

This discussion paper is/has been under review for the journal The Cryosphere (TC).  
Please refer to the corresponding final paper in TC if available.

# Of the gradient accuracy in Full-Stokes ice flow model: basal slipperiness inference

N. Martin<sup>1,2</sup> and J. Monnier<sup>1,2</sup>

<sup>1</sup>INSA, 135 Avenue de Rangueil, 31077, Toulouse Cedex 4, France

<sup>2</sup>IMT, 118, Route de Narbonne, 31062 Toulouse Cedex 9, France

Received: 1 July 2013 – Accepted: 24 July 2013 – Published: 5 August 2013

Correspondence to: N. Martin (nmartin@insa-toulouse.fr)

Published by Copernicus Publications on behalf of the European Geosciences Union.

TCD

7, 3853–3897, 2013

Of the gradient  
accuracy in  
Full-Stokes ice flow  
model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Abstract

This work focuses on the numerical assessment of the accuracy of an adjoint-based gradient in the perspective of variational data assimilation and parameter identification in glaciology. We quantify the ability to identify the basal slipperiness for such methods with a non-linear friction law. The complete adjoint problem is solved and a comparison with the so called “self-adjoint” method, neglecting the viscosity dependency to the velocity, common in glaciology, is carried out. A lower bound of identifiable wavelengths of 10 ice thickness in the friction coefficient is established, when using the full adjoint method, while the “self-adjoint” method is limited to a maximum of 20 ice thickness wavelengths. In addition, the full adjoint method demonstrates a better robustness and reliability for the parameter identification process. The derivation of the adjoint model using algorithmic differentiation leads to formulate a generalization of the “self-adjoint” approximation towards an *incomplete adjoint method*, adjustable in precision and computational burden.

1 Introduction

The main available observations of the cryosphere are generally obtained from remote-sensed techniques and are thus essentially surface observations. In other respects, the ice dynamic is known to be highly sensitive to the state of the bedrock and therefore to its modeling. The basal slipperiness is consequently a crucial parameter in the perspective of controlling ice flows.

This raises questions about, on the one hand, the ability for the surface data to encompass basal conditions and, on the other hand, the potential for inverse methods to recover information.

The first question has been treated by many authors through the problem of the effect of the bedrock topography on the surface. Balise and Raymond (1985) propose one of the earlier study concerning the transmission at surface of basal slipperiness

Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

fluctuations in the case of a Newtonian fluid, using perturbation methods. The non-local aspect of the transmission of the variations of the friction coefficient at surface is established in Raymond (1996) depending on the slip ratio (ratio between mean sliding velocities and mean ice deformation velocities). These queries are extended in Gudmundsson (2003), still under the Newtonian hypothesis using perturbation methods. One of the main conclusions is the increase of the transmission of basal variability at surface with an increased sliding.

The question of the representability of the friction coefficient through surface velocity observations (horizontal and vertical) using an inverse method is studied by Gudmundsson and Raymond (2008). The method, based on a Bayesian approach, is used to study the effect of density and quality of surface velocity data on the estimation of the friction coefficient for a Newtonian fluid and a linear sliding law. A limit wavelength in the reconstruction of small amplitude variations of the friction coefficient is found to be around 50 times the ice thickness. A similar method in the case of a non-Newtonian fluid and a non-linear sliding law is developed in Raymond and Gudmundsson (2009).

In other respects, the identification method based on Macayeal (1993) and widely used (see e.g. Larour et al., 2005; Joughin et al., 2004; Morlighem et al., 2010) makes the assumption that viscosity is independent of the velocity and a limited attention has been paid to the quality of the resulting estimations in terms of spatial variability of the friction coefficient (see Gudmundsson and Raymond, 2008). Comparisons with this so called “self-adjoint” method and the use of an exact adjoint are made by Goldberg and Sergienko (2011), based on a vertically integrated approximation. Limitations for the minimizing process are highlighted when using the “self-adjoint” method. To the best of our knowledge, the use of an exact adjoint for the non-Newtonian Stokes problem has been done only by Petra et al. (2012). A comparison to the results of Gudmundsson and Raymond (2008) on an academic problem bring them to conclude that the exact adjoint method is able to recover wavelengths in the friction coefficient of approximately 20 times the ice thickness in the case of a linear sliding law.

## Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



The purpose of this study is the numerical evaluation of the limitations of the “self-adjoint” method compared to the exact method. The “self-adjoint” approximation is presented as a limit case of the reverse accumulation method used to compute the adjoint when obtained using source-to-source automatic differentiation. From a strictly numerical perspective, numerical tests on the reachable accuracy of the gradients for both methods are performed, demonstrating an important limitation for the gradient computed by the “self-adjoint” method. We then study the identifiability, for a non-linear sliding law, of high frequencies in the friction coefficient depending on the level of noise considered on synthetic data. The quality of the estimations provided by both methods is compared in the case of dense horizontal surface velocity observations for a quasi-uniform flow and then for a realistic flow presenting an important spatial variability. The realistic case is eventually treated for less dense data.

## 2 Forward and backward model

In this section, we briefly present what shall be referred to hereafter as the forward model and describe the derivation of the adjoint model and the computation of the adjoint state.

### 2.1 Forward model

The flow model considered here is the bidimensional power-law Stokes model applied to a gravity driven flow, discretized using triangular Taylor–Hood finite-elements and solved on a given domain  $\Omega$ :

$$\text{div}(\mathbf{u}) = 0 \quad \text{in } \Omega \quad (1)$$

$$-\text{div}(2\eta(\mathbf{u})\underline{D}) + \nabla p = \rho \mathbf{g} \quad \text{in } \Omega \quad (2)$$

$$\eta(\mathbf{u}) = \eta_0 \|\underline{D}\|_{\text{F}}^{\frac{1-n}{n}} \quad (3)$$

where  $\underline{\sigma} = \eta(\mathbf{u})\underline{D} - p\mathbf{I}$  represents the Cauchy stress tensor,  $\eta(\mathbf{u})$  the viscosity,  $\eta_0$  the consistency of the fluid,  $n$  the power-law exponent,  $\underline{D}$  the strain rate tensor,  $\mathbf{u} = (u_x, u_z)$  the velocity field,  $p$  the pressure field,  $\rho$  the ice density,  $\mathbf{g}$  the gravity and  $\|\underline{D}\|_F^2 = \underline{D} : \underline{D}$  the Frobenius matrix norm.

5 A Weertman-type sliding law is prescribed at the bedrock boundary  $\Gamma_{\text{fr}}$ :

$$|\underline{\sigma}_{nt}|^{m-1} \underline{\sigma}_{nt} = \beta \mathbf{u} \cdot \mathbf{t} \quad \text{on } \Gamma_{\text{fr}} \quad (4)$$

$$\mathbf{u} \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_{\text{fr}} \quad (5)$$

10 A velocity profile corresponding to the solution of the Stokes problem for a uniform steady flow in a flat channel with non-linear friction defined by Eq. (4) at the bottom is prescribed on the inflow boundary. This solution  $\mathbf{u} = (u_{\bar{x}}, u_{\bar{z}})$ , expressed in the “mean-slope” reference frame  $(\bar{x}, \bar{z})$ , writes (see e.g. Martin and Monnier, 2013):

$$u_{\bar{x}}(\bar{z}) = \frac{(-\rho g \sin(\theta) h)^m}{\beta} + \frac{1}{1+n} (2\eta_0)^{-n} (\rho g \sin(\theta))^{1+n} (h^{1+n} - (H - \bar{z})^{1+n}) \quad (6)$$

$$u_{\bar{z}} = 0 \quad (7)$$

$$15 \quad p(\bar{z}) = \rho g \cos(\theta) (H - \bar{z}) \quad (8)$$

with  $\theta$  defining the slope of the channel,  $H$  the height of the upper-surface and  $h$  the thickness of the channel.

20 A hydrostatic pressure is considered on the outflow. All the simulations are performed with an exponent  $m = 3$  for the sliding law. The solution of the continuous forward problem is obtained using a classical fixed point algorithm.

The sensitivities and identifications carried out in this work use adjoint-based computation and thus require the solution of the adjoint problem associated with the full-Stokes model.

## Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## 2.2 The basic principles of the adjoint model

The output of the forward model is represented by a scalar valued function  $j$  called *cost function*, depending on the parameters of the model, which evaluate a quantity to minimize. In presence of observations, part of the cost measures the discrepancy between the computed state and an observed state (through any type of data).

The parameters of interest are called *control variables* and constitute a control vector  $\mathbf{k}$ . The minimizing procedure operates on this control vector to generate a set of parameters which allows a computed state closer to the observations to be obtained. In the sequel, the control vector includes only the spatially variable friction coefficient  $\beta$ . The optimal control problem we solve reads:

$$\min_{\mathbf{k}} j(\mathbf{k}) \quad (9)$$

This optimization problem is solved numerically by a descent algorithm. Thus, we need to compute the gradient of the cost function. This is done by introducing the adjoint model.

## 2.3 Cost function, twin experiments and Morozov's discrepancy principle

The cost function used for the identification is defined by:

$$j(\beta) = \int_{\Gamma_s} \|u_s^{\text{obs}} - u_s(\beta)\|_2^2 dx + \mathcal{T}(\nabla\beta) \quad (10)$$

where the synthetic data  $u_s^{\text{obs}}$  are the horizontal surface velocities perturbed with a random Gaussian noise of varying level  $\delta$ . The term  $\mathcal{T}(\nabla\beta)$  called Tykhonov's regularization controls the oscillations of the control variable gradient. It is defined by:

$$\mathcal{T}(\nabla\beta) = \gamma \int_{\Gamma_s} \|\partial_x \beta\|_2^2 ds \quad (11)$$

## Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



where the parameter  $\gamma$  quantifies the strength of the imposed smoothness. This term plays a role of convexification for the optimal control problem and thus restricts the region of admissible parameters to smoothly varying fields. The tuning of these weights can be achieved from various considerations generally related to the quality of the data (or the noise level) and the degree of smoothness sought on the control variable. A classical approach, referred to as the Morozov's discrepancy principle (see e.g. Vogel, 2002; Martin, 2013), consists in choosing  $\gamma$  such that the final cost  $j(\beta; \gamma)$  matches the noise level on the data.

The methodology of twin experiments is used as a first step towards real data processing and calibration model. Twin experiments are defined as follows: reference parameters of the model  $\mathbf{k}_{\text{ref}}$  are used to generate observations  $u_{\text{obs}}$ . Then, the goal is to retrieve the set of parameters  $\mathbf{k}_{\text{ref}}$  starting from an initial guess  $\mathbf{k} \neq \mathbf{k}_{\text{ref}}$  using the minimization of the cost function  $j$ . Thus, the algorithm compute at each iteration  $i$  a new set of parameters  $k_i$  according to the gradient  $\frac{\partial j}{\partial \mathbf{k}}$  in order to make the cost  $j_i = j(k_i)$  decrease.

Twin experiments are an ideal framework in which the only source of uncertainty comes from the ignorance of the parameters. It plays a role of validation of the minimization procedure and gives information on its efficiency and on the robustness of the inverse problem.

## 2.4 Derivation of the adjoint model

In order to efficiently compute all partial derivatives of the cost function  $j(\mathbf{k})$  with respect to the components of the control vector  $\mathbf{k}$ , we introduce the adjoint model (see e.g. Lions, 1971).

In DassFlow software, the adjoint model is obtained by using algorithmic differentiation of the source code (see Honnorat et al., 2007; DassFlow Software). This last approach ensures a better consistency between the computed cost function (including all types of errors i.e. discretization errors, rounding errors, iterative algorithms etc.) and its gradient, since it is the computed cost function that is differentiated. A large part

of this extensive task can be automated using automatic differentiation (see Griewank, 1989). In the case of DassFlow-Ice, the direct code is written in Fortran 95 and is derived using the automatic differentiation tool Tapenade (see Hascoët and Pascual, 2004).

- 5 Let  $\mathcal{K}$  be the space of control variables and  $\mathcal{Y}$  the space of the forward code response. In the present case, we have:

$$\mathbf{k} = (\beta)^T \text{ and } Y = (y, j)^T$$

where  $\beta$  is defined by Eq. (4).

- 10 Let us point out that we include both the state and the cost function in the response of the forward code. The direct code can be represented as an operator  $\mathcal{M} : \mathcal{K} \rightarrow \mathcal{Y}$  such that:

$$\mathcal{Y} = \mathcal{M}(\mathcal{K})$$

The tangent model becomes  $\frac{\partial \mathcal{M}}{\partial \mathbf{k}}(\mathbf{k}) : \mathcal{K} \rightarrow \mathcal{Y}$ . As input variable, it takes a perturbation of the control vector  $d\mathbf{k} \in \mathcal{K}$ , it then gives the variation  $dY \in \mathcal{Y}$  as output variable:

$$15 \quad dY = \frac{\partial \mathcal{M}}{\partial \mathbf{k}}(\mathbf{k}) \cdot d\mathbf{k}$$

The adjoint model is defined as the adjoint operator of the tangent model. This can be represented as follows:

$$\left( \frac{\partial \mathcal{M}}{\partial \mathbf{k}}(\mathbf{k}) \right)^* : \mathcal{Y}' \rightarrow \mathcal{K}'$$

- 20 It takes  $dY^* \in \mathcal{Y}'$  as an input variable and provides the adjoint variable  $d\mathbf{k}^* \in \mathcal{K}'$  at output:

$$d\mathbf{k}^* = \left( \frac{\partial \mathcal{M}}{\partial \mathbf{k}}(\mathbf{k}) \right)^* \cdot dY^*$$

Now, let us make the link between the adjoint code and the gradient  $\frac{dj}{dk}$  we seek to compute. By definition of the adjoint, we have:

$$\left\langle \left( \frac{\partial \mathcal{M}}{\partial \mathbf{k}} \right)^* \cdot dY^*, d\mathbf{k} \right\rangle_{\mathcal{K}' \times \mathcal{K}} = \left\langle dY^*, \left( \frac{\partial \mathcal{M}}{\partial \mathbf{k}} \right) \cdot d\mathbf{k} \right\rangle_{\mathcal{Y}' \times \mathcal{Y}} \quad (12)$$

It reads, using the relations presented above:

$$\langle d\mathbf{k}^*, d\mathbf{k} \rangle_{\mathcal{K}' \times \mathcal{K}} = \langle dY^*, dY \rangle_{\mathcal{Y}' \times \mathcal{Y}}. \quad (13)$$

If we set  $dY^* = (0, 1)^T$  and by denoting the perturbation vector  $d\mathbf{k} = (\delta\beta)^T$ , we obtain:

$$\left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} dy^* \\ dj^* \end{pmatrix} \right\rangle_{\mathcal{Y}' \times \mathcal{Y}} = \langle (\delta\beta^*), (\delta\beta) \rangle_{\mathcal{K}' \times \mathcal{K}}$$

Furthermore, we have by definition:

$$dj = \frac{\partial j}{\partial \beta}(\mathbf{k}) \cdot \delta\beta \quad (14)$$

Therefore, the adjoint variable  $d\mathbf{k}^*$  (output of the adjoint code with  $dY^* = (0, 1)^T$ ) corresponds to the partial derivatives of the cost function  $j$ :

$$\frac{\partial j}{\partial \beta}(\mathbf{k}) = \beta^* \quad (15)$$

A single integration of the forward model followed by a single integration of the adjoint model allow to compute all components of the gradient of the cost function.

The optimal control problem (Eq. 9) is solved using a local descent algorithm, more precisely the L-BFGS algorithm (a quasi-Newton method), implemented in the M1QN3 routine (see Gilbert and Lemaréchal, 1989). Thus, these partial derivatives are used as input to the minimization algorithm M1QN3. The global optimization process is represented in Fig. 1.

## Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





### 3 Exact adjoint, reverse accumulation and “self-adjoint” approximation

The current model has been obtained using algorithmic (or automatic) differentiation of the source code. Automatic differentiation of a fixed point type iterative routine of the form  $y = \Phi(y, u)$  (such as the solution of the non-linear Stokes problem using a Picard method) is carried out by reverse accumulation (see Griewank, 1989; Griewank et al., 1993). The reverse accumulation technique consists in building a computational graph for the function evaluation where the nodes of the graph represent every value taken by the function. To every node, an adjoint quantity containing the gradient of the function  $\Phi$  with respect to the node, is associated.

The adjoint values are computed in reverse order. The final value of the gradient is given by the sum of the partial derivatives of the function of the nodes of the computational graph. This result is a consequence of the chain rule. This process a priori requires to store as many state of the system as iterations performed by the forward solver to reach the converged state.

It is shown by Christianson (1994) that, in the case of a forward computation carried out by a fixed point method, the adjoint quantity also satisfies a fixed point problem whose rate of convergence is at least equal to the rate of convergence of the forward fixed point. Based on this result, it is a priori necessary to retain every iteration of the forward run to evaluate the gradient. In practice, as further detailed in Sect. 3.2, the number of reverse iterations required to obtain an adjoint state with the same precision of the forward state can be adjusted depending on the convergence speed of the direct construction.

The “self-adjoint” approximation in glaciology, applied to the shallow-ice system, has been proposed by Macayeal (1993). The approximation consists of deriving the adjoint equation system without taking into account the explicit dependency of the viscosity  $\eta$  to the velocity field  $\mathbf{u}$ . It seems important to recall that the terminology “self-adjoint” only makes sense in the Newtonian case ( $n = 1$ ) due to a resulting symmetrical stiffness matrix.

## Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



In general, the procedure consists in calculating a mechanical equilibrium based on the complete non-linear system and to obtain the gradient by simply transposing the final computed state. The extension of this method to the non-Newtonian Stokes problem can be done without difficulty (see Morlighem et al., 2010). The mechanical meaning of this method for  $n \neq 1$  is unclear.

In contrast, in the automatic differentiation context, the meaning of this approximation can be easily deduced. Indeed, the use of such a method comes to retain, in the reverse accumulation process, only the gradient computed from the final evaluation of the function  $\Phi$ . The quality of such an approximation is thus questionable and will strongly depend on the problem one considers and the required accuracy on the gradient.

The quality of this approximation (compared to the exact adjoint state) for parameter identification is assessed by Goldberg and Sergienko (2011) for depth-integrated shallow-ice type equations but has never been treated for the full-Stokes equations.

### 3.1 Precision of the “self-adjoint” approximation

We consider the flow described in Sect. 2.1. The domain is a channel with a flat bottom and constant thickness with an aspect ratio of 1/10 on a 10% slope. The friction condition at the bottom is given by Eq. (4) with a constant  $\beta$  and an exponent  $m = 3$ . A stationary free surface flow, uniform with respect to  $x$ , is thus obtained.

The cost function is defined by:

$$j(\mathbf{k}) = \frac{1}{2} \int_{\Gamma_s} |u(\mathbf{k}, \bar{z}) - u^{\text{obs}}|^2 d\bar{x} \quad (18)$$

where the observations  $u^{\text{obs}}$  are the horizontal velocities at the surface  $\Gamma_s$ ,  $(\bar{x}, \bar{z})$  designates the mean-slope frame and the control vector  $\mathbf{k}$  is restricted to the discrete basal slipperiness  $\beta$ .

The gradient tests carried out for the “self-adjoint” and exact adjoint methods, using cost function (18), are plotted on Fig. 2. The tests are performed for various levels of

precision of the forward problem  $\nu = \|u_{k+1} - u_k\| / \|u_k\|$  in order to quantify the best attainable precision by the backward problem with respect to  $\nu$ . This precision is explicitly given to the direct solver through a convergence threshold but can be seen as the available accuracy on the data  $u^{\text{obs}}$ ; data presenting a noise of 0.01 % corresponds to a direct solution accuracy  $\nu = 10^{-4}$ . The gradient test compares the gradient computed by the adjoint code to a reference gradient. For these tests, the reference gradient is obtained using a centered finite difference approximation (of order 2) computed for a precision on the function evaluation of  $10^{-12}$ . This precision being considerably higher than those considered for the solution of the forward problem, the finite difference gradient plays the role of an “exact” value (see Sect. 2.5).

The exact adjoint shows the expected theoretical behavior. We recover the slope of 2 (in logarithmic scale) associated with the order of convergence of the finite difference approximation (16). Figure 2 thus shows that the precision of the adjoint state is of the same order as the one of the direct solver.

On the contrary, the precision of the gradient provided by the “self-adjoint” approximation is rather limited. The best reachable precision is slightly smaller than 1 irrespective of the direct solver precision  $\nu$  (and thus, only one gradient test curve is plotted on Fig. 2, for the case  $\nu = 10^{-8}$ ,  $\nu$  being the precision of the forward solution).

The “self-adjoint” approximation used within a parameter identification process is thus not able to provide an accurate gradient. However, as further discussed thereafter, numerical tests demonstrate a certain ability for this approximation to partially reconstruct the friction coefficient (for a computational cost well below the one of the exact adjoint). Nevertheless, significant weaknesses for the reconstruction of high frequencies as well as the reconstruction of the main frequency of the friction coefficient signal, specifically for extreme situations of sliding (very slow or very fast), are brought to the fore.

TCD

7, 3853–3897, 2013

## Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

### 3.2 Truncation of the reverse accumulation

This section focuses on the effect of a truncation of the reverse accumulation process. Figure 3 plots gradient test results obtained for a truncated evaluation of the adjoint state. To do so, we run the gradient test on adjoint states successively cut of one additional iteration. We thus obtain all the levels of precision attainable by all the intermediary adjoint states between the full adjoint and the “self-adjoint” approximation. This test is carried out for various levels of precision  $\nu$  of the direct solver. The number of iterations performed by the direct solver to reach the required accuracy  $\nu$  depends on this precision.

The previous results concerning the precision of the gradient presented previously are well recovered (see Fig. 2). The lowest precision, identical for every  $\nu$  and equal to 0.6, is obtained from the “self-adjoint” approximation (corresponding to 1 reverse iteration) and the highest precision is reached by the full adjoint (corresponding to the last point of each curve).

A linear decreasing of the error (in logarithmic scale) presenting a slope of 3.7 is observed. This error behavior is coherent with the result of Christianson (1994) who states that the computation of the adjoint state by reverse accumulation amount to a fixed point computation. In the present case, we have a reverse accumulation algorithm presenting a rate of convergence of 3.7. Yet, the convergence speed of the forward fixed point (not plotted here) leads to a slope of 3. The convergence of the adjoint state computation is therefore higher than the one of the direct state computation. This result explains the *plateau* observed for the final iterations; indeed, a faster convergence of the reverse accumulation algorithm allows to reach the converged adjoint state with fewer iterations.

Again, the accuracy of the “self-adjoint” approximation appears strongly limited and the possibility of an incomplete method, intermediary between the complete adjoint and the retention of only one iteration could bring an important gain of precision; taking into account the linearly decreasing error (in logarithmic scale) leads to significantly

## Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



improved accuracy for each additional iteration retained during the computation of the adjoint state.

Furthermore, the faster convergence of the reverse accumulation algorithm compared to the direct solver allows, in any case, to spare few iterations during the computation of the adjoint state without any loss of precision. The number of unnecessary iterations is a priori strongly dependent on the situation and must be studied in every case.

For the present test case, we observe that the 5 last iterations during the reverse accumulation are useless whatever the level of precision of the forward run (see the *plateau* on Fig. 3). These 5 last iterations correspond to the 5 first iterations carried out by the direct solver. Avoiding the accumulation of these iterations for the adjoint state evaluation amounts to starting the reverse accumulation from a residual on the forward run of 0.1 (i.e. a relative variation between two successive iterates of 0.1). This observation, although dependent on the considered case, can be seen as an empirical method to define a criteria on the number of direct iterations that should be accumulated to obtain the best accuracy on the adjoint state. In the present case, it amounts to initiate the memory storage of direct iterations once the direct solver residual is lower than 0.1.

## 4 Basal slipperiness identifiability

This section focuses on the practical limits of identifiability of the basal slipperiness by both the full adjoint and the “self-adjoint” method.

The main goal is to draw conclusions on the possibility of using the “self-adjoint” method (which brings an important time and memory saving) and then on the quality of the results it provides in the perspective of realistic identification of the basal slipperiness. The quality of the results is evaluated in terms of frequencies and amplitudes of the reconstructed friction coefficients compared to the target ones.

## Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



As presented before, the precision of the “self-adjoint” gradient is bounded, whatever the level of precision of the direct solver  $\nu$ . This level of precision can be seen as an a priori accuracy on the data considered in the cost function. Only synthetic data are used in the present work. A Gaussian noise of level  $\delta$  is thus added a posteriori to emulate real data. The precision of the full adjoint gradient depending on  $\nu$  (and equivalently on the level of noise  $\delta$  on the data), we seek to observe which value of  $\nu$  is required to observe the limit of precision of the self-adjoint method.

To this end, we consider three noise levels  $\delta$  of 0.01 %, 0.1 % and 1 % representing respectively very low, low and realistic noise. In all cases, to just observe the final cost reached by both methods does not allow their precision, for a noise level greater or equal to 1 %, to be distinguished. On the contrary, the frequency analysis demonstrates that an identical final cost does not mean an identical reconstruction of the friction coefficient. It shows the equifinality issue on this type of ill-posed inverse problem.

We first consider the idealized case of a quasi-uniform flow in a sloppy channel with flat bottom with very low and low noise level in order to highlight the numerical limits of the “self-adjoint” method.

We then perform pseudo-realistic, spatially variable, flow experiments with a realistic noise for various density of the surface data. All the identifications presented hereafter use, as an initial guess for the friction coefficient, the average value  $a$  of the target coefficient.

### 4.1 Quasi-uniform flow

The following experiments are performed on the same flat bottom channel as in Sect. 3.1. A non-linear friction law, defined by Eq. (4) is considered at the bottom with an exponent  $m = 3$ . The target friction coefficient, variable in  $x$ , is given by:

$$\beta_r^n(x) = a + \frac{a}{2} \sin\left(\frac{2\pi x}{20dx}\right) + \frac{a}{5} \sum_{i=1}^n f_i(x) \quad (19)$$

and

$$f_1(x) = \sin\left(\frac{2\pi x}{10dx}\right) \quad (20)$$

$$f_2(x) = \sin\left(\frac{2\pi x}{4dx}\right) \quad (21)$$

$$f_3(x) = \sin\left(\frac{2\pi x}{2dx}\right) \quad (22)$$

where  $a$  is the average value of the basal slipperiness in  $\text{Pa} \cdot \text{s} \cdot \text{m}^{-1}$ . By extension, we set:

$$f_0(x) = \sin\left(\frac{2\pi x}{20dx}\right) \quad (23)$$

where  $dx = 0.2 \text{ m}$  denotes the length of a basal edge or, in other words, the sharpness of the bedrock discretization.

We set  $\beta_r = \beta_r^3$ , the friction coefficient resulting from the sum of 4 frequencies corresponding to wavelengths of 20, 10, 5 and 2 edge length  $dx$ . The low frequency  $f_0$  represents a carrier wave for the 3 higher frequencies  $f_i, 1 \leq i \leq 3$ . In terms of thickness of the domain  $h$ , frequencies  $f_i, 0 \leq i \leq 3$  correspond to wavelengths of  $4h, 2h, 0.8h$  and  $0.4h$  respectively. The coefficients  $\beta_r^n, 1 \leq n \leq 3$  are plotted on Fig. 4 for the case  $a = 1$ . These properties are summarized in Table 1.

The flow is uniform when the basal slipperiness is constant along the domain and can be described as quasi-uniform when the basal slipperiness is given by Eq. (19).

We seek to determine the level of spatial variability of the basal slipperiness the full adjoint and the “self-adjoint” method can provide through the identification process, based on surface velocity observations, with respect to the degree of slip. The degree of slip depends on the value of parameter  $a$  and will be described thereafter in terms of *slip ratio*  $r$ . The slip ratio is a dimensionless quantity that quantifies how slippery the

TCD

7, 3853–3897, 2013

## Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



bedrock is. It is calculated as the ratio of the mean sliding velocity  $\overline{u_b}$  to the difference between mean surface velocity  $\overline{u_s}$  and mean basal velocity  $\overline{u_b}$  (cf. Hindmarsh, 2004). It leads to:

$$r = \overline{u_b} / |\overline{u_s} - \overline{u_b}| \quad (24)$$

A slip ratio  $r = 1$  represents a situation where surface velocities result half from the sliding and half from the deformation.

We consider 6 different slip ratios ranging from very high friction (close to adherence) to very rapid sliding. The slip ratios  $r = 0.005$ ,  $r = 0.05$  and  $r = 0.5$  can be described as moderate sliding and the slip ratios  $r = 5$ ,  $r = 50$  and  $r = 500$  as rapid sliding.

In order to bring to the front the limitations of the “self-adjoint” approximation, the following identifications of  $\beta$  are performed for noise levels  $\delta = 0.1\%$  and  $\delta = 0.01\%$  on the surface velocity data. Let us point out that the “self-adjoint” method provides very similar results to the full adjoint in terms of final cost when  $\delta = 1\%$  (not plotted on Fig. 5) and the distinction between both methods clearly appears for lower noises.

The cost function is defined by:

$$j(\beta) = \frac{1}{2} \int_{\Gamma_s} \left| u_s(n) - u_s^{\text{obs}}(n_t) \right|^2 d\mathbf{x} + \gamma \int_{[0,L]} \nabla \beta \cdot \nabla \beta d\mathbf{x} \quad (25)$$

The tuning of the regularization parameter  $\gamma$  is achieved according to the Morozov discrepancy principle (see Sect. 2.3). We plot on Fig. 5 the application of this method to the identifications performed with both methods (full adjoint and “self-adjoint”) in the case of an intermediate friction ( $r = 0.5$ ). The corresponding curves for other slip ratios are identical and consequently not plotted.

Figure 5 clearly demonstrates the inability, for the “self-adjoint” method, to provide a gradient accurate enough for sufficiently low noise. As a matter of fact, the “self-adjoint” gradient does not allow the optimal misfit, corresponding to noise levels  $\delta = 0.1\%$  and  $\delta = 0.01\%$ , to be reached. In this situation, the “self-adjoint” approximation is a priori not valid. This observation is independent of the degree of slip.

## Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



In order to study the effects of the approximation on the gradient computation, we compare, in the following, the friction coefficient inferred by both methods for  $\delta = 0.01\%$  and  $\delta = 0.1\%$ .

The best inferred friction coefficient (according to Morozov) are noted  $\beta_f$  for the *full-adjoint* and  $\beta_s$  for the *self-adjoint*. The quantities  $\widehat{\beta}_f$  and  $\widehat{\beta}_s$  thus denote their associated Discrete Fourier Transform (DFT). We denote by  $\widehat{\beta}_r$  the DFT associated to the target friction coefficient (Eq. 19). The DFT  $\widehat{\beta}_r$ ,  $\widehat{\beta}_f$  and  $\widehat{\beta}_s$  obtained for the three small slip ratios (moderate sliding) are plotted on Fig. 6 and those obtained for the three high slip ratios (rapid sliding) are plotted on Fig. 7.

#### 4.1.1 Moderate sliding

One observes first that frequencies  $f_0$  and  $f_1$  (see Table 1) are globally well reproduced by both methods for  $\delta = 0.01\%$  whatever the slip ratio. Namely the carrier frequency  $f_0$  is very well reconstructed by both methods and this property seems desirable. The full adjoint method shows a greater robustness when identifying these two low frequencies with respect to the slip ratio whereas a noticeable deterioration for the identification of frequency  $f_1$  occurs for the “self-adjoint” method when slip ratio decreases.

However, frequency  $f_2$  appears correctly captured by the full adjoint method while it does not appear in the spectrum of the “self-adjoint” one. An increased difficulty in capturing this frequency occurs with slip ratio decreasing.

Finally, the highest frequency  $f_3$  does not appear in any of the spectrums of both  $\beta_f$  and  $\beta_s$  whatever the degree of slip.

For a noise level  $\delta = 0.1\%$ , one loses the ability to retrieve frequency  $f_2$  using the exact method. The identification of frequency  $f_1$  is accurately obtained for the slip ratio  $r_1 = 0.5$  but we observe a deterioration of the result when slip ratio decreases. The “self-adjoint” method captures almost none of frequency  $f_1$  whatever the slip ratio.

Concerning the carrier frequency, one observes difficulties for the “self-adjoint” method to reconstruct it accurately even for the slip ratio  $r_1 = 0.5$ . The frequency dis-

tinctly appears on the spectrum but only 80 % of the target amplitude is recovered. The decreasing of the slip ratio deteriorates, for both methods, the identification of  $f_0$ . In the case  $r_3 = 0.005$ , the full adjoint method recovers 70 % of the target amplitude where the “self-adjoint” method recovers 50 %.

5 **4.1.2 Rapid sliding**

Again, low frequencies  $f_0$  and  $f_1$  are well retrieved with the full adjoint method for every noise level. The carrier wave reconstruction is nevertheless diminished (around 80 % of the target amplitude) compared to the moderate sliding situation  $r_1 = 0.5$  but stable with the increasing of  $r$ . Similarly, frequency  $f_1$  is rather well represented by the exact method for all the situations despite a certain degradation with increasing  $r$ . However, frequency  $f_2$  does not appear in any spectrum irrespective of both slip ratio and method, contrarily to the moderate sliding situations. Again, frequency  $f_3$  is never captured. A small but noticeable noise appears for the case  $r = 500$  for the full adjoint method, particularly when  $\delta = 0.01$  %.

15 The “self-adjoint” method shows a relatively good reconstruction of  $f_0$  and  $f_1$  for the case  $r = 5$  but introduces a noises between frequencies  $f_1$  and  $f_2$ . A strong deterioration of the reconstruction occurs when  $r$  increases; for a noise level  $\delta = 0.1$  %, the “self-adjoint” identification is almost unable to recover the signal for  $r \geq 50$ .

**4.1.3 Assessments**

20 From these observations, we draw the following conclusions. Firstly, the degree of slip of the target plays a strong role for the limit of identifiability of the basal slipperiness in terms of frequencies; a smaller slip ratio induces a lower sensitivity of the flow to the basal slipperiness and consequently a higher filtering on the transmission of information from the bedrock to the surface.

Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



This result is in coherence with the result presented in Martin and Monnier (2013) where one observes a propagation of the sensitivity along the thickness when the vertical velocity profile approaches the one of a *plug* flow.

A strong friction induces a vertical velocity profile rather convex with velocity gradients (shearing) mostly concentrated close to the bottom and thus a lower transmission of the information from the bottom to the surface.

Similarly to strong frictions, low frictions also reduce the quality of the reconstruction. Again it comes from a reduced sensitivity of the flow to the basal slipperiness when rapid sliding occurs but this lower sensitivity appears for different reasons. Intuitively, the case of a very low friction leads to lower local topographical effects and the resistance to the ice flow acts through an equivalent global topography at larger scale. This characteristic appears in the explicit solution of the uniform flow (Eq. 6): in order for the mathematical expression to makes sense when  $\beta$  tends to 0, it requires the slope parameter  $\theta$  to tend to 0 as well. This phenomena is physically observed: in the presence of an extended sub-glacial lake, one observes a signature of this lake at the surface as a very flat surface topography over the lake. This interpretation is retrieved in the normalized sensitivities plotted on Fig. 10.

These two observations allow to state the existence of a numerical identifiability maximum for the friction coefficient using adjoint-based method; the best situation to carry out quality identifications corresponds to the intermediate friction range where sliding effects and deformation effects on the dynamics are balanced (typically  $0.5 < r < 5$ ). The low accuracy of the “self-adjoint” gradient appears to be a strong limitation in the case of rapid sliding ( $r > 5$ ).

For the current quasi-uniform flow, for a noise level  $\delta = 0.1\%$ , a limit on identifiable wavelength using full adjoint, for any degree of slip, is  $2h$ ,  $h$  the thickness of the domain. More accurate data could allow to infer higher frequencies in the case of moderate sliding ( $r \geq 0.5$ ).

## Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



For the “self-adjoint” method, for a slip ratio  $r \leq 5$ , a wavelength of  $4h$  is well inferred and a wavelength of  $2h$  is captured for  $r = 0.5$  and  $r = 5$ . For a slip ratio  $r > 5$ , the frequencies considered in the experiment are inappropriate.

In other respects, a tendency for the “self-adjoint” method to introduce non-physical interferences within the inferred coefficient for very low noise appears. This non desirable phenomena increases when the slip ratio takes extreme values. Beyond the approximation aspect, one can deduce a lack of robustness of the “self-adjoint” method for very low noises. It seems coherent regarding the low precision the “self-adjoint” gradient provides. On the contrary, the full adjoint method provides a less accurate identification when the slip ratio goes away from 1 without introducing non physical effects in the inferred parameter.

It is of interest to notice that the inability to recover frequency  $f_3$  is not a numerical limitation. For sufficiently accurate data, it is also identifiable using the exact adjoint.

Similar experiments are performed in the next section for a pseudo-realistic flow ran on a radar vertical profile of the grounded part of the Mertz glacier in Antarctica for surface velocity data with different density and a 1 % noise.

## 4.2 Real topography flow: the Mertz glacier

The flow considered in this section is identical to the one presented in Eq. (2.1). The computational domain is built from real field data; topography of the bedrock and of the surface are bidimensional radar-sensed layer of the *Mertz ice tongue* in East Antarctica. These layers have been measured along a flowline of this outlet glacier (American program ICECAP 2010, see Greenbaum et al., 2010). Our study focuses on the grounded part of the glacier.

Synthetic data are obtained using the following friction coefficient:

$$\beta_r^n(x) = a + \frac{a}{2} \sin\left(\frac{2\pi x}{50dx}\right) + \frac{a}{5} \sum_{i=1}^n f_i(x) \quad (26)$$

## Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



and

$$f_1(x) = \sin\left(\frac{2\pi x}{20dx}\right) \quad (27)$$

$$f_2(x) = \sin\left(\frac{2\pi x}{10dx}\right) \quad (28)$$

$$f_3(x) = \sin\left(\frac{2\pi x}{5dx}\right) \quad (29)$$

with  $a$  the average basal slipperiness,  $dx = 100$  m the bedrock edge length and one set:

$$f_0(x) = \sin\left(\frac{2\pi x}{50dx}\right) \quad (30)$$

The context of a non-uniform flow on a complex topography allows to carry on the comparison between both methods in the case of a realistic flow simulation. We can then draw practical conclusions on the validity of using the “self-adjoint” approximation. Frequency  $f_0$  is a carrier wave with  $50dx$  wavelength corresponding to  $5h$ ,  $h \sim 1$  km the average thickness of the domain. Frequencies  $f_1$ ,  $f_2$  and  $f_3$  correspond then to wavelengths of  $2h$ ,  $h$  and  $h/2$  respectively, providing a situation similar to the channel test case (see Table 2).

In the present case of a non uniform flow with complex topography, it is not feasible to simulate an average slip ratio  $r = 500$ . Given the important spatial variability, we are able to achieve a maximum average slip ratio  $r = 50$ . In the following identification, we consider only 5 slip ratios ranging from  $r = 0.005$  to  $r = 50$ .

The Morozov’s discrepancy principle applied to these 5 situations is plotted on Fig. 8.

The observed behavior is similar to the one noticed (but not plotted) for the previous academical situation. Both methods behave identically in terms of cost decreasing for a 1 % noise level on the data. They demonstrate a robust behavior providing in all cases an optimal discrepancy (according to Morozov). The trend of over-fitting of the cost (i.e.

## Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



to realize a discrepancy smaller than the reachable one when using the target coefficient with perturbed data) already observed (for the exact adjoint) demonstrates that the gradient accuracy provided by both methods is a priori adapted to the noise level considered on the data. The peculiar behavior for the case  $r = 50$  where the discrepancy remains lower than the optimal one regardless of the regularization parameter  $\gamma$  is detailed thereafter.

Figure 9 plots the DFT of the friction coefficients inferred by both methods and of the target coefficient (Eq. 19) for a noise level of 1 % on the data and for the 5 slip ratios  $r$ .

While the wavelengths considered in the friction coefficient (Eq. 19) are similar (in terms of thickness ratio) to those considered for the quasi-uniform test case, the use of a higher noise on a non-uniform flow deteriorated the reconstruction at all levels. The carrier frequency amplitude (of wavelength  $5h$ ) is never fully recovered by any method but clearly appears for  $r \leq 5$ . Likewise, frequency  $f_1$  (of wavelength  $2h$ ), well captured in previous simulations by the full adjoint method, is fairly well reconstructed only for  $0.05 \leq r \leq 5$ . Again the “self-adjoint” method is able to recover it only very partially. However, the interferences introduced by the “self-adjoint” method within the inferred friction coefficient do not appear anymore for this level of noise on the surface data. It therefore seems coherent with the limited accuracy of the gradient provided by this method.

As a consequence, the chosen frequencies for these simulations are too high to be recovered in this non-uniform flow with realistic data. Numerical experiments show that an accurate reconstruction can be obtained, for the exact method, for a carrier wave of wavelength  $10h$  and a perturbation of wavelength  $5h$ ; shorter wavelengths are not accessible.

What is of further interest is that the exact method brings, in all cases, an enhanced and more faithful reconstruction of the friction coefficient for both the carrier wave and the first perturbation.

The pattern of behavior of the rapid sliding case ( $r = 50$ ) is different compared to the other cases. The full adjoint method retrieves roughly the carrier frequency with very

## Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

high interferences (including one low frequency high amplitude interference) and the “self-adjoint” method does not capture any information of the target signal in addition to the initial guess.

In order to understand this phenomena, we plot on Fig. 10 the gradients  $\partial j / \partial \beta(\beta_0)$  with  $\beta$  defined by (Eq. 26) for several average value  $a$  of the basal slipperiness, described in terms of slip ratio  $r$ . The computed gradients are evaluated around  $\beta_0 = a$ .

Increasing the slip ratio has a very clear effect on the sensitivities. For slip ratios  $r < 1$ , the sensitivities include the local effects of the high frequencies contained in  $\beta$ , thus providing a highly variable gradient around an average behavior. The fact that the sensitivity decreases with  $r$ , due to poorer information transmission between the bottom and the surface, is recovered. It follows that, in the cases  $r < 1$ , the limitations in the identification of all the frequencies of the basal slipperiness come from the precision on the data.

The situations  $r > 1$  bring significantly smoother gradients. The cases  $r = 6$  and  $r = 13$ , that still represent moderate slip ratios, contain a certain local variability but their rather smooth appearance shows a strong correlation with the global topography and the high frequencies of  $\beta$  seem already erased from the gradient. In these situations, the main component resisting the flow is more the large scale (or equivalent) topography than the friction itself.

For higher slip ratios, the topographical effects seem to vanish as well, and the gradient only grows from the inflow boundary to the outflow boundary to reach a maximum value close to the right border. In the present case, one can deduce that the only effect resisting the flow is the hydrostatic pressure considered on the right boundary.

A global decreasing of the sensitivity with increasing  $r$  is also observed, reinforcing the existence of a sensitivity peak for in-between  $r$ . For  $r > 1$ , it is not the quality of the data that prevents an accurate reconstruction of  $\beta$  but the non-local behavior of the flow. When basal slipperiness vanishes, it does not embody more than a small fraction of the global resistance to the flow. An extreme example is the progress of an ice-shelf on water where the friction resistance is close to zero. In the case of a tridi-

## Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

mensional solution, stresses would be taken over by lateral shearing. In our case, these effects do not exist and it is the hydrostatic pressure boundary condition that resists the flow. These clearly non-local effects demonstrate the inability to control otherwise than globally the flow, thus limiting the range of identifiable frequencies regardless of data accuracy.

These phenomena imply a strong equifinality for basal slipperiness lower than a certain value. This observation appears in the Morozov's curves (see Fig. 8) for the case  $r = 50$ . Indeed, the discrepancies for both methods are smaller than the theoretical optimal one, even for very strong regularization ( $\gamma$  large) providing almost constant  $\beta$  around  $\beta_0$ . The initial cost itself, evaluated for a constant  $\beta$  equal to the average value, is barely higher than the theoretical optimal cost.

The associated minimization problem is ill-posed and the Tykhonov regularization on the gradient of  $\beta$  does not allow to overcome this problem.

For the case  $r = 50$  and a regularization small enough (considering that the Morozov's principle does not allow the optimal  $\gamma$  value to be selected) it is noticeable that the full adjoint method is able to retrieve a small quantity of information, along with a large noise (optimal control problem obviously ill-posed) whereas the "self-adjoint" method does not provide anything else than the initial guess, irrespective of the value of  $\gamma$ .

## 5 Density of the data

The previous simulations has been performed using surface velocity data quite dense (one measure every  $dx$ ). This section deals with identical test cases to the previous section but using sparser (one measure point every 1 km) and thus more realistic data (corresponding to one ice thickness, see e.g. Gudmundsson and Raymond, 2008). This density corresponds to approximately 10 times less measure points than the previous

### Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



case. We consider thereafter the following friction coefficient for the synthetic data:

$$\beta_r^n(x) = a + \frac{a}{2} \sin\left(\frac{2\pi x}{200dx}\right) + \frac{a}{5} \sum_{i=1}^n f_i(x) \quad (31)$$

with  $dx = 100$  m the edge length at the bottom and with:

$$f_1(x) = \sin\left(\frac{2\pi x}{100dx}\right) \quad (32)$$

$$f_2(x) = \sin\left(\frac{2\pi x}{50dx}\right) \quad (33)$$

$$f_3(x) = \sin\left(\frac{2\pi x}{20dx}\right) \quad (34)$$

and:

$$f_0(x) = \sin\left(\frac{2\pi x}{200dx}\right) \quad (35)$$

10 The friction coefficient chosen for these simulation contains lower frequencies than the previous one, simulating a carrier wave of wavelength  $20h$  perturbed by high frequencies of wavelengths  $10h$ ,  $5h$  and  $2h$ . These characteristics are summarized in Table 3. Results are plotted on Fig. 11 for a noise level of 1 %.

15 As a consequence, the level of identifiability assessed for dense data in the previous section is no longer valid. However, considering that one out of ten points has been retained, results seem rather convincing. The exact method is able to accurately recover frequencies of wavelengths  $20h$  and  $10h$  (corresponding to  $f_0$  and  $f_1$ ) for all degrees of slip. The “self-adjoint” method recovers the carrier wave quite well although a stronger friction ( $r \leq 0.05$ ) significantly degrades the reconstruction of the amplitude. Frequency  
20  $f_1$  is well captured for propitious situations ( $0.5 \leq r \leq 5$ ). Frequency  $f_2$  (of wavelength

## Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

5*h*, the lowest frequency considered in the dense data situation) is partially reconstructed by the exact method for  $r \leq 5$  and never captured by the “self-adjoint” method.

The case  $r = 50$  is a lot less problematic than previously, due to lower frequencies and subsequently less local effects regarding the sharpness of the bed discretization.

- 5 A pronounced difficulty appears for the identification of frequency  $f_1$  (of wavelength 10*h*). The case  $r = 50$  is the only one where frequency  $f_2$  does not appear in the spectrum of  $\hat{\beta}_f$  (consistently to the previous simulations).

## 6 Conclusions

10 The significant CPU-time saving brought by the “self-adjoint” method represents an important asset in its favor. However, its reliability is questionable and it seems important to know its limitations in order to perform realistic experiments.

15 The realistic simulation (low density data, 1 % noise, real topography, non-linear friction) allows us to assess the full adjoint method’s ability to identify accurately wavelength greater or equal to 10 ice thickness and to capture effects of wavelength up to 5 thickness for slip ratio lower than 5.

The “self-adjoint” method is able to reconstruct wavelengths greater than 20 thickness (with noticeable difficulties for strong friction). Wavelengths of 10 ice thickness can be captured in propitious situations of intermediate sliding ( $0.5 \leq r \leq 5$ ).

20 The results provided by the exact method are significantly better than those given by Petra et al. (2012) (who assess a limit of 20 ice thickness for a non-linear rheology). It is difficult to compare considering that the authors provide neither their slip ratio nor the density of the data. In addition, the authors of Petra et al. (2012) use a linear friction law.

25 The use of a non-linear friction allows to simulate complex behaviors of the ice-bedrock interaction. This type of law can describe a non-linear deformation of the basal substrate or a non-linear response of the sliding velocity to the water pressure of subglacial cavities. The former reconstructions focus on the identification of a generic  $\beta$ .

## Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



However one may confidently generalize these results to more complex sliding laws where  $\beta$  would be identified through its parameterization (by a water pressure, a contact surface with sub-glacial cavities, a sedimentary roughness, a geothermal flux,). It is important to recall that over-parameterization is hardly ever in favor of an accurate identification and that the identification of several parameters simultaneously would strongly reinforce the problem of equifinality (i.e. the ill-posedness of the optimization problem).

This work focuses on the identification of the basal slipperiness that plays a major role to control the flow (i.e. the model shows a great sensitivity to the friction). The identification of a parameter such as the consistency  $\eta_0$ , for which the model sensitivity is significantly lower, needs to be done with precaution for the exact adjoint method (see Martin and Monnier, 2013) and thus with increased precaution for the “self-adjoint” method.

Finally, the adjoint obtained from source-to-source algorithmic differentiation allows to simulate every level of needed precision between the best precision of the exact adjoint to the lowest one of the “self-adjoint” approximation. It leads to the consideration of an *incomplete adjoint* methodology where the approximation is completely adjustable, thus allowing the right compromise between CPU-time, memory burden and required accuracy to be achieved. Numerical experiments demonstrate that the retention of the last two states within the gradient computation significantly improves its precision while maintaining a quite small computational burden.

**Acknowledgements.** This work was partially supported by PRES Toulouse, with the PhD fund of the first author. The authors want to thank Benoit Legrésy (LEGOS) for his data on the Mertz glacier and Ronan Madec (IMT) for his help on the development of the adjoint model. This work was also supported by Agence Nationale de la Recherche through ADAGe project No. ANR-09-SYSC-001.

## References

- Balise, M. J. and Raymond, C. F.: Transfer of basal sliding variations to the surface of a linearly viscous glacier, *J. Glaciol.*, 31, 308–318, 1985. 3854
- Christianson, B.: Reverse accumulation and attractive fixed points, *Optim. Method. Softw.*, 3, 311–326, 1994. 3863, 3866
- DassFlow Software: Data assimilation for free-surface Flows, open source freeware, available at: <http://www-gmm.insa-toulouse.fr/~monnier/DassFlow>, June 2013. 3859
- Gilbert, J. C. and Lemaréchal, C.: Some numerical experiments with variable-storage quasi-Newton algorithms, *Math. Programm.*, 45, 407–435, 1989. 3861
- Goldberg, D. N. and Sergienko, O. V.: Data assimilation using a hybrid ice flow model, *The Cryosphere*, 5, 315–327, doi:10.5194/tc-5-315-2011, 2011. 3855, 3864
- Greenbaum, J. S., Blankenship, D. D., Young, D. A., Richter, T. G., Legresy, B., Galton-Fenzi, B., and Gim, Y.: Basal characteristics and inferred bathymetry beneath the Mertz Glacier Tongue, Antarctica from coupled airborne radar sounding and gravity prior to the February 12th 2010 breakup event, in: 4th SCAR Open Science Conference – Antarctica: Witness to the Past and Guide to the Future, Submitted Abstracts, 2010. 3874
- Griewank, A.: On automatic differentiation, *Math. Programm.*, 6, 83–107, 1989. 3860, 3863
- Griewank, A., Bischof, C., Corliss, G., Carle, A., and Williamson, K.: Derivative convergence for iterative equation solvers, *Optim. Method. Softw.*, 2, 321–355, 1993. 3863
- Gudmundsson, G. H.: Transmission of basal variability to a glacier surface, *J. Geophys. Res.-Solid Earth*, 108, 2253, doi:10.1029/2002JB002107, 2003. 3855
- Gudmundsson, G. H. and Raymond, M.: On the limit to resolution and information on basal properties obtainable from surface data on ice streams, *The Cryosphere*, 2, 167–178, doi:10.5194/tc-2-167-2008, 2008. 3855, 3878
- Hascoët, L. and Pascual, V.: TAPENADE 2.1 User's Guide, available at: <http://www-tapenade.inria.fr:8080/tapenade/index.jsp>, 2004. 3860
- Hindmarsh, R. C. A.: A numerical comparison of approximations to the Stokes equations used in ice sheet and glacier modeling, *J. Geophys. Res.*, 109, F01012, doi:10.1029/2003JF000065, 2004. 3870
- Honnorat, M., Marin, J., Monnier, J., and Lai, X.: Dassflow v1.0: a variational data assimilation software for 2-D river flows, INRIA Research Report, RR-6150, available at: <http://hal.inria.fr/inria-00137447>, 2007. 3859, 3862

### Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Joughin, I., MacAyeal, D. R., and Tulaczyk, S.: Basal shear stress of the Ross ice streams from control method inversions, *J. Geophys. Res.*, 109, B09405, doi:10.1029/2003JB002960, 2004. 3855
- Larour, E., Rignot, E., Joughin, I., and Aubry, D.: Rheology of the Ronne Ice Shelf, Antarctica, inferred from satellite radar interferometry data using an inverse control method, *Geophys. Res. Lett.*, 32, L05503, doi:10.1029/2004GL021693, 2005. 3855
- Lions, J. L.: *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, 1971. 3859
- Macayeal, D. R.: A tutorial on the use of control methods in ice-sheet modeling, *J. Glaciol.*, 39, 91–98, 1993. 3855, 3863
- Martin, N.: Direct and Inverse modeling of geophysical flows usgin variational methods – Application to glaciology, Ph.D. thesis, Institut National des Sciences Appliquées, Université de Toulouse, 2013. 3859
- Martin, N. and Monnier, J.: Inverse rheometry for power-law geophysical flows investigated using variational sensitivities and data assimilation, in review, 2013. 3857, 3873, 3881
- Morlighem, M., Rignot, E., Seroussi, H., Larour, E., Dhia, H. B., and Aubry, D.: Spatial patterns of basal drag inferred using control methods from a full-Stokes and simpler models for Pine Island Glacier, West Antarctica, *Geophys. Res. Lett.*, 37, L14502, doi:10.1029/2010GL043853, 2010. 3855, 3864
- Petra, N., Zhu, H., Stadler, G., Hughes, T. J. R., and Ghattas, O.: An inexact Gauss-Newton method for inversion of basal sliding and rheology parameters in a nonlinear Stokes ice sheet model, *J. Glaciol.*, 58, 889–903, doi:10.3189/2012JoG11J182, 2012. 3855, 3880
- Raymond, C.: Shear margins in glaciers and ice sheets, *J. Glaciol.*, 42, 90–102, 1996. 3855
- Raymond, M. J. and Gudmundsson, G. H.: Estimating basal properties of ice streams from surface measurements: a non-linear Bayesian inverse approach applied to synthetic data, *The Cryosphere*, 3, 265–278, doi:10.5194/tc-3-265-2009, 2009. 3855
- Vogel, C. R.: *Computational Methods for Inverse Problems*, SIAM, 2002. 3859

## Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



**Of the gradient  
accuracy in  
Full-Stokes ice flow  
model**

N. Martin and J. Monnier

**Table 1.** Characteristics of signal  $\beta$  given by Eq. (19).

	$f_0$	$f_1$	$f_2$	$f_3$
Wavelengths w.r.t $h = 1$ m (thickness)	$4h$	$2h$	$0.8h$	$0.4h$
Wavelengths w.r.t. $dx = 0.2$ m (edge length)	$20dx$	$10dx$	$4dx$	$2dx$
Wavelengths w.r.t. $L = 10$ m (domain length)	$2.5$ Hz	$5$ Hz	$12.5$ Hz	$25$ Hz

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

I◀

▶I

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

## TCD

7, 3853–3897, 2013

# Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

**Table 2.** Characteristics of signal  $\beta$  given by Eq. (26).

	$f_0$	$f_1$	$f_2$	$f_3$
Wavelengths w.r.t $h = 1$ km (thickness)	$5h$	$2h$	$h$	$0.5h$
Wavelengths w.r.t $dx = 100$ m (edge length)	$50dx$	$20dx$	$10dx$	$5dx$
Wavelengths w.r.t $L = 33.3$ km (domain length)	6.6 Hz	16.6 Hz	33.3 Hz	66.6 Hz

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

I◀

▶I

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

**Of the gradient  
accuracy in  
Full-Stokes ice flow  
model**

N. Martin and J. Monnier

**Table 3.** Characteristics of signal  $\beta$  given by Eq. (31).

	$f_0$	$f_1$	$f_2$	$f_3$
Wavelengths w.r.t $h = 1$ km	$20h$	$10h$	$5h$	$2h$
Wavelengths w.r.t $dx = 100$ m	$200dx$	$100dx$	$50dx$	$2dx$
Wavelengths w.r.t $L = 33.3$ km	1.66 Hz	3.33 Hz	6.66 Hz	16.6 Hz

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

I◀

▶I

◀

▶

Back

Close

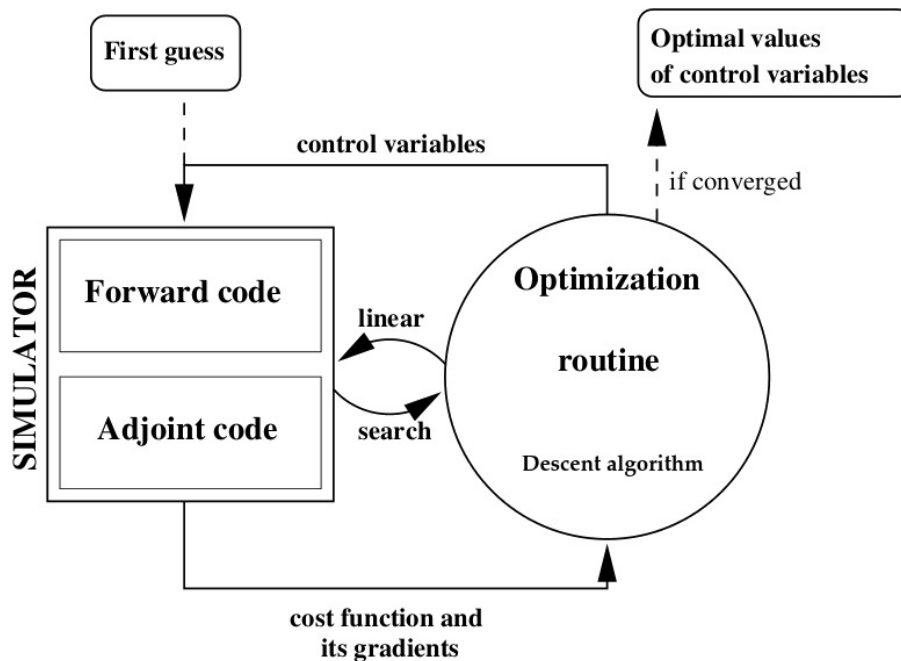
Full Screen / Esc

Printer-friendly Version

Interactive Discussion

# Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

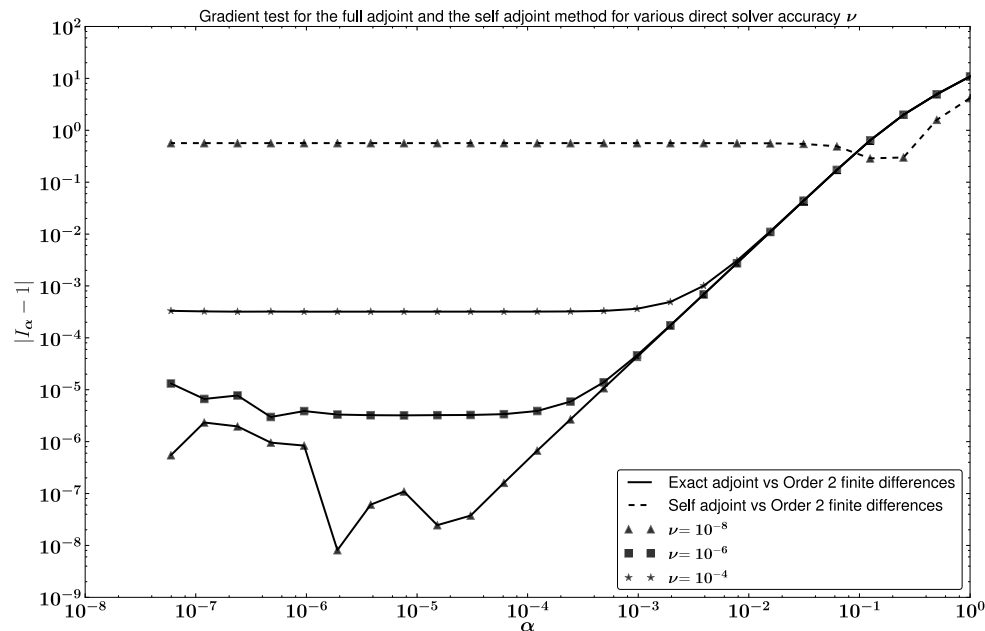


**Fig. 1.** Principle of a 3-D-Var type variational data assimilation algorithm.

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)

# Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

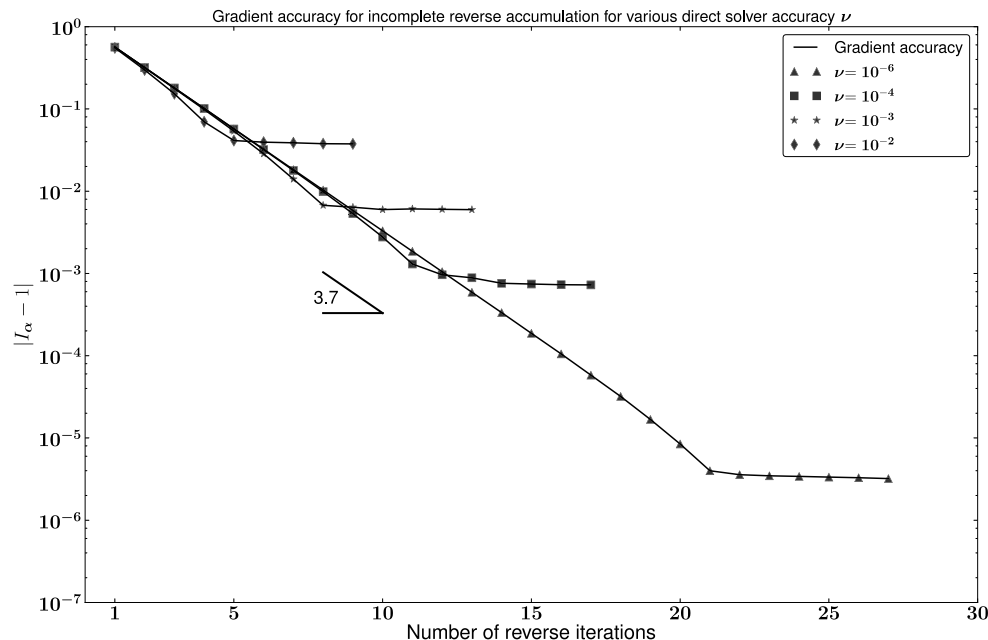


**Fig. 2.** Gradient test for the exact adjoint method the “self-adjoint” method for various levels of precision  $\nu$  of the forward solution. The quantity  $I_\alpha$  is defined by Eq. (17).

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)


# Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier



**Fig. 3.** Accuracy of the gradient for incomplete reverse accumulation for various levels of precision of the direct solution  $\nu$ .

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

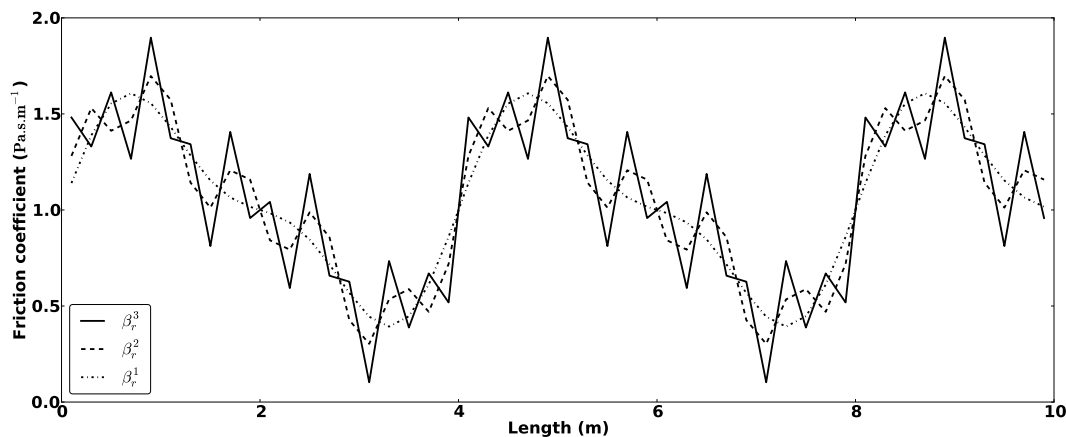
Printer-friendly Version

Interactive Discussion



# Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

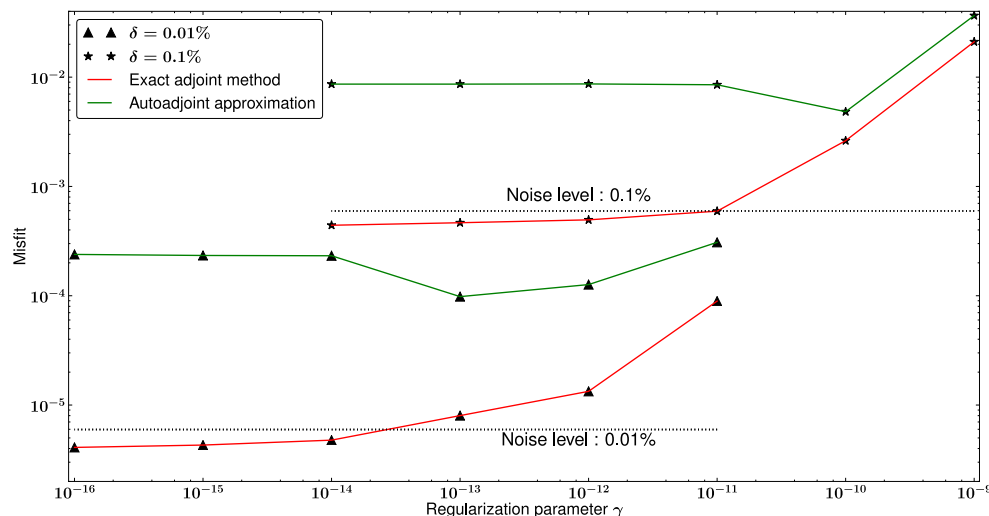


**Fig. 4.** Friction coefficient  $\beta_r^n$ ,  $1 \leq n \leq 3$  given by Eq. (19) with  $a = 1$ .

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)

# Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier



**Fig. 5.** Application of Morozov's discrepancy principle. Final discrepancy (or misfit)  $j_1 = \int_{\Gamma_s} \left| u_s(n) - u_s^{\text{obs}}(n_t) \right|^2 dx$  with respect to the value of the regularization parameter  $\gamma$ . The horizontal line represents the level of noise corresponding to the optimal discrepancy obtained from the target coefficient.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

# Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

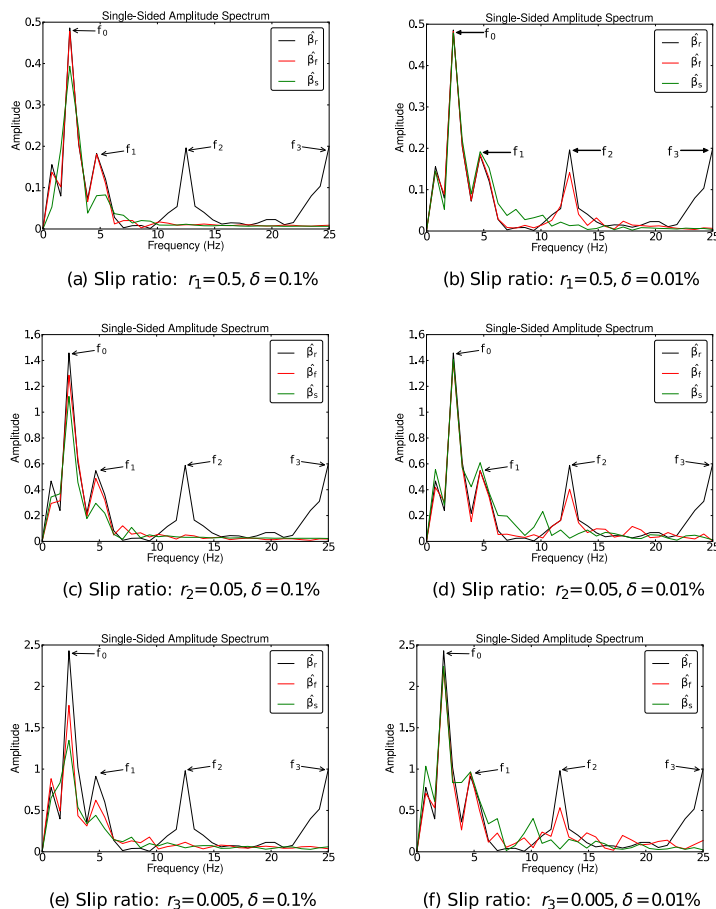
Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



**Fig. 6.** Discrete Fourier Transform of inferred friction coefficient  $\beta_t$  and  $\beta_s$  and of the target friction coefficient  $\beta_r$ . Moderate sliding.

# Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

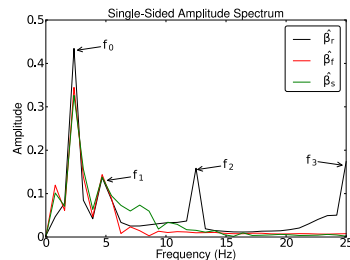
Back

Close

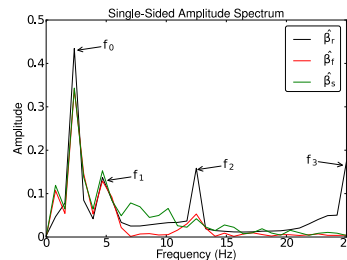
Full Screen / Esc

Printer-friendly Version

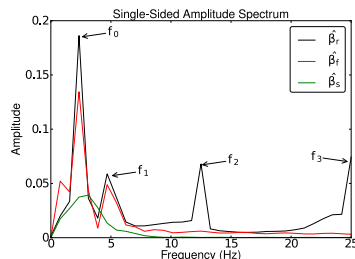
Interactive Discussion



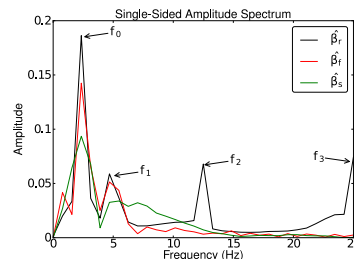
(a) Slip ratio:  $r_1 = 5$ ,  $\delta = 0.1\%$



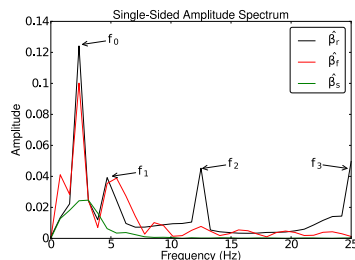
(b) Slip ratio:  $r_1 = 5$ ,  $\delta = 0.01\%$



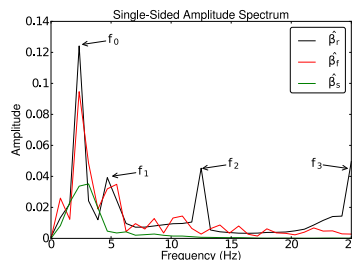
(c) Slip ratio:  $r_2 = 50$ ,  $\delta = 0.1\%$



(d) Slip ratio:  $r_2 = 50$ ,  $\delta = 0.01\%$



(e) Slip ratio:  $r_3 = 500$ ,  $\delta = 0.1\%$

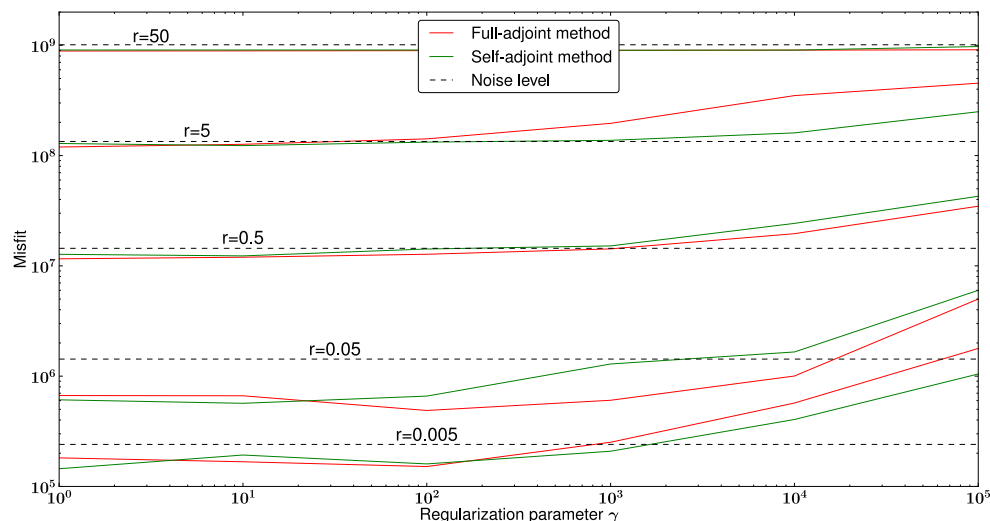


(f) Slip ratio:  $r_3 = 500$ ,  $\delta = 0.01\%$

**Fig. 7.** Discrete Fourier Transform of inferred friction coefficient  $\hat{\beta}_f$  and  $\hat{\beta}_s$  and of the target friction coefficient  $\hat{\beta}_r$ . Rapid sliding.

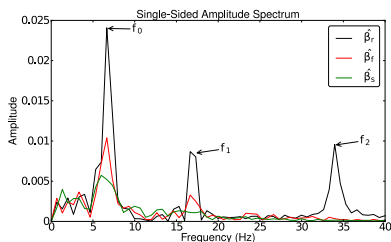
# Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

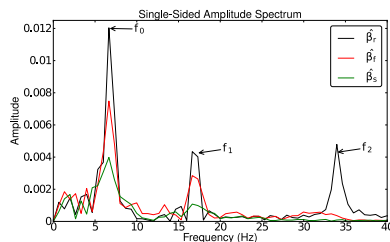


**Fig. 8.** Morozov's discrepancy principle applied to slip ratios  $r = 0.005$ ,  $r = 0.05$ ,  $r = 0.5$ ,  $r = 5$  and  $r = 50$  on the realistic flow. Absolute values of the discrepancy correspond to the real value obtained during the simulations. The range of parameter  $\gamma$  has been modified to remain between 1 and  $10^5$  for the sake of readability.

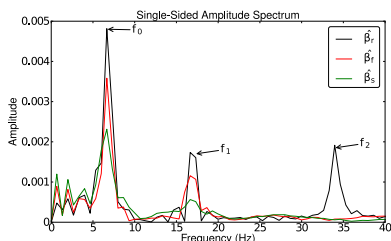
[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)



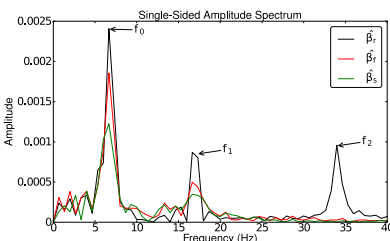
(a)  $\delta = 1\%$ ,  $r = 0.005$



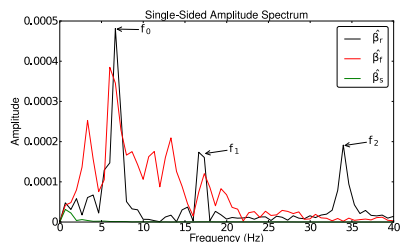
(b)  $\delta = 1\%$ ,  $r = 0.05$



(c)  $\delta = 1\%$ ,  $r = 0.5$



(d)  $\delta = 1\%$ ,  $r = 5$

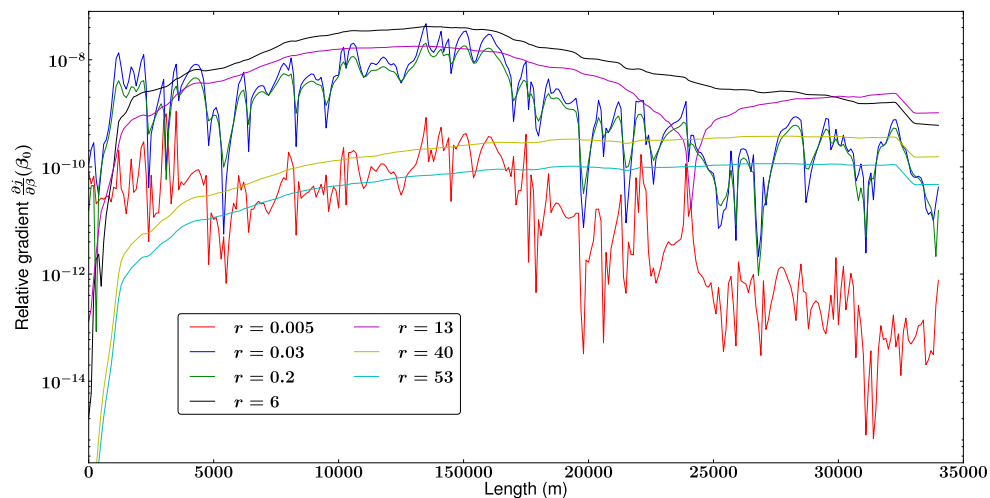


(e)  $\delta = 1\%$ ,  $r = 50$

**Fig. 9.** Discrete Fourier Transform for inferred friction coefficients  $\beta_f$  and  $\beta_s$  and for the target one  $\beta_r$ . Frequency  $f_3$  is never captured by any method and is thus not plotted on the curves. A noise level  $\delta = 1\%$  is used in all situations.

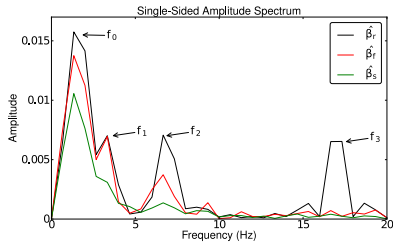
# Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

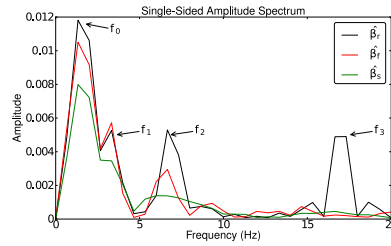


**Fig. 10.** Relative sensitivities to the friction coefficient with respect to the abscissa  $x$  for various slip ratios, evaluated around the average value  $a$ .

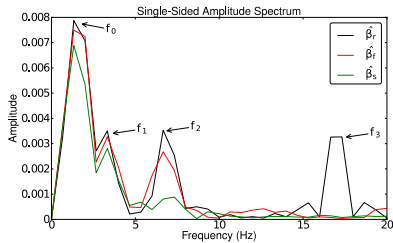
[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)



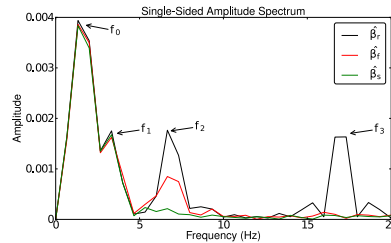
(a)  $\delta = 1\%$ ,  $r = 0.005$



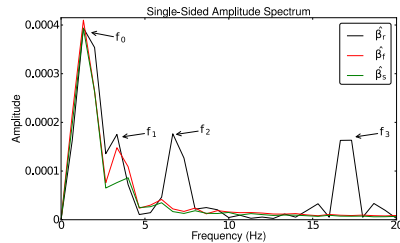
(b)  $\delta = 1\%$ ,  $r = 0.05$



(c)  $\delta = 1\%$ ,  $r = 0.5$



(d)  $\delta = 1\%$ ,  $r = 5$



(e)  $\delta = 1\%$ ,  $r = 50$

**Fig. 11.** Discrete Fourier Transform for inferred friction coefficients  $\beta_f$  and  $\beta_s$  and for the target one  $\beta_r$  for sparse data a 1 % noise level.

TCD

7, 3853–3897, 2013

# Of the gradient accuracy in Full-Stokes ice flow model

N. Martin and J. Monnier

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion