Forward and adjoint quasi-geostrophic models of the geomagnetic secular variation

Elisabeth Canet, Alexandre Fournier, Dominique Jault

Forward and adjoint quasi-geostrophic models of the geomagnetic secular variation

Elisabeth Canet[1]*, Alexandre Fournier[2], Dominique Jault[1]

[1] Laboratoire de Géophysique Interne et Tectonophysique, Université Joseph-Fourier, CNRS, Grenoble, France

[2] Équipe de Géomagnétisme, Institut de Physique du Globe de Paris Université Paris Diderot, INSU/CNRS, Paris, France

* Corresponding author: Elisabeth.Canet@obs.ujf-grenoble.fr

Abstract

We introduce a quasi-geostrophic model of core dynamics, which aims at describing core processes on geomagnetic secular variation timescales. It extends the formalism of Alfvén torsional oscillations by incorporating non-zonal motions. Within this framework, the magnetohydrodynamics takes place in the equatorial plane; it involves quadratic magnetic quantities, which are averaged along the direction of rotation of the Earth. In addition, the equatorial flow is projected on the core-mantle boundary. It interacts with the magnetic field at the core surface, through the radial component of the magnetic induction equation. That part of the model connects the dynamics and the observed secular variation, with the radial component of the magnetic field acting as a passive tracer. We resort to variational data assimilation to construct formally the relationship between model predictions and observations. Variational data assimilation seeks to minimize an objective function, by computing its sensitivity to its control variables. The sensitivity is efficiently calculated after integration of the adjoint model. We illustrate that framework with twin experiments, performed first in the case of the kinematic core flow inverse problem, and then in the case of Alfvén torsional oscillations. In both cases, using the adjoint model allows us to retrieve core state variables which, while taking part in the dynamics, are not directly sampled at the core surface. We study the effect of several factors on the solution (width of the assimilation time window, amount and quality of data), and we discuss the potential of the model to deal with real geomagnetic observations.

1 Introduction

Current descriptions of core dynamics rely on two sources of information: observations of the magnetic field, and physical laws governing the evolution of the state of the core. The Earth's magnetic field is assumed to have an internal origin through the process of

geodynamo; it is generated and sustained by fluid motions in the metallic liquid outer core, and varies on a wide range of time scales reflecting the various time and space scales of core magnetohydrodynamics.

The quality of observations of the Earth's magnetic field has much improved since the set-up of the first network of magnetic observatories by Gauss and co-workers in 1834, which was followed by the large increase in the number of observatories at the beginning of the twentieth century. Other turning points have occurred since: the introduction of the proton precession magnetometer, the development of declination/inclination magnetometers (DIflux) widely used in observatories by the 1970's, and finally the rise of the Intermagnet network of digital observatories sharing modern measurement practices after 1990 (see, e.g., the review by Turner et al., 2007). The good temporal coverage of observatory data has now been supplemented by the excellent spatial coverage of satellite data. Following the launch of three low Earth orbiting satellites -Oersted, CHAMP and SAC-C- supplying geomagnetic data, a continuous satellite time series extends now to 10 years.

The magnetic field can be downward continued throughout the solid mantle to the fluid core surface. Field models are built, describing the radial component of the main field and its time evolution at the core-mantle boundary (CMB) (Hulot et al., 2007; Jackson and Finlay, 2007). Calculating the geomagnetic secular variation, the first time derivative of the main field time series, emphasizes rapid changes of the magnetic field, on characteristic time scales ranging from years to centuries. Inversions of a snapshot of the geomagnetic secular variation can be performed using the radial induction equation at the CMB, in order to retrieve the large-scale part of the flow beneath it (Eymin and Hulot, 2005; Holme and Olsen, 2006; Pais and Jault, 2008; Olsen and Mandea, 2008). The root mean square (rms) speed of these flows is typically of the order on 15 km/y. Such inversions, however, face non-uniqueness problems (Backus, 1968). Further assumptions are thus required to remove the non-uniqueness, and a great part of the work consists in finding constraints and regularizations to specify the flow (Holme, 2007, section 8.04.2). Alongside these kinematic inversions, numerical models of the geodynamo have been available for more than 10 years, since the pioneering work of Glatzmaier and Roberts (1995). The magnetic field generated by those dynamical models explains features of the Earth's magnetic field (dipolar geometry, spatial spectrum); yet their parameters are far from those of the Earth's core (Christensen and Wicht, 2007, section 8.08.4). Rau et al. (2000) and Amit et al. (2007) tried to connect these two approaches (core flow inversion and forward numerical modelling) when inverting synthetic data from dynamo models. They found their core flow inversion method and the additional regularization to be adequate for the retrieval of large-scale flow and magnetic field patterns.

A quality-control of core flow models is the angular momentum they carry (Jault et al., 1988). Comparison of these estimates with core angular momentum changes inferred from decadal length-of-day variations is encouraging, yet discrepancies remain. Angular momentum series are also derived from atmospheric and oceanic flow models, based upon the data assimilation methodology (Kalnay et al., 1996). Those variations of angular momentum account very well for the observed seasonal and interannual changes in length-of-day (Chen, 2005; Gross, 2007, section 3.09.4).

Data assimilation, routinely used in atmospheric science and more recently in oceanography, is now in early stages of use in the field of core physics. Applied to the core, this technique should allow us to interpret the secular variation in terms of dynamics, thereby enlarging the work done on kinematic core flow inversion. Resorting to a toy model, Fournier et al. (2007) assimilated synthetic data in a one-dimensional model that retains characteristic features of the induction and Navier-Stokes equations. They concluded that a good knowledge of the observed magnetic field can be translated into a good knowledge of core flow, through the process of data assimilation, which takes explicitly into account the dynamical relationship that exists between magnetic and velocity fields. This conclusion was also drawn by Sun et al. (2007), using a much similar toy model and a different implementation of data assimilation (sequential as opposed to variational). In a preliminary study, Liu et al. (2007) applied an optimal interpolation scheme to a three-dimensional dynamo model, using a synthetic set of observations. They showed in particular that assimilation of observations could partially alleviate the negative impact of wrong model parameter values. Still, the values of the parameters used typically in that class of simulations are far from being Earth-like, due to the numerical cost of their integration. There is hope, however, that systematic and appropriate exploration of the parameter space of those models will eventually yield scaling laws of the kind proposed by Christensen and Aubert (2006), which will ultimately permit a reliable extrapolation between their output and the observed secular variation. In the context of geomagnetic data assimilation, the current situation is even worse, though, since one assimilation run requires several tens of forward realisations.

A solution to this conundrum is to construct a simplified dynamical model, tailored to the study of the secular variation. In this paper, we introduce a model which relies on quasi-geostrophic dynamics. As a matter of fact, on rapid time scales as those characterizing the secular variation, rotational forces prevail on magnetic forces in the bulk of the fluid (Jault, 2008). The resulting two-dimensional flow interacts with the radial component of the magnetic field (in that instance, a passive tracer) at the core-mantle boundary. The quasi-geostrophic assumption has recently been used in the framework of kinematic core flow inversions (Pais and Jault, 2008; Gillet et al., 2009). It also provides us with the tools necessary to build a dynamical model of the geomagnetic secular variation.

Being quasi-geostrophic, this model comprises the equations for torsional Alfvén waves, for which the dynamics is axisymmetric; Alfvén torsional waves are associated with geostrophic (zonal) motions in the core (Braginsky, 1970). The frequency of these oscillations is proportional to the rms strength of the magnetic field B_s perpendicular to the rotation axis (s denotes the cylindrical radius). Accordingly, using a database of computed core flows, Zatman and Bloxham (1997, 1999) and Buffett et al. (2009) have calculated radial profiles of the quadratic cylindrical radial component of the magnetic field averaged on geostrophic cylinders, $\{B_s^2\}$, within the core.

Our quasi-geostrophic model generalizes that axisymmetric approach by adding non-zonal motion and magnetic field. Theoretical solutions of the model include various families of diffusionless hydromagnetic waves, some of which were first studied by Hide (1966) in order to explain the observed secular variation.
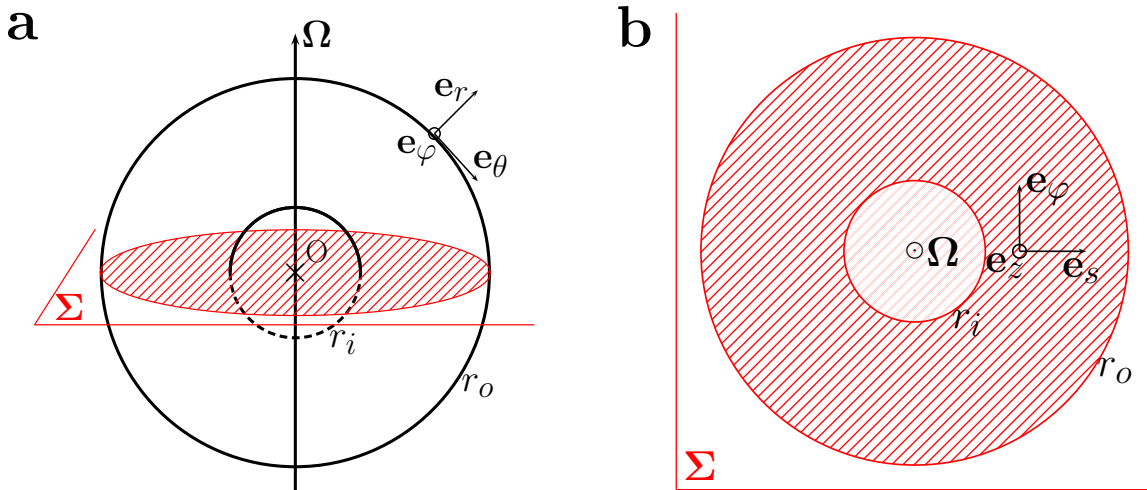
3

Figure 1. Geometry of the system and notations; a: Side view, b: Equatorial section. Σ is the equatorial plane and the CMB is the outer sphere, located at $r = r_o$. r_i is the radius of the inner core. Spherical (O, $\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\varphi$) and cylindrical (O, $\mathbf{e}_s, \mathbf{e}_\varphi, \mathbf{e}_z$) coordinate system bases are defined at the core surface and in Σ respectively.

The goal of this paper is thus to describe a quasi-geostrophic forward model of the Earth's core fast dynamics, and to place it at the heart of a geomagnetic data assimilation process. In section 2, we derive that quasi-geostrophic model, along with its link to the observations at the CMB. Variational data assimilation is introduced in section 3 and its principles are illustrated in section 4 with twin experiments. That section begins with the study of the classical kinematic inversion of a steady core flow, set within that framework. A second illustration is dedicated to the retrieval of the magnetic field sheared by Alfvén torsional waves. Results are summarized and discussed in section 5.

2   Quasi-geostrophic forward model

We shall model the Earth's outer core as a spherical fluid shell of inner radius r_i and outer radius r_o. The fluid has density ρ, and it is electrically conducting. Its magnetic diffusivity is η. The system is rotating at angular velocity Ω around the z-axis. Figure 1 sketches the geometry and summarizes the notations.

The fluid is characterized by its magnetic field $\mathbf{B}$, its velocity $\mathbf{u}$, and the reduced pressure Π, that includes pressure and the centrifugal potential. We choose r_o as length scale and B_0, a typical magnetic field intensity in the core interior, as the magnetic field scale. Velocities are scaled with the Alfvén waves speed

$$V_A = \frac{B_0}{\sqrt{\mu_0\rho}}, \tag{1}$$

in which μ_0 is the magnetic permeability of free space. The pressure scale is ρV_A^2, and the

4

time scale is the Alfvén waves period: $T_A = r_o/V_A$.

The evolution of the magnetic field in the core is governed by the induction equation, which has the dimensionless form,

$$\partial_t \mathbf{B} = \boldsymbol{\nabla} \times (\mathbf{u} \times \mathbf{B}) + S^{-1} \nabla^2 \mathbf{B}. \tag{2}$$

The Lundquist number S characterizes the ratio between the magnetic diffusion time and the period of Alfvén waves (e.g. Roberts, 1967),

$$S = \frac{r_o V_A}{\eta}. \tag{3}$$

For the Earth's core, S is of the order of 10^4 to 5×10^4 (Jault, 2008). On secular variation time scales, diffusion becomes negligible compared to induction, hence the large value of S, which yields the frozen-flux approximation.

Assuming that the mantle is electrically insulating on secular variation time scales, the magnetic field can be downward continued to the core-mantle boundary. In the case of a perfectly conducting fluid (the frozen-flux limit), the radial component of the magnetic field B_r is the sole magnetic component continuous across the spherical CMB. At the top of the core, B_r interacts with core motions by means of the radial component of the diffusionless version of equation (2) at $r = r_o$,

$$\partial_t B_r = -\nabla_H \cdot (\mathbf{u} B_r), \tag{4}$$

with the horizontal divergence operator $\nabla_H \cdot$ defined as

$$\nabla_H \cdot \mathbf{v} = (\sin \theta)^{-1} \partial_\theta (\sin \theta v_\theta) + (\sin \theta)^{-1} \partial_\varphi v_\varphi, \tag{5}$$

where (r, θ, φ) are the spherical coordinates. It is that equation at the core-mantle boundary that connects our model to the observations. The time-varying B_r acts as a passive tracer (a drifting buoy), because it interacts with the velocity field at the core surface and does not affect the dynamics that sets up in the interior of the core (see below).

On secular variation time scales rotation forces are much larger than magnetic forces in the bulk of the fluid. Jault (2008) suggests that rapidly rotating motions of lengthscale l are axially invariant if the non-dimensional Lehnert number, λ_l, is small enough. That number measures the ratio between the period of inertial waves, $1/\Omega$, and the period of Alfvén waves, l/V_A (Lehnert, 1954):

$$\lambda_l = \frac{B_0}{\Omega(\mu_0 \rho)^{1/2} l} \quad ; \quad \lambda_{r_o} = \frac{1}{\Omega T_A}, \tag{6}$$

Note that λ_l is a decreasing function of l. In his calculations, the flow appears to be invariant in the direction parallel to the rotation axis, provided $\lambda_l \ll 1$. For the Earth's core, with B_0 of the order of 2 mT (Christensen et al., 2009) and $l \approx 10^6$ m, $\lambda_l \approx 10^{-4}$. Therefore, we shall assume that the flow is geostrophic at leading order. Working in the

equatorial plane Σ (crosshatched in Figure 1), a cylindrical set of coordinates (s, φ, z), with $\mathbf{e}_z$ parallel to the axis of rotation, is well-suited to study the resulting columnar patterns.

The main force balance involves the Coriolis force and the pressure gradient

$$2\mathbf{e}_z \times \mathbf{u}^0 = -\nabla\Pi^0, \tag{7}$$

where the superscript 0 denotes the main order. Taking the curl of equation (7) yields the Proudman-Taylor theorem, namely the z-invariance of the flow.

Within a spherical container, $\mathbf{u}^0$ does not satisfy the non-penetration boundary condition at the CMB, except if it consists of cylindrical flows organized around the rotation axis. Thus we have to add the first-order contribution in λ_{r_o} of the Coriolis force, leading to

$$\mathrm{D}_t\mathbf{u}^0 + 2\lambda_{r_o}^{-1}\mathbf{e}_z \times \mathbf{u}^1 = -\boldsymbol{\nabla}\Pi^1 + (\boldsymbol{\nabla} \times \mathbf{B}) \times \mathbf{B}, \tag{8}$$

where D_t denotes the material derivative $\mathrm{D}_t = \partial_t + (\mathbf{u}^0 \cdot \boldsymbol{\nabla})$. At first-order, magnetic forces are a natural candidate to trigger a departure from geostrophy, since magnetic energy is large compared to kinetic energy in Earth's core. Buoyancy forces are another candidate that we could additionally take into account, which we discard for now for the sake of simplicity. Viscous forces are neglected, while equation (8) shows that the Coriolis force is scaled with the inverse of the Lehnert number λ_{r_o}. The non-penetration boundary condition at the CMB: $\mathbf{u}^1 \cdot \mathbf{e}_r = 0$ at $z = \pm h$, yields a linear dependence of $\mathbf{u}^1$ with respect to z

$$u_z^1(s > r_i, \varphi, z) = z\beta u_s^0(s, \varphi). \tag{9}$$

If $h = \sqrt{r_o^2 - s^2}$ denotes the half-height of the column located at a given cylindrical radius s, the slope of the upper surface is $\mathrm{d}h/\mathrm{d}s$, and we can define

$$\beta(s) = h^{-1}\mathrm{d}h/\mathrm{d}s. \tag{10}$$

The notation β has been chosen in reference to the β-plane approximation. This approximation -with uniform β- is ubiquitous in geophysical fluid dynamics (e.g. Vallis, 2006, section 2.3). It is convenient, indeed, to study planetary Rossby waves assuming that the Coriolis parameter $(f_0 = 2\Omega\cos\theta)$ varies linearly with latitude; β is then the northward gradient of the Coriolis parameter.

According to our quasi-geostrophic approach, the flow in the outer core is nearly two-dimensional, which makes it natural to take the vertical average of the Navier-Stokes equation (8). The vertical average $\langle\cdot\rangle$ of a quantity X is defined as

$$\langle X \rangle (s, \varphi) = \frac{1}{2h(s)} \int_{-h}^{h} X(s, \varphi, z)\,\mathrm{d}z. \tag{11}$$

In a multiply-connected domain, the φ-averaged vorticity equation is not equivalent to the φ-averaged Navier-Stokes equation (Plaut, 2003), as the former does not ensure the

existence of the pressure field; accordingly, we describe the evolution of the non-zonal flow $\mathbf{u}^{NZ}$ by means of the axial vorticity equation, while the φ-averaged momentum equation directly provides us with the time changes of the zonal velocity $\mathbf{u}^Z = u_\varphi^Z \mathbf{e}_\varphi$. In the remainder of this paper, the superscript capital Z marks zonal quantities. It should not be confused with small z, which refers to the direction of rotation.

The non-zonal (NZ) velocity field $\mathbf{u}^{NZ}$ is written as the curl of a z-independent non-zonal streamfunction Ψ,

$$\mathbf{u}^{NZ}(s,\varphi) = \boldsymbol{\nabla} \times \Psi(s,\varphi)\mathbf{e}_z. \tag{12}$$

The non-zonal vorticity field $\boldsymbol{\zeta}$ is defined by $\boldsymbol{\zeta} = \boldsymbol{\nabla} \times \mathbf{u}^{NZ}$, and its vertical component is

$$\zeta_z(s,\varphi) = -\nabla_E^2 \Psi(s,\varphi), \tag{13}$$

in which the equatorial Laplacian operator is defined by

$$\nabla_E^2 = s^{-1}\partial_s\left(s\partial_s\right) + s^{-2}\partial_\varphi^2. \tag{14}$$

If we now curl the non-zonal part of the z-averaged Navier-Stokes equation (8), we find that the vertical component of the vorticity equation is then identical to equation (17) of Pais and Jault (2008), with an explicit right-hand side term,

$$\begin{aligned}
\mathrm{D}_t\zeta_z - 2\lambda_{r_o}^{-1}\beta s^{-1}\partial_\varphi\Psi &= \left(s^{-1}\partial_s\partial_\varphi + s^{-2}\partial_\varphi\right)\left(\langle B_\varphi^2\rangle - \langle B_s^2\rangle\right) \\
&\quad + \left(3s^{-1}\partial_s - s^{-2}\partial_\varphi^2 + \partial_s^2\right)\langle B_s B_\varphi\rangle.
\end{aligned} \tag{15}$$

The magnetic surface terms, which appear when taking the z-average of the Lorentz force, are neglected because we assume the magnetic field at the core surface to be much smaller than in the bulk of the fluid. The non-penetration boundary condition, at $s = r_o$ and at the tangent cylinder $s = r_i$, imposes $\Psi = 0$ at both boundaries.

The time evolution of the zonal velocity $u_\varphi^Z(s) = s\omega_g(s)$ is governed by

$$\mathrm{D}_t\omega_g = \left(s^3 h\right)^{-1}\partial_s\left(s^2 h\langle B_s B_\varphi\rangle\right). \tag{16}$$

The two flow equations (15) and (16) contain z-averaged squared magnetic quantities $\langle B_s^2\rangle$, $\langle B_\varphi^2\rangle$ and $\langle B_s B_\varphi\rangle$, whose time evolution is in turn derived from the diffusionless version of the induction equation (2)

$$\partial_t\langle B_s^2\rangle = -\left[\mathbf{u}^0\cdot\boldsymbol{\nabla}\right]\langle B_s^2\rangle + 2\langle B_s^2\rangle\partial_s u_s^0 + 2s^{-1}\langle B_s B_\varphi\rangle\partial_\varphi u_s^0, \tag{17}$$

$$\partial_t\langle B_\varphi^2\rangle = -\left[\mathbf{u}^0\cdot\boldsymbol{\nabla}\right]\langle B_\varphi^2\rangle - 2\langle B_\varphi^2\rangle\partial_s u_s^0 + 2s\langle B_s B_\varphi\rangle\partial_s\left(s^{-1}u_\varphi^0\right), \tag{18}$$

$$\partial_t\langle B_s B_\varphi\rangle = -\left[\mathbf{u}^0\cdot\boldsymbol{\nabla}\right]\langle B_s B_\varphi\rangle + s\langle B_s^2\rangle\partial_s\left(s^{-1}u_\varphi^0\right) + s^{-1}\langle B_\varphi^2\rangle\partial_\varphi u_s^0, \tag{19}$$

where we have made use of the solenoidal character of $\mathbf{B}$ and $\mathbf{u}$.

The vertical averaging naturally sets the magnetohydrodynamics in the equatorial plane Σ (Figure 1b). The flow is then projected at the CMB, where it interacts with the radial magnetic field B_r via equation (4) above.

An alternative model, where the velocity field entering the set of equations (17) to (19) has a z-component given by (9) and a φ-component modified in order to ensure that the total velocity field remains solenoidal, is discussed in appendix A.

3   Variational data assimilation framework

In this section, we introduce the geomagnetic secular variation data assimilation problem with the notations suggested by Ide et al. (1997). In comparison with the 4DVar label commonly used in data assimilation (e.g. Courtier, 1997), our framework is 1+2DVar, since state variables are defined in two-dimensional spaces to which a third (temporal) dimension is added - the 3DVar label is traditionally reserved for three-dimensional (in space) static problems (Courtier, 1997).

The state vector $\mathbf{x}$ for the Earth's core gathers the variables involved in the description of the system state

$$\mathbf{x} = \left[\omega_g, \Psi, \left\langle B_s^2 \right\rangle, \left\langle B_\varphi^2 \right\rangle, \left\langle B_s B_\varphi \right\rangle, B_r \right]^T, \tag{20}$$

where superscript T means transpose. Observations $\mathbf{y}$ are available at $T_{\mathbf{y}}$ different epochs and $N_{\mathbf{y}}$ spatial locations during the assimilation time window $[0, T]$; the size of $\mathbf{y}$ is typically smaller than the dimension of $\mathbf{x}$. The observation vector is related to the true core state $\mathbf{x}^t$ via the observation operator H:

$$\mathbf{y} = \mathsf{H}\mathbf{x}^t + \boldsymbol{\epsilon}, \tag{21}$$

in which $\boldsymbol{\epsilon}$ is the observation error.

Variational data assimilation aims at adjusting a model solution $\mathbf{x}^f$ to the observations (Talagrand, 1997), by minimizing a misfit function which comprises the quadratic discrepancy -if H is linear and errors are Gaussian- between the true observations and those predicted by the computed state, J_H (Courtier, 1997):

$$J_H = \sum_{j=1}^{T_{\mathbf{y}}} \left[\mathsf{H}_j \mathbf{x}_j^f - \mathbf{y}_j \right]^T \mathbf{R}_j^{-1} \left[\mathsf{H}_j \mathbf{x}_j^f - \mathbf{y}_j \right], \tag{22}$$

where j is the discrete time index and $\mathbf{R} = \mathsf{E}\left(\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T \right)$ is the observation error covariance matrix, $\mathsf{E}(\cdot)$ denoting statistical expectation. The matrix $\mathbf{R}$ describes the level of confidence we set in the observations.

It might be necessary to constrain the solution sought in order to enforce its uniqueness, especially if the problem is non-linear. Constraining the assimilation refers to either adding a background state $\mathbf{x}^b$ from which the estimate shall not strongly deviate, or applying additional constraints on the core state. Imposing a penalty, the goal of which is to favor a solution with a moderate level of complexity (e.g. Courtier and Talagrand, 1987), described by a matrix C applied to the state vector, consists in adding a second term to the objective function, of the form

$$J_C = \sum_{j=0}^{T} \mathbf{x}_j^T \mathbf{W}_j \mathbf{x}_j, \tag{23}$$

where $\mathbf{W} = \mathsf{C}^T\mathsf{C}$. The total misfit function J to minimize is then given by the sum

$$J = \frac{\alpha_H}{2}J_H + \frac{\alpha_C}{2}J_C, \tag{24}$$

in which the two contributions are weighted by two scalars, α_H and α_C (Fournier et al., 2007).

In practice, $\mathbf{x}^f$ is the solution of a numerical non-linear model M, that describes the temporal evolution of $\mathbf{x}$ at any discrete time $t_j \in [0, T]$,

$$\mathbf{x}^f_{j+1} = \mathsf{M}_j\mathbf{x}^f_j; \tag{25}$$

that notation symbolically summarizes the set of equations developed in section 2. Since the temporal history of the state is constructed with a dynamical model that couples its various components, we can, in principle, retrieve information about every state variable, even if only part of the state vector is directly probed by the observations on hand (Fournier et al., 2007). Moreover, the initial condition $\mathbf{x}_0$, termed the control vector, uniquely defines the model trajectory in state space. That reduces the dimension of the associated inverse problem: we only seek the initial condition, $\mathbf{x}_0$, starting from which the temporal evolution $\mathbf{x}^a$ will best fit the observations; in assimilation parlance, this best solution is called the analysis. The minimization of J (that is the search for $\mathbf{x}^a$) is performed with a descent algorithm that involves the sensitivity of J to its control variables, $\mathbf{x}_0$: $\nabla_{\mathbf{x}_0}J$. Its transpose is efficiently estimated with the use of the adjoint model M^T (Le Dimet and Talagrand, 1986). For a given $\mathbf{x}_0$, one couples a forward integration of M with a backward integration of M^T to express the gradient of J. The adjoint model is that of the local tangent linear equations (see e.g. Talagrand and Courtier, 1987; Giering and Kaminski, 1998). Introducing the adjoint variable $\mathbf{a}$ of $\mathbf{x}$, the adjoint equation imposed by the cost function (equation (24)) is (Fournier et al., 2007)

$$\mathbf{a}_{j-1} = \mathsf{M}^T_{j-1}\mathbf{a}_j + \alpha_H\mathsf{H}^T_{j-1}\mathbf{R}^{-1}_{j-1}\left(\mathsf{H}_{j-1}\mathbf{x}_{j-1} - \mathbf{y}_{j-1}\right) + \alpha_C\mathbf{W}_{j-1}\mathbf{x}_{j-1}, \tag{26}$$

where H^T is the transpose of the observation operator (equation (21)), which projects a vector from observation space to state space. Through equation (26), the adjoint field is fed by observation residuals (the difference $\mathsf{H}\mathbf{x} - \mathbf{y}$), as soon as there is an observation available.

The backward integration requires the knowledge of the model trajectory over the assimilation time window. The storage of the complete trajectory may cause memory issues, which are traditionally resolved using a checkpointing strategy. The state of the system is stored at a limited number of discrete times, termed checkpoints. Over the course of the backward integration of the adjoint model, these checkpoints are used to recompute local portions of the trajectory on-the-fly, whenever those portions are needed (e.g. Hersbach, 1998).

Non-linear optimization is performed with the M1QN3 routine (Gilbert and Lemaréchal, 1989), which implements a limited memory quasi-Newton algorithm that approximates the inverse Hessian (second derivative) of J.

4 Applications

We show two illustrations of variational data assimilation applied to fast core dynamics, as described by our quasi-geostrophic model, with synthetic data. The methodology of twin experiments is explained and applied to a steady non-zonal flow problem and, in a second step, to a dynamical zonal model of torsional Alfvén waves. These two problems correspond to two subsets of the model presented in section 2.

4.1 Twin experiments

Instead of being satellite or observatory data, observations in our twin experiments are created from a synthetic true state, which is the result of the integration of the forward model for a given set of initial conditions, $\mathbf{x}_0^t$. Synthetic data have the advantage of representing only the physics involved in the model and are, in a first step, appropriate to test the implementation of the variational data assimilation algorithm. A database of observations is produced with equation (21).

To construct the observation catalog, we include some geophysical realism by averaging the state at the core-mantle boundary. We choose the averaging window so that it corresponds to the ignorance of the spherical harmonic coefficients of degree $n > L$. Then, the product $\mathsf{H}\mathbf{x}^t$ corresponds to the convolution over the core-mantle boundary of the true state with a Jacobi polynomial of degree L (Backus et al., 1996, paragraph 4.4.4):

$$(\mathsf{H}\mathbf{x}^t)(\theta_o, \varphi_o) = \frac{L+1}{4\pi} \int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} \mathbf{x}^t(\theta, \varphi) P_L^{(1,0)}(\cos\alpha) \sin\theta \mathrm{d}\theta \mathrm{d}\varphi, \qquad (27)$$

in which (θ_o, φ_o) are the coordinates of the observation locations, α is the angular distance between the points (θ_o, φ_o) and (θ, φ), and $P_L^{(1,0)}$ is a Jacobi polynomial (Abramowitz and Stegun, 1964, p. 773). In the following experiments, we set $L = 15$ and observations are made at a fixed temporal frequency.

Next, we start the assimilation with a different set of initial conditions, called initial guess, $\mathbf{x}_0^g$. After a forward integration, the computed observable, $\mathsf{H}\dot{B}_r^{\,f}$ in section 4.2 and $\mathsf{H}B_r^f$ in section 4.3, is compared (over the entire time window) with the observations, and the discrepancy between the two gives the initial misfit (see equation (24)). After assimilation, the decrease of the misfit and the relative difference between $\mathbf{x}^t$ and $\mathbf{x}^a$, in the l^2-sense, are used to assess the quality of the recovered state.

4.2 The kinematic core flow problem

In this section, a connection is made between core flow kinematic inversions and data assimilation. The steady flow hypothesis has been previously used by Voorhies and Backus (1985) and Waddington et al. (1995) to break the non-uniqueness of the kinematic inversion problem. Here, we study the effect of a steady non-zonal and equatorially symmetric flow

on the evolution of the radial magnetic field, and more particularly its secular variation $\dot{B}_r$,

$$\dot{B}_r = -\nabla_H \cdot \left(\mathbf{u}^{NZ} B_r \right). \tag{28}$$

Symmetry with respect to the equator ensures uniqueness of the solution when B_r and $\dot{B}_r$ are perfectly known. The time scale characterizing this problem is the advection time, t_{adv}. The non-zonal velocity effectively enters equation (28) via the non-zonal streamfunction $\Psi(s,\varphi)$ (see equation (12)), projected at the core-mantle boundary. The state vector $\mathbf{x}$ includes a parameter, the streamfunction Ψ, and a variable, the radial magnetic field B_r. Here, Ψ is called a parameter because it is steady in that experiment. We seek the distribution of $\Psi(s,\varphi)$ which best explains the observed synthetic database of secular variation $\dot{B}_r^o$ at the top of the core.

The tangent linear form of (28) is

$$
\begin{align}
\delta \dot{B}_r &= \mathsf{P}\left(\delta\Psi, B_r \right) + \mathsf{Q}\left(\Psi, \delta B_r \right), \tag{29}\\
\partial_t \delta B_r &= \delta \dot{B}_r, \tag{30}
\end{align}
$$

where $\delta\Psi$, δB_r and $\delta \dot{B}_r$ are the differentials of Ψ, B_r and $\dot{B}_r$, respectively, and P and Q the differentials of the right-hand side term of equation (28) with respect to Ψ and B_r, respectively (they are developed in appendix C). Let us introduce Ψ^T, B_r^T and $\dot{B}_r^{\,T}$ as the adjoint variables of Ψ, B_r and $\dot{B}_r$. The adjoint model of equations (29) and (30) is then

$$
\begin{align}
\Psi^T &= \sum_{j=0}^{T} \mathsf{P}^T \left(\dot{B}_{r_j}^{\,T}, B_{r_j} \right), \tag{31}\\
B_r^T &= \mathsf{Q}^T \left(\dot{B}_r^{\,T}, \Psi \right), \tag{32}
\end{align}
$$

in which P^T, Q^T are the adjoint functions of P and Q (see detailed equations in appendix C). The link to the observations has been obtained from equation (24), and is computed at each time step as in equation (26):

$$\dot{B}_r^{\,T} = \alpha_H \left(\mathsf{H}^T \, \mathsf{H}\dot{B}_r - \mathsf{H}^T \, \dot{B}_r^o \right). \tag{33}$$

The trajectory of the true state is computed from the following set of initial conditions:

1. $B_r^t(\theta, \varphi, t=0)$ is obtained from the CHAOS main field model (Olsen et al., 2006) for epoch 2002, truncated at spherical harmonic degree and order 12. It is taken outside the tangent cylinder and multiplied by a sine function of θ in order to have $B_r^t(\theta, \varphi, t=0) = 0$ at the tangent cylinder (see Figure 2),

2. $\Psi^t(s,\varphi)$ is shown in Figure 3a, it is the non-zonal part of an inverted flow from Pais et al. (2004) truncated at degree and order 4, and multiplied by $\cos^2\theta = (1 - s^2)$ and a function of s, $(s - r_i)$, in order to satisfy the flow boundary conditions at $s = r_i, r_o$. It is normalized in order to have a dimensionless rms velocity of order 1; the scaling V_{adv} is such that $\int_0^{2\pi} \int_{\theta_c}^{\pi-\theta_c} (u_s^2 + u_\varphi^2) \sin\theta \mathrm{d}\theta \mathrm{d}\varphi = V_{adv}^2 \int_0^{2\pi} \int_{\theta_c}^{\pi-\theta_c} \sin\theta \mathrm{d}\theta \mathrm{d}\varphi$, with $\theta_c = \mathrm{asin}(r_i/r_o)$.
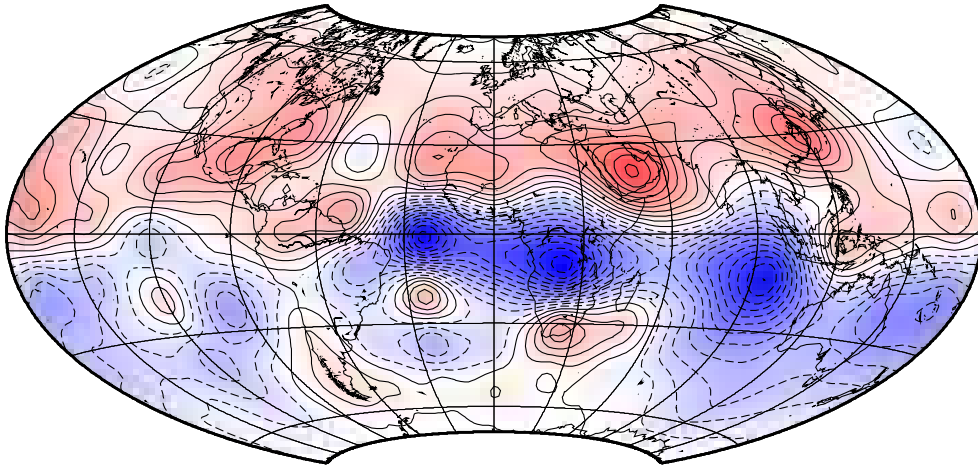
11

Figure 2. Map of the main field at epoch 2002 truncated at spherical harmonic degree and order 12 modified from the CHAOS model (Olsen et al., 2006). True radial component of the magnetic field B_r^t at initial time for the twin experiments. Contours are drawn each $0.5 B_r^{rms}$ (solid (resp. dashed) for positive (resp. negative) values). Note that the problem is linear in B_r.

We consider perfect observations, setting $\boldsymbol{\epsilon} = \mathbf{0}$ in equation (21). For the following simulations, the numerical time step is $6 \times 10^{-5} \, t_{adv}$ for integration times ranging from $0.03 \, t_{adv}$ to $0.57 \, t_{adv}$. Other numerical parameters relative to the simulations are given in appendix B.

In this problem of seeking a steady streamfunction that explains the observations, we want to show the benefit of including the temporal dimension (data assimilation) instead of relying on a single observation epoch, as it is the case for a standard kinematic inversion.

We first consider solutions obtained with only one observation epoch and less observation locations ($N_\theta^O = 50, N_\phi^O = 11$) than grid points ($N_\theta = 200, N_\phi = 15$). We start the assimilation with a first guess Ψ^g corresponding to the minimal hypothesis: $\Psi^g = 0$. The solution we obtain gives us some information about what could be achieved within the kinematic framework in that configuration. Figure 3b shows in particular that the true state (Figure 3a) is not completely recovered, due to the truncation used in the construction of H and the limited number of observation locations, compared to the total number of grid points.

Using this solution obtained from a single epoch inversion as a reference solution, we can study the benefit we get resorting to a timeseries of observations, as opposed to a single snapshot. In other words, we investigate whether the issue of spatial subsampling can be partially fixed by considering the temporal dimension. To that end, we do experiments with assimilation time windows ranging from $0.03 \, t_{adv}$ to $0.57 \, t_{adv}$, at a given temporal frequency of observation $f_{\mathbf{y}} = 100 \, t_{adv}^{-1}$, keeping the same number of virtual magnetometers.

Results of a typical experiment are shown in Figure 3c, for which $T = 0.57 \, t_{adv}$. The large scale pattern is retrieved, but the solution is polluted by small spatial scales (no extra smoothing term is added to the misfit function). We find, however, that the consideration
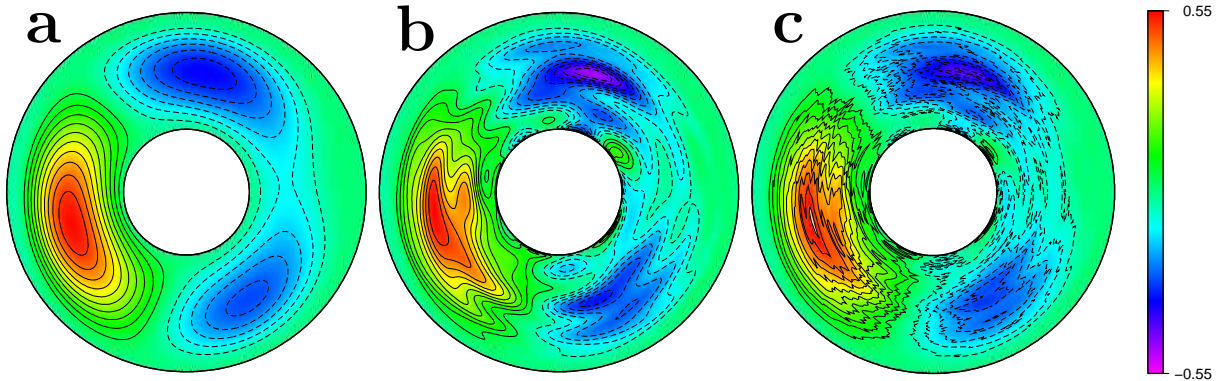
12

Figure 3. Maps in the equatorial plane Σ of the steady streamfunction Ψ: a: true state, b: analyzed state with a single epoch inversion, c: analyzed state with $T = 0.57\ t_{adv}$. Contours are drawn each 0.05 (solid (resp. dashed) for positive (resp. negative) values). Extrema are $-0.33; 0.54$ (a), $-0.49; 0.54$ (b) and $-0.43; 0.56$ (c).

of the temporal dimension improves the solution. Moreover, the distance between the true streamfunction and the retrieved streamfunction becomes smaller when the assimilation time window is widened (see Figure 4).

As described by Evensen (2007, chapter 6), if one enlarges the width of the assimilation time window in a non-linear context, the misfit function presents more and more spikes and minima. A very good first guess is thus needed to converge to the global minimum. To circumvent this issue, we decided to use the results obtained over short assimilation time windows as initial guesses for assimilation over longer time windows. That strategy is analogous to the approach used in the atmospheric variational assimilation community, which consists in solving a series of strong constraint inverse problems, defined for separate subintervals in time (e.g. Evensen, 2007).

4.3   Forward and adjoint modeling of Alfvén torsional waves

In non-rotating magnetized flows, classical Alfvén waves result from the balance between inertial and magnetic forces.

In the Earth's core, where the Coriolis force is large, Braginsky (1970) showed that a special class of Alfvén waves comes into play, in which only the component of the magnetic field normal to the axis of rotation, B_s, participates. Associated motions are geostrophic; they are organized in axial cylinders about the axis of rotation, hence the name torsional oscillations. The period of torsional waves depends on the strength and distribution of B_s inside the core.

In order to study these waves, one can consider a subset of the complete dynamical model. Since torsional waves are geostrophic and axisymmetric motions, let us discard the non-zonal part of the flow and magnetic induction in equations (15) to (19).

In addition to the vertical average, $\langle \cdot \rangle$, introduced in equation (11), we now define the
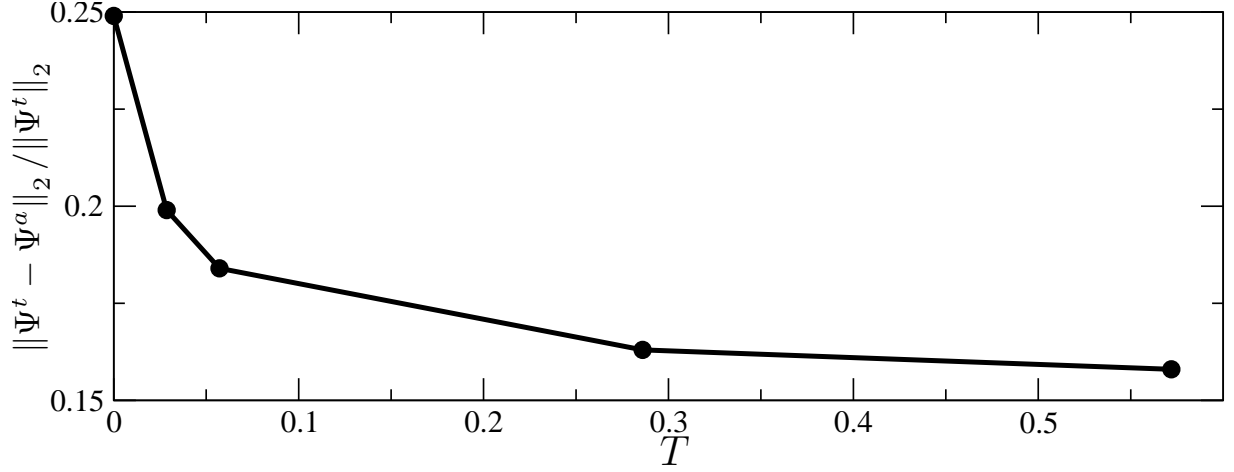
Figure 4. Effect of the assimilation time T on the analyzed state at a fixed observation frequency $f_{\mathbf{y}} = 100\ t_{adv}^{-1}$, measured with the norm of the relative error $\left\| \Psi^t - \Psi^a \right\|_2 / \left\| \Psi^t \right\|_2$. This norm is equal to one when the initial guess $\Psi^g = 0$. The simulation at $T = 0$ is referred in the text as the single epoch inversion.

average on a geostrophic cylinder, $\{\cdot\}$, by

$$\{X\}(s) = \frac{1}{2\pi} \int_0^{2\pi} \langle X \rangle (s, \varphi)\, \mathrm{d}\varphi. \tag{34}$$

In the following, we will not indicate explicitly the dependence on s of quantities averaged on a geostrophic cylinder. Application of $\{\cdot\}$ to equation (16) yields

$$\partial_t \omega_g = \left(s^3 h\right)^{-1} \partial_s \left(s^2 h \left\{B_s B_\varphi\right\}\right). \tag{35}$$

Similarly, the equations governing the evolution of magnetic quantities become

$$\partial_t \left\{B_s^2\right\} = 0, \tag{36}$$
$$\partial_t \left\{B_s B_\varphi\right\} = s \left\{B_s^2\right\} \partial_s \omega_g. \tag{37}$$

Written in terms of the geostrophic angular velocity ω_g, the torsional wave equation is

$$\partial_t^2 \omega_g = \left(s^3 h\right)^{-1} \partial_s \left(s^3 h \left\{B_s^2\right\} \partial_s \omega_g\right). \tag{38}$$

Equation (38) can be transformed into a set of two first-order equations

$$\partial_t \omega_g = \left(s^3 h\right)^{-1} \partial_s \tau, \tag{39}$$
$$\partial_t \tau = s^3 h \left\{B_s^2\right\} \partial_s \omega_g, \tag{40}$$

in which $\tau = s^2 h \left\{B_s B_\varphi\right\}$ is an auxiliary variable.

We have taken as boundary condition for the angular velocity:

$$\partial_s \omega_g = 0, \quad \text{at } s = r_o. \tag{41}$$

14

Then, the boundary condition,

$$\partial_s \omega_g = 0, \quad \text{at } s = r_i, \tag{42}$$

ensures the conservation of the total angular momentum carried by the fluid in the computational domain.

Projected at the CMB, those motions interact with B_r through equation (4), which simplifies here into

$$\partial_t B_r = -\omega_g \partial_\varphi B_r. \tag{43}$$

The system state gathers the geostrophic angular velocity ω_g, the variable τ, the cylindrical average of the s-component of the magnetic field, $\{B_s^2\}$, and the radial component of the magnetic field B_r at the CMB.

We define $\omega_g^T, \tau^T, \{B_s^2\}^T, B_r^T$ as the adjoint variables of $\omega_g, \tau, \{B_s^2\}, B_r$ respectively. The torsional oscillations adjoint model is

$$-\partial_t \omega_g^T = \partial_s^T \left[s^3 h \{B_s^2\} \tau^T \right] - B_r^T \partial_\varphi B_r, \tag{44}$$

$$-\partial_t \tau^T = \partial_s^T \left[(s^3 h)^{-1} \omega_g^T \right], \tag{45}$$

$$\{B_s^2\}^T (s) = \sum_{j=0}^{T} s^3 h \left(\partial_s \omega_g \right)_j \tau_j^T + \alpha_C \mathbf{W} \{B_s^2\}, \tag{46}$$

$$-\partial_t B_r^T = -\partial_\varphi^T \left[\omega_g B_r^T \right], \tag{47}$$

where ∂_s^T and ∂_φ^T are the adjoints of the differential operators ∂_s and ∂_φ (see appendix D for more details on the adjoint system). The link to the constraint on $\{B_s^2\}$ has been obtained from equation (24). The model is completed by the information supplied by the observations, as in equation (33). The boundary conditions for the adjoint model are $\tau^T = 0$, at both $s = r_i$ and $s = r_o$. The temporal and spatial discretizations of this problem are described in appendix B.

For the experiments that follow, the set of initial profiles, which define the true state, is:

1. the same $B_r^t(\theta, \varphi, t = 0)$ as in the kinematic core flow problem of section 4.2,

2. a Gaussian function for the angular velocities: $\omega_g^t(s, 0) = \omega_0 \exp\left[-\sigma_\omega^{-2}(s - s_\omega)^2\right]$, with $\sigma_\omega^{-2} = 150$ and $s_\omega = 0.65$; it satisfies the boundary conditions (41) and (42), its amplitude being scaled by ω_0 (discussed hereafter),

3. $\tau^t(s, 0) = 0$,

4. an arbitrary function $\{B_s^2\}^t$ (see the black curve in Figure 6, right) given by:
   $\{B_s^2\}^t (s) = c_1 + c_2 \sin(\pi/2 - L) + c_3 \exp\left[-\sigma_B^{-2} (s - s_B)^2\right]$, with $c_1 = 0.1$, $c_2 = 0.02$, $c_3 = 1$, $\sigma_B^{-2} = 20$, $s_B = 0.8$ and $L = 14s$. It is normalized in order to have a dimensionless rms magnetic field of unity inside the core; the scaling B_0 is such that $\int_0^{2\pi} \int_{r_i}^{r_o} B_s^2 \mathrm{d}s \mathrm{d}\varphi = B_0^2 \int_0^{2\pi} \int_{r_i}^{r_o} \mathrm{d}s \mathrm{d}\varphi$.
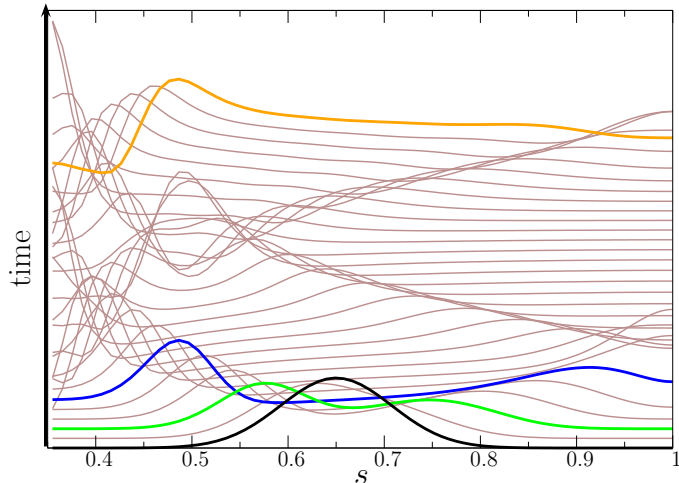
15

Figure 5. Torsional wave twin experiments results: successive profiles of the angular velocity ω_g, showing the propagation of a torsional wave in the computational domain $s \in [0.35, 1]$ during 1.16 Alfvén time T_A. Initial condition (black curve) for ω_g, and snapshots, at 0.12 (green), 0.18 (blue) and 1.16 T_A (orange).

As the velocity has been scaled by the Alfvén velocity, ω_0 is the ratio between the Alfvén time and the advection time, $\omega_0 = T_A/t_{adv}$.

Assimilation is performed on both ω_g and $\{B_s^2\}$. We seek the steady profile of $\{B_s^2\}$ and the initial profile of ω_g which best explain the synthetic database of B_r^o. Our first guess consists of flat profiles: $\omega_g^g(s, 0) = 0.1\omega_0$ and $\{B_s^2\}^g = 0.6$ (see the red curves in Figure 6).

We show here experiments with a fixed frequency of observations $f_\mathbf{y} = 20\ T_A^{-1}$ and as many observations locations as grid points. The observations are blurred by the averaging kernel (equation (27)), which causes errors. Consequently, the analysis can develop small scales, which are not very well constrained by the observations. We choose to reduce the complexity of the solution by adding a smoothing term to the cost function, taking $\alpha_C = 10^{-8}$ in equation (24). Here we penalize only the strong spatial gradients of $\{B_s^2\}$. The reference case (with $\omega_0 = 0.34$ and $T = 1.16\ T_A$, see Figures 6 and 7) shows that both the angular velocity and interior magnetic field are well recovered. As shown in Figure 7, the error field is substantially weaker after assimilation.

On a technical note, the M1QN3 algorithm (Gilbert and Lemaréchal, 1989), used in the optimization loop, stops in that case when the initial misfit is divided by a factor of 4×10^5, which is reached in 214 iterations.

In order to assess the effect of the width of the assimilation window on the retrieved state variables, we vary the assimilation time T between 0.12 and 1.16 T_A, keeping ω_0 constant, equal to 0.34 as above. The geostrophic angular velocity is in all cases completely recovered (not shown) with similar spurious oscillations as in Figure 6 (left) near the outer boundary. On the other hand, the area over which $\{B_s^2\}$ is correctly retrieved increases with T, indicating that the assimilated area is controlled by the distance over which the initial pulse has propagated (Figure 8, left). Figure 5 shows that, if $T = 1.16\ T_A$, the
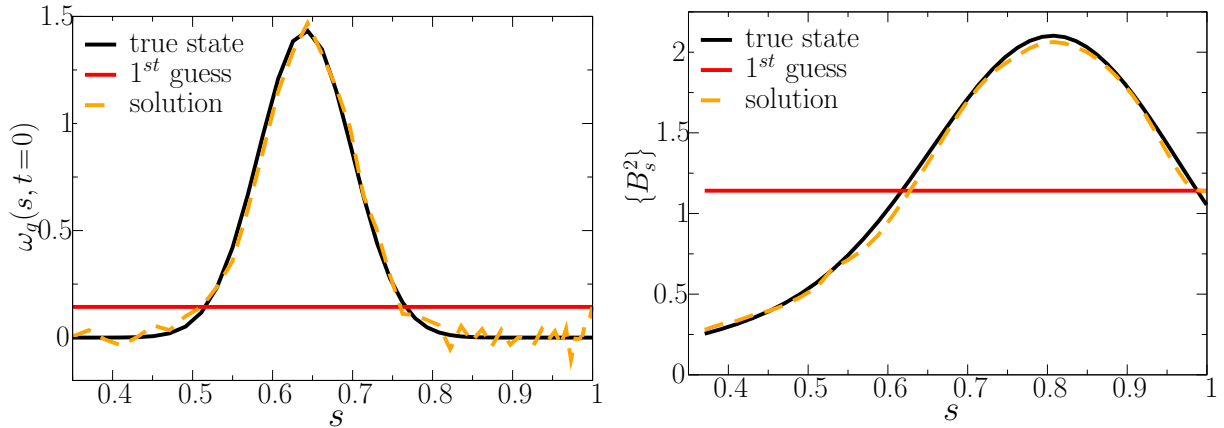
16

Figure 6. Torsional wave twin experiments results: true state (black), profile before assimilation (red) and solution after assimilation (dashed-orange) for ω_g (left) and $\left\{B_s^2\right\}(s)$ (right). The parameters for that reference case are $\omega_0 = 0.34$ and $T = 1.16\ T_A$. Regularization has been added to the spatial derivative of $\left\{B_s^2\right\}$, of amplitude $\alpha_C = 10^{-8}$ (See text for details).
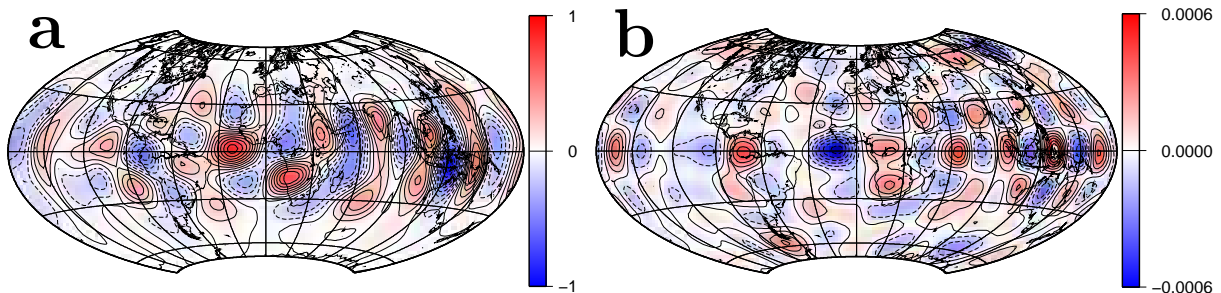

Figure 7. Relative difference between observed and computed $\mathsf{H}B_r$ at final time, $\left[B_r^o(T) - \mathsf{H}B_r^f(T)\right] / \|B_r^o(T)\|_2$ before assimilation (**a**) and after assimilation (**b**) for the reference case (same parameters as in Figure 6). Contours are drawn each 0.1 (**a**) and 10^{-1} (**b**) (solid (resp. dashed) for positive (resp. negative) values). Extrema are $-0.94; 0.86$ (**a**) and $[-6.4; 7.2] \times 10^{-4}$ (**b**).
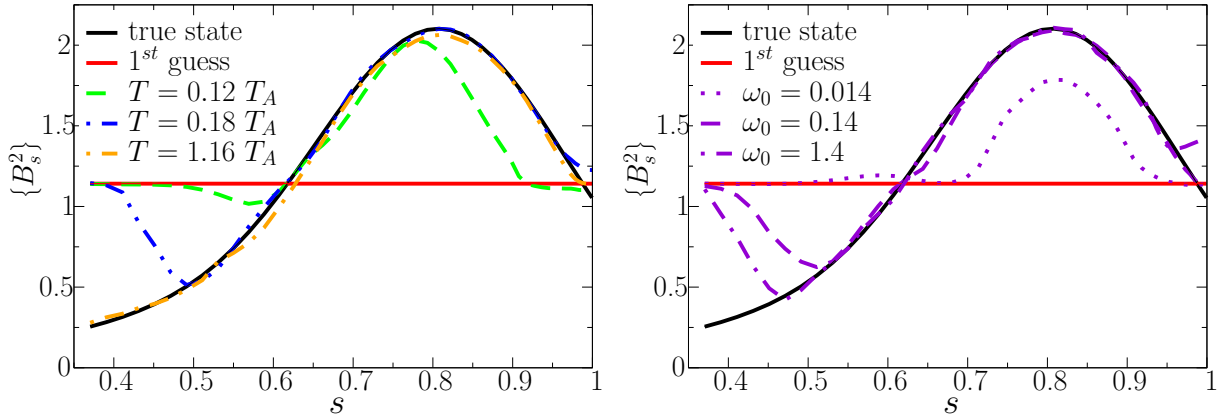
17

Figure 8. Effect on the analysis of the assimilation time window of width T (left) and of the amplitude of the initial pulse ω_0 (right). Left: true state (black), profile before assimilation (red) and solution after assimilation (dashed-green, dashed-double-dotted-blue, dashed-dotted-orange) for $\{B_s^2\}(s)$, for different values of T (0.12, 0.18 and 1.16 T_A respectively) for a fixed $\omega_0 = 0.34$. Right: true state (black), profile before assimilation (red) and solution after assimilation (violet) for $\{B_s^2\}(s)$. Dotted, dashed and dashed-dotted violet curves are obtained for different values of the amplitude of the initial pulse ω_0 (0.014, 0.14 and 1.4 respectively), keeping T fixed, equal to 0.18 T_A.

wave has enough time to explore the whole domain. On the contrary, if $T = 0.12$ or 0.18 T_A, a lesser portion of the domain is sampled by the wave, over which $\{B_s^2\}$ has been effectively retrieved. The angular velocity is better recovered than $\{B_s^2\}$ because it is directly connected to the observations, as opposed to $\{B_s^2\}$. That can be seen in the adjoint equations: ω_g^T (equation (44)) depends on B_r^T that contains B_r^o, the observed quantity, whereas $\{B_s^2\}^T$ (equation (46)) is only directly connected to τ^T. In turn, τ^T sees ω_g^T, which is ultimately linked to the observations.

For a fixed T, the dependence on the amplitude ω_0 has been studied (Figure 8, right). For $\omega_0 = 0.014$, ω_g and $\{B_s^2\}$ are not well recovered. Starting from $\omega_0 = 0.14$, however, increasing ω_0 by more than one order of magnitude has little effect on the retrieval of $\{B_s^2\}$ and no effect at all on ω_g^a.

Let us stress that the convergence of the calculations presented here is also sensitive to the profile of the initial condition (for both true state and first guess), and to the amount of measurements, as in the steady case of section 4.2. Moreover, we have observed in other instances (not shown) that convergence is sped up if the acceleration τ is not zero at initial time. More generally, that particular example shows that the success of assimilation is controlled by the intrinsic dynamics of the system under study, as well as a good guess of its state.

Until now, we have assumed perfect observations: $\boldsymbol{\epsilon} = \mathbf{0}$ in equation (21). In the prospect of future applications, observation error should be considered. In the next experiments, observations contaminated by errors are assimilated. Centered, normally dis-
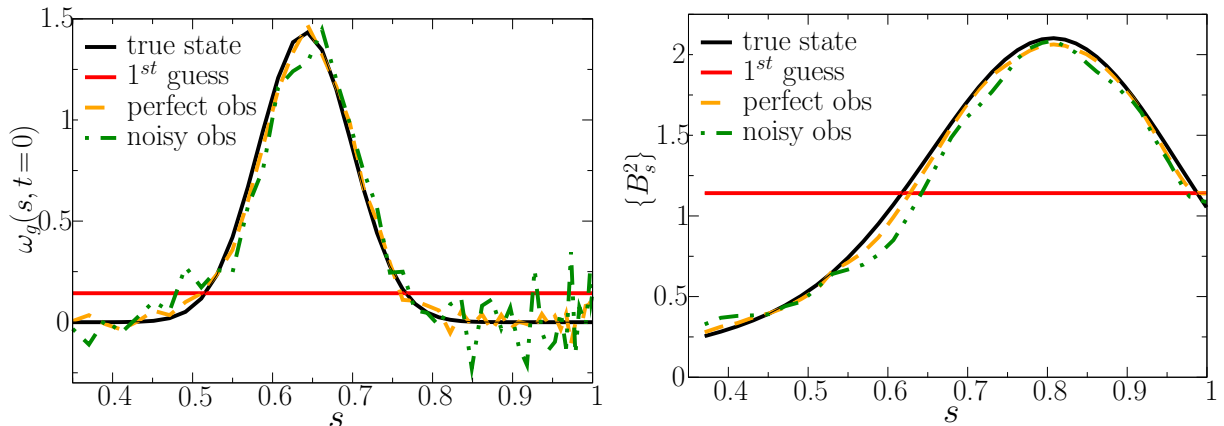
Figure 9. Torsional wave twin experiments results obtained with noisy data: True profiles (black), profile before assimilation (red) and solution after assimilation (orange, green) for ω_g (left) and $\{B_s^2\}$ (right). Dashed-orange curves have been computed with perfect observations, dashed-dotted-green curves considering centered, normally distributed Gaussian observation errors of standard deviation $10^{-1} B_r^{rms}$ added to the database. Regularization has been added on the spatial derivative of $\{B_s^2\}$, of amplitude $\alpha_C = 10^{-8}$. The assimilation window width is $T = 1.16\ T_A$, and $\omega_0 = 0.34$.

tributed Gaussian observation errors of standard deviation $10^{-1} B_r^{rms}$ are added to the previous database. That experiment is carried out with an assimilation window width of 1.16 T_A. Even with a database contaminated by observation errors, it is still possible to recover the shape and strength of the true state (see Figure 9). $\{B_s^2\}$ seems less sensitive to observation errors than ω_g: we still have an extra penalty term in the misfit function as in the perfect observations case, but the overall effect of small scales is actually to degrade the solution for both fields.

For completeness, we have also decreased the temporal frequency of observation and observed that the recovering of the true state was possible provided that the frequency of observation, $f_\mathbf{y}$, was greater than $3\ T_A^{-1}$.

5 Discussion

We have derived a quasi-geostrophic model of core dynamics, which aims at describing core processes on geomagnetic secular variation timescales. Under the quasi-geostrophic assumption, the magnetohydrodynamics takes place in the equatorial plane and is written outside the tangent cylinder. The flow is defined by its zonal velocity and non-zonal axial vorticity. The magnetic induction appears through z-averaged quadratic magnetic quantities, while we assume that the magnitude of the magnetic field at the core surface is smaller than in the core interior. In addition, the equatorial flow is projected on the core-mantle boundary. It interacts with the magnetic field at the core surface, through the radial component of the magnetic induction equation, in the frozen-flux approximation. That part

of the model connects the dynamics and the observed secular variation, with the radial component of the magnetic field acting as a passive tracer. We have resorted to variational data assimilation to construct formally the relationship between model predictions and observations. The use of an imperfect observation operator mimics our truncated vision of the reality. We have extensively tested the variational data assimilation algorithm with twin experiments owing to the non-prohibitive numerical costs of the computations. Assimilation was controlled by the initial state and possibly some static model parameters.

Let us stress some important results from our numerical simulations. In our time-dependent framework, we have found that increasing the time window width T always improves the solution. In the steady core flow experiment case, that property has proven useful in constraining intermediate flow length scales otherwise unconstrained by a sparse distribution of observations. Here, the benefit originates from the dynamical relationship which exists between successive observations through the radial component of the magnetic induction equation. In the torsional oscillation case (an illustration of the dynamical model presented in this study), the same property allowed us to retrieve the z-averaged quadratic product of B_s (which is only remotely linked to the observed quantity) over the entire domain provided that T was large enough for the wave to propagate over (and effectively sample) the whole domain. Interestingly, it has not been necessary to include a dissipation term in the forward model. We have investigated the sensitivity of the solution to the frequency of observation in the presence of observation errors. Adding an extra smoothing term to the cost function proved an efficient way to produce analyses with a moderate level of complexity.

From the geophysical point of view, with a typical estimate of the magnetic field strength in the core interior, 2 mT, one Alfvén time, T_A, amounts to 6 years. We have varied the frequency of observation from 2.5 to 20 T_A^{-1}, 3 T_A^{-1} appearing as a minimum to recover the fields, which represents a two-year interval between observations. Therefore, we expect that we may be able to resolve properly torsional waves using the last 10 years of satellite measurements, since the most recent magnetic field models have a resolution of a fraction of year (Olsen and Mandea, 2008).

Pais and Jault (2008) and Gillet et al. (2009) have recently used magnetic field models obtained from satellite data in kinematic inversions of quasi-geostrophic core flows. Their calculated core flows are dominated by a giant retrograde gyre. Gillet et al. (2009) suspect that the weaker momentum of the gyre for the period 1960-1980, compared to the period 1990-2008, is an artifact produced by the lesser data quality before 1980. Resorting to a toy model, Fournier et al. (2007) have demonstrated that the benefit due to the existence of a denser network at the end of an assimilation window, is manifest over a substantial part of the window, thanks to the variational data assimilation approach. In other words, the recent high quality of observatory and satellite measurements can be in principle backward propagated in order to reassimilate historical data series. It is then possible to imagine that the refined series could contribute to a more precise description of both small and large scales of the fluid circulation in the core.

The interplay between magnetic and rotation forces in a two-dimensional model has been investigated in other contexts. Tobias et al. (2007) have recently studied a local two-

dimensional β-plane numerical model to show the impact of a weak large-scale magnetic field on the dynamics of the solar tachocline. Instead of using quadratic products of the magnetic field as variables, they have written the magnetic field as a function of a unique scalar potential A:

$$\mathbf{B}(s, \varphi) = \boldsymbol{\nabla} \times [A(s, \varphi)\mathbf{e}_z].$$ (48)

Then, the magnetic term in the vorticity equation becomes $[\boldsymbol{\nabla} \times (A\mathbf{e}_z)] \cdot \boldsymbol{\nabla} [\nabla_E^2 A]$ (compare with the right-hand side term of equation (15)), and the induction equation is

$$\partial_t A = -\mathbf{u} \cdot \boldsymbol{\nabla} A + S^{-1}\nabla_E^2 A.$$ (49)

The ansatz (48) is restrictive, as axial invariance of the magnetic field is assumed. It enables the inclusion of magnetic diffusion, the effect of which cannot be rigorously introduced in the set of equations (17) to (19). The model given by equations (48) and (49) is also attractive, since it is still able to describe a variety of physical situations. As an example, Diamond et al. (2005) mention the transition from two-dimensional magnetohydrodynamic turbulence at small length scales, to turbulence controlled by Rossby wave interactions at larger length scales. Solutions of (48) and (49) (without the diffusion term) are also solutions permitted by our equations (15) to (19). Investigating that simplified set of equations thus appears as an appealing intermediate step before the actual implementation of the less restrictive equations based upon quadratic magnetic quantities.

A   A first-order variant of the induction equation

To write equations (17) to (19), we have retained only the zeroth-order part of the flow. We may wish to take into account, in these equations, the z-component of the flow that enters in the Coriolis term in equation (15) and in the induction equation at the core-mantle boundary (equation (4)).

Thus, in order to ensure incompressibility, we define

$$\mathbf{u}_E^{NZ}(s,\varphi) = \gamma(s)\left[\boldsymbol{\nabla} \times \Psi(s,\varphi)\mathbf{e}_z\right];\tag{50}$$

the continuity equation for $\mathbf{u}_E^{NZ} + u_z^1\mathbf{e}_z$ yields $\gamma(s) = h^{-1}(s)$.

The non-zonal vorticity field $\boldsymbol{\zeta}$ is then defined by $\boldsymbol{\zeta} = \boldsymbol{\nabla} \times \mathbf{u}_E^{NZ}$, and its vertical component is

$$\zeta_z(s,\varphi) = -\nabla_E^2\left[h^{-1}(s)\Psi(s,\varphi)\right].\tag{51}$$

Finally, the set of equations (17) to (19) becomes

$$\partial_t\left\langle B_s^2\right\rangle = -[\mathbf{u}\cdot\nabla]\left\langle B_s^2\right\rangle - 2s^{-1}\left\langle B_s^2\right\rangle u_s - 2s^{-1}\left\langle B_s^2\right\rangle\partial_\varphi u_\varphi + 2s^{-1}\left\langle B_sB_\varphi\right\rangle\partial_\varphi u_s,\tag{52}$$

$$\partial_t\left\langle B_\varphi^2\right\rangle = -[\mathbf{u}\cdot\nabla]\left\langle B_\varphi^2\right\rangle - 2\left\langle B_\varphi^2\right\rangle\partial_s u_s + 2s\left\langle B_sB_\varphi\right\rangle\partial_s\left(s^{-1}u_\varphi\right),\tag{53}$$

$$\partial_t\left\langle B_sB_\varphi\right\rangle = -[\mathbf{u}\cdot\nabla]\left\langle B_sB_\varphi\right\rangle + s\left\langle B_s^2\right\rangle\partial_s\left(s^{-1}u_\varphi\right) + s^{-1}\left\langle B_\varphi^2\right\rangle\partial_\varphi u_s$$
$$- [\nabla_E\cdot\mathbf{u}]\left\langle B_sB_\varphi\right\rangle.\tag{54}$$

B Numerical model

Fields are discretized in radius on a (possibly irregular) staggered grid (see Figure 10), $s = i_s\Delta_s(s); i_s \in [0, N_s]$; $\omega_g(i_s + 1/2, j)$ and $\tau(i_s, j)$. Ψ, $\{B_s^2\}$ and τ are calculated on the same spatial grid (note that $\{B_s^2\}$ is not defined on the endpoints). The latitudinal part of B_r is discretized on a meridian and every grid point is mapped on the CMB from the grid point on Σ, except at the equator (see Figure 10); thus $\theta = i_\theta\Delta_\theta(\theta); i_\theta \in [0, 2N_s - 1]$.

The azimuthal part of non-zonal fields, Ψ and B_r is expanded in Fourier series:

$$\Psi(s,\varphi,t) = \sum_{m=0}^{m_{max}}\left[a_m(s,t)\cos\left(m\varphi\right) + b_m(s,t)\sin\left(m\varphi\right)\right],\tag{55}$$

$$B_r(\theta,\varphi,t) = \sum_{m=0}^{m_{max}}\left[c_m(\theta,t)\cos\left(m\varphi\right) + d_m(\theta,t)\sin\left(m\varphi\right)\right],\tag{56}$$

in which the number of azimuthal Fourier mode, m_{max}, is related to the number of equidistant grid points in longitude, N_φ: $m_{max} = (N_\varphi - 1)/2$.

The time step being Δ_t, time is discretized using finite differences $t = j\Delta_t; j \in [0, N_t]$.

Spatial derivatives are computed with a finite difference scheme, except for the longitudinal derivatives for which we use the Fast Fourier Transform in order to compute them in spectral space.

For the simulations, the cylindrical radius is discretized in $N_s = 100$ grid points (including the boundaries), the CMB in $N_\theta = 200$ grid points in latitude and $N_\varphi = 33$ (unless otherwise specified) grid points in longitude.
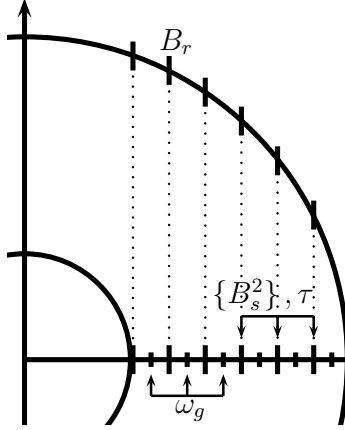
Figure 10. Sketch of a regular staggered radial grid in the equatorial plane, and its projection on the CMB.

C   Steady non-zonal flow model

We define the streamfunction at the top of the core, Ψ_o, as $\Psi_o = \mathsf{M}_s^\theta \Psi$, where subscript o refers to the outer boundary. Let M_s^θ be the operator which projects a vector from the equatorial plane to the top of the core, and let M_θ^s be its transpose. The forward model is the radial component of the induction equation at the top of the core,

$$
\begin{aligned}
\dot{B}_r &= -\nabla_H \cdot \left(\mathbf{u}^{NZ} B_r \right), && (57) \\
&= (\sin\theta\cos\theta)^{-1} \left[\partial_\theta \Psi_o \partial_\varphi B_r - \partial_\varphi \Psi_o \partial_\theta B_r \right] + (\cos\theta)^{-2} \Psi_o \partial_\varphi B_r. && (58)
\end{aligned}
$$

The tangent linear equation is

$$
\begin{aligned}
\delta \dot{B}_r &= (\sin\theta\cos\theta)^{-1} \left[\partial_\theta \Psi_o \partial_\varphi \delta B_r - \partial_\varphi \Psi_o \partial_\theta \delta B_r \right] + (\cos\theta)^{-2} \Psi_o \partial_\varphi \delta B_r && (59) \\
&+ (\sin\theta\cos\theta)^{-1} \left[\partial_\theta \delta\Psi_o \partial_\varphi B_r - \partial_\varphi \delta\Psi_o \partial_\theta B_r \right] + (\cos\theta)^{-2} \delta\Psi_o \partial_\varphi B_r, \\
\partial_t \delta B_r &= \delta \dot{B}_r. && (60)
\end{aligned}
$$

We introduce the adjoint variables Ψ^T, B_r^T and $\dot{B}_r^{\,T}$ for Ψ, B_r and $\dot{B}_r$ respectively. The adjoint model is

$$
\begin{aligned}
\Psi^T &= \mathsf{M}_\theta^s \sum_{j=0}^{T} \left\{ \partial_\theta^T \left[(\sin\theta\cos\theta)^{-1} \partial_\varphi B_{r_j} \dot{B}_{r_j}^{\,T} \right] - \partial_\varphi^T \left[(\sin\theta\cos\theta)^{-1} \partial_\theta B_{r_j} \dot{B}_{r_j}^{\,T} \right] \right. && (61) \\
&\quad \left. + (\cos\theta)^{-2} \partial_\varphi B_{r_j} \dot{B}_{r_j}^{\,T} \right\} \cos^2\theta, \\
B_r^T &= \partial_\varphi^T \left[(\sin\theta\cos\theta)^{-1} \partial_\theta \Psi_o \dot{B}_r^{\,T} \right] - \partial_\theta^T \left[(\sin\theta\cos\theta)^{-1} \partial_\varphi \Psi_o \dot{B}_r^{\,T} \right] && (62) \\
&\quad + \partial_\varphi^T \left[(\cos\theta)^{-2} \Psi_o \dot{B}_r^{\,T} \right], \\
\dot{B}_r^{\,T} &= -\partial_t B_r^T,
\end{aligned}
$$

23

where ∂_φ^T and ∂_θ^T are the adjoints of ∂_φ and ∂_θ^T, and $-\partial_t$ indicates that the integration is performed backward in time.

D   Alfvén waves model

The forward equations are:

$$\partial_t \omega_g = \left[s^3 h \right]^{-1} \partial_s \tau, \tag{63}$$

$$\partial_t \tau = s^3 h \left\{ B_s^2 \right\} \partial_s \omega_g, \tag{64}$$

$$\partial_t B_r = -\mathsf{M}_s^\theta \mathsf{M}_{\mathrm{zonal}}^{\mathrm{total}} \omega_g \partial_\varphi B_r, \tag{65}$$

where $\mathsf{M}_{\mathrm{zonal}}^{\mathrm{total}}$ transforms a one-dimensional zonal vector into a two-dimensional one, by duplicating it in each meridional plane. Let $\mathsf{M}_{\mathrm{total}}^{\mathrm{zonal}}$ be its transpose. The forward model is completed by the boundary conditions (41) and (42).

We define the adjoint variables $\omega_g^T, \tau^T, \{B_s^2\}^T, B_r^T$ for $\omega_g, \tau, \{B_s^2\}, B_r$ respectively. The adjoint model is

$$\partial_t \omega_g^T = \partial_s^T \left[s^3 h \left\{ B_s^2 \right\} \tau^T \right] - \mathsf{M}_{\mathrm{total}}^{\mathrm{zonal}} \mathsf{M}_\theta^s \left[\partial_\varphi B_r B_r^T \right], \tag{66}$$

$$\partial_t \tau^T = \partial_s^T \left[\left(s^3 h \right)^{-1} \omega_g^T \right], \tag{67}$$

$$F^T(s) = \sum_j s^3 h \tau_j^T \left\{ B_s^2 \right\} (\partial_s \omega_g)_j + \alpha_C \mathbf{W} \left\{ B_s^2 \right\}, \tag{68}$$

$$\partial_t B_r^T = -\partial_\varphi^T \left[\mathsf{M}_s^\theta \mathsf{M}_{\mathrm{zonal}}^{\mathrm{total}} \omega_g B_r^T \right], \tag{69}$$

where ∂_s^T is the adjoint of the operator ∂_s and the term in α_C corresponds to the extra penalty term in the misfit function (see also equation 23); $\mathbf{W} = \partial_s^T \partial_s$ in the experiments. In order to enforce its positivity during the optimization phase, $\{B_s^2\}$ is rather written $\{B_s^2\} = \exp[F(s)]$, with $F \in \mathbb{R}$ and F^T computed as indicated in equation (68).

References

M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover, New York, 9th Dover printing, 10th GPO printing edition, 1964.

H. Amit, P. Olson, and U. Christensen. Tests of core flow imaging methods with numerical dynamos. *Geophys. J. Int.*, 168:27–39, 2007.

G. Backus, R.L. Parker, and C. Constable. *Foundations of Geomagnetism.* Cambridge University Press, 1996.

G.E. Backus. Kinematics of geomagnetic secular variation in a perfectly conducting core. *Phil. Trans. R. Soc. Lond. Series A*, 263:239–266, 1968.

S.I. Braginsky. Torsional magnetohydrodynamic vibrations in the Earth's core and variations in day length. *Geomag. Aeron.*, 10:1–10, 1970.

B.A. Buffett, J. Mound, and A. Jackson. Inversion of torsional oscillations for the structure and dynamics of Earth's core. *Geophys. J. Int.*, 177:878–890, 2009.

J. Chen. Global mass balance and the length-of-day variation. *J. Geophys. Res.*, 110, 2005. doi: 10.1029/2004JB003474.

U.R. Christensen and J. Aubert. Scaling properties of convection-driven dynamos in rotating spherical shells and application to planetary magnetic fields. *Geophys. J. Int.*, 166: 97–114, 2006.

U.R. Christensen and J. Wicht. Numerical Dynamo Simulations. In *Treatise on Geophysics*, volume 8: Core Dynamics, pages 245–282. ed. Olson, P., Elsevier, Oxford, 2007.

U.R. Christensen, V. Holzwarth, and A. Reiners. Energy flux determines magnetic field strength of planets and stars. *Nature*, 457:167–169, 2009. doi: 10.1038/nature070626.

P. Courtier. Variational methods. *J. of the Meteor. Soc. Jap.*, 75:101–108, 1997.

P. Courtier and O. Talagrand. Variational assimilation of meteorological observations with the adjoint vorticity equation. II: Numerical results. *Quarterly Journal of the Royal Meteorological Society*, 113(478), 1987.

P.H. Diamond, S.I. Itoh, K. Itoh, and T.S. Hahm. Zonal flows in plasma: a review. *Plasma Phys. Control. Fusion*, 47:35–161, 2005.

G. Evensen. *Data assimilation: The ensemble Kalman filter.* Springer Verlag, 2007.

C. Eymin and G. Hulot. On core surface flows inferred from satellite magnetic data. *Phys. Earth Planet. Inter.*, 152:200–220, 2005.

A. Fournier, C. Eymin, and T. Alboussière. A case for variational geomagnetic data assimilation: insights from a one-dimensional, nonlinear, and sparsely observed MHD system. *Nonlin. Proc. in Geophysics*, 14:163–180, 2007.

R. Giering and T. Kaminski. Recipes for adjoint code construction. *ACM Transactions on Mathematical Software*, 24:437–474, 1998.

J.C. Gilbert and C. Lemaréchal. Some numerical experiments with variable-storage quasi-Newton algorithms. *Mathematical Programming*, 45:407–435, 1989.

N. Gillet, M.A. Pais, and D. Jault. Ensemble inversion of time-dependent core flow models. *Geochem. Geophys. Geosyst.*, 10,Q06004, 2009. doi: 10.1029/2008GC002290.

G.A. Glatzmaier and P.H. Roberts. A three-dimensional self-consistent computer simulation of a geomagnetic field reversal. *Nature*, 377:203–209, 1995.

R.S. Gross. Earth Rotation Variations - Long Period. In *Treatise on Geophysics*, volume 0: Geodesy, pages 239–294. ed. Herring, T., Elsevier, Oxford, 2007.

H. Hersbach. Application of the adjoint of the WAM model to inverse wave modeling. *J. Geophys. Res.*, 103(C5):10469–10488, 1998.

R. Hide. Free hydromagnetic oscillations of the Earth's core and the theory of the geomagnetic secular variation. *Phil. Trans. R. Soc. Lond. Series A, Mathematical and Physical Sciences*, 259:615–647, 1966.

R. Holme. Large-Scale Flow in the Core. In *Treatise on Geophysics*, volume 8: Core Dynamics, pages 107–130. ed. Olson, P., Elsevier, Oxford, 2007.

R. Holme and N. Olsen. Core surface flow modelling from high-resolution secular variation. *Geophys. J. Int.*, 166:518–528, 2006.

G. Hulot, T. Sabaka, and N. Olsen. The Present Field. In *Treatise on Geophysics*, volume 5: Geomagnetism, pages 33 – 75. ed. Kono, M., Elsevier, Oxford, 2007.

K. Ide, P. Courtier, M. Ghil, and C. Lorenc. Unified notations for data assimilation: Operational, sequential and variational. *J. of the Meteor. Soc. Jap.*, 75:181–189, 1997.

A. Jackson and C. Finlay. Geomagnetic Secular Variation and its Applications to the Core. In *Treatise on Geophysics*, volume 5: Geomagnetism. ed. Kono, M., Elsevier, Oxford, 2007.

D. Jault. Axial invariance of rapidly varying diffusionless motions in the Earth's core interior. *Phys. Earth Planet. Inter.*, 166:67–76, 2008.

D. Jault, C. Gire, and J. L. Le Mouël. Westward drift, core motions and exchanges of angular momentum between core and mantle. *Nature*, 333:353–356, 1988.

E. Kalnay, M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, et al. The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, 77:437–471, 1996.

F-X. Le Dimet and O. Talagrand. Variational methods for analysis and assimilation in meteorological observations. *Tellus*, 38:97–110, 1986.

B. Lehnert. Magnetohydrodynamic waves under the action of the Coriolis force. *Astrophys. J.*, 119:647–654, 1954.

D. Liu, A. Tangborn, and W. Kuang. Observing system simulation experiments in geomagnetic data assimilation. *J. Geophys. Res.*, 112, 2007. doi: 10.1029/2006JB004691.

N. Olsen and M. Mandea. Rapidly changing flows in the Earth's core. *Nature Geoscience*, 1:390, 2008.

N. Olsen, H. Lühr, T.J. Sabaka, M. Mandea, M. Rother, L. Tøffner-Clausen, and S. Choi. CHAOS-a model of the Earth's magnetic field derived from CHAMP, Ørsted, and SAC-C magnetic satellite data. *Geophys. J. Int.*, 166:67–75, 2006.

M.A. Pais and D. Jault. Quasi-geostrophic flows responsible for the secular variation of the Earth's magnetic field. *Geophys. J. Int.*, 173:421–443, 2008.

M.A. Pais, O. Oliveira, and F. Nogueira. Nonuniqueness of inverted core-mantle boundary flows and deviations from tangential geostrophy. *J. Geophys. Res.*, 109, 2004. doi: 10.1029/2004JB003012.

E. Plaut. Nonlinear dynamics of traveling waves in rotating Rayleigh-Bénard convection: Effects of the boundary conditions and of the topology. *Phys. Rev. E*, 67:046303, 2003.

S. Rau, U. Christensen, A. Jackson, and J. Wicht. Core flow inversion tested with numerical dynamo models. *Geophys. J. Int.*, 141:485–497, 2000.

P.H. Roberts. *An introduction to magnetohydrodynamics*. Longmans, London, 1967.

Z. Sun, A. Tangborn, and W. Kuang. Data assimilation in a sparsely observed one-dimensional modeled MHD system. *Nonlin. Proc. in Geophysics*, 14:181–192, 2007.

O. Talagrand. Assimilation of observations, an introduction. *J. of the Meteor. Soc. Jap.*, 75:191–209, 1997.

O. Talagrand and P. Courtier. Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Q. J. R. Meteorol. Soc.*, 113:1311–1328, 1987.

S.M. Tobias, P.H. Diamond, and D.W. Hughes. β-plane magnetohydrodynamic turbulence in the solar tachocline. *Astrophys. J.*, 667:L113–L116, 2007.

G.M. Turner, J.L. Rasson, and C.V. Reeves. Observation and Measurement Techniques. In *Treatise on Geophysics*, volume 5: Geomagnetism, pages 33 – 75. ed. Kono, M., Elsevier, Oxford, 2007.

G. K. Vallis. *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-Scale Circulation.* Cambridge University Press, Cambridge, U.K., 2006.

C.V. Voorhies and G.E. Backus. Steady flows at the top of the core from geomagnetic field models: The steady motions theorem. *Geophys. Astrophys. Fluid Dyn.*, 32:163–173, 1985.

R. Waddington, D. Gubbins, and N. Barber. Geomagnetic field analysis-V. Determining steady core-surface flows directly from geomagnetic observations. *Geophys. J. Int.*, 122: 326–350, 1995.

P. Wessel and W.H.F. Smith. Free software helps map and display data. *Eos, Transactions American Geophysical Union*, 72:441–441, 1991.

S. Zatman and J. Bloxham. Torsional oscillations and the magnetic field within the Earth's core. *Nature*, 388:760–763, 1997.

S. Zatman and J. Bloxham. On the dynamical implications of models of B_s in the Earth's core. *Geophys. J. Int.*, 138:679–686, 1999.