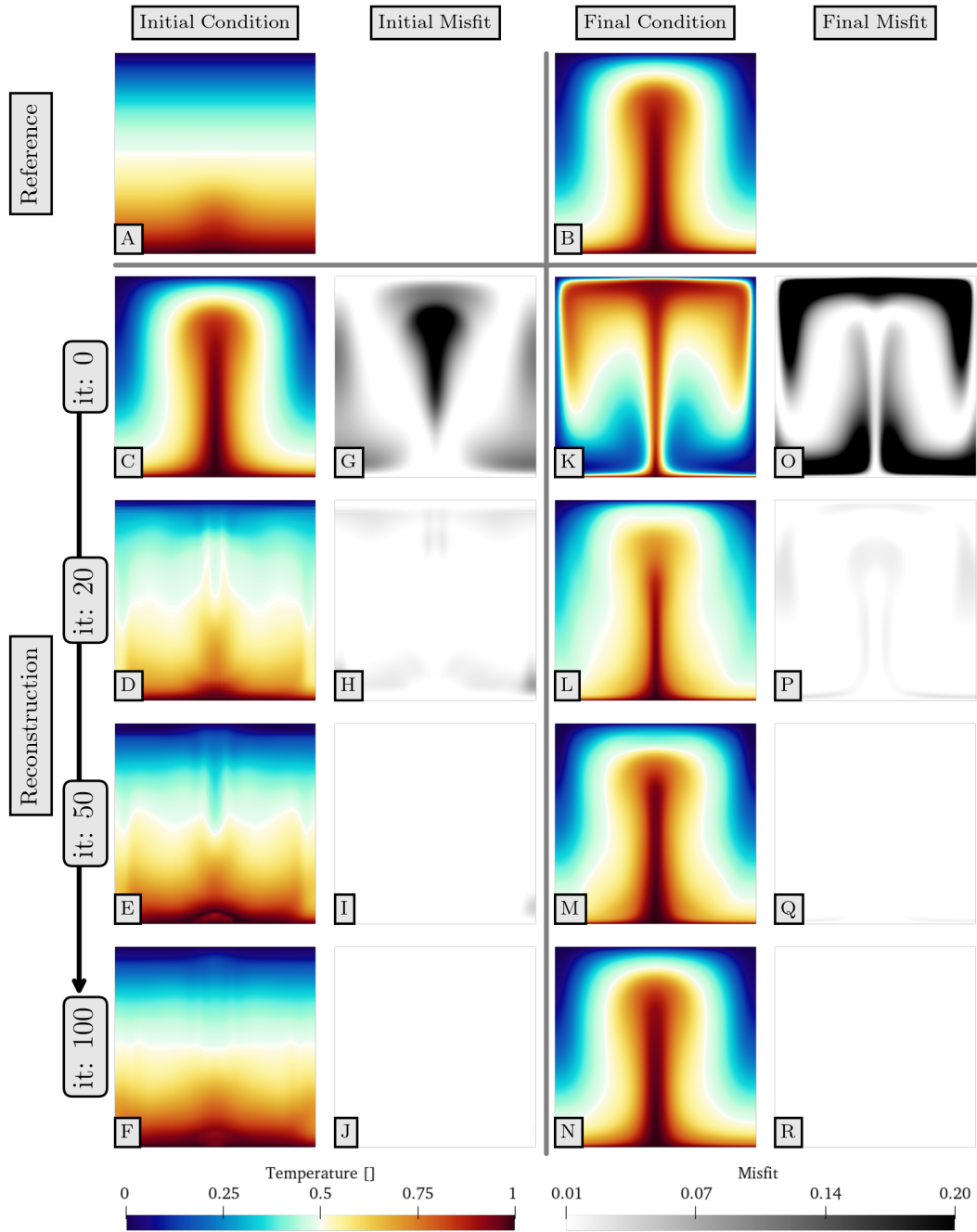
Figure 6. Iterative optimisation process visualised: A and B depict the reference initial and final conditions. C-F present the reconstructed initial conditions at the 0th, 20th, 50th, and 100th iterations. G-J highlight the misfits, representing the squared differences between the reconstructed initial temperature fields and the reference temperature. Similarly, K-N display the reconstructed final temperature fields after the 0th, 20th, 50th, and 100th iterations, with their respective misfits demonstrated in O-R.
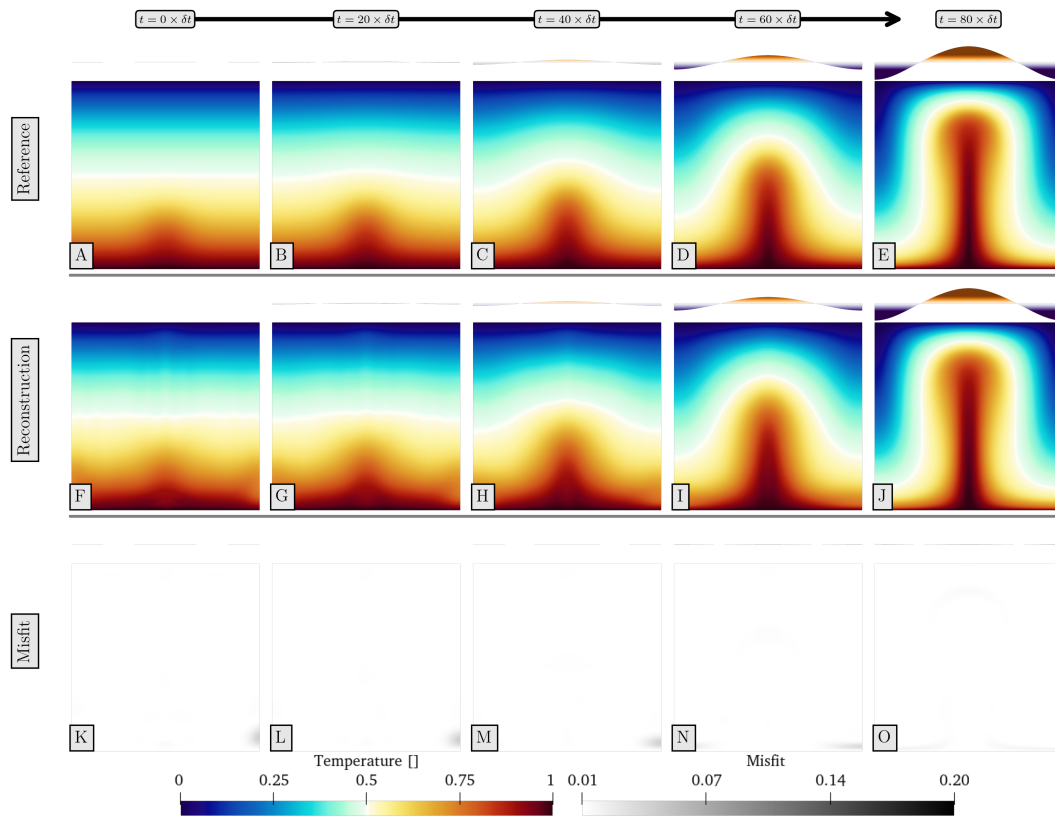
Figure 7. Comparison of reference forward and reconstructed simulations over time: A-E present the evolution in the reference forward, while F-J depict the evolution in the reconstructed simulations. The misfits between the reference and reconstructed scenarios at each time step are illustrated in K-O. Note that surface dynamic topography is represented by visualising the normal stresses at the top boundary.
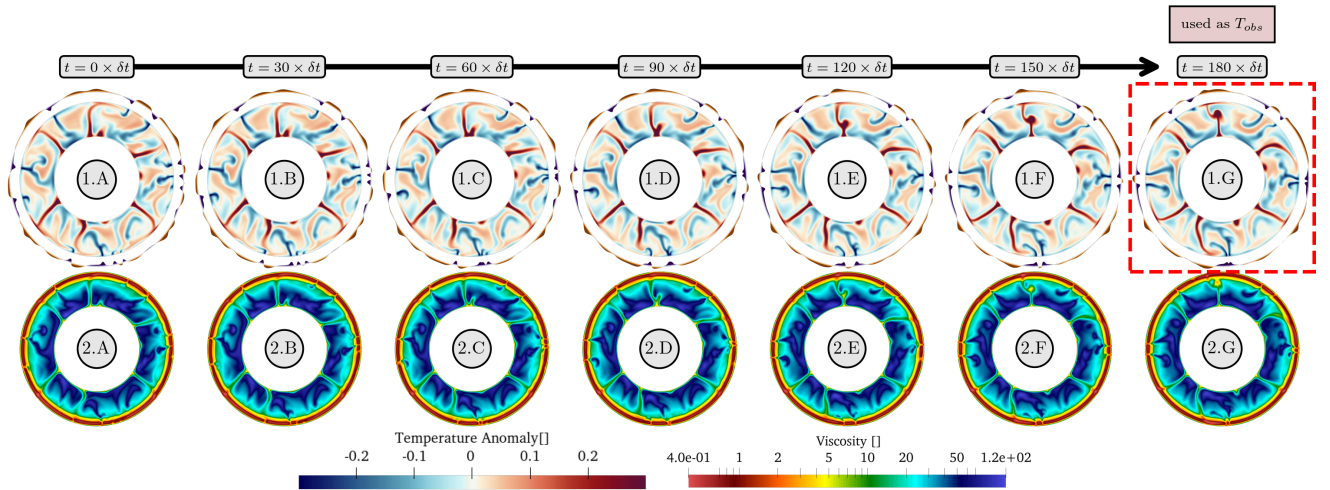
Figure 8. Reference forward simulation spanning a duration of $t = 200 \times (5 \times 10^{-6})$, as depicted in 1.A to 1.G. Figs. 1 (upper panel) detail the temporal evolution of the reference temperature field, while Figs. 2 (lower panel) show the viscosity field at each time. The viscosity demonstrates a variation of nearly three orders of magnitude: approximately 1.0 in the asthenosphere, 140 within colder lower mantle slabs, and 0.4 in convergent regions exhibiting elevated strain-rates.
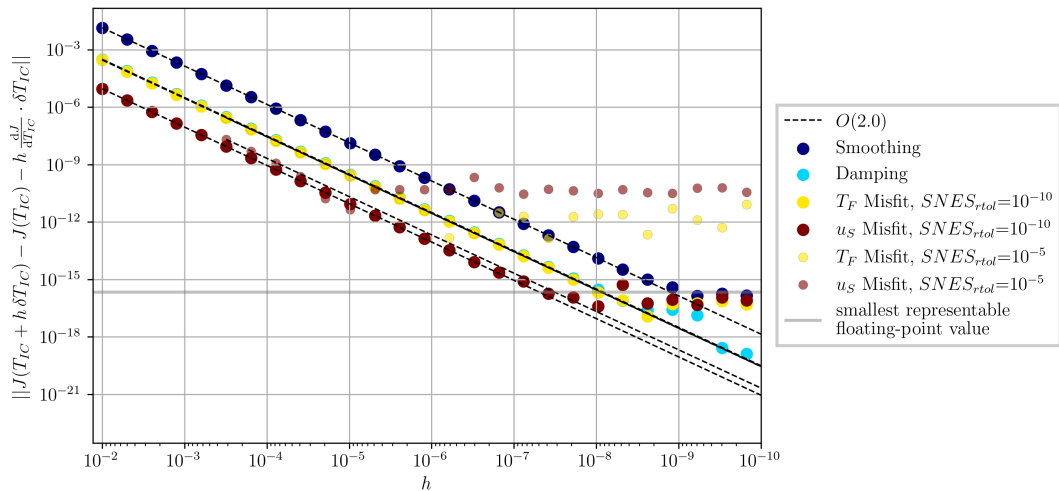
Figure 9. Second-order Taylor remainder test for convection with a temperature, depth and stress-dependent rheology. The semi-transparent markers are cases where the relative solver tolerance of Newton solves (SNES) is set to 10^{-5}. With such tolerances, perturbations beyond solver tolerance exhibit divergence from the $O(h^2)$ trend. When setting this tolerance to 10^{-10} (larger solid markers), residuals show accuracy of gradient information down to floating-point precision.
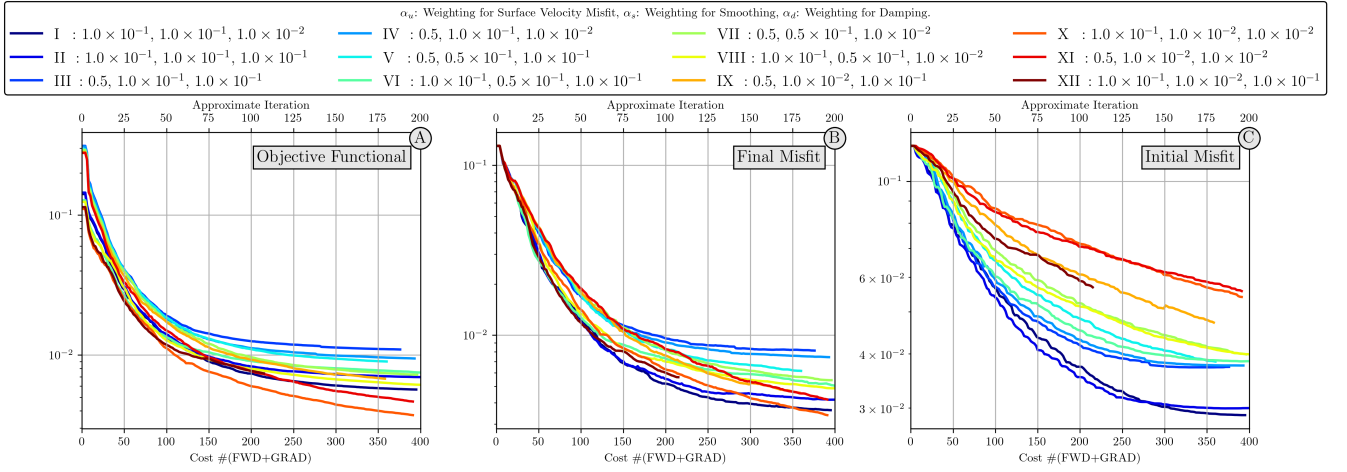
Figure 10. Summary of optimisation outcomes from 2-D annulus experiments. (A) visualises the process of objective functional minimisation. (B) illustrates the final misfit, representing the misfit between the reconstructed final temperature field and T_{obs}. (C) depicts the initial misfit, indicating the difference between the reconstructed initial condition T_{IC} and the reference initial condition. Notably, despite significant reduction in the objective functional and final misfit in simulations XI and XII, these simulations do not perform as well in terms of the initial misfit, which is a key measure in our experiments.
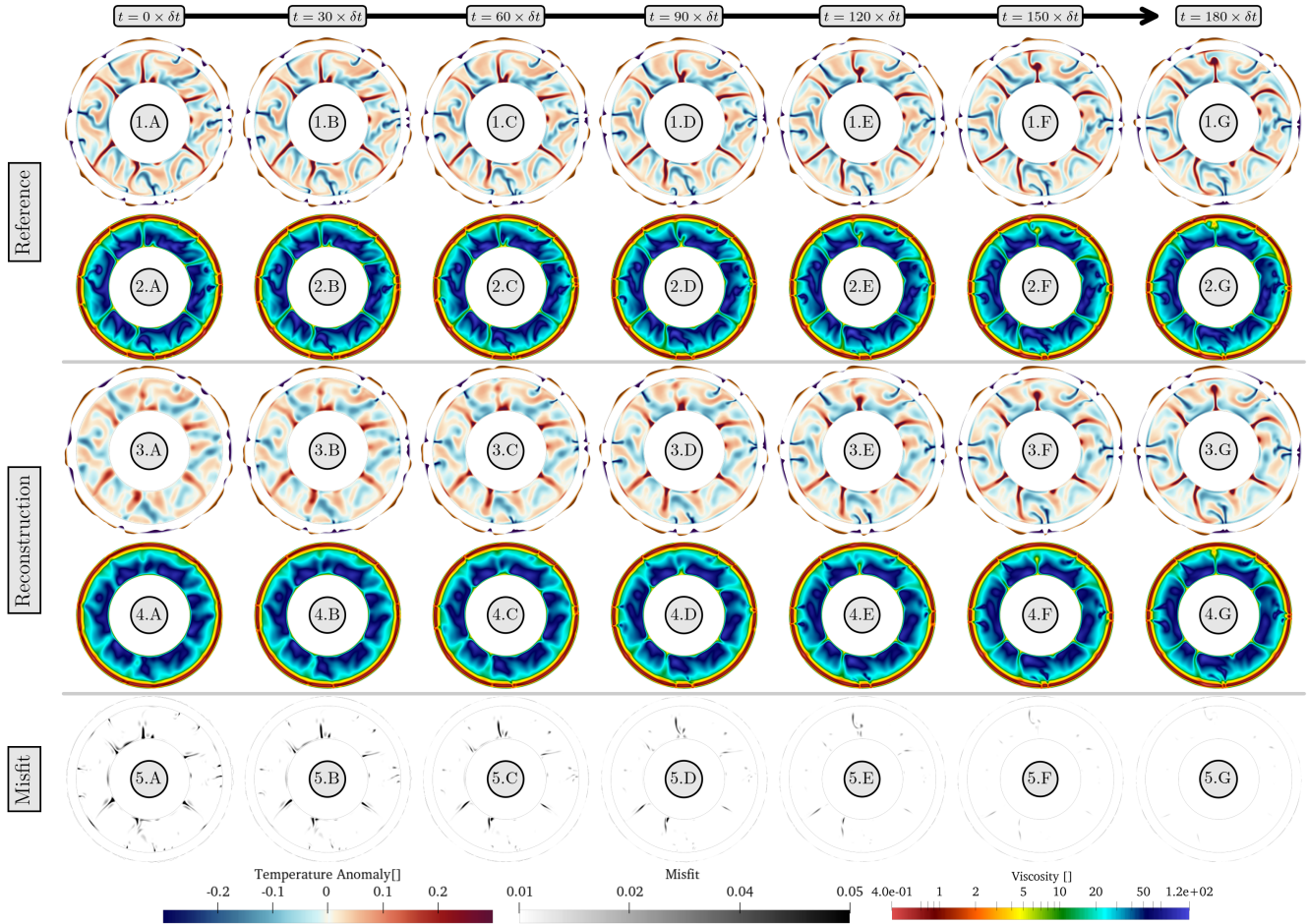
Figure 11. Comparing the reference (1 and 2) and the best reconstruction simulation (3 and 4). Temperature and viscosity fields are shown in panels 1,3 and 2,4 respectively. The misfit, which is the squared difference between the reconstructed and reference temperatures, is shown in panel 5. To highlight the effectiveness of the reconstruction of the evolution of surface dynamic topography, a field representing the normal stresses acting on the top boundary is visualised alongside the temperature fields in 1 and 3. The reconstruction simulation employs values of $\alpha_u = 10^{-1}$, $\alpha_s = 10^{-1}$ and $\alpha_d = 10^{-2}$, and is marked with I in Fig. 10.