

Chapter 3: The inverse of mantle flow with velocity and topography constraints

August 23, 2017

Abstract

Constraining the rheological parameters of the mantle is important to estimate the broad-scale forces in the mantle. Moreover, constraining the mantle's rheology is important as it allows for better estimation of the shear and normal stresses at plate boundaries. Inferring the constitutive parameters of the mantle requires minimizing a misfit between model output and observed data such as plate motion. However, plate velocities are sensitive to the reference frame and assume the observational area is rigid, which poses a problem near and at plate boundaries. Therefore, some form of surface deformation data needs to be incorporated into a misfit to better infer plate coupling. Dynamic topography is an available piece of data that can aid in inferring plate coupling. With this data in hand, we will formulate the cost function and derive the adjoint system with dynamic topography as a constraint. We will derive the adjoint forcing term with surface velocity and surface normal stress. Furthermore, we analyze simple case studies of a sinking mass anomaly where we infer the a weak layer prefactor, strain rate exponent and activation energy and discuss the limitations of this method in reference to subduction zones when using a free-slip boundary condition.

Introduction

Slab pull is likely the primary force driving plate motions [11, 7] and is estimated to account for approximately 70% of the driving forces in the mantle [8]. An important constraint in models is reproducing the asymmetric motion at subduction zones. Without the direct coupling of slabs to the overriding plate, there would be symmetric convergence at subduction zones [8]. Ridge push and slabs in the lower mantle (slab suction) may also be important as they act as a 'push force' and a resisting force respectively. Some key resisting forces include a bending force, when slabs first bend during subduction, and the frictional resistance from faults that all act to reduce plate motion. Slabs may act as stress-guides [20], allowing stresses to propagate from the lower mantle to the oceanic lithosphere. Nevertheless, the question as to how strong slabs are compared to the ambient mantle remains open.

Accurately estimating broad-scale forces requires models that contain the salient physics in addition to the necessary resolution to resolve the fine scale features of mantle flow models. The correct rheology would incorporate shear thinning due to dislocation creep in the upper mantle [14] and dynamic weakening, which is controlled by the yield stress [20, 2, 1, 5, 4]. Each of the rheological parameters, (strain rate exponent, yield stress and plate coupling), plays a key role in the amount of viscous dissipation in the mantle. Viscous dissipation occurs as plates overcome the bending force and is also tied to how strong slabs are (weak slabs promote less dissipation compared to strong slabs).

In addition to a non-Newtonian rheology, thermal boundary layers, slabs and fault zones need to have the appropriate resolution. Properly resolving models needs requires either refining with a uniform mesh, which can be very costly, or using adaptive mesh refinement (AMR). Recently, with the use of adaptive mesh refinement (AMR), there have been spherical models that incorporate non-Newtonian rheology along with fine-scale resolution of fault zones and slabs [20, 2] which can produce the complex motions of both large-scale and micro plates.

While high-resolution models with the correct rheology are important to constraining broad-scale forces, they still fall short due to a mismatch in predicted plate motion and observed motion. To minimize the misfit between models and observations requires solving an optimization problem with plate motion data [18] where the constitutive parameters are inferred. It was shown that using plate motion data, we can constrain the model parameters such as the plate couplings, strain rate exponent and yield stress. However, there is a limit to the amount of information that can be gleaned from plate motions [18] due to sensitivity between rheological parameters and plate motion.

Using plate motion data can give first order limits on the mantle rheology since there is a first order relationship between plate motions and the rheology. However, plate motion data is not unique with regard to the reference frames [12, 3] and can potentially change the inferences of parameters. A caveat of using plate motion data is the assumption of rigid plates, that is there is no deformation (strain rate is negligible), which implies that plate motion data cannot be used near trenches since there is significant deformation there [16]. Therefore, solely using plate motion data can potentially lead to poorly constrained plate couplings.

While plate motion data can fall short of sufficiently constraining plate coupling, there are other available observations to constrain the effective viscosity such as dynamic topography, free-air gravity anomalies, and plate strain. Dynamic topography is one of main observations that is used to compare model results with observations. It is known that dynamic topography is correlated with density anomalies such as plumes and slabs [13]. In addition to the positive correlation with the geoid, the magnitude of dynamic topography is less than 1km, but is very wavelength dependent [10]. Using dynamic topography as a constraint, there have been studies that have focused on minimizing the misfit between predicted from dynamic topography and residual topography constraints [23].

While it is important to include the correct density distribution to capture the long wavelength structure, it is still important to have the correct effective viscosity structure to constrain the surface normal stresses. To this end, the incorporation of lateral variations in

viscosity is important to reproduce the dynamic topography signal [17]. While the lateral variations in viscosity are important, the radial distribution of the effective viscosity is just as important due to the 30-100x increase in viscosity from the upper to lower mantle that provides a good fit to the geoid [15].

There have been studies to constrain the short-wavelength signal at subduction zones, such as using a weak mantle wedge [6]; however, the rheology does not include dynamic weakening and shear thinning. Recently, there have been advances in numerical solutions that incorporate rheology with the salient physics and they are able to produce they dynamic weakening at hinge zones and shear thinning in the upper mantle. However, with the correct physics, the mantle properties have not been constrained [20, 2] which leads to data misfit in both plate motions and dynamic topography. There have been strides to constrain the rheology of the mantle using data in an optimization framework [22, 18] to recover the rheological parameters. While constraining plate motion would constrain the rheology to a first order such as the strain rate exponent and yield stress, the plate coupling between plates may not be well constrained. Therefore, the incorporation of dynamic topography is an important constraint as this piece of data reflects the surface fluctuation due to mantle flow. Therefore, dynamic topography data is important because it gives a stronger local constraint on the deforming region at trenches unlike plate motions which can only inform rigid areas. Better estimates of plate coupling which can contribute to a better understanding of coupling and where great earthquakes occur [19].

In this chapter we will derive the adjoint system with plate velocities and surface normal stress data. We will implement this new adjoint systems for a simple test case of a sinking mass anomaly where there is a smooth surface normal stress signal. Furthermore, will present examples of the recovery various parameters such as the strain rate exponent, weak factor and activation energy. We will show that the gradients for this new adjoint formulation are consistent and can thus be used for consistent plate coupling inferences. Lastly, we will present the current limitation of applying this method to realistic subduction zone models.

Forward Model

We model mantle flow with the infinite Prandtl-number Boussinesq approximation, which leads to the following non-dimensional Stokes equations:

$$\nabla \cdot \boldsymbol{\sigma} = -\text{Ra} T \mathbf{e}_r \quad \text{on } \Omega, \quad (1a)$$

$$\nabla \cdot \mathbf{u} = 0 \quad \text{on } \Omega, \quad (1b)$$

where Ω is the mantle domain,

$\boldsymbol{\sigma} = 2\eta\dot{\boldsymbol{\epsilon}}(\mathbf{u}) - p\mathbf{I}$ is the stress tensor with the viscosity $\eta = \eta(\dot{\epsilon}_{\text{II}}, \Gamma, T, \sigma_y)$, which depends on the velocity \mathbf{u} (through the second invariant of the strain tensor $\dot{\epsilon}_{\text{II}}$ defined below), on multiplicative factor Γ modelling plate boundaries, on the temperature T and on the yield stress σ_y , (a global quantity). Moreover, $\dot{\boldsymbol{\epsilon}}(\mathbf{u}) := \frac{1}{2}(\nabla\mathbf{u} + \nabla\mathbf{u}^T)$ is the strain rate tensor, p is the pressure and \mathbf{I} the identity tensor. The Stokes equations (1) are driven by thermal

buoyancy. Here, $Ra = \frac{\rho g \alpha \Delta T D^3}{\kappa \eta_{\text{ref}}}$ is the thermal Rayleigh number, where ρ is the density of the mantle, g is the gravitational acceleration, α is the thermal diffusivity, ΔT is the temperature difference, D is the length scale, η_{ref} is the reference viscosity, and κ is the thermal diffusivity. We define the second invariant of the strain rate tensor as $\dot{\epsilon}_{\text{II}} = \frac{1}{2}[\text{tr}(\dot{\epsilon}^2(\mathbf{u})) - \text{tr}(\dot{\epsilon}(\mathbf{u}))]$.

No normal flow and free-slip tangential conditions on the boundary $\partial\Omega$ of Ω are used, i.e.,

$$\mathbf{u} \cdot \mathbf{n} = 0, \quad \mathbf{T}(\boldsymbol{\sigma}\mathbf{n}) = 0. \quad \text{on } \partial\Omega \quad (2)$$

Here, we use the tangential operator for the Neumann condition defined as $\mathbf{T} = \mathbf{I} - \mathbf{n} \otimes \mathbf{n}$ is the projection onto the tangential direction. In particular, plate velocities on the top are not imposed but are an outcome of model calculations.

In the following, we prefer to work with the weak (variational) form of the Stokes equations (1). This weak form is derived by multiplying (1a) and (1b) by arbitrary functions \mathbf{v} and q , respectively, which are assumed to be sufficiently smooth, and satisfy the equivalent Dirichlet boundary condition, $\mathbf{v} \cdot \mathbf{n} = 0$. Using integration by parts and the boundary conditions (2), this results in

$$\int_{\Omega} 2\eta(\dot{\epsilon}_{\text{II}}, \Gamma, n, \sigma_y) \dot{\epsilon}(\mathbf{u}) : \dot{\epsilon}(\mathbf{v}) d\Omega - \int_{\Omega} p \nabla \cdot \mathbf{v} d\Omega - \int_{\Omega} q \nabla \cdot \mathbf{u} d\Omega = \int_{\Omega} Ra T \mathbf{e}_r \cdot \mathbf{v} d\Omega. \quad (3)$$

On geological time scales, the mantle behaves like a viscous fluid from thermally activated creep. The viscosity strongly depends on temperature, and this dependence can be represented by an Arrhenius-type law. In the upper mantle, dislocation creep likely dominates over diffusion creep [21]. Although one can prescribe the rheology as composite [5, 20] such that both, diffusion and dislocation creep can play a role depending on the state of stress and the strain rate, we have found that dislocation creep dominates within the plates and slabs and hence is the deformation mechanism which likely controls plate motions. Thus, underlying our models is a temperature-dependent shear-thinning rheology,

$$\tilde{\eta}(\dot{\epsilon}_{\text{II}}, T) = \Gamma a(T) \dot{\epsilon}_{\text{II}}^{\frac{1-n}{2n}}, \quad \text{with } \Gamma(x) = 1 - \sum_i (1 - \Gamma_i) \chi_i(x),$$

where $a(T) := A_{\text{rad}} \exp(\beta(0.5 - T))$, and $\chi_i(\cdot)$ are characteristic functions for individual plate boundaries, i.e., a function with value 1 at the (volumetrically modeled) plate boundary, and a value of 0 away from the plate boundary. The strength/weakness of the coupling along plate boundaries is controlled by the weakening factors $\Gamma_i > 0$. Plate decoupling occurs over long time scales within seismogenic zones, where great earthquakes typically occur. The degree of frictional resistance that occurs along the seismogenic zone is controlled by the factors Γ_i : small values of Γ_i give rise to weakly coupled plate boundaries, while larger values enforce stronger coupling. Plate boundaries require high spatial resolution in computational models, and the coupling factors Γ_i will act as parameters in the inversion.

An important aspect of the mantle rheology is dynamic weakening through shear thinning, in particular near hinge zones. Thus, we use a rheology that involves plastic yielding additionally to polynomial shear thinning. For computational reasons we also incorporate

lower and upper viscosity bounds $0 < \eta_{\min} < \eta_{\max}$ in the rheology, such that the (effective) viscosity is

$$\eta(\dot{\epsilon}_{\text{II}}, \Gamma, n, \sigma_y) = \eta_{\min} + \min(\Gamma \min(\eta_{\max}, a(T)(\dot{\epsilon}_{\text{II}} - d)^{\frac{1}{2n}} \dot{\epsilon}_{\text{II}}^{-\frac{1}{2}}), \frac{1}{2} \sigma_y \dot{\epsilon}_{\text{II}}^{-\frac{1}{2}}). \quad (4)$$

The choice (4) for the effective viscosity corresponds to first applying the upper viscosity bound to the temperature and strain rate dependent viscosity. This is followed by the multiplication with $\Gamma(x)$, a function describing plate boundaries through low viscosity zones. Finally, the plastic yielding condition is imposed. Adding η_{\min} enforces a lower bound on the viscosity, as well as a one-to-one correspondence between strain rate and stress in the case of plastic yielding. In (4), we use a shift $d \geq 0$ to ensure that η_{\max} is incorporated in a way that the viscosity is continuously differentiable with respect to $\dot{\epsilon}_{\text{II}}$, and thus also with respect to the velocity.

An important model outcome of mantle flow models is the surface normal stress (σ_{rr}), or dynamic topography at trenches. We obtain the dynamic topography from the solution of (1) as a post-processing step in (5),

$$h = \frac{\sigma_{rr}}{\rho g} \quad (5)$$

where h is the dynamic topography, g gravity, and ρ the density.

Bayesian Inverse Problem

With the addition of the cost function of the surface normal stress, we can also formulate the Bayesian Inverse problem as follows:

$$\boldsymbol{\pi}_{\text{post}} \propto \boldsymbol{\pi}_{\text{likelihood}} \boldsymbol{\pi}_{\text{prior}} \quad (6)$$

with the likelihood distribution given,

$$\begin{aligned} \boldsymbol{\pi}_{\text{likelihood}} \propto & \exp(-\mathcal{J}) \\ & \exp\left\{-\frac{1}{2} \int_{\Omega_1} (\mathcal{O}\mathbf{u} - \mathbf{u}_{\text{obs}})^T \mathcal{C}_{vel}^{-1} (\mathcal{O}\mathbf{u} - \mathbf{u}_{\text{obs}}) d\Omega_1 \right. \\ & \left. + \frac{1}{2} \int_{\Omega_2} (\mathcal{O}\boldsymbol{\sigma}_n - \boldsymbol{\sigma}_{\text{obs}})^T \mathcal{C}_{stress}^{-1} (\mathcal{O}\boldsymbol{\sigma} - \boldsymbol{\sigma}_{\text{obs}}) d\Omega_2 \right\} \end{aligned} \quad (7)$$

while the prior distribution $\boldsymbol{\pi}_{\text{prior}}$ is

$$\boldsymbol{\pi}_{\text{prior}} \propto \exp\left\{-\frac{1}{2} (\mathbf{m} - \mathbf{m}_0)^T \mathcal{C}_{\text{prior}}^{-1} (\mathbf{m} - \mathbf{m}_0)\right\}. \quad (8)$$

Typically, \mathbf{m}_0 is the mean value, (usually chosen as a reasonable parameter value), while $\mathcal{C}_{\text{prior}}^{-1}$ is the covariance distribution of each parameter. In most cases (especially in this

chapter), the prior term is a Gaussian distribution due to the ease of drawing samples and the smoothness of the distribution. An important aspect of the likelihood model is the incorporation of noise in the data or,

$$\mathbf{f} = \mathbf{u}_{obs} + \mathcal{N}(0, \mathcal{C}_{noise}) \quad (9)$$

where we assume a normal distribution for the noise in the data (zero mean and covariance \mathcal{C}_{noise}). Our data misfit function will now include both plate velocities and surface normal stresses, with a total cost function

$$\begin{aligned} \mathcal{J}(\mathbf{u}, \mathbf{m}, p) &:= \mathcal{J}_u + \mathcal{J}_\sigma \\ \mathcal{J}(\mathbf{u}, \mathbf{m}, p) &:= \frac{1}{2} \int_{\partial\Omega_1} (\mathcal{O}\mathbf{u} - \mathbf{u}_{obs})^T \mathcal{C}_{vel}^{-1} (\mathcal{O}\mathbf{u} - \mathbf{u}_{obs}) d\partial\Omega_1 \\ &\quad + \frac{1}{2} \int_{\partial\Omega_2} (\mathcal{O}\boldsymbol{\sigma}_n - \boldsymbol{\sigma}_{obs})^T \mathcal{C}_{stress}^{-1} (\mathcal{O}\boldsymbol{\sigma} - \boldsymbol{\sigma}_{obs}) d\partial\Omega_2 \end{aligned} \quad (10)$$

where first term on the right hand side is the surface velocities misfit as previously used in [18]. The cost function for the second term on the right hand side is for surface normal stress with $\boldsymbol{\sigma}_n = \mathbf{n}(\boldsymbol{\sigma}\mathbf{n})$ being the normal stress, and \mathcal{O} is the observation operator (\mathcal{O} retrieves the model observations at physical points in space). For our test problems, \mathcal{O} is the same for \mathbf{u} and $\boldsymbol{\sigma}_{rr}$; however, this is not necessarily the case for geophysical problems, as they could be different.

We have shown that constraining plate motion can give a strong constraint on the rheological properties of the mantles such as the strain rate exponent, yield stress and plate couplings. However, the surface normal stress at trenches might provide refined estimates as velocity data was not included in the deforming regions near the trench. To constrain the rheological parameters, we first solve for the maximum a posteriori point (MAP) by solving the PDE-constrained optimization problem,

$$\min_{\mathbf{m}} \mathcal{J}(\mathbf{u}, \mathbf{m}, p) \quad (11)$$

subject to (1). We first construct the Lagrangian,

$$\begin{aligned} \mathcal{L}(\mathbf{u}, p, \mathbf{v}, q, \mathbf{m}) &= \mathcal{J}(\mathbf{u}, \mathbf{m}, p) + \int_{\Omega} 2\eta(\dot{\epsilon}_{II}, \Gamma, n, \sigma_y) \dot{\epsilon}(\mathbf{u}) : \dot{\epsilon}(\mathbf{v}) d\Omega \\ &\quad + \int_{\Omega} p \nabla \cdot \mathbf{v} d\Omega - \int_{\Omega} q \nabla \cdot \mathbf{u} d\Omega - \int_{\Omega} \text{Ra} T \mathbf{e}_r \cdot \mathbf{v} d\Omega. \end{aligned} \quad (12)$$

Taking variations of (12) with respect to the adjoint variables (\mathbf{v}, q) recovers the forward problem, while derivatives with respect to the forward variables (\mathbf{u}, p) yields the adjoint equations,

$$\begin{aligned} \nabla \cdot \mathbf{v} &= 0 & \text{on } \Omega, \\ \nabla \cdot \hat{\boldsymbol{\sigma}}_{\mathbf{u}} &= 0 & \text{on } \Omega, \end{aligned} \quad (13)$$

with boundary conditions

$$\begin{aligned} \mathbf{v} \cdot \mathbf{n} &= 0 \quad \text{on } \partial\Omega, \\ \mathbf{T}(\hat{\boldsymbol{\sigma}}_{\mathbf{u}}\mathbf{n}) &= \begin{cases} 0 & \text{on } \partial\Omega \setminus \partial\Omega_t, \\ -\mathcal{O}^T \mathcal{C}_{\text{noise}}^{-1}(\mathcal{O}\mathbf{u} - \mathbf{u}_{\text{obs}}) & \text{on } \partial\Omega_t, \end{cases} \end{aligned}$$

where $\hat{\boldsymbol{\sigma}}_{\mathbf{u}} = \hat{\boldsymbol{\sigma}}_{\mathbf{u}}(\mathbf{v}, q)$ is the adjoint stress tensor defined by

$$\hat{\boldsymbol{\sigma}}_{\mathbf{u}} = 2 \left(\eta(\dot{\epsilon}_{\text{II}}, \Gamma, n, \sigma_y) \mathbb{I} + \frac{1}{2} \eta_{,\dot{\epsilon}_{\text{II}}} [\dot{\boldsymbol{\epsilon}}(\mathbf{u}) \otimes \dot{\boldsymbol{\epsilon}}(\mathbf{u})] \right) \dot{\boldsymbol{\epsilon}}(\mathbf{v}) - q \mathbf{I} \quad (14)$$

with \mathbb{I} being the fourth-order identity tensor, and $\eta_{,\dot{\epsilon}_{\text{II}}}$ given by

$$\eta_{,\dot{\epsilon}_{\text{II}}} = \begin{cases} \min \left(0, \frac{1}{2} \Gamma a(T) (\dot{\epsilon}_{\text{II}} - d)^{\frac{1}{2n}} \dot{\epsilon}_{\text{II}}^{-\frac{1}{2}} \frac{\dot{\epsilon}_{\text{II}} - (\dot{\epsilon}_{\text{II}} - d)n}{\dot{\epsilon}_{\text{II}}(\dot{\epsilon}_{\text{II}} - d)n} \right) & \text{in } \Omega \setminus \Omega_y \\ -\frac{1}{2} \sigma_y \dot{\epsilon}_{\text{II}}^{-\frac{3}{2}} & \text{in } \Omega_y. \end{cases} \quad (15)$$

Incorporating surface normal stress into the adjoint formulation requires taking variations of (16) with respect to (\mathbf{u}, p) would add an extra forcing term to the adjoint system since there is a misfit in surface normal stress that needs to be minimized.

$$\mathcal{J}_{\boldsymbol{\sigma}} := \frac{1}{2} (\mathcal{O}\boldsymbol{\sigma}_n - \boldsymbol{\sigma}_{\text{obs}})^T \mathcal{C}_{\text{topog}}^{-1} (\mathcal{O}\boldsymbol{\sigma}_n - \boldsymbol{\sigma}_n). \quad (16)$$

The derivative of (16) is

$$\frac{\partial \mathcal{J}_{\boldsymbol{\sigma}}}{\partial(\mathbf{u}, p)} = \mathcal{O}^T \frac{\partial \boldsymbol{\sigma}_n}{\partial(\mathbf{u}, p)} (\mathcal{O}\boldsymbol{\sigma}_n - \boldsymbol{\sigma}_{\text{obs}}) \quad (17)$$

The gradient of (16) w.r.t. \mathbf{u} in a direction $\tilde{\mathbf{u}}$ is,

$$\begin{aligned} \frac{\partial \mathcal{J}_{\boldsymbol{\sigma}}}{\partial \mathbf{u}} \tilde{\mathbf{u}} &= 2 [\eta \dot{\boldsymbol{\epsilon}}(\tilde{\mathbf{u}}) + \frac{\partial \eta}{\partial \mathbf{u}} \dot{\boldsymbol{\epsilon}}(\mathbf{u}) \tilde{\mathbf{u}}] \\ &= 2 \left(\eta(\dot{\epsilon}_{\text{II}}, \Gamma, n, \sigma_y) \mathbb{I} + \frac{1}{2} \eta_{,\dot{\epsilon}_{\text{II}}} [\dot{\boldsymbol{\epsilon}}(\mathbf{u}) \otimes \dot{\boldsymbol{\epsilon}}(\mathbf{u})] \right) \dot{\boldsymbol{\epsilon}}(\tilde{\mathbf{u}}) \end{aligned} \quad (18)$$

while the gradient of (16) w.r.t. the forward pressure p in the direction \tilde{p} is,

$$\frac{\partial \mathcal{J}_{\boldsymbol{\sigma}}}{\partial p}(\tilde{p}) = -\tilde{p} \mathbf{I} \quad (19)$$

and

$$\frac{\partial \mathcal{J}_{\boldsymbol{\sigma}}}{\partial(\mathbf{u}, p)}(\tilde{\mathbf{u}}, \tilde{p}) = 2 \left(\eta(\dot{\epsilon}_{\text{II}}, \Gamma, n, \sigma_y) \mathbb{I} + \frac{1}{2} \eta_{,\dot{\epsilon}_{\text{II}}} [\dot{\boldsymbol{\epsilon}}(\mathbf{u}) \otimes \dot{\boldsymbol{\epsilon}}(\mathbf{u})] \right) \dot{\boldsymbol{\epsilon}}(\tilde{\mathbf{u}}) - \mathbf{I} \tilde{p} \quad (20)$$

We will make use of the following identities:

$$\begin{aligned} \mathbb{I} \cdot \mathbf{n} &= \mathbf{n} \\ \mathbf{n}(\mathbb{I} \cdot \mathbf{n}) &= \mathbf{n} \cdot \mathbf{n} \\ (\mathbf{A} \otimes \mathbf{B})\mathbf{n} &= \mathbf{A}(\mathbf{B}\mathbf{n}) \\ \mathbf{n}(\mathbf{A} \otimes \mathbf{B})\mathbf{n} &= (\mathbf{n}\mathbf{A}) \cdot (\mathbf{B}\mathbf{n}) \end{aligned} \quad (21)$$

where \mathbb{I} is the fourth order Identity tensor, \mathbf{A} and \mathbf{B} are second order tensors and \mathbf{n} is a vector. Thus, $\mathbf{n} \frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{u}, p} \mathbf{n}$ is,

$$\begin{aligned}
\mathbf{n} \frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{u}, p} \mathbf{n} &= \mathbf{n} \left[2 \left(\eta(\dot{\epsilon}_{\text{II}}, \Gamma, n, \sigma_y) \mathbb{I} + \frac{1}{2} \eta_{, \dot{\epsilon}_{\text{II}}} [\dot{\epsilon}(\mathbf{u}) \otimes \dot{\epsilon}(\mathbf{u})] \right) - \mathbf{I} \right] \mathbf{n} \\
&= 2 \left(\eta(\dot{\epsilon}_{\text{II}}, \Gamma, n, \sigma_y) \mathbf{n} \cdot \mathbf{n} + \frac{1}{2} \eta_{, \dot{\epsilon}_{\text{II}}} \mathbf{n} [\dot{\epsilon}(\mathbf{u}) \otimes \dot{\epsilon}(\mathbf{u})] \mathbf{n} \right) - \mathbf{n} \cdot \mathbf{n} \\
&= 2 \left(\eta(\dot{\epsilon}_{\text{II}}, \Gamma, n, \sigma_y) \mathbf{n} \cdot \mathbf{n} + \frac{1}{2} \eta_{, \dot{\epsilon}_{\text{II}}} (\mathbf{n} \dot{\epsilon}(\mathbf{u})) \cdot (\dot{\epsilon}(\mathbf{u}) \mathbf{n}) \right) - \mathbf{n} \cdot \mathbf{n} \\
&= 2 \left([\eta(\dot{\epsilon}_{\text{II}}, \Gamma, n, \sigma_y) - \frac{1}{2}] \mathbf{n} \cdot \mathbf{n} + \frac{1}{2} \eta_{, \dot{\epsilon}_{\text{II}}} (\mathbf{n} \dot{\epsilon}(\mathbf{u})) \cdot (\dot{\epsilon}(\mathbf{u}) \mathbf{n}) \right)
\end{aligned} \tag{22}$$

Note, $\mathbf{n} \frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{u}, p} \mathbf{n}$ is a 0^{th} order tensor (scalar). Therefore the adjoint w/the misfit in surface normal stresses is,

$$\begin{aligned}
\mathbf{v} \cdot \mathbf{n} &= 0 \quad \text{on } \partial\Omega, \\
\mathbf{T}(\hat{\boldsymbol{\sigma}}_{\mathbf{u}} \mathbf{n}) &= \begin{cases} 0 & \text{on } \partial\Omega \setminus \partial\Omega_t, \\ -\mathcal{O}^T \mathcal{C}_{\text{noise}}^{-1} (\mathcal{O} \mathbf{u} - \mathbf{u}_{\text{obs}}) & \text{on } \partial\Omega_t, \end{cases} \\
\mathbf{n}(\hat{\boldsymbol{\sigma}}_{\mathbf{u}} \mathbf{n}) &= \begin{cases} 0 & \text{on } \partial\Omega \setminus \partial\Omega_t, \\ -\mathcal{O}^T \mathcal{C}_{\text{noise}}^{-1} \mathbf{n} \frac{\partial \boldsymbol{\sigma}}{\partial (\mathbf{u}, p)} \mathbf{n} (\mathcal{O} \boldsymbol{\sigma}_n - \boldsymbol{\sigma}_{\text{obs}}) & \text{on } \partial\Omega_t, \end{cases}
\end{aligned}$$

With the addition of the misfit in the surface normal stresses, we now have an additional contribution to the gradient with the misfit in the surface velocities.

$$\mathcal{G}(\mathbf{m})_i = \int_{\Omega} 2\eta_{,i}(\dot{\epsilon}_{\text{II}}, \Gamma, n, \sigma_y) \dot{\epsilon}(\mathbf{u}) : \dot{\epsilon}(\mathbf{v}) d\Omega + \int_{\partial\Omega} \mathcal{O}^T \mathcal{C}_{\text{noise}}^{-1} \mathbf{n} \frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{m}_i} \mathbf{n} (\mathcal{O} \boldsymbol{\sigma}_n - \boldsymbol{\sigma}_{\text{obs}}), \tag{23}$$

where $\mathbf{m} = \log(\Gamma)$

$$\eta_{,i}(\dot{\epsilon}_{\text{II}}, \Gamma, n, \sigma_y) = \min(\Gamma_i \chi_i \min(\eta_{\text{max}}, a(T)(\dot{\epsilon}_{\text{II}} - d)^{\frac{1}{2n}} \dot{\epsilon}_{\text{II}}^{-\frac{1}{2}}), 0) \tag{24}$$

while $\mathbf{m} = \log(n)$ is given by,

$$\eta_{,i}(\dot{\epsilon}_{\text{II}}, \Gamma, n, \sigma_y) = \begin{cases} \Gamma a(T) \omega (\dot{\epsilon}_{\text{II}} - d)^{\frac{1}{2n}} \dot{\epsilon}_{\text{II}}^{-\frac{1}{2}} & \text{in } \Omega_w, \\ 0 & \text{in } \Omega \setminus \Omega_w, \end{cases}$$

where $\omega = \log((\dot{\epsilon}_{\text{II}} - d)^{-\frac{1}{2n^2}})$. The derivative corresponding to the yield stress is,

$$\eta_{,i}(\dot{\epsilon}_{\text{II}}, \Gamma, n, \sigma_y) = \min(0, \frac{1}{\sigma_y} \dot{\epsilon}_{\text{II}}^{-\frac{1}{2}}). \tag{25}$$

For the contribution from the dynamic topography (namely the total surface normal stresses), we need the derivative, $\frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{m}}$ given below.

$$\frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{m}} = 2\eta_{,i}(\dot{\epsilon}_{\text{II}}, n, \Gamma, \sigma_y) \dot{\epsilon}(\mathbf{u}) \tag{26}$$

where $\eta_{,i}(\dot{\epsilon}_{\text{II}}, n, \Gamma, \sigma_y)$ is given in (33) to (34). Furthermore,

$$\mathbf{n} \frac{\partial \boldsymbol{\sigma}}{\partial \mathbf{m}} \mathbf{n} = 2\eta_{,i}(\dot{\epsilon}_{\text{II}}, n, \Gamma, \sigma_y) \mathbf{n}(\dot{\epsilon}(\mathbf{u})\mathbf{n}) \quad (27)$$

Previously, we computed the Gauss-Newton version of the Hessian of the misfit. We will formally derive the Hessian system, which involves taking the second order variations with respect to the Lagrangian, using the surface velocities (plate motion data). The Hessian system is given,

$$\mathcal{H} = \begin{bmatrix} \mathcal{L}_{uu} & \mathcal{L}_{um} & \mathcal{L}_{uv} \\ \mathcal{L}_{mu} & \mathcal{L}_{mm} & \mathcal{L}_{mv} \\ \mathcal{L}_{vu} & \mathcal{L}_{vm} & 0 \end{bmatrix} \quad (28)$$

To solve the incremental forward equations, we look at the third (last row). The first term is formally the adjoint operator. The strong form of the adjoint equation is

$$\begin{aligned} \nabla \cdot \mathbf{u} &= 0 \\ \nabla \cdot \boldsymbol{\sigma}_u &= -\nabla \cdot \boldsymbol{\tau}_u \end{aligned} \quad (29)$$

The incremental adjoint equations is given by solving the first row of the Hessian,

$$\begin{aligned} \nabla \cdot \mathbf{v} &= 0 \\ \nabla \cdot \boldsymbol{\sigma}_v &= -\nabla \cdot \Psi \end{aligned} \quad (30)$$

with BC's,

$$\begin{aligned} \tilde{\mathbf{v}} &= 0 \\ \mathbf{T}\boldsymbol{\sigma}_u\mathbf{n} &= -\mathcal{O}^T \mathcal{C}_{\text{noise}}^{-1} \mathcal{O}\tilde{\mathbf{u}} - \mathbf{T}(\Psi\mathbf{n}) \end{aligned} \quad (31)$$

The action of the Hessian on a direction $\tilde{\mathbf{m}}$ is

$$\mathcal{H}\tilde{\mathbf{m}} = 2\eta_{,i}[\dot{\epsilon}(\mathbf{u})\dot{\epsilon}(\mathbf{v}) + \dot{\epsilon}(\mathbf{u})\dot{\epsilon}(\tilde{\mathbf{v}}) + \dot{\epsilon}(\tilde{\mathbf{u}})\dot{\epsilon}(\mathbf{v})]\tilde{\mathbf{m}} \quad (32)$$

where $\mathbf{m} = \log(\Gamma)$

$$\eta_{,i}(\dot{\epsilon}_{\text{II}}, \Gamma, n, \sigma_y) = \min(\Gamma_i \chi_i \min(\eta_{\text{max}}, a(T)(\dot{\epsilon}_{\text{II}} - d)^{\frac{1}{2n}} \dot{\epsilon}_{\text{II}}^{-\frac{1}{2}}), 0) \quad (33)$$

while $\mathbf{m} = \log(n)$ is given by,

$$\eta_{,i}(\dot{\epsilon}_{\text{II}}, \Gamma, n, \sigma_y) = \begin{cases} \Gamma a(T)[\omega_1^2 + \omega_2](\dot{\epsilon}_{\text{II}} - d)^{\frac{1}{2n}} \dot{\epsilon}_{\text{II}}^{-\frac{1}{2}} & \text{in } \Omega_w, \\ 0 & \text{in } \Omega \setminus \Omega_w, \end{cases}$$

where $\omega_1 = \log((\dot{\epsilon}_{\text{II}} - d)^{-\frac{1}{2n}})$ and $\omega_2 = \log((\dot{\epsilon}_{\text{II}} - d)^{\frac{1}{2n}})$. The derivative corresponding to the yield stress is,

$$\eta_{,i}(\dot{\epsilon}_{\text{II}}, \Gamma, n, \sigma_y) = \min(0, \frac{1}{2}\sigma_y \dot{\epsilon}_{\text{II}}^{-\frac{1}{2}}). \quad (34)$$

Model Setup

We consider a simple system where there is a sinking mass anomaly as shown in Fig.1, along with the surface normal stress signal (Fig.1). This model is temperature dependent and has strain rate weakening, however, it lacks dynamic weakening because we do not include plates idealized as thermal boundary layers. An advantage of not using a subduction zone model is the smoothness of the surface normal stress that is normally not present at the trenches. Furthermore, we find that we have reasonable trench depths for a sinking mass anomaly.

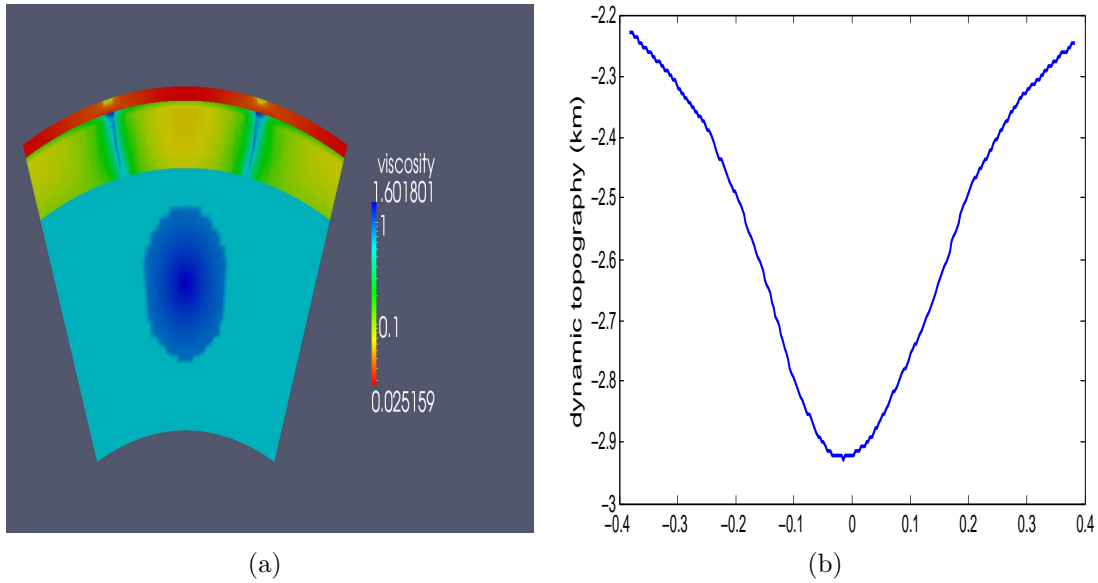


Figure 1: (a)Effective viscosity (b)Dynamic topography (km).

The surface normal stress signal in Fig.1b will be used as the observation for the inversions for the resulting cases. Using this piece of data is advantageous because there is a strong response of the signal to the effective viscosity structure.

Results

For the inversions, we investigate how well we can recover the parameters with just surface normal stress data. The parameters that we will investigate how well we can recover the strain rate exponent, the weak factor of the top layer and the activation energy. The summary of case studies are shown in Table 1.

Table 1: Case study summary

Case	n (Guess/Recov.)	γ (Guess/Recov.)	Activ. Energy (Guess/Recov.)	Iter.
1	assumed	$10^{-2}/10^{-1}$	assumed	4
2	assumed	$10^{-3}/10^{-1}$	assumed	4
3	2.85/3.0	assumed	assumed	4
4	2.0/3.0	assumed	assumed	4
5	3.5/3.0	assumed	assumed	4
6	assumed	assumed	0.1/1.0	4
7	assumed	assumed	0.4/1.0	4
8	2.8/3.0 (both data)	assumed	assumed	4
9	assumed	$10^{-2}/10^{-1}$ (both data)	assumed	4

For Cases 1 and 2 we invert for the prefactor of the weak layer near the top of the system. The true prefactor γ of the top layer is $\gamma = 1.0 \cdot 10^{-1}, n = 3$ while the guessed value is $\gamma = 10^{-3}$ or 10^{-2} . We infer the weak layers in two separate inversions with two different initial guesses to test if we can arrive at the correct value regardless of initial guess.

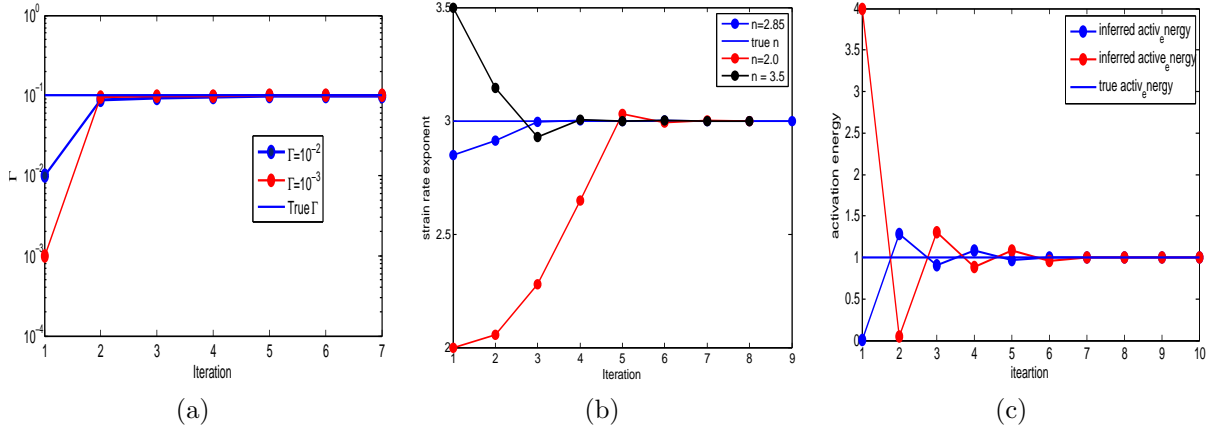


Figure 2: (a)Inferred γ (Cases 1 and 2) (b)strain rate exponent (Cases 3, 4 and 5) and (c)Activation energy (Cases 6 and 7) vs. Gauss Newton iteration.

We find that regardless of the initial guess, we can recover the true value of the prefactor as shown in Fig.2. An important aspect to these inversions is the gradient consistency from the adjoints and the convergence of the Hessians. To this end, we test both gradient and Hessian convergence as shown in Tables 2 and 3.

Similarly, we invert for the strain rate exponent while keeping the weak factor (γ) fixed in Cases 3,4 and 5. The true strain rate exponent is $n = 3$, while the guessed value is $n = 2.85, 3.5$ or 2.0 . We find that we can correctly infer the strain rate exponent independent of the initial guess in all three cases as shown in Fig.2. Furthermore, the gradient and Hessians for each of those inversions are consistent as shown in Tables 4, 5, and 6.

While the strain rate exponent and weakfactor are important parameters in determining the governing physics, the activation energy is a key parameter that strongly influences the temperature dependent part of the rheology, which in turn would affect the surface normal stress signal. To this end, we seek to infer the activation energy (Cases 6-7) in a similar way the strain rate exponent and weakfactors were inverted for in Cases 1-5. We similarly test how sensitive the inversion is to the initial guess by performing two inversions with a different starting guesses ($E_a = 0.1$ in Case 6 and $E_a = 4$ in Case 7) for the activation energy. We find that we can recover the true activation energy and that the Hessian converges to the same value regardless of initial guess.

We also test how well we can infer either the weak layer factor Γ or strain rate exponent using both surface normal stress and surface velocity. In Fig.3 shows that we can correctly infer either the weakfactor layer or the strain rate exponent using both pieces of data. Furthermore, we infer either rheological parameter in about 4 iterations.

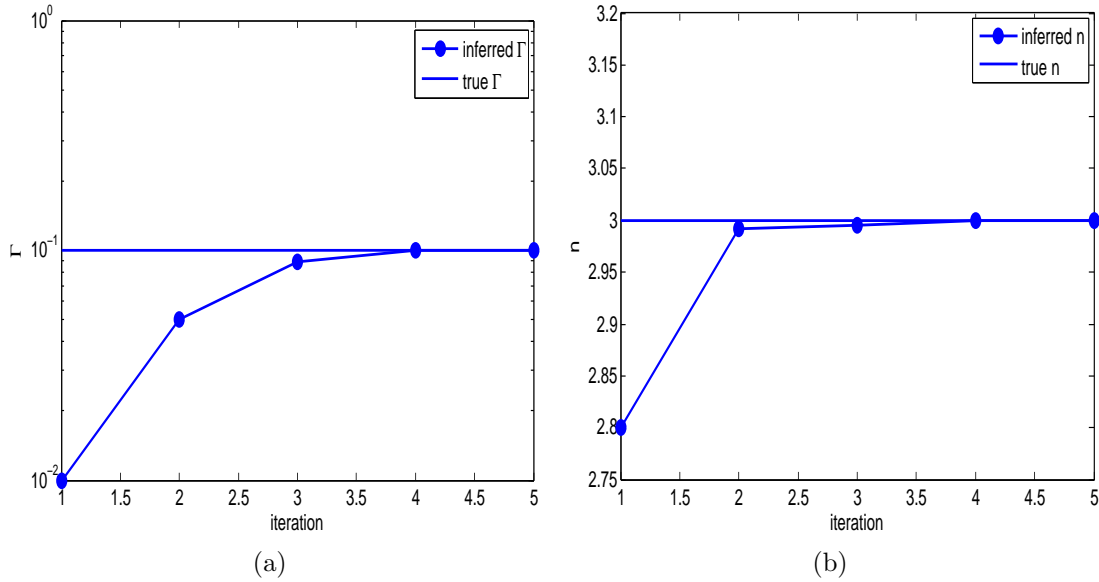


Figure 3: (a) Inferred values for the prefactor (Case 9) (b) strain rate exponent (Case 8) using both pieces of data.

Limitations at subduction zones

We were able to infer various rheological parameters for the sinking mass anomaly using this new adjoint formulation. We show that in Cases 1-2, we can correctly infer the weak layer prefactor, (independent of the initial guess), within 3 iterations. We similarly find that we can infer the strain rate exponent (Cases 3-5) and activation energy (Cases 6-7) within 4-5 iterations, independent of the initial guess. In Cases 8 and 9, we use both surface velocities and surface normal stresses to infer the strain rate exponent or the weakfactor, and find that with both these pieces of data, we can infer the correct parameters.

While we were able to implement and prove the validity of using the surface normal stress in the adjoint formulation, there still are issues in applying this method to subduction zones. One of these issues lies in the large variations in viscosity and the boundary conditions on the surface [9]. The problem is that there are large dynamic topography near the trench that exceed 10km in forward models, which is certainly not found in the observations. Therefore, the adjoint formulation with surface normal stress will not work when no forward model is able to produce reasonable short wavelength topographic signals.

The problematic issue of large topographic magnitudes can be traced to the large variations in viscosity of the forward model between plate and the weakzone ($\mathcal{O}(10^6)$). We find that when the viscosity variations are reduced, the magnitude of the topography was also reduced. However, this also caused an increase in the fore-bulge. Additionally, when using a sticky-air surface (weak viscosity layer at the surface), we find that we could reduce the amplitude of the topography signal at the trench. The sticky-air method brings an added complication of measuring the topography due to the interface between the sticky-air layer and the oceanic plates. It is certainly clear from those investigations that to remedy the large trench depths requires a different formulation of the effective viscosity and/or surface boundary conditions. However, after the issue of large topographic signals is resolved, the use of the adjoint with surface normal stress can be readily used.

References

- [1] L. Alisic, M. Gurnis, G. Stadler, C. Burstedde, L. C. Wilcox, and O. Ghattas. Slab stress and strain rate as constraints on global mantle flow. *Geophysical Research Letters*, 37(22), 2010. ISSN 1944-8007. URL <http://dx.doi.org/10.1029/2010GL045312>. DOI: 10.1029/2010GL045312.
- [2] L. Alisic, M. Gurnis, G. Stadler, C. Burstedde, and O. Ghattas. Multi-scale dynamics and rheology of mantle flow with plates. *Journal of Geophysical Research: Solid Earth*, 117(B10), 2012. ISSN 2156-2202. URL <http://dx.doi.org/10.1029/2012JB009234>. DOI: 10.1029/2012JB009234.
- [3] D. F. Argus, R. G. Gordon, and C. DeMets. Geologically current motion of 56 plates relative to the no-net-rotation reference frame. *Geochemistry, Geophysics, Geosystems*, 12(11), 2011. ISSN 1525-2027. URL <http://dx.doi.org/10.1029/2011GC003751>. DOI: 10.1029/2011GC003751.
- [4] M. I. Billen and G. Hirth. Newtonian versus non-Newtonian upper mantle viscosity: Implications for subduction initiation. *Geophysical Research Letters*, 32(19), 2005. ISSN 1944-8007. URL <http://dx.doi.org/10.1029/2005GL023457>. DOI: 10.1029/2005GL023457.
- [5] M. I. Billen and G. Hirth. Rheologic controls on slab dynamics. *Geochemistry, Geophysics, Geosystems*, 8(8), 2007. ISSN 1525-2027. URL <http://dx.doi.org/10.1029/2007GC001597>. DOI: 10.1029/2007GC001597.

- [6] Magali I Billen and Michael Gurnis. A low viscosity wedge in subduction zones. *Earth and Planetary Science Letters*, 193(1):227–236, 2001.
- [7] W. M. Chapple and T. E. Tullis. Evaluation of the forces that drive the plates. *Journal of Geophysical Research*, 82(14):1967–1984, 1977. ISSN 2156-2202. URL <http://dx.doi.org/10.1029/JB082i014p01967>. 10.1029/JB082i014p01967.
- [8] C. P. Conrad and C. Lithgow-Bertelloni. How mantle slabs drive plate tectonics. *Science*, 298(5591):207–209, 2002. URL <http://www.sciencemag.org/content/298/5591/207.abstract>. DOI: 10.1126/science.1074161.
- [9] F Cramer, CR Lithgow-Bertelloni, and PJ Tackley. The dynamical control of subduction parameters on surface topography. *Geochemistry, Geophysics, Geosystems*, 2017.
- [10] Nicolas Flament, Michael Gurnis, and R Dietmar Müller. A review of observations and models of dynamic topography. *Lithosphere*, 5(2):189–210, 2013.
- [11] D. Forsyth and S. Uyeda. On the relative importance of the driving forces of plate motion. *Geophysical Journal International*, 43(1):163–200, 1975. URL <http://gji.oxfordjournals.org/content/43/1/163.abstract>. DOI: 10.1111/j.1365-246X.1975.tb00631.x.
- [12] A. E. Gripp and R. G. Gordon. Current plate velocities relative to the hotspots incorporating the nuvel-1 global plate motion model. *Geophysical Research Letters*, 17(8):1109–1112, 1990. ISSN 1944-8007. URL <http://dx.doi.org/10.1029/GL017i008p01109>. DOI: 10.1029/GL017i008p01109.
- [13] Bradford H Hager, Robert W Clayton, Mark A Richards, Robert P Comer, and Adam M Dziewonski. Lower mantle heterogeneity, dynamic topography and the geoid. 1984.
- [14] S. Karato and P. Wu. Rheology of the upper mantle: A synthesis. *Science*, 260(5109):771–778, 1993.
- [15] Georg Kaufmann and Kurt Lambeck. Mantle dynamics, postglacial rebound and the radial viscosity profile. *Physics of the Earth and Planetary Interiors*, 121(3):301–324, 2000.
- [16] Corné Kreemer, William E Holt, and A John Haines. An integrated global model of present-day plate motions and plate boundary deformation. *Geophysical Journal International*, 154(1):8–34, 2003.
- [17] Louis Moresi and Michael Gurnis. Constraints on the lateral strength of slabs from three-dimensional dynamic flow models. *Earth and Planetary Science Letters*, 138(1-4):15–28, 1996.
- [18] Vishagan Ratnaswamy, Georg Stadler, and Michael Gurnis. Adjoint-based estimation of plate coupling in a non-linear mantle flow model: theory and examples. *Geophysical Journal International*, 202(2):768–786, 2015.

- [19] C. H. Scholz and J. Campos. The seismic coupling of subduction zones revisited. *Journal of Geophysical Research: Solid Earth (1978–2012)*, 117(B5), 2012.
- [20] G. Stadler, M. Gurnis, C. Burstedde, L. C. Wilcox, L. Alisic, and O. Ghattas. The dynamics of plate tectonics and mantle flow: From local to global scales. *Science*, 329(5995):1033–1038, 2010. URL <http://www.sciencemag.org/content/329/5995/1033.abstract>. DOI: 10.1126/science.1191223.
- [21] R. L. Stocker and M. F. Ashby. On the rheology of the upper mantle. *Reviews of Geophysics*, 11(2):391–426, 1973.
- [22] J. Worthen, G. Stadler, N. Petra, M. Gurnis, and O. Ghattas. Towards an adjoint-based inversion for rheological parameters in nonlinear viscous mantle flow. *Physics of Earth and Planetary Interiors*, 234:23–34, 2014. DOI: 10.1016/j.pepi.2014.06.006.
- [23] Ting Yang and Michael Gurnis. Dynamic topography, gravity and the role of lateral viscosity variations from inversion of global mantle flow. *Geophysical Journal International*, 207(2):1186–1202, 2016.

A Gradient and Hessian Tests

A.1 Cases 1 and 2

Table 2: Inversion statistics for $\gamma_{\text{guess}} = 10^{-2}$

<i>Iteration</i>	\mathcal{H}	$ \mathcal{G} $
1	$2.02 \cdot 10^9$	$3.675 \cdot 10^{10}$
2	$3.47 \cdot 10^{10}$	$2.948 \cdot 10^9$
3	$4.95 \cdot 10^{10}$	$1.899 \cdot 10^8$
4	$5.46 \cdot 10^{10}$	$9.91 \cdot 10^6$
5	$5.72 \cdot 10^{10}$	$5.06 \cdot 10^5$

Table 3: Inversion statistics for $\gamma_{\text{guess}} = 10^{-3}$

<i>Iteration</i>	\mathcal{H}	$ \mathcal{G} $
1	$1.248 \cdot 10^8$	$8.452 \cdot 10^{10}$
2	$5.53 \cdot 10^{10}$	$3.7718 \cdot 10^9$
3	$5.775 \cdot 10^{10}$	$1.531 \cdot 10^8$
4	$5.917 \cdot 10^{10}$	$9.91 \cdot 10^6$
5	$6.002 \cdot 10^{10}$	$7.97 \cdot 10^5$
6	$6.05 \cdot 10^{10}$	$6.374 \cdot 10^5$
7	$6.08 \cdot 10^{10}$	$5.06 \cdot 10^4$

A.2 Cases 3-5

Table 4: Inversion statistics for $n_{\text{guess}} = 2.85$

<i>Iteration</i>	\mathcal{H}	$ \mathcal{G} $
1	$3.977 \cdot 10^{12}$	$1.717 \cdot 10^{11}$
2	$3.0604 \cdot 10^{12}$	$8.663 \cdot 10^{10}$
3	$2.104 \cdot 10^{12}$	$3.356 \cdot 10^9$
4	$2.057 \cdot 10^{12}$	$4.461 \cdot 10^8$
5	$2.0638 \cdot 10^{12}$	$6.985 \cdot 10^7$
6	$2.06281 \cdot 10^{12}$	$1.044 \cdot 10^7$
7	$2.06296 \cdot 10^{12}$	$1.567 \cdot 10^6$
8	$2.06294 \cdot 10^{12}$	$2.342 \cdot 10^5$
9	<i>N/A</i>	$3.5047 \cdot 10^4$

Table 5: Inversion statistics for $n_{\text{guess}} = 2.0$

<i>Iteration</i>	\mathcal{H}	$ \mathcal{G} $
1	$2.7665 \cdot 10^{13}$	$3.488 \cdot 10^{12}$
2	$3.944 \cdot 10^{13}$	$2.912 \cdot 10^{12}$
3	$3.092 \cdot 10^{13}$	$1.257 \cdot 10^{12}$
4	$8.617 \cdot 10^{12}$	$2.675 \cdot 10^{11}$
5	$1.797 \cdot 10^{12}$	$1.1778 \cdot 10^{10}$
6	$2.146 \cdot 10^{12}$	$3.647 \cdot 10^9$
7	$2.0496 \cdot 10^{12}$	$5.832 \cdot 10^8$
8	$2.0655 \cdot 10^{12}$	$1.137 \cdot 10^8$
9	$2.06245 \cdot 10^{12}$	$2.159 \cdot 10^7$

Table 6: Inversion statistics for $n_{\text{guess}} = 3.5$

<i>Iteration</i>	\mathcal{H}	$ \mathcal{G} $
1	$3.658 \cdot 10^{11}$	$3.1396 \cdot 10^{11}$
2	$1.2446 \cdot 10^{12}$	$1.76443 \cdot 10^{11}$
3	$2.64779 \cdot 10^{12}$	$1.351 \cdot 10^{11}$
4	$2.03321 \cdot 10^{12}$	$6.92 \cdot 10^9$
5	$2.06992 \cdot 10^{12}$	$1.386 \cdot 10^9$
6	$2.06271 \cdot 10^{12}$	$2.525 \cdot 10^8$
7	$2.06406 \cdot 10^{12}$	$6.143 \cdot 10^7$
8	$2.06372 \cdot 10^{12}$	$2.221 \cdot 10^7$
9	$2.06381 \cdot 10^{12}$	$7.9271 \cdot 10^6$

Table 7: Gradient check for activation energy with $E_a^0 = 4.0$

$iteration$	$\mathcal{G}_{Adjoint}$	\mathcal{G}_{FD}
1	$3.812 \cdot 10^8$	$3.892 \cdot 10^8$
2	$-4.622 \cdot 10^7$	$-4.632 \cdot 10^7$
3	$2.114 \cdot 10^6$	$2.101 \cdot 10^6$
4	$-5.44 \cdot 10^5$	$-5.342 \cdot 10^5$
5	$6.9301 \cdot 10^4$	$6.922 \cdot 10^4$
6	$-5.10 \cdot 10^4$	$-5.102 \cdot 10^4$
7	$2.110 \cdot 10^4$	$2.115 \cdot 10^4$
8	$-1.01 \cdot 10^4$	$-1.02 \cdot 10^4$