Inference of velocity pattern from isochronous layers in firn, using an inverse method

Olaf EISEN *

Versuchsanstalt für Wasserbau, Hydrologie und Glaziologie (VAW), ETH Zürich, Gloriastr. 37-39, CH-8092 Zürich
E-mail: oeisen@awi-bremerhaven.de

ABSTRACT. This study investigates the suitability of a kinematic approach to find the velocity field from dated internal-layer architecture in firn. Internal layers are isochrones and the depositional age of a layer particle is treated as a tracer. The forward problem uses two-dimensional steady-state advection of age, and conservation of mass to predict layer architecture. Different combinations of constraints on horizontal or vertical velocity properties are added. The inverse problem can be formulated as the solution of underdetermined and overdetermined systems of equations . The systems are solved using singular-value decomposition, allowing analysis of the singular-value spectrum, model **resolution,** and data resolution. **Solutions of the inverse problem** are evaluated by comparing the velocity-field solutions with synthetic input velocity data. Compared to conventional accumulation estimates, **the new approach takes lateral advection into account,** enabling improved separation of spatial and temporal variations in accumulation. **Two glaciological applications are presented: the determination of the migration velocity of a spatially non-stationary accumulation pattern, and reconstruction of past accumulation and its stationarity over time.**

1. INTRODUCTION

Internal layering is widely observed by radar sounding in cold firn and ice, on high alpine and polar glaciers as well as ice sheets. Layer architecture results from the interplay of spatio-temporal variation of surface accumulation, bottom melting, and advection caused by ice dynamics. Most layers are isochrones, **i.e. surfaces of equal age.** Whereas age information retrieved from ice cores is representative only for the immediate vicinity of the drilling location, the layer architecture provides a spatial picture. It represents an integrated view of the temporal evolution of an ice mass.

Several studies exploited this property to enhance the view of past conditions and to understand present conditions. The simplest application is the one-dimensional direct inversion of layer depth and density distribution for accumulation, covering shallow depth and a view millennia at most (see Annals of Glaciology 39 and 41, and references therein, for a summary of studies). However, effects of horizontal advection are not considered; these effects can introduce errors into the inferred accumulation. Recently, Arcone and others (2005) used an accumulation-rate model to investigate how accumulation-rate anomalies and ice velocity affect stratigraphic variations of internal layers. Other approaches utilize forward modeling [illegible] and least-squares techniques to solve for the accumulation rate by minimizing differences between calculated and measured internal layer architecture (Siegert and others, 2003; Jacobel and Welch, 2005). Parrenin and Hindmarsh (2007) gave analytical solutions for layer stratigraphy, depending on mass balance, flow field, and ice thickness. Of special interest is the reconstruction of trajectories of particle flow to improve firn and ice-core dating and separate spatial from temporal variations. Based on observed thickness anomalies between isochrones, Leonard and others (2004) identified a high-accumulation region upstream of the Vostok ice core and quantified its effect on the paleoclimatic reconstruction. Morse (1997) iteratively solved a non-linear least-squares minimisation problem to invert the surface velocity field at Taylor Dome for ice rheology and flow parameters. **Waddington and others (2007) used a forward model** for calculating surface height, particle paths, and internal layer shapes to infer **an** accumulation pattern that reproduces observed layer architecture. They apply the method to the area around Taylor Dome.

In this study I **formulate a formal inverse approach using** observed and dated layer architecture in firn, i.e. the age–depth distribution, to kinematically determine horizontal and vertical velocities. The direct solution for the flow field from internal layers in the firn column with depth-dependent density poses a problem that has not been investigated previously. Because of the variation of density with depth, the modeling of firn rheology is much more difficult than that of solid ice. Studies concerned with deeper layers (below a few hundred meters depth) therefore usually consider density to be constant over the whole ice column. The kinematic approach has the advantage that no assumptions about firn rheology are needed and a true density distribution can be utilised.

2. INFERRING VELOCITIES FROM TRACER FIELDS

The debate in the oceanographic community on the question "Can a tracer field be inverted for velocity?", as formulated by Wunsch (1985) two decades ago, showed that it is in principle possible. Without going into details here, it can be said that useful information about the underlying flow field can be extracted from a tracer distribution, even for underdetermined problems **(that is, there are less known equations than unknowns; see appendix).** A number of

*Now at Alfred-Wegener-Institut für Polar- und Meeresforschung, Bremerhaven, Germany physical and chemical parameters can be used as tracers in ice masses. Of particular interest is the age of deposition at the ice-sheet surface of a certain material particle, hereafter simply referred to as age. In comparison to physical or chemical tracers, such as isotopic composition or aerosols, age can definitely be considered a conservative tracer in the sense that it is subject to neither diffusion nor reaction. In the context of ice-core deep drilling for paleoclimate research, glaciological applications focused mainly on forward modeling of this tracer under estimated environmental and dynamical conditions (e.g. Nereson and Waddington, 2002; Clarke and others, 2005). Typical application examples are reconnaissance for suitable drilling sites, or ice-core dating by flow modeling.

Before ⌈**solving the inverse problem for the kinematic model with real field data**⌉, it is important to understand strengths and to identify pitfalls ⌈**of the kinematic model**⌉. This can best be achieved by creating synthetic data to test algorithms, because all parameter fields are known beforehand, ⌈**and as a result, the solution of the inverse problem can be checked**⌉. I use a simple prognostic forward model to create synthetic stationary age distributions under prescribed conditions for a range of flow scenarios of varying complexity for the upper 100 m of the ice sheet, i.e. the firn column. Subsequently I apply a diagnostic ⌈**inverse approach**⌉to the synthetic age distribution to solve for the velocity field. The inversion is based on ⌈⌉singular-value decomposition (SVD). SVD has several advantages over other schemes, such as e.g. least squares ⌈**normal equations**⌉, especially in terms of analysing the inversion results (as summarized, for instance, by Wunsch, 1996). Various combinations of boundary conditions and constraints are used to set up systems of equations to be solved, covering the full range from under- to overdetermined systems. Comparison of reference velocities calculated by the prognostic model with the inferred velocities from the inverse problem then provides a means to evaluate the performance and reliability of the SVD for different constraints. I introduce the flow scenarios, the ⌈⌉inversion formalism, and ⌈**constraints**⌉in the next sections. The main ⌈**body**⌉of the paper (section 5) exploits SVD properties for interpreting the results. ⌈**Finally, I apply the kinematic model to two glaciological problems (section 6): the first problem deals with application of the inverse approach to determine the migration velocity of an accumulation pattern from the age–depth distribution and an accumulation proxy at the surface. The second problem aims at reconstructing the past distribution of accumulation and determine its stationarity over time.**⌉

2.1. Kinematic Equations

The approach presented here is based on a kinematic consideration of the firn volume; therefore the equations for conservation of energy and momentum are not taken into account. In general, the distribution of any tracer in a medium can be described by an advection-diffusion equation. (Details on the tracer transport and formulation in ice sheets are discussed extensively by Clarke and others (2005).) In our case, the corresponding tracer is depositional age, $A = A(\mathbf{r}, t)$, a non-diffusive property, which obeys

$$\partial_t A + \mathbf{v} \cdot \nabla A = 1. \tag{1}$$

All calculations are carried out in two-dimensional (2D) space, $\mathbf{r} = (x, z)$ (z positive and increasing downward), and the velocity $\mathbf{v} = (u, w) = \mathbf{v}(\mathbf{r}, t)$. ∂_t denotes the partial derivative with respect to the subscript variable, here time t. ⌈**(See appendix for conventions and a list of symbols.)**⌉Equation (1) is sometimes referred to as the age equation (e.g. Hindmarsh and others, 2006). The right-hand side represents a source term, which is responsible for the actual aging of the firn with time.

The second governing equation is the conservation of mass,

$$\partial_t \rho + \nabla \cdot (\rho \mathbf{v}) = 0. \tag{2}$$

where $\rho = \rho(\mathbf{r}, t)$ is the density. These two equations form the fundamental system of linear equations used in the forward ⌈**problem**⌉.

2.2. Assumptions and boundary conditions

A number of assumptions are employed for the sake of simplicity;⌈**howev they do not depreciate**⌉the general applicability of the ⌈**inverse-problem formulation**⌉. The considered firn volume extends from the surface ($z = 0$) to an arbitrary depth ($z = z_{max}$). The density distribution is taken to be laterally homogeneous and time-independent, ⌈**i.e.**⌉ $\partial_x \rho = \partial_y \rho = \partial_t \rho = 0$ (Sorge's law), but depth-dependence is maintained ($\partial_z \rho \neq 0$). This assumption is well justified on a regional scale for ice-sheet plateaus (e.g. Frezzotti and others, 2004; Richardson-Näslund, 2004; Rotschky and others, 2004; Arcone and others, 2005), but has to be considered with care on cold alpine glaciers. Note that the depth-dependency of density is a prominent ⌈**deviation from**⌉the incompressibility assumption often used in ice-sheet modeling. Time-dependence of equations (1) and (2) is maintained in the prognostic forward model. The system of equations to be ⌈**solved**⌉, however, is formulated in a time-independent way so that $\partial_t(\cdot) = 0$ (where $(\cdot)$ denotes any term to be differentiated), because the forward model produces a steady-state age distribution as output.

No forces appear in the above equations, simplifying matters such that the upper boundary can be taken as a horizontal surface, i.e. parallel to x. Position and direction of scalar and vector quantities then always refer to this surface. (Consider a radargram as an illustrative example. It contains records of the reflector depth with respect to the relative surface. A topographic correction is applied only during data processing.) The kinematic boundary condition at the surface is $w(x, z = 0) = \dot{b}(x)/\rho_0$, where $\dot{b}(x)$ is the surface accumulation and $\rho_0 = \rho(z = 0)$ is the density at the surface. Additional constraints are introduced later, primarily as prescribed velocity properties.

2.3. Prognostic forward model

The forward model runs under prescribed stationary allocations of density, horizontal velocity, and accumulation on an ordinary grid, discretised with finite differences. It calculates the vertical velocity from the combined effect of accumulation at the surface, advection, and densification, and yields the synthetic age–depth distribution. Starting from an initial laterally homogeneous, vertically increasing age distribution, the prognostic model runs in a transient mode until a steady state is reached, i.e. when the particles from the surface at $t = 0$ reach the edge of the domain. ⌈**As a boundary condition at the surface the age is set to zero. At the**

inflow of the model domain the horizontal age gradient is set to zero. ⌉Details on grid parameters are listed in Table 1. The age–depth distribution constitutes the essential output, which is passed to the ⌈**inverse problem.**⌉The prescribed horizontal velocities $\mathbf{u}^{ref}$ and calculated vertical velocities $\mathbf{w}^{ref}$ of the forward model are defined for all grid points. We later refer to them as the reference-velocity field, denoted by the superscript *ref*, against which the inferred velocity field, denoted by the superscript *est*, is compared.

2.4. Linear system for ⌈inverse model⌉

The time-independent forms of equations (1) and (2) read

$$u\partial_x A + w\partial_z A = 1 \tag{3a}$$

$$\rho\partial_x u + \rho\partial_z w + w\partial_z \rho = 0. \tag{3b}$$

The discretisation schemes for solving this linear system on a triplex-staggered grid (a grid consisting of three subgrids shifted⌈**relative to**⌉each other) are taken in an adapted form from Fiadeiro and Veronis (1982) and Wunsch (1985). The input fields of age and density are prescribed on a rectangular grid, the A-grid, with a grid spacing of Δx and Δz in x- and z-direction, respectively. The A-grid has $I \times K$ nodes. Corresponding indices for the gridded variables are $i = 1, \ldots, I$ for the horizontal coordinate (increasing downstream, ⌈**left to right**⌉) and $k = 1, \ldots, K$ for the vertical (increasing downward, ⌈**top to bottom**⌉) coordinate, as indicated in Figure 1(a). The grid nodes representing u and w (u- and w-grid) are shifted by half the grid spacing in the horizontal and vertical direction, respectively, relative to the nodes on which the input parameters for age A and density ρ are prescribed (Figure 1). Application of staggered-grid differences to Equation (3) leads to a discrete system, which for a unit cell (Figure 1(a)) can be expressed as

$$\begin{pmatrix} c^{\alpha}_{i-1,k} & c^{\beta}_{i,k} & c^{\gamma}_{i,k-1} & c^{\delta}_{i,k} \\ c^{\kappa}_{i-1,k} & c^{\lambda}_{i,k} & c^{\mu}_{i,k-1} & c^{\nu}_{i,k} \end{pmatrix} \begin{pmatrix} u_{i-1,k} \\ u_{i,k} \\ w_{i,k-1} \\ w_{i,k} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{4}$$

Detailed expressions of the staggered-grid differences and coefficients $\{c^{\alpha,\ldots,\nu}_{i,k}\} = f(A,\rho)$ are given in the appendix. As ⌈⌉sketched in Figure 1(b) for the node labeled $A_{2,4}$, ⌈**five A-nodes are involved in the discretised representation of**⌉the age equation for a single node. Consequently, the $u_{i,k}, w_{i,k}$ for a unit cell always ⌈**depend on the values of A and ρ at the neighbouring nodes. These values**⌉are contained in the $c_{i,k}$-coefficients in (4). The $u_{i,k}, w_{i,k}$ can thus not be fully determined on the boundaries, but only within the dashed region shown in Figure 1. This region is termed the solution domain. This formulation has the advantage that no other specific conditions are necessary at the boundaries of the domain where the⌈**inverse problem**⌉is solved with SVD. As can also be seen in Figure 1(b), in each dimension, x and z, the total number n of nodes for unknown variables u and w differs. Within the solution domain, the number n^x_u of variables u in a row (x-direction) is $n^x_u = I - 1$. Analogously, along a column (z-direction) $n^z_u = K - 2$. For the variable w, $n^x_w = I - 2$ and $n^z_w = K - 1$. The total number of elements of each variable within the solution domain is $n_u = n^x_u n^z_u$ and $n_w = n^x_w n^z_w$. Defining the following vectors and matrix,

$$\begin{aligned} \mathbf{d} &= \{d_p\} = (1, 1, \ldots, 0, 0)^T \in \mathcal{R}^M, M = 2n^z_u n^x_w, \\ \mathbf{v} &= \{v_q\} = (\{u_{i,k}\}, \{w_{i,k}\})^T \\ &= (\mathbf{u}^T, \mathbf{w}^T)^T \in \mathcal{R}^N, N = n_u + n_w, \\ \mathbf{M} &= \{M_{p,q}\} = (\{c^{\alpha}_{i,k}\}, \ldots, \{c^{\nu}_{i,k}\}) \in \mathcal{R}^{M \times N}, \end{aligned} \tag{5}$$

allows one to set up the matrix equation

$$\mathbf{Mv} = \mathbf{d}. \tag{6}$$

⌈**The variables p, q are merely indices of vector and matrix elements, to be distinguished from the coordinate indices i, k of the actual grid. The vector**⌉$\mathbf{d}$ represents the data in data space $\mathcal{R}^M$, and ⌈**the vector**⌉$\mathbf{v}$ represents the model parameters in model space $\mathcal{R}^N$. M is the number of (known) equations, N is the number of unknowns, in our case the velocities within the solution domain. The relationship between model parameters and data is described by the model matrix $\mathbf{M}$, sometimes referred to as the data kernel (Menke, 1989, p.9). ⌈**The reader might wonder how it is actually possible to define uncertainties of the data vector d, which contains only ones and zeros. For this particular inverse problem, the actually measurable quantities, age and density, appear on the left-hand side in the matrix elements of the data kernel. The uncertainty of the data vector is thus a measure how the uncertainties of the data kernel cause the vector on the right-hand side of (6) to differ from values being exactly ones and zeros, even for exact velocities v. This point will be explored later in more detail by using a Monte Carlo-based approach.**⌉

3. SINGULAR VALUE DECOMPOSITION

3.1. ⌈Principles⌉

The SVD of a matrix $\mathbf{M}$ is a generalisation of the spectral decomposition of a square to a rectangular matrix. The spectral decomposition of a rectangular matrix always exists. Here we apply SVD to calculate the pseudo-inverse (or generalised inverse) of $\mathbf{M}$, mainly following the notation of Wunsch (1996). Any rectangular matrix $\mathbf{M}$ can be decomposed into a factorisation of the form

$$\mathbf{M} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V}^T, \tag{7}$$

where $\mathbf{U}$ and $\mathbf{V}$ are both unitary rectangular matrices, $\mathbf{U} \in \mathcal{R}^{M \times M}$, $\mathbf{V} \in \mathcal{R}^{N \times N}$, and $\mathbf{V}^T$ denotes the transpose of $\mathbf{V}$. The generally non-square matrix $\boldsymbol{\Lambda} \in \mathcal{R}^{M \times N}$ contains the singular values (square root of eigenvalues) of $\mathbf{M}$ in decreasing order on the main diagonal, $\Lambda_{p,q} = \delta_{pq}\lambda_p$, with the Kronecker symbol δ_{pq}. The matrix $\mathbf{V}$ contains a set of orthogonal base-vectors of $\mathbf{M}$, spanning the N-dimensional model (or solution) space, whereas the matrix $\mathbf{U}$ contains a set of orthogonal base-vectors spanning the M-dimensional data (or observation) space. The number R of non-zero singular values is the rank of $\mathbf{M}$. If some singular values are zero or $M \neq N$, one or more of the rows or columns of $\boldsymbol{\Lambda}$ must all be zeros. One can then drop those columns of $\mathbf{U}$ and $\mathbf{V}$ that are multiplied by zeros only, thus reducing the matrices in (7) to the expression

$$\mathbf{M} = \mathbf{U}_R\boldsymbol{\Lambda}_R\mathbf{V}_R^T\ , \tag{8}$$

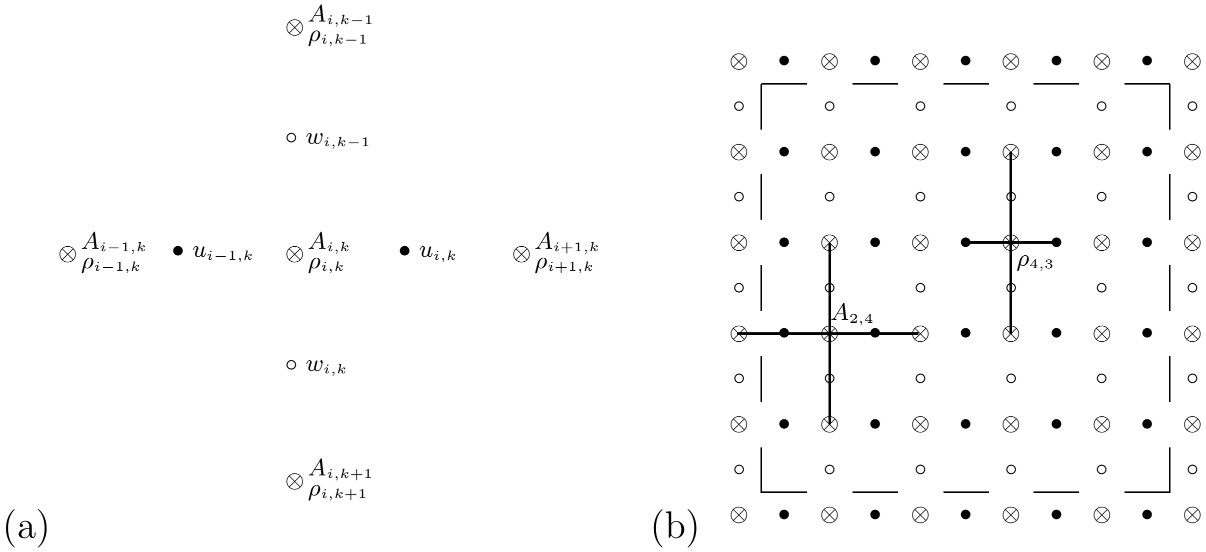

Fig. 1. (a) Unit-cell scheme of the numerical grid used for solving the linear system of equation (3). (b) Scheme of the triplex-staggered numerical grid for $I = K = 6$. The uppermost row corresponds to the surface. Distance between nodes of similar type is Δx and Δz, between nodes of different type $\Delta x/2$ and $\Delta z/2$ in the horizontal and vertical direction, respectively. The thick cross centered on the $\otimes$-node labeled $A_{2,4}$ represents the unit cell in (a) and strikes all nodes involved in the age equation for the $A_{2,4}$-node. Likewise, the ⌜**thick cross**⌝ labeled $\rho_{4,3}$ strikes all nodes involved in the conservation of mass equation for the $\rho_{4,3}$-node. ⌜**Both equations**⌝ can therefore only be solved for those A-nodes within the region bounded by the dashed line, referred to as solution domain. The $\otimes$-nodes on the corners are displayed for completeness, but not used in the inverse problem.

where the subscript R indicates the number of columns, with $\mathbf{U}_R \in \mathcal{R}^{M\times R}$ and $\mathbf{V}_R \in \mathcal{R}^{N\times R}$. $\mathbf{\Lambda}_R \in \mathcal{R}^{R\times R}$ is the square submatrix of $\mathbf{\Lambda}$ with non-vanishing singular values. It can be shown (e.g. Wunsch, 1996) that $\mathbf{V}_R\mathbf{\Lambda}_R^{-1}\mathbf{U}_R^T$ is the pseudo-inverse of $\mathbf{M}$, which we use to solve (6) for the unknown model vector, the solution

$$\mathbf{v} = \mathbf{V}_R\mathbf{\Lambda}_R^{-1}\mathbf{U}_R^T\ \mathbf{d}\ , \tag{9}$$

where $\mathbf{\Lambda}_R^{-1}$ is the inverse of $\mathbf{\Lambda}_R$, i.e. with λ_p^{-1} on the main diagonal ($\lambda_p \neq 0$) and zeros elsewhere. The above expressions for $\mathbf{M}$, $\mathbf{U}$, and $\mathbf{V}$ define four spaces,⌜**which are explained further below:**⌝ the model range $\mathbf{V}_R \in \mathcal{R}^{N\times R}$ (column space of $\mathbf{M}$), the model nullspace $\mathbf{V}_0 \in \mathcal{R}^{N\times(N-R)}$, the data range $\mathbf{U}_R \in \mathcal{R}^{M\times R}$ (row space of $\mathbf{M}$), and the data nullspace $\mathbf{U}_0 \in \mathcal{R}^{M\times(M-R)}$. Depending on the size of M, N, and R, not of all of these spaces need to exist (in the sense that they are not empty sets). ⌜**Conditions for existence of these spaces, definition for over- and underdetermined systems of equations, and combinations of these are listed in the appendix.**⌝ If there is a data nullspace $\mathbf{U}_0$ ($R < M$), and if the data have components in it, then it will be impossible to fit the data exactly. This data mismatch between true data and estimated data, referred to as ⌜**the**⌝ residual norm, will then be different from zero. (As a norm we will use the L_2 norm or Euclidean length of a vector, later denoted by the operator $||\cdot||$. See appendix for definition and further information.) On the other hand, if the model has components in the model nullspace $\mathbf{V}_0$ ($R < N$), then it will be impossible to determine the model exactly⌜**(hence the term model nullspace).**⌝ In that case, the model solution can be presented as a sum of the particular solution given by (9), which contains only range vectors⌜**and solves (6), and an arbitrary homogeneous solution $\mathbf{V}_0\boldsymbol{\alpha}$, which solves the homogeneous system of equations $\mathbf{Mv} = \mathbf{0}$.**⌝ The vector $\boldsymbol{\alpha}$ contains $(N-R)$ coefficients for the linear combination of the $(N-R)$ column vectors of $\mathbf{V}_0$ in the model nullspace, about which the equations provide no information.

The SVD is related to the least-squares approach. All of the structure imposed by SVD is also present in least-squares solutions. One commonality is that the SVD simultaneously minimises the residual and solution norms (minimum norm property, e.g. Scales and others (2001, p. 66)). However, the SVD solution generalises the least-square solution to the case where the matrix inverses of $\mathbf{M}^T\mathbf{M}$ or $\mathbf{M}\mathbf{M}^T$, the simplest forms, do not exist, for instance if the system is not full rank (Wunsch, 1996, 157f). An important advantage for the application of SVD and the interpretation of the solution is that only a single algebraic formulation is necessary, ⌜**for**⌝ over-, under-, or just-determined systems. The SVD provides its control over the solution norms, uncertainties, and covariances through choice of the effective rank $\hat{R} \leq R$, which leads to the so-called truncated SVD, demonstrated later. The truncated form makes a clear separation between range and nullspace in both solution and data spaces.

3.2. Resolution

A useful feature of the SVD is that it provides direct access to the resolution ⌜**obtainable when mapping**⌝ between model and data spaces (for discussions see Menke (1989, p.62f) and Wunsch (1996, p.165)). The model resolution matrix, defined as

$$\mathbf{T}_\mathbf{V} = \mathbf{V}_R\mathbf{V}_R^T, \tag{10}$$

determines the relationship between the general solution and the particular solution. If no model nullspace exists ($R = N$), the general and particular solution are equal. Then $\mathbf{T_V} = \mathbf{I}_N$, the $N \times N$-dimensional identity matrix, meaning that the model is completely resolved. If a nullspace exists, non-zero terms will appear off the main diagonal in (10), so only averages of some model parameters can be resolved. Analogously, the data resolution matrix

$$\mathbf{T_U} = \mathbf{U}_R \mathbf{U}_R^T \tag{11}$$

provides information on how well the observed data are estimated when the model solution obtained with the generalised inverse is used in the forward model to predict observable quantities. Both resolution matrices are functions of the data kernel $\mathbf{M}$, which contains the a-priori information about the physical representation of the problem, i.e. by the time-independent equations (3). When the problem is linear, ⌈**resolution matrices depend on neither the model parameters v nor the data d.**⌉

3.3. Error covariance and uncertainty

Solving the inverse problem yields an estimate of model parameters, denoted $\mathbf{v}^{est}$, which are subject to uncertainties. Using the estimated $\mathbf{v}^{est}$ in the forward problem (6) yields a prediction of the data vector, $\mathbf{d}^{est}$, which differs from the true data vector $\mathbf{d}$ by some residuals, denoted $\mathbf{n} = \mathbf{d} - \mathbf{d}^{est}$. The residuals can in general arise from two contributions: noise from errors in the measurement of data, and inadequacy of the forward algorithm to describe the problem exactly. The covariance $\mathbf{C_{vv}}$ of the estimated model parameters depends on the residual covariance $\mathbf{R_{nn}}$ (the second-moment or covariance matrix of $\mathbf{n}$, see appendix for details). It can be shown to be (Wunsch, 1996, p.143)

$$\mathbf{C_{vv}} = \mathbf{V}_R \mathbf{\Lambda}_R^{-1} \mathbf{U}_R^T \mathbf{R_{nn}} \mathbf{U}_R \mathbf{\Lambda}_R^{-1} \mathbf{V}_R^T \ . \tag{12}$$

In the case of uncorrelated uniform variance σ_n^2 of the data, (12) simplifies to

$$\mathbf{C_{vv}} = \sigma_n^2 \mathbf{V}_R \mathbf{\Lambda}_R^{-2} \mathbf{V}_R^T \ . \tag{13}$$

The covariance of the model parameters arises from uncertainties present in the data and generates uncertainty in the coefficients of the model range vectors. Data covariance is thus mapped onto model covariance. To obtain the complete solution uncertainty $\mathbf{P_{vv}}$ of the model parameters, the influence of the missing nullspace contribution has to be taken into account as well. It follow as (Wunsch, 1996, p.151)

$$\mathbf{P_{vv}} = \mathbf{C_{vv}} + \mathbf{V}_0 \mathbf{R}_{\alpha\alpha} \mathbf{V}_0^T , \tag{14}$$

where $\mathbf{R}_{\alpha\alpha}$ is the second-moment matrix (or covariance matrix, see appendix) of the coefficients $\boldsymbol{\alpha}$ of the model nullspace $\mathbf{V}_0$, forming the homogeneous solution $\mathbf{V}_0\boldsymbol{\alpha}$. ⌈**The matrix**⌉ $\mathbf{R}_{\alpha\alpha}$ may be entirely unknown, or an estimate from a-priori information might be available. The uncertainty of the residuals, $\mathbf{P_{nn}}$, follows from the variance of the estimated residuals about their mean (Wunsch, 1996, p.117), which can be written as

$$\mathbf{P_{nn}} = \mathbf{U}_0 \mathbf{U}_0^T \mathbf{R_{nn}} (\mathbf{U}_0 \mathbf{U}_0^T)^T . \tag{15}$$

The covariance (12) of the estimated model parameters is very sensitive to small non-zero singular values. Solution variance can be reduced by choosing an effective rank $\hat{R} < R$ to exclude small λ_p. Inspecting the singular-value spectrum of the data kernel enables one to choose an appropriate cut-off size for contributing singular values (Menke, 1989, p.122). This artificial reduction of model- and data-space dimensions leads to rank deficiency, and thus worse resolution, and increased dimensions of the nullspaces, but decreases model covariance. ⌈**The choice of**⌉ the effective rank $\hat{R}$ therefore provides a means to trade off variance and resolution, or solution norm and residual norm, respectively.

3.4. Scaling and weighting

Weighting is in general used to give more importance to certain observations than to others, mainly to correct for uncertainty. An undesired weighting effect occurs if a system consist of different physical equations, involving different physical quantities. In our case, the conservation of mass and the age equation involve the quantities age and density. In the linear system (6), the rows of $\mathbf{M}$ represent these equations. Their different physical origin leads to different norms of the row vectors (i.e. Euclidean length) of the matrix $\mathbf{M}$. To correct for this effect, we first perform row scaling of the matrix $\mathbf{M}$ by multiplying each row with the reciprocal of its row norm (see appendix for details). This is carried out below by operations with the matrix $\mathbf{W}$, which contains the row norms of $\mathbf{M}$ on its diagonal. Likewise, the column vectors of $\mathbf{M}$ have different norms. Therefore we require column scaling after the row scaling is performed. This is done by operations with the matrix $\mathbf{S}$. Performing row scaling first, and column scaling second, transforms our linear system (6) from the original space to the so-called scaled space, denoted by the tilde attribute $\sim$. The transformation has the form

$$\mathbf{W}^{-T/2} \mathbf{M} \mathbf{S}^{T/2} \mathbf{S}^{-T/2} \mathbf{v} = \mathbf{W}^{-T/2} \mathbf{d}, \tag{16}$$

⌈**which we abbreviate as**

$$\widetilde{\mathbf{M}} \tilde{\mathbf{v}} = \tilde{\mathbf{d}}. \tag{17}$$

The⌉ notation for $\mathbf{W}$ stems from its Cholesky decomposition $\mathbf{W} = \mathbf{W}^{T/2}\mathbf{W}^{1/2}$ (Wunsch, 1996, p.159). Similarly, $\mathbf{S}$ has the Cholesky decomposition $\mathbf{S} = \mathbf{S}^{T/2}\mathbf{S}^{1/2}$ and contains the column norms of the already row-scaled matrix $\mathbf{W}^{-T/2}\mathbf{M}$ on its diagonal.

The SVD is applied in the scaled space. Back transformation of the solution $\tilde{\mathbf{v}}$ in the scaled space to the desired solution $\mathbf{v}$ in the original space is carried out by $\mathbf{v} = \mathbf{S}^{T/2}\tilde{\mathbf{v}}$. It can be shown that for a full-rank underdetermined (overdetermined) system, row (column) scaling is irrelevant, ⌈**as the respective scaling matrix is not present in the solution anymore (Wunsch, 1996, p.161 and 164). Despite this fact,**⌉ we always apply both scalings to cover all general cases. In addition to scaling, the use of $\mathbf{W}$ and $\mathbf{S}$ allows a degree of control of the relative norms of solution and residual.⌈⌉

3.5. Separation of mean and variation

⌈**Depending on the problem we are dealing with, information about the variations of the velocity around an average is more interesting than the average velocity, as the velocity variations tell us more about the processes occurring at the ice-sheet surface and their interaction with ice dynamics.**⌉ Unfortunately, the minimum-norm property of the SVD will result in a solution that is smallest, in the sense of being closest to zero. ⌈**This means that we might get a wrong structure of the velocity field. It is therefore**⌉ feasible to consider only the

Table 1. Simulation parameters

Scenario*	$\bar{u}^{ref}$ (m a^{-1})	$\partial_x u^{ref}$ (a^{-1})	$\breve{u}^{ref}$
NF	0	0	0
SF	1	0	0
MF	10	0	0
MDF	10	$4 \cdot 10^{-5}$	$\neq 0$

dimension	min	max	increment
Prognostic Forward Model			
x	0	5 km	100 m
z	0	100 m	1 m
SVD solution			
x	0	5 km	500 m
z	0	50 m	5 m

*NF: no flow; SF: slow flow; MF: moderate flow; MDF: moderate divergent flow with a 20% increase in u over the x-domain. $\bar{u}$ is the mean horizontal velocity averaged over the entire domain.

variations of the flow field on a homogeneous background. Hence we separate the mean flow from its spatial variations by

$$\mathbf{v} = \overline{\mathbf{v}} + \breve{\mathbf{v}}, \tag{18}$$

where $\overline{\mathbf{v}} = (\overline{\mathbf{u}}^T, \overline{\mathbf{w}}^T)^T$ is the mean flow field and $\breve{\mathbf{v}} = (\breve{\mathbf{u}}^T, \breve{\mathbf{w}}^T)^T$ is the spatial variation. Separate mean values $\bar{u} = \langle \mathbf{u} \rangle, \bar{w} = \langle \mathbf{w} \rangle$, each averaged over the entire domain, are used for horizontal and vertical velocities, respectively, and $\overline{\mathbf{u}} = \bar{u}\, \mathbf{i}_{n_u}, \overline{\mathbf{w}} = \bar{w} \mathbf{i}_{n_w}$, where $\mathbf{i}_n$ is a vector of length n with all ones. Our linear system (3) can then be reformulated as

$$\mathbf{M}\breve{\mathbf{v}} = \breve{\mathbf{d}} = \mathbf{d} - \mathbf{M}\overline{\mathbf{v}}. \tag{19}$$

⌈**In case that the mean velocities used for this separation are incorrect, then the SVD-solution of the inverse problem will try to correct this error, e.g. by providing a velocity variation very different from zero on average.**⌉For the rest of the paper we drop the tilde attribute ∼. We assume that separation of mean and variation and subsequent scaling has been applied prior to SVD. The results are then discussed in terms of the variational component of the velocity field $\breve{\mathbf{v}}$, as well as the complete velocity field $\mathbf{v}$.

4. SIMULATIONS AND INVERSE PROBLEMS

4.1. Scenarios

⌈⌉Synthetic scenarios of flow are created with the forward model, with physical parameters chosen to mimic real conditions. The horizontal flow field $\mathbf{u}^{ref}$ is prescribed. A Gaussian variation in surface accumulation $\dot{b}(x)$ is superimposed,

$$\dot{b}(x) = \dot{b}_0 \left(1 + \exp\left[-\frac{(x - x_\mu)^2}{x_\sigma^2}\right]\right), \tag{20}$$

where $\dot{b}_0 = 50$ kg m^{-2} a^{-1}is the background accumulation, a value typical for the Antarctic plateau. The maximum ac-

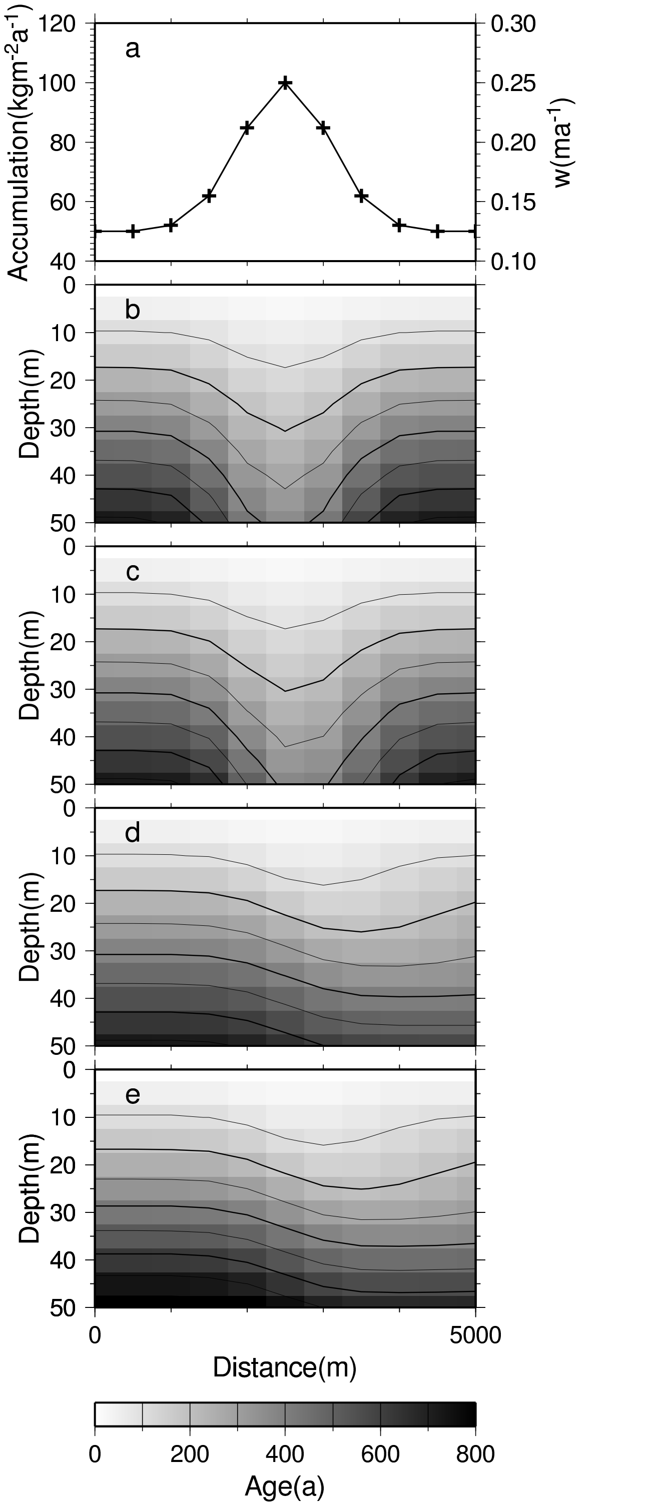

Fig. 2. Accumulation forcing (a) and resulting age–depth distributions using different horizontal velocities of scenario (b) no flow, NF,(c) slow flow, SF, (d) moderate flow, MF, (e) moderate divergent flow, MDF, for the upper 50 m of the firn column (Table 1). ⌈**Colorscale represents age values at grid nodes, with the spatial resolution of the colorscale corresponding to the resolution used for discretising the inverse problem. Contours are lines of equal age.**⌉Horizontal flow is from left to right. Crosses in (a) indicate position of nodes on A-grid, scale on the right is vertical velocity at the surface.

cumulation occurs at $x_\mu = 0.5(x_{min} - x_{max})$, the center of the x-domain, with $\dot{b}(x_\mu) = 2\dot{b}_0$. $x_\sigma = x_\mu/6$ determines the width of the distribution (Figure 2a). Following Richardson and Holmlund (1999), density is parameterised as

$$\rho(z) = \rho_i + (\rho_0 - \rho_i)e^{-c_\rho z}. \qquad (21)$$

⌈**The variables $\rho_0 = 400$ kg m^{-3}and $\rho_i = 900$ kg m^{-3}represent the density at the surface, and the density of solid ice**⌉, respectively, and $c_\rho = 0.05$ m^{-1}. Such a density distribution is commonly observed in Antarctica.

⌈**For the numerical forward model and the inverse problem, the continuous functions defined in (20) and (21) are discretised onto the respective grids. The triplex-staggered grid used in the inverse problem of the linear system (3) has been explained above, with more specifications given below. The forward model is implemented on a grid spanning 5 km in the horizontal and 100 m the vertical direction, containing 51×101 nodes (Table 1). This volume suffices to cover the firn region of cold polar or high-altitude sites and also comprises those length scales which show prominent variations in internal layer architecture over short distances, as imaged by radar at various places in Antarctica (Rotschky and others, 2004; Arcone and others, 2005; Anschütz and others, 2006).**

The effect of four different flow regimes of firn with prescribed horizontal velocity field (Table 1) on the age–depth distribution are displayed in Figure 2. ⌉In the ⌈**simplest**⌉case, no horizontal advection takes place (scenario “no flow”, NF). This could be considered the case on a broad ice dome or along an ice divide. The other cases consider constant slow flow (SF), $\bar{u} = 1$ m a^{-1}, and constant moderate flow (MF), $\bar{u} = 10$ m a^{-1}, which are also typical for polar ice sheets (Xiaolan and Jezek, 2004; Bamber and others, 2000) or high-altitude alpine glaciers (e.g. Lüthi and Funk, 2001; Schwerzmann and others, 2006). For these three scenarios the prescribed velocity variation $\breve{\mathbf{u}}^{ref} = \mathbf{0}$. For the moderate velocity of $\bar{u} = 10$ m a^{-1}, a fourth scenario considers divergent flow (MDF) of the form $u(x) = \bar{u} + c_u(x - x_\mu)$, with c_u such that $u(x)$ increases by 20% from $0.9\bar{u}$ to $1.1\bar{u}$ over the x-domain, and thus $\breve{\mathbf{u}}^{ref} \neq \mathbf{0}$. A scenario with non-constant horizontal velocities is the most likely case to encounter in reality, so it will be ⌈⌉the special focus of the later analysis. Typical velocities for fast ice-stream flow are not taken into account in the main part of this feasibility study,⌈**but a set-up with a higher flow velocity of 50 m a^{-1}will be treated in the application of the inverse approach to glaciological problems in section 6.**⌉The scenarios clearly show how the varying horizontal advection affects the resulting age–depth distribution (Figure 2). For scenario SF, the effect of the accumulation variation tapers off before an affected ice particle leaves the model domain. For both MF-scenarios, advection is larger, so the accumulation effect is still present at the ⌈**outflow of model boundary.**⌉⌈⌉

4.2. Additional constraints

A standard approach to determine the parameters of a physical model, assumed to be a compatible description of a system, is to minimise an objective function that gauges the misfit between measurements and model results. Model physics are usually enforced as constraints on the minimisation in the form of exact equations, so-called hard constraints (e.g. Wunsch, 1996). For ice-flow modeling this was for instance presented by MacAyeal (1993) in the case of estimating the basal friction of an ice stream and applied to real data later (MacAyeal and others, 1995; Vieli and Payne, 2003; Joughin and others, 2004; Larour and others, 2005), and Truffer (2004) estimated the basal velocity of valley glaciers. In addition to the basic physical description of a system, certain aspects of a solution such as structure, norm, or boundary values are also sometimes known a-priori. This information is valuable and helps to restrict the non-uniqueness in solutions of inverse problems. It can be included in the objective function either as a hard constraint by Lagrange multipliers, or as a soft constraint by trade-off between the norm of the solution and the norm of the data mismatch. The trade-off can be implemented in several ways, e.g. by weighting, tapered least squares, or damped least-squares (Menke, 1989, p.52). Although the SVD does not explicitly employ an objective function, constraints can likewise be imposed. ⌈**An example is provided by Waddington and others (2007), who also use SVD to invert a linear system of equations representing a thermomechanical ice-flow model.**⌉

Table 2. Prescribed ⌈**constraints**⌉and system properties

Strategy*	u, w	$\partial_x u$	$\partial_z u$	M	R	$\hat{R}$
Plain	–	–	–	162	162	90
⌈***Bw***⌉	$w_{i,0}$	–	–	171	171	90
Bu	$u_{i,0}$	–	–	172	172	100
Pf	–	–	$0\,\forall(i,k)$	242	180	170
Du	–	$\Delta_x u\,\forall(i,k)$	–	243	180	171
⌈***BwPf*** ⌉	$w_{i,0}$	–	$0\,\forall(i,k)$	251	180	180
BuPf	$u_{i,0}$	–	$0\,\forall(i,k)$	252	180	180
⌈***BwDu***⌉	$w_{i,0}$	$\Delta_x u\,\forall(i,k)$	–	252	180	171
BuDu	$u_{i,0}$	$\Delta_x u\,\forall(i,k)$	–	253	180	172

*Age advection and conservation of mass are considered for all cases. Strategy coding: *Plain*: no additional constraints; *Bw*: boundary conditions of w at surface prescribed; *Bu*: boundary conditions of u at surface prescribed; *Du*: horizontal divergence of u prescribed⌈**at all depth**⌉; *Pf*: plug flow (no shear) prescribed. Constraints are enforced by additional equations to model matrix $\mathbf{M}$. Symbols: $\forall(i,k)$: prescribed for all nodes (i,k); dimension of data space M ⌈**(number of equations)**⌉; dimension of model space $N = 180$ (⌈**number of unknowns, equal**⌉for all ⌈**inverse problems**⌉); R mathematical rank; $\hat{R}$ effective (reduced) rank used for⌈**the inverse problem.**⌉

Each of the different sets of constraints applied in the following exercises with a synthetic scenario can in reality also be determined from measured data. For the problem I address here, the flow and deformation of firn, one usually has⌈**a first guess**⌉of the flow field at the surface. Horizontal surface velocities can be measured directly (e.g. ground-based global-positioning-system surveys of stakes) or indirectly (e.g. observations with satellite-based interferometric synthetic aperture radar). Here, the reference velocity field $\mathbf{v}^{ref}$ represents possible measurements, and thus provides a-priori information about various velocity characteristics. ⌈**(For real field applications these $\mathbf{v}^{ref}$ would be subject to measurements errors. For the synthetic scenario, however, they are the true values.)**⌉It is thus possible to prescribe the horizontal velocity at one or more positions at the surface ($z = 0$). For the rest of the paper I will use the discrete index notation. The surface corresponds to index $k = 1 = k_0$, so that

$$u_{i,k_0} = u^{ref}_{i,k_0}, \tag{22}$$

can be prescribed on one or more horizontal nodes i at the surface. In addition to the velocity, other properties ⌈**such as**⌉the derivative of horizontal velocity, e.g. uniform, divergent, or convergent flow, can be prescribed as well. With $\Delta_x u^{ref}_{i,k}$ denoting the horizontal difference of the horizontal reference velocity at the node (i, k) between neighboring nodes, we can constrain

$$u_{i-1,k_0} - u_{i,k_0} = \Delta_x u^{ref}_{i,k_0}. \tag{23}$$

Distribution of horizontal velocities with depth are deducible from measurements of borehole deformation, enabling us to also use $k \neq k_0$ in (22) for values at depth at the borehole location ($i = i_b$), ⌈**and**⌉also to infer properties ⌈**of surface-parallel**⌉shearing,

$$u_{i_b,k} = u^{ref}_{i_b,k}, \tag{24}$$

$$u_{i_b,k-1} - u_{i_b,k} = \Delta_z u^{ref}_{i_b,k}. \tag{25}$$

where $\Delta_z u^{ref}_{i_b,k}$ is the vertical difference of horizontal reference velocity at (i_b, k). The case $\Delta_z u^{ref}_{i_b,k} = 0$, i.e. constant horizontal velocity along the vertical, is commonly ⌈**called**⌉plug flow. This case will be used later.

Not only can horizontal deformations be deduced from borehole deformation, it is also possible to directly determine the vertical velocities by different methods. One way is to observe the movement of markings in a borehole wall (Hawley and others, 2004; Schwerzmann and others, 2006). This provides similar information for the vertical velocities,

$$w_{i_b,k} = w^{ref}_{i_b,k}, \tag{26}$$

$$w_{i_b,k-1} - w_{i_b,k} = \Delta_z w^{ref}_{i_b,k}. \tag{27}$$

To infer information about the properties of the problem posed here, such as stability of the solution and general solution structure, I will employ different combinations of the ⌈**equations constraining**⌉the linear system (3) to increase the degree of determinacy. The constraints are enforced by expanding the number of rows of the model matrix $\mathbf{M}$ and the data vector $\mathbf{d}$ in equation (6). Each combination of constraints will be referred to as an⌈**inverse problem**⌉, which is then applied to a simulation scenario (Table 2). ⌈**The simplest case (denoted *Plain*) does not employ further constraints and considers just equations for advection and conservation of mass.**⌉Other constraints are set up by prescribing conditions for u ⌈**or w:**⌉the horizontal⌈**or vertical**⌉velocity at the surface as boundary condition (denoted *Bu*⌈**or *Bw*, respectively),**⌉plug flow (*Pf*), and horizontal divergence (*Du*). Moreover, combinations of these constraints are also used in the inverse problems ⌈***BwPf*, *BuPf*, *BwDu*, *BuDu*, and *BwPf*.**⌉

The inverse problem *Plain* shows that the principal property of the kinematic approach is underdeterminacy⌈**, i.e. there are less known equations M than unknowns N (M= 162 < N = 180)**⌉. All other ⌈**inverse problems**⌉with constraining equations are less underdetermined, with the majority being overdetermined systems (Table 2). Only the rather complex MDF-scenario (moderate flow with divergence) will be solved with ⌈**several constraints**⌉and will be used later to discuss the solution properties in detail.

The SVD inversion is implemented with the linear algebra package (LAPACK) routines integrated in MATLAB. As most of the densification of snow takes place in the upper part of the firn column, the inverse problems address only the upper 50 m. The grid used for the inverse problems spans 11 × 11 nodes, with increments of 500 m and 5 m in the horizontal and vertical, respectively. The grid used for the ⌈**inverse problems**⌉has a five-fold lower resolution, but its nodes coincide with a subset of the grid used in the forward ⌈**problem**⌉. As a result⌈⌉, the fields of age and density input to the ⌈**inverse problems**⌉do not have to be interpolated. A linear interpolation of the u- and w-reference-velocity fields⌈**($\mathbf{u}^{ref}$ and $\mathbf{w}^{ref}$)**⌉is carried out to project these values onto the triplex-staggered grid (Figure 1). Evidently, the lower resolution and the interpolation will have some influence on the results. However, this effect could be considered equivalent to⌈**small**⌉measurement errors for real data. The influence of data errors on the results will be considered at the end of the following analysis section.

5. RESULTS AND ANALYSIS

This section compares the solutions of the different⌈**inverse problems**⌉for ⌈**the MDF**⌉scenario. I first illustrate the advantages of SVD-based concepts for comprehensive analyses by investigating the singular-value spectrum (Figure 3) together with some norm properties (Figure 4), and resolution matrices (Figure 5)⌈⌉for the MDF scenario. ⌈**For three inverse problems the velocity fields of the solutions are presented (Figure 6).**⌉Subsequently I discuss the distribution of several norms ⌈**(Figure 7)**⌉, which enable us to evaluate the solutions and compare the results for the ⌈**inverse problems.**⌉The first norm type is the L_2-norm (see appendix) of the residual and solution vectors, $||\check{\mathbf{n}}||$ and $||\check{\mathbf{v}}||$, respectively. (We consider the solution of the velocity variation $\check{\mathbf{v}}$, our main interest, instead of the complete velocity field $\mathbf{v}$.) The residual norm is a measure of the mismatch between the data and the model predictions of the data by the estimated model parameters $\check{\mathbf{v}}$. The solution norm is a measure of the length of the solution vector $\check{\mathbf{v}}$. As discussed above, the SVD simultaneously minimises these norms to produce the particular solution, with the rank $\hat{R}$ determining the trade-off between residual and solution norm. The second type of norm is the norm of the difference between the reference velocity field $\check{\mathbf{v}}^{ref}$ ⌈**(which is the right "answer"**⌉from the prognostic forward model linearly interpolated to the u- and w-grid) and the velocity-field solution $\check{\mathbf{v}}^{est}$, separately for horizontal and vertical velocities ($||\Delta\check{\mathbf{u}}|| = ||\check{\mathbf{u}}^{ref} - \check{\mathbf{u}}^{est}||$ and $||\Delta\check{\mathbf{w}}|| = ||\check{\mathbf{w}}^{ref} - \check{\mathbf{w}}^{est}||$). Hereafter, these are referred to as velocity-difference norms. They provide a measure of how well the inversion ⌈**for a specific inverse problem**⌉performed with respect to the known reference data set.

5.1. Singular-value spectrum

We focus on four ⌈**inverse problems with different constraints** ⌉to determine the solution for the velocity field of the MDF scenario: the underdetermined and ⌈**simplest**⌉case *Plain*, the⌈**almost-determined inverse problem *Bw***⌉(boundary conditions of w at surface prescribed), and the overdetermined inverse problems ⌈***BwPf* and *BwDu* (as *Bw***⌉, but

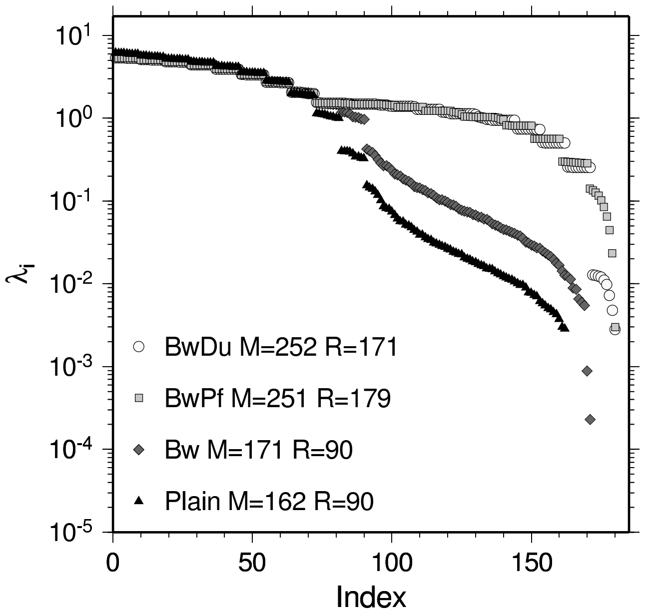

Fig. 3. Singular-value spectrum for⌈**four inverse problems with different constraints applied to the MDF scenario (see Table 1)**⌉. The number of unknown variables $N = 180$ for all inverse problems.

additionally plug flow (*Pf*) or horizontal divergence (*Du*) prescribed as constraints, respectively (Table 2)).

The first third of the ordered singular values (up to index 72 ⌈**in Figure 3**⌉), is basically identical for all⌈**inverse problems.**⌉Beyond this index, up to index 170, the spectra of the overdetermined ⌈**inverse problems fall off slowly in several steps up to index 170**⌉, whereas the underdetermined inverse problems show only one or two further steps before falling steadily.⌉⌈**All spectra show a final discrete drop at singular values of ∼0.25–0.5.**⌉⌈**Such an abrupt**⌉and final discrete drop in a singular value spectrum is a typical phenomenon for various problems (Menke, 1989). Beyond the final discrete drop, all spectra fall continuously ⌈**on the log-scale.** ⌉The spectra for underdetermined ⌈**inverse problems decrease faster with increasing index than the overdetermined inverse problems**⌉. Whereas the rate of decrease of the spectrum for *Plain* does not change significantly, the other⌈**inverse problems**⌉show an increasing rate of decrease for the smallest singular values on the log-scale. ⌈**In general, the spectra differ from one another the most for approximately the smallest 20–30% of the singular values. This has important implications for the residual norms and solution norms.**⌉Using the untruncated spectra for estimating the model parameters usually results in very small residual norms, equivalent to high ⌈**parameter**⌉resolution, but larger solution norms. The corresponding velocity fields show very detailed velocity structures, which, however, need not to be the correct.

⌈**To demonstrate the influence of the the choice of the reduced rank $\hat{R}$, Figure 4 displays the resulting difference norms for the inverse problem *Bw* of the whole range of possible values for $\hat{R}$. The difference norm of vertical velocities, $||\Delta\breve{\mathbf{w}}||$, weighted with the square root of its mean, is constant at about 1 for $\hat{R} \leq 80$, then falls off rapidly to steady values around 0.26, before it rapidly increases for $\hat{R} > 169$. This distribution indicates that for $90 \leq \hat{R} \leq 169$, the vertical reference-velocity structure is approached best, although not exactly matched. This can be confirmed by checking the complete velocity structure for other $\hat{R}$ in figures comparable to Figure 6, but these are omitted here for brevity.**

The distribution of $||\Delta\breve{\mathbf{u}}||$ for *Bw*, likewise weighted with the square root of its mean, is constant around 1.7–1.8 for $\hat{R} < 120$. Two plateaux are present for $130 \leq \hat{R} \leq 160$. In this region the mismatch of horizontal reference and solution velocities are at their minimum. For $\hat{R} > 160$, $||\Delta\breve{\mathbf{u}}||$ increases with rank $\hat{R}$.

Adding both velocity-difference norms, each weighted with the square root of its mean velocity, a broad minimum with two plateaux for $||\Delta\breve{\mathbf{v}}||$ is apparent again for $130 \leq \hat{R} \leq 160$. This range corresponds to the tail of the singular-value spectrum (Figure 3), where the singular values fall continuously. Similar analysis for the variation of difference norms with reduced rank for the other inverse problems yield equivalent findings: the velocity-difference norms always show a minimum for a range of singular values before these show the tendency to decrease more rapidly with larger index. Within this minimum region the choice of $\hat{R}$ leads to only little differences of the final velocity solution. The inverse problems which constrain the horizontal velocity at the surface, i.e. *Bu*, *BuPf*, and *BuDu*, basically display the same features.

⌈**One choice for $\hat{R}$**⌉is the index of the last step-like drop-off as the lower bound of the singular value spectrum to be used for estimating the solution of our inverse problem in equation (9). The continuously and rapidly falling part of the singular spectra is thus truncated, a common practice when employing SVD for solving inverse problems (e.g. Wunsch,

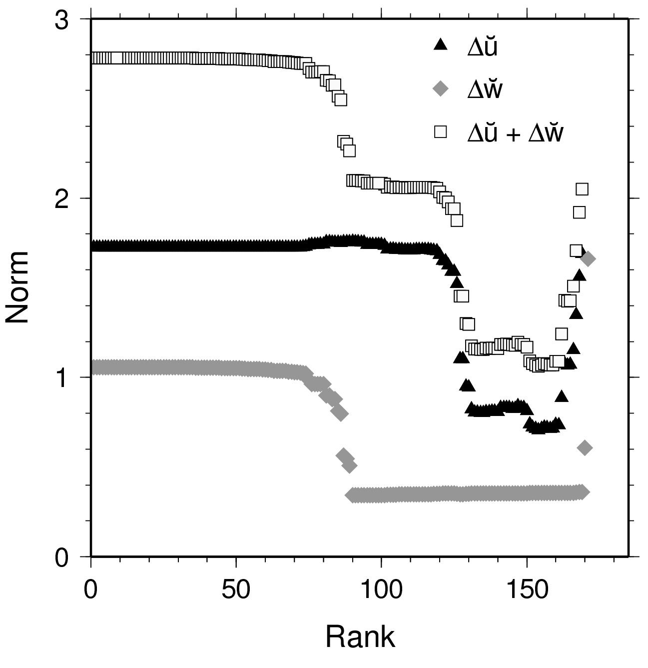

Fig. 4. Distribution of velocity-difference norms $||\Delta\breve{\mathbf{u}}||$, $||\Delta\breve{\mathbf{w}}||$, and $||\Delta\breve{\mathbf{u}}||+||\Delta\breve{\mathbf{w}}||$ as a function of reduced rank $\hat{R}$ for inverse problem *Bw* for the MDF scenario. The norms are scaled with the ⌈**square root**⌉of their mean.

1996). This leads to poorer resolution, but smaller solution norms and velocity-difference norms, and yields sufficiently realistic results for most ⌈**inverse problems**⌉(Figure 6). Although the resulting smallest singular value of the truncated spectra ⌈**are about the same order of magnitude for all inverse problems**⌉, the corresponding reduced rank $\hat{R}$ differs significantly (Table 2). This results from the fact that, depending on the number and type of ⌈**constraints**⌉, the equations show a varying degree of linear independence. The smaller the singular values, the more linearly dependent are the equations. ⌈**For *BwPf*, however, this choice of $\hat{R} = 171$ at the final discrete drop produces a field of almost constant horizontal velocities, implying that important information is still present in the tail of small singular values for larger $\hat{R}$. For *BwPf* it is actually possible to maintain the full rank and obtain realistic solutions. To accommodate this observation, another, however more subjective choice for $\hat{R}$, would be to choose a singular value of 0.2 as cut-off value for all** ⌈**inverse problems.**⌉Which choice too make is in general a difficult one, especially if no further information is available from a-priori information. For all inverse problems I decide to choose the final discrete step-like drop-off, apart from *BuPf* and *BwPf*, for which full rank is maintained. This is justified as *BuPf* does not show falling tail of singular values at all (not shown), and the drop for *BwPf* occurs only for a very large index and less severe than for the comparable spectrum of *BwDu* (Figure 3).

5.2. Model and data resolution

⌈**The resolution matrices $\mathbf{T_U}$ and $\mathbf{T_V}$ provide another means to judge the solution of an inverse problem.**⌉If non-diagonal elements are non-zero, the related main-diagonal element must be less than unity, indicating that this parameter is not fully resolved, i.e. only averages⌈**of nearby parameters**⌉can be determined. I now discuss⌈**the solution of three inverse problems with different constraints**⌉for the MDF scenario. At full rank, the data are fully resolved for all underdetermined ⌈**inverse problems**⌉, and the model parameters are fully resolved for all overdetermined ⌈**inverse problems**⌉. ⌈**The latter is the case for *BwPf*, for which the full rank $R = 180$ is maintained (Figure 5).**⌉For the truncated underdetermined solutions ⌈**(*Plain* and *Bu*)**⌉⌈⌉the model resolution matrix $\mathbf{T_V}$ indicates that the horizontal velocities are only poorly resolved (Figure 5b). ⌈⌉The vertical velocities are equally well resolved for both ⌈**inverse problems**⌉. It will become evident that this is in accordance with comparison of the actual velocity fields shown in Figure 6 discussed below. Without checking the reference-velocity field it is thus possible to assume that ⌈**in the underdetermined cases the vertical-velocity solutions are more reliable than the horizontal-velocity solutions.**⌉

For all ⌈**overdetermined or truncated underdetermined cases,**⌉the data cannot be fitted exactly, giving rise to larger residuals. ⌈**The order of the diagonal elements of the data-resolution matrix $\mathbf{T_U}$ in Figure 5a follows from that of the structure of M of the linear system (6), rearranged into a vector. This vector represents groups of equations or constraints (as indicated on the abscissa of Figure 5a for *BwPf*: element 1–81: age equation, 82–162: conservation of mass, etc.). Within each group of equations, the elements are sequentially ordered by horizontal rows of grid nodes.**⌉The diagonal elements now indicate that especially the data predicted by the age equations are only poorly resolved for all ⌈**inverse problems**⌉. The equations for conservation of mass are better resolved, though not fully. Especially without further constraints (*Plain*), they show a decreasing resolution trend with depth (larger element index). For *BwPf*, the plug-flow constraint is well resolved. This result will be evident in the horizontal-velocity structure discussed later. ⌈**For both, *Bw* and *BwPf* the equations representing the vertical velocities at the surface are very badly resolved.**⌉The "oscillations" in data resolution are not arbitrary. The variations visible in Figure 5a seem to systematically depend on the position of the underlying node. The variations are smaller in the horizontal than in the vertical direction. ⌈⌉Overall, the model-parameter and data-resolution matrices allow us to judge and improve the quality of the solution by inspecting the residual and solution norms and the singular-value spectrum without requiring a reference-velocity field.

5.3. Solution- versus reference-velocity fields

The principal results obtained in the last section are clearly seen in the velocity distribution (Figure 6). ⌈**The reference velocity fields $\mathbf{u}^{ref}$ and $\mathbf{w}^{ref}$, which are the correct solutions being sought, are shown in Figure 6a and a'.**⌉The underdetermined solution *Plain* ⌈**without constraints**⌉does not reproduce the horizontal velocity, but gives an ⌈⌉idea what the vertical velocity field might look like. In the almost-determined case *Bw* the vertical structure is reproduced correctly, but the vertical velocities in the solution are smaller than the reference velocities. ⌈**The horizontal velocities again do not show the expected divergence.**⌉The vertical velocities in the overdetermined case *BwPf* are ⌈**very similar to the almost-determined case *Bw*, but differ slightly more from the reference values. Because plug flow was used as a constraint for this** ⌈**inverse problem**⌉, the horizontal velocities $\mathbf{u}^{est}$ are now in very good agreement with the reference field $\mathbf{u}^{ref}$, although ⌈**with overall smaller values.**⌉The better agreement of the horizontal velocities of the solution and reference is consistent with the fact that the horizontal velocities are well-resolved for this ⌈**inverse problem**⌉(see diagonal elements of model resolution matrix in Figure 5b). ⌉

5.4. Norm properties of solutions

⌈**We next discuss the different norm properties of the inverse problem with different constraints**⌉as listed in Table 2.⌉The difference norm for horizontal velocities, $||\Delta\check{\mathbf{u}}||$, is very sensitive to the choice of the mean velocity $\bar{u}$. To provide a similar foundation for all ⌈**inverse problems**⌉, the mean velocity $\bar{u}$ is always provided as the mean of the reference-velocity field for each scenario, such that only the variations in the velocity solutions are compared (Table 1). The influence of zero-mean velocities will be discussed later in this section.

For full-rank SVD, ordering the ⌈**different inverse problems**⌉with increasing M (the number of equations), as done in Figure 7, would generally illustrate the dependence of the residual norm on determinacy. Naturally, for full-rank underdetermined systems ($M < N$) the data can be fit exactly, resulting in $||\check{\mathbf{n}}|| = 0$. For reduced rank, however, the residual norm $||\check{\mathbf{n}}||$ increases, but yields a ⌈**smaller**⌉solution norm $||\check{\mathbf{v}}||$. ⌈⌉For the ⌈**MDF**⌉scenario, the residual norm $||\check{\mathbf{n}}||$ is more than a factor

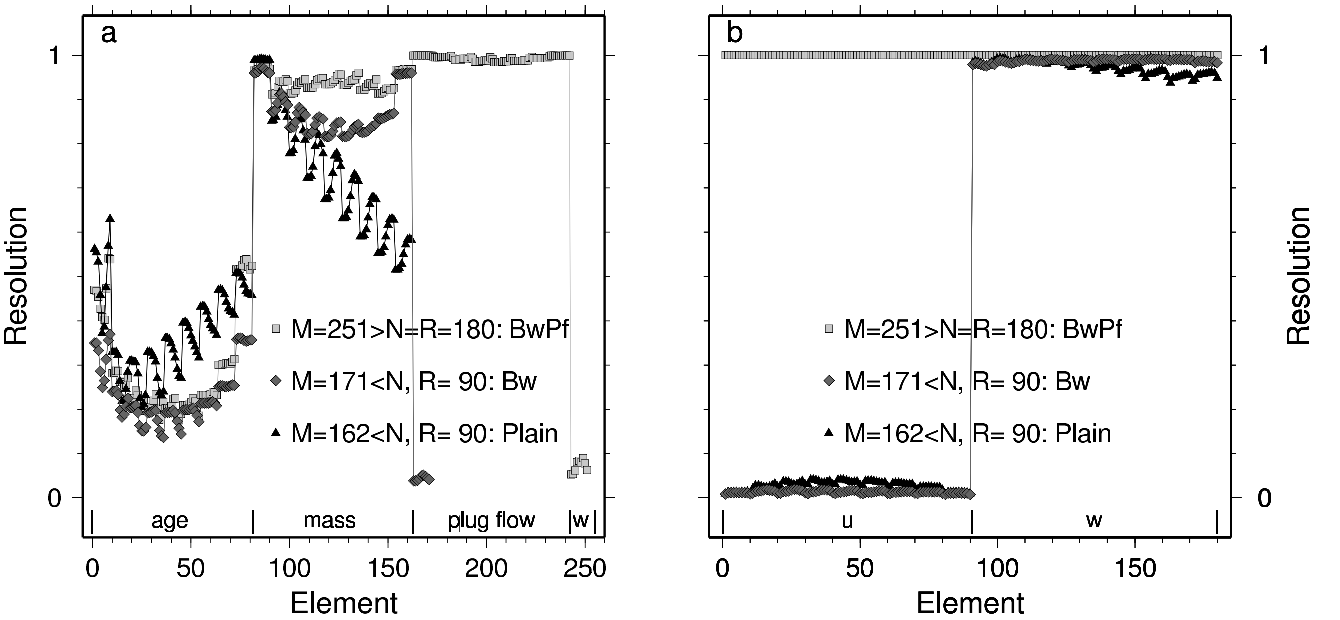

Fig. 5. Diagonal elements of (a)⌈**data resolution matrix $\mathbf{T_U}$, and (b) model resolution matrices $\mathbf{T_V}$,**⌉for the ⌈**inverse problems**⌉*BwPf*, *Bu*, and *Plain*, given in the legend, applied to the MDF scenario with $N = 180$. ⌈**In (a)**⌉, components of $\check{\mathbf{d}}$ (element of data⌈⌉) for *BwPf* correspond to the age equation, conservation of mass equation, plug-flow constraint, and constraint of vertical velocity at the surface, as indicated on the abscissa. ⌈**In (b),**⌉components of $\check{\mathbf{v}}$ (element of model parameter⌈⌉) correspond to u and w, respectively, as also indicated.

of two larger for the underdetermined problems ⌈**(*Plain, Bw, Bu*)**⌉than for the overdetermined problems ⌈**(*BwPf, BwDu, BuPf, BwDu*) (Figure 7a).**⌉In each of these two groups the residual norm is ⌈**quite**⌉constant. The velocity norm $||\check{\mathbf{v}}||$ spans an order of magnitude (Figure 7b), with opposite ratio for under- and overdetermined ⌈**inverse problems**⌉than for the residual norm, as expected.

⌈**More interesting from an application point of view is the residual between reference and solution velocities (Figure 7c and d). The difference norms of horizontal velocities drop from $||\Delta\check{\mathbf{u}}|| \approx 0.5 - 0.6$ m a^{-1}for underdetermined problems to values close to zero for the overdetermined problems. The difference norms of vertical velocities vary around $||\Delta\check{\mathbf{w}}|| \approx 0.10 - 0.11$ m a^{-1} for the over- and underdetermined problems, except for the cases *Plain* and *BuDu*, which are only slightly larger with $||\Delta\check{\mathbf{w}}|| \approx 0.12$ m a^{-1}.** ⌉

⌈**In some experiments, a-priori information on horizontal velocity fields may be unavailable.**⌉In those cases, $\bar{u} = 0$ would have to be used. Employing this case for the MDF scenario, the velocity-difference norm remains quasi constant, but the residual norm significantly increases for those ⌈**inverse problems**⌉that do not incorporate boundary values for $\mathbf{u}$ at the surface. Without a non-zero estimate for mean velocities, the solution produces the smallest velocity norm as a consequence of the minimum-norm property of the SVD. Reducing the rank does not provide a remedy in this case.

5.5. Error and covariance estimates

The last point to investigate, fundamental to all inverse problems, is the solution uncertainty. The quantities density ρ and age A are part of the data kernel $\mathbf{M}$. Density measurements along ice cores are very accurate, usually with an uncertainty <2%. However, our assumption of a ⌈**laterally**⌉homogeneous density distribution might be wrong, even if mean distributions are considered. The uncertainty of the age–depth distribution determined from radar surveys depends on numerous factors: converting radar travel time to depth based on integrated density, estimating age from ice cores, transferring the ice-core age to the internal horizons, tracking of individual horizons, and interpolation of the age distribution onto the SVD grid. From analysis of Antarctic field data, Eisen and others (2004) found a maximum error of approximately 2% for the age–depth distribution in firn. In alpine regions, or regions with a ⌈**laterally**⌉inhomogeneous density distribution, this error might be larger.

An error estimate of the model parameters $\check{\mathbf{v}}$ requires knowledge ⌈**of**⌉the data covariance $\mathbf{C_{vv}}$, according to equations (12) and (14). For the linear system considered here, uncorrelated uniform variance for the data cannot be assumed, as different physical equations are taken into account. Instead of prescribing an arbitrary data covariance, we perform a Monte Carlo-based estimate of covariances, using perturbed reference velocities, age, and density distributions as input to a forward calculation using equation (6). A total of 10^3 experiments, each of which uses a normally distributed random error of 10% for A, 2% for ρ, and 1% for $\mathbf{v}^{ref}$, results in a distribution of estimated data vectors. From this the corresponding distribution of residuals $\mathbf{n}$ follows. Subsequent analysis finally yields an estimate of the residual covariance $\mathbf{R_{nn}}$. As could be expected from the numerical setup, the different equations are not uncorrelated. Although the main diagonal dominates $\mathbf{R_{nn}}$, secondary diagonals also exhibit significant components. The contribution of the covariance of the nullspace vectors through $\mathbf{R_{\alpha\alpha}}$ to the model uncertainty is neglected, as a-priori information ⌈**about its structure is not**⌉available.

⌈**We compare the model uncertainty for the solution obtained with** ⌈**inverse problems**⌉*BuPf* and *BwPf*. Following ⌈**Equations**⌉(18) and (19), both ⌈**inverse prob-**

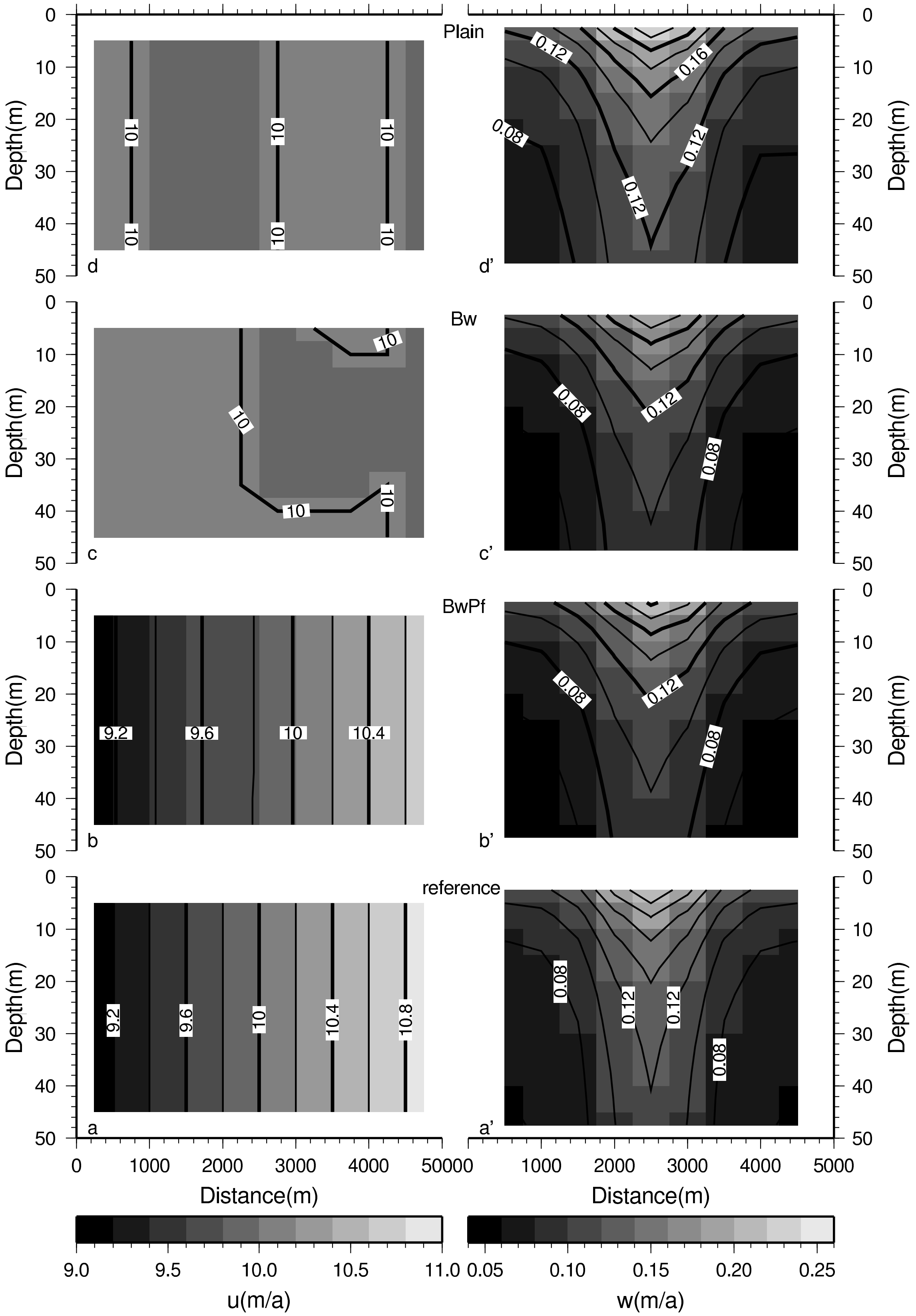

Fig. 6. Solutions for the horizontal (left, a–d) and vertical (right, a'–d') velocity fields for the MDF-scenario ⌈**of inverse problems**⌉ *Plain* (d, d'), *Bw* (c, c'), *BwPf* (b, b') , and the reference fields (a, a'). The different horizontal and vertical spatial domain of **u** and **v** result from the different grids used (Figure 1).

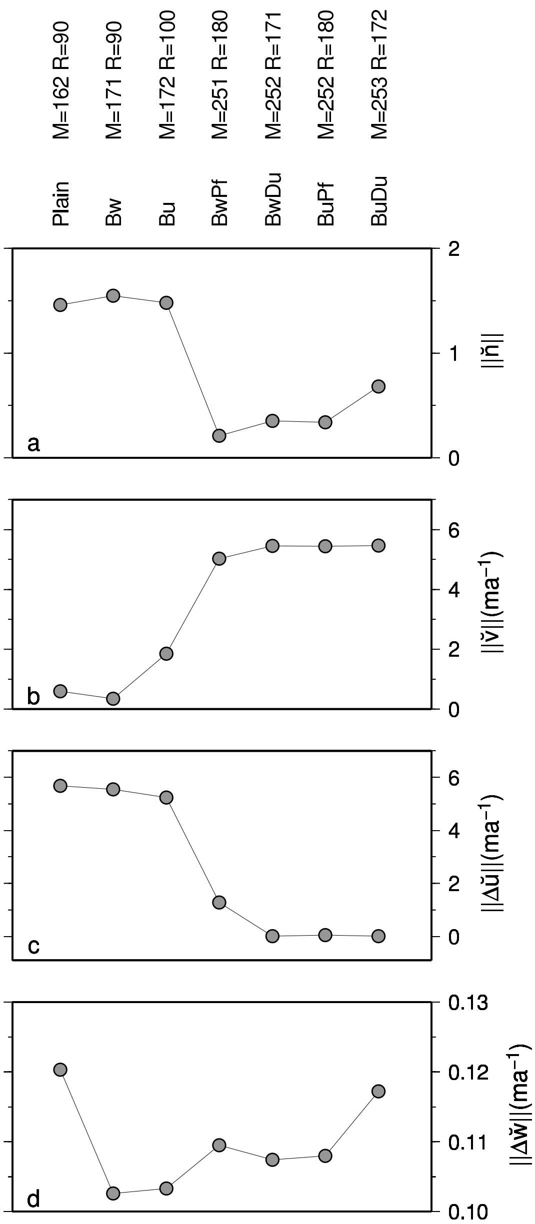

Fig. 7. (a) Residual norm $||\check{\mathbf{n}}||$, (b) velocity norm $||\check{\mathbf{v}}||$, (c) horizontal velocity-difference norm $||\Delta\check{\mathbf{u}}||$, (d) vertical velocity-difference norm $||\Delta\check{\mathbf{w}}||$ of the ⌈**MDF scenario (Table 1). The inverse problems are indicated on the top abscissa, ordered with increasing number of equations M. For all** ⌈**inverse problems**⌉the number of unknowns $N = 180$.⌉

lems with constraints for the MDF scenario⌉solve for the velocity variation $\check{\mathbf{v}}$ on a background velocity of $\bar{u} = 10$ m a^{-1}and $\bar{w} = 0.1$ m a^{-1}. ⌈**The model uncertainties $\mathbf{P}_{vv}$ for $\check{u}$ and $\check{w}$ of the solution of *BuPf* at full rank increase with element, i.e. depth (Figure 8a and b). (Note that according to the definition in equation (5), the elements of the vector $\check{\mathbf{v}}$ are sequentially ordered by horizontal rows of grid nodes.) For the horizontal velocity variation $\check{u}$, maximum uncertainties occur at larger depth and are about the equal to the maximum velocity variation (Figure 8a). For the vertical velocity variation $\check{w}$, uncertainties for near-surface nodes are an order of magnitude smaller than the velocity variation, and at larger depth they are about equal to the maximum variation (Figure 8b). The uncertainty estimates for $\check{u}$ are always at least one order of magnitude larger than the actual residual between the solution of *BuPf* and the reference velocity. For $\check{w}$ residuals and uncertainties are of comparable magnitude.** ⌉⌈**For the inverse problem *BwPF*, also solved at full rank $\hat{R} = 180$, the uncertainty of the horizontal velocities is about two orders of magnitude larger than the maximum velocity variation (Figure 8c). The uncertainty for the vertical velocity variation is comparable to those of *BuPf* (Figure 8d).** ⌉⌈**Although *BuPf* and *BwPf* produce very similar solutions for the velocity field, the uncertainties of their horizontal velocities of the solution are very different. This can be attributed to the different constraints for the horizontal velocity. For *BuPf* the horizontal surface velocities are prescribed as a constraint, and thus by constraining plug flow, also the horizontal velocities at larger depth are constraint. For *BwPf*, merely plug flow is constraint. The actual value of the horizontal velocities is thus more influenced by the age–depth field for *BwPf* than for *BuPf*, and thus subject to larger uncertainties.** ⌉

The uncertainty of the residuals $\mathbf{n}$,⌈**and thus of the model covariance C, depends significantly**⌉on the rank chosen. Generally, for $\hat{R}$ close to the full rank R, the uncertainties of the solution are larger than for smaller $\hat{R}$. For instance, for $\hat{R} = 178$, the uncertainties for the horizontal velocities of *BwPf* are comparable for those of *BuPf* for full rank with $\hat{R} = 180$. Reducing the rank used for the solution leads to smaller uncertainties, but decreases the resolution of the model parameters. Again, this is the manifestation of the trade-off between resolution and model covariance. Moreover, for $\hat{R} < R$ the covariance of the null-space vectors $\mathbf{R}_{\alpha\alpha}$ contributes to the model uncertainty ⌈**of equation (14),**⌉but cannot be estimated without a-priori information.

6. APPLICATION TO TWO GLACIOLOGICAL PROBLEMS

⌈**In this final section I apply the inverse approach to answer two fundamental questions, which emerge from the analysis of radar data: 1. What is the migration velocity of an accumulation pattern relative to that of the underlying ice? 2. Was an accumulation pattern constant over time?** ⌉

6.1. Variation of the migration velocity

⌈**Under certain conditions an accumulation pattern is migrating at a different velocity than the underlying ice. For instance, this is the case for megadunes on the Antarctic plateau (Frezzotti and others, 2002; Fahnestock and others, 2000) and smaller dune-like features in coastal areas (Anschütz and others, 2006). Although estimates of the horizontal velocity of the ice might be available, we cannot use it to deduce the migration velocity of the accumulation pattern. The internal layer structure, however, provides the key to the answer, as it is influenced by the relative velocity between the accumulation pattern and the ice, and not by the absolute velocity of the ice itself. Using the surface ice-velocity constraints under**

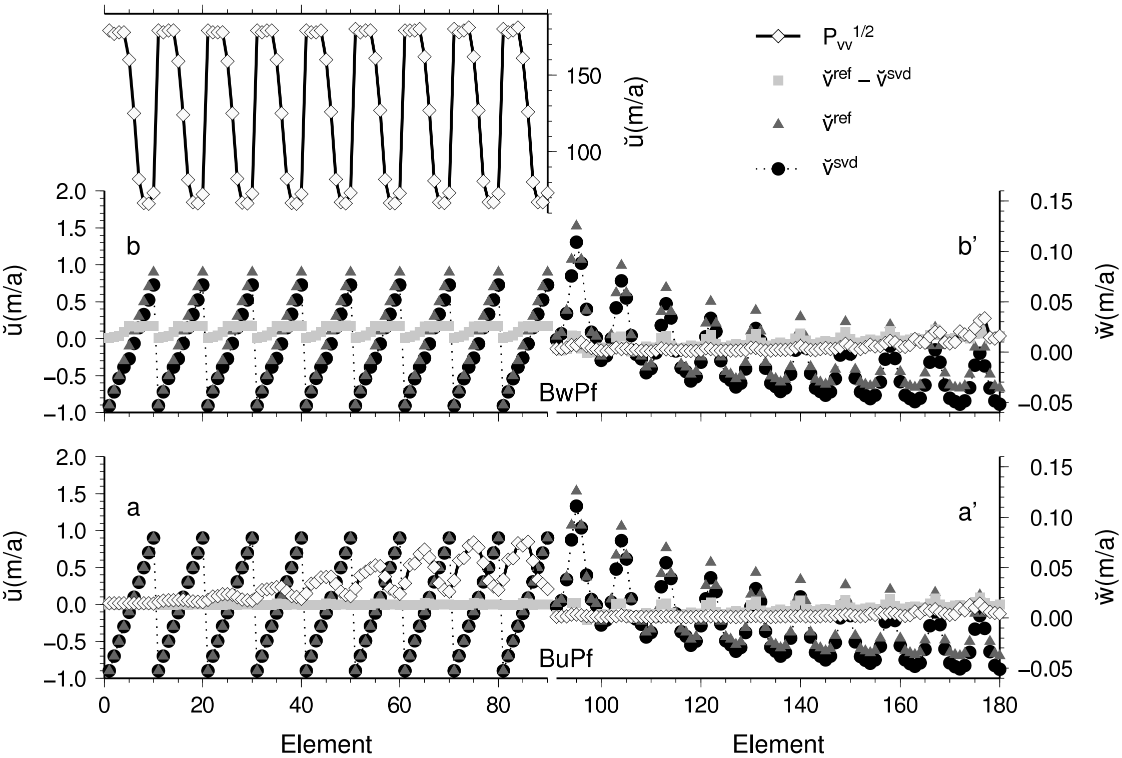

Fig. 8. Elements of the solution vector $\check{\mathbf{v}}^{est}$ and the reference $\check{\mathbf{v}}^{ref}$ for velocity variation, the solution and reference residual vector $\Delta\check{\mathbf{v}} = \check{\mathbf{v}}^{ref} - \check{\mathbf{v}}^{est}$, and solution uncertainty, the diagonal of $\mathbf{P}_{\mathbf{vv}}^{1/2}$, for the **inverse problems** (a, a') *BuPf*, (b, b') *BwPf* at full rank $\hat{R} = N = 180$ for the MDF scenario (Table 1). The display is split into (a, b) horizontal velocities $\check{\mathbf{u}}$, and (a', b') vertical velocities $\check{\mathbf{w}}$. In b) the y-axis on the right corresponds to the elements of $\mathbf{P}_{\mathbf{vv}}$, as they are two orders of magnitude larger than the velocity variation $\check{\mathbf{u}}$. **The components of each vector correspond to sequentially ordered horizontal rows of grid nodes. For instance, the uppermost horizontal velocities of the solution domain correspond to elements 1–11, the nodes in the row below to elements 12–22, etc.**

such conditions would not result in a realistic pattern of vertical velocity and accumulation. It would be more reasonable to prescribe additional flow conditions and determine the migration velocity of the accumulation pattern relative to the ice surface by solving the resulting inverse problem.

Assume that from a field survey, GPR data and dated firn or shallow ice cores are available. The age–depth distribution results from merging the GPR profile with the age and density profiles of the core. Let us assume for this example, for the sake of brevity, that the true distribution of the migration velocity and other physical properties correspond to the MDF scenario as treated before. We now take the shallowest internal layer as a proxy for surface accumulation and use it as the first constraint (*Bw*). Although this internal layer is subject to advection relative to the accumulation pattern, as are the deeper layers, the advected distance will in general be small enough to provide a first guess of the surface accumulation. As we only have shallow GPR data covering the firn column, we can assume plug flow in the firn and use this as the second constraint (*Pf*). We thus get the inverse problem with constraints *BwPf*, different properties of which have been determined and discussed for the MDF scenario above already. The horizontal and vertical velocity fields of the solution to the inverse problem are the ones shown in Figure 6.

For the glaciological problem assumed here, the horizontal velocity field now corresponds to the relative horizontal migration velocity of the accumulation field in respect to the ice. If the ice velocity is available in addition, the absolute migration velocity of the accumulation pattern can be calculated. This example shows how useful a kinematic inverse approach can be to provide an estimate of the horizontal advection velocity field, even if no further velocity information is available.

6.2. Estimation and stationarity of an accumulation pattern over time

The reader may wonder why it is actually necessary to use a mathematically complex inversion scheme under the simplifying assumption of plug flow in firn. If the flow is indeed plug flow, then all information on the horizontal field could be deduced from measurements at the surface. However, determination of accumulation from the age distribution produces significantly different results for conventional techniques and for inverse-problem solutions. With the conventional technique, accumulation is estimated as the quotient of cumulative mass difference and age difference between two isochronous layers. The effect of advection on layer architecture for an inhomogeneous accumulation pattern **can lead** to non-intuitive results, as demonstrated for a number of cases by Arcone and others (2005). A spatially varying accumulation pattern and strong advection cause ⌈**convolution of surface signals**⌉to appear in the internal layer structure. From the conventional approach, even if a correction for advection is included, it is impossible to tell whether the accumulation pattern and value was constant in the past. Misinterpretations of internal layer data are therefore possible. ⌈**To demonstrate the capability of the kinematic inverse approach to provide the answer for this case, I now discuss the problem of a spatially oscillating accumulation pattern and a constant advection velocity presented by Arcone and others (2005) in their figure 10c. This problem is comparable to the MF scenario (Table1) with a higher horizontal velocity. For the analysis, the age–depth field (Figure 9a) is produced by the forward model for a model domain of 30 km in x-direction and 100 m in z-direction. At the inflow of the model domain, cyclic boundary conditions are used, mimicking an infinite extension of the accumulation pattern at the surface. Accumulation and flow parameters are comparable to the scenario of Arcone and others (2005, figure 10c): a constant horizontal velocity $u = 50$ m a^{-1}, a stationary cosine-like accumulation pattern with a wavelength of 10 km, a mean accumulation $\dot{b}_0 = 225$ kg m^{-2} a^{-1}, and a spatial amplitude of $0.55\dot{b}_0$ (Figure 9b). The density–depth function is the same as before. We only consider the first 25 km as the model domain of the inverse problem.**

For the conventional accumulation estimate, advection can be simply taken into account by shifting the accumulation distribution determined from neighbouring internal layers upstream by the distance covered with the mean horizontal flow velocity since the layers were deposited at the surface (Figure 9d). The result shows that, apart from the accumulation pattern derived from the layer closest to the surface, the accumulation values calculated from deeper layers vary considerably from the actual accumulation pattern at the surface. For accumulation minima at the surface, the accumulation derived from the deepest layers determined with the conventional approach is up to 70 kg m^{-2} a^{-1}higher than the actual accumulation at the surface, equivalent to $\sim 70\%$ of the reference value. For accumulation maxima, it varies about ± 30 kg m^{-2} a^{-1}(8.5% of the reference value). In addition, the conventional accumulation pattern cannot be reconstructed over the complete x-domain, because the layer architecture essentially necessary for a complete reconstruction has been partly advected outside the domain of the known age–depth distribution (the deepest continuous layer has an age of about 340 a, corresponding to an advection of 17 km). For a detailed analytical discussion on the related topic of causal relations between changes in accumulation, layer architecture, and particle trajectories, see Parrenin and Hindmarsh (2007).

For solving the kinematic inverse problem, we assume that the horizontal surface velocities are known from measurements and that plug flow prevails. We can then use the constraints *BuPf*. To provide enough numerical nodes per wavelength of the accumulation pattern, it is necessary to increase the resolution of the grid 25×25 nodes. This yields a total of $M = 1609$ equations and $N = 1104$ unknowns. As for the comparable case mentioned earlier, the full rank $R = 1104$ is applicable to the inverse problem *BuPf*. From the vertical velocities of the solution the accumulation can be determined. It is provides a very congruent distribution for the accumulation derived from vertical velocities at all depth. For accumulation maxima the solution is about 5 kg m^{-2} a^{-1}smaller than the actual accumulation pattern, equivalent to -1.4% of the reference values. For minima, it is about 5 kg m^{-2} a^{-1}larger, equivalent to $+5.2\%$. The congruent shape of the accumulation pattern derived by the inverse approach implies that the assumption of steady state is correct. This, in turn, tells us that the accumulation was constant over time. Again, this result can not be achieved from by the conventional accumulation estimates alone. ⌉

7. SUMMARY

In this paper I investigated the feasibility of inferring the velocity field in an advective flow regime in firn by employing age–depth data and a kinematic inverse-problem approach. The inverse problem was solved by means of a singular-value decomposition of a linear system of equations. The comparison of ⌈**inverse problems with different**⌉constraints shows that all kinematic systems provide a generally stable solution, given that the singular spectrum is adequately truncated, and that the choice of the reduced rank can be based on objective criteria. For the underlying system of equations, the given advection scenario, and the prescribed spatially inhomogeneous accumulation, the inverted ⌈**horizontal-velocity is much more sensitive to the employed constraints than the vertical-velocity solution.**

The amount of information retrieved about the velocity field naturally varies with the degree of determinacy of the underlying linear system. For all ⌈**inverse problems**⌉, the prescription of ⌈**some**⌉surface velocities seems necessary to retrieve small velocity variations superimposed on a mean flow field. Without any quantitative information on horizontal velocity, the minimum-norm property of the SVD makes realistic solutions difficult. ⌈⌉

A detailed investigation of the solution is possible by exploiting the mathematical advantages of the SVD. The solutions were examined in terms of resolution, error estimates, and trade-off of resolution and solution covariance. ⌈⌉The⌈**inverse-problem**⌉approach is likely applicable to other flow scenarios as well. Two applications to realistic scenarios were presented. Interaction of a spatially constant accumulation pattern with a high-velocity flow field was analysed to ⌈**exclude**⌉temporal variations in accumulation by removing the advective components of accumulation estimates. Although the approach presented here assumes a steady-state pattern, larger temporal variations in accumulation derived from layer ages at different depth ⌈**reveal**⌉temporally varying accumulation.

A possible extension of the kinematic inversion approach ⌈**presented would be the use of**⌉more unknown parameters, ⌈**e.g. we could use**⌉a certain density parameterisation and ⌈**solve**⌉for those parameters as well. Another possibility is some form of ⌈**time dependence.**⌉⌈**We could also include**⌉dynamical equations and then ⌈**solve**⌉an inverse problem to find parameters for a flow law of firn.

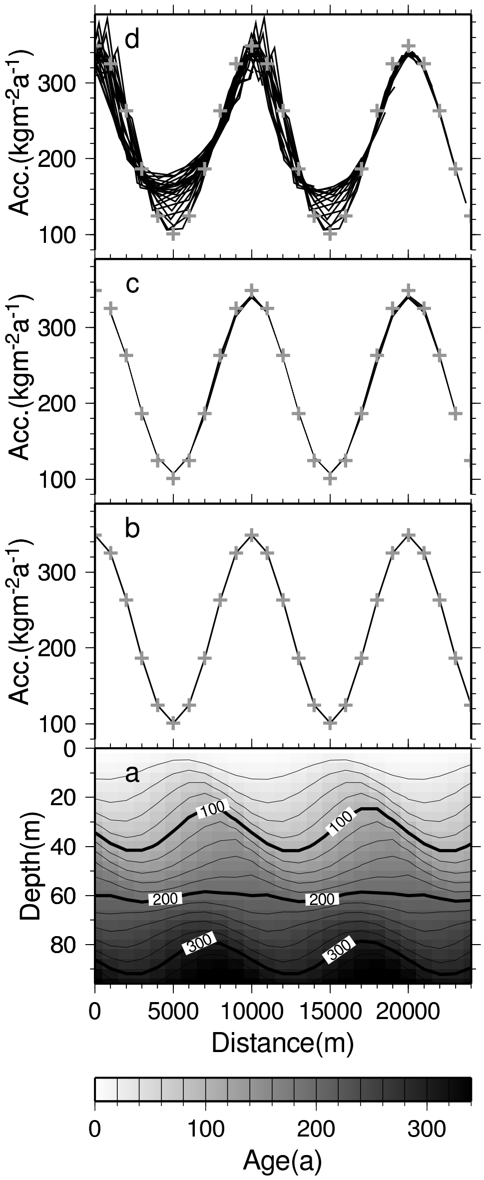

Fig. 9. ⌜**SVD solution vs. conventional accumulation estimates and prescribed values. a) Age–depth distribution according to the scenario presented by Arcone and others (2005, figure 10c) and discussed in the text, with ice flow $u = 50$ m a^{-1}from left to right. Note the almost horizontal isochrones for an age of around 200 a. b) prescribed surface accumulation (black line and grey crosses) producing the age–depth distribution of a); c) accumulation solution for the inverse problem *BuPf* calculated from the vertical-velocity solution and the prescribed density–depth distribution at different depth indicated as lines; d) conventional accumulation estimates with correction for horizontal advection according to layer age indicated by lines (see text for details). As the correction for advection corresponds to an upstream shift (to the left), the inferred accumulation distributions can thus cover only parts of the x-dimension, as further information is outside the domain for which the age–depth distribution is available (beyond 25 km). Grey crosses in b)–d) indicate reference values for accumulation at numerical nodes at the surface.** ⌝

ACKNOWLEDGEMENTS

I am deeply indebted to Edwin Waddington. Without the effort he put into critical and constructive reviews this work would not have come to this state. His linguistic and stylistic recommendations improved the text tremendously, and I hope that his mission is not as quixotic as it might seem. The efforts by scientific editor Ralf Greve and an anonymous reviewer are likewise acknowledged. Martin Losch deserves thanks for pointing out oceanographic fundamentals and analogies in the very beginning of this study. Preparation of this work was supported by the “Emmy Noether”-programme of the Deutsche Forschungsgemeinschaft.

References

Anschütz, Helgard, Olaf Eisen, Wolfgang Rack and Mirko Scheinert, 2006. Periodic surface features in coastal East Antarctica, *Geophys. Res. Letters*, **33**(22), L22501.

Arcone, Steven A., Vandy B. Spikes and Gordon S. Hamilton, 2005. Stratigraphic variation within polar firn caused by differential accumulation and ice flow: interpretation of a 400 MHz short-pulse radar profile from West Antarctica, *J. Glaciol.*, **51**(174), 407–422.

Bamber, J. L., R. J. Hardy and I. Joughin, 2000. An analysis of balance velocities over the Greenland ice sheet and comparison with synthetic aperture radar interferometry, *J. Glaciol.*, **46**(152), 67–74.

Clarke, G. K. C., N. L'homme and S. J. Marshall, 2005. Tracer transport in the Greenland ice sheet: three-dimensional isotopic stratigraphy, *Quat. Sci. Rev.*, **24**(1-2), 155–171.

Eisen, O., U. Nixdorf, F. Wilhelms and H. Miller, 2004. Age estimates of isochronous reflection horizons by combining ice core, survey, and synthetic radar data, *J. Geophys. Res.*, **109**(B04106), doi: 10.1029/2003JB002858.

Fahnestock, Mark A., Ted A. Scambos, Christopher A Shuman, R. J. Arthern, D. P. Winebrenner and R. Kwok, 2000. Snow megadune fields on the East Antarctic plateau: Extreme atmosphere-ice interaction, *Geophys. Res. Letters*, **27**(22), 3719 – 3722.

Fiadeiro, M. E. and G. Veronis, 1982. On the determination of absolute velocities in the ocean, *J. Mar. Res.*, **40**(Suppl.), 159–182.

Frezzotti, M., S. Gandolfi and S. Urbini, 2002. Snow megadunes in Antarctica: Sedimentary structure and genesis, *J. Geophys. Res.*, **107**(0), ACL X-1–X-12.

Frezzotti, Massimo, Michel Pourchet, Onelio Flora, Stefano Gandolfi, Michel Gay, Stefano Urbini, Christian Vincent, Silvia Becagli, Roberto Gragnani, Marco Proposito, Mirko Severi, Rita Traversi, Roberto Udisti and Michel Fily, 2004. New estimations of precipitation and surface sublimation in East Antarctica from snow accumulation measurements, *Climate Dynamics*, **23**(7-8), 803 – 813.

Hawley, R. L., E. D. Waddington, K. C. Taylor and G. W. Lamorey, 2004. Vertical-strain measurements at Siple Dome, Antarctica, *J. Glaciol.*, **50**(170), 447–452.

Hindmarsh, R.C.A, G.J.M.C. Leysinger-Vieli, M.J. Raymond and G.H. Gudmundsson, 2006. Draping or overriding: The effect of horizontal stress gradients on internal layer architecture in ice-sheets, *J. Geophys. Res.*, **111**(02), F02018.

Jacobel, R. W. and B. C. Welch, 2005. A Time Marker at 17.5 KYBP Detected Throughout West Antarctica, *Ann. Glac.*, **41**, 47–51.

Joughin, I., D. R. MacAyeal and S. Tulaczyk, 2004. Basal shear stress of the Ross ice streams from control method inversions, *J. Geophys. Res.*, **109**(B9).

Larour, E., E. Rignot, I. Joughin and D. Aubry, 2005. Rheology of the Ronne Ice Shelf, Antarctica, inferred from satellite radar interferometry data using an inverse control method, *Geophys. Res. Letters*, **32**, L05503.

Leonard, K. L., R. E. Bell, M. Studinger and B. Tremblay, 2004. Anomalous accumulation rates in the Vostok ice-core resulting from ice flow over Lake Vostok, *Geophys. Res. Letters*, **31**, L24401.

Lüthi, M. P. and M. Funk, 2001. Modelling heat flow in a cold, high-altitude glacier: interpretation of measurements from Colle Gnifetti, Swiss Alps, *J. Glaciol.*, **47**(157), 314–324.

MacAyeal, D. R., 1993. A tutorial on the use of control methods in ice-sheet modeling, *J. Glaciol.*, **39**(131), 91–98.

MacAyeal, D. R., R. A. Bindschadler and T. A. Scambos, 1995. Basal friction of Ice Stream E, West Antarctica, *J. Glaciol.*, **41**, 247–262.

Menke, W., 1989. Geophysical Data Analysis: Discrete Inverse Theory, Academic, San Diego.

Morse, David L., 1997. Glacier Geophysics at Taylor Dome, Antarctica, (Ph.D. thesis, University of Washington.

Nereson, N. A. and E.D. Waddington, 2002. Isochrones and isotherms beneath migrating ice divides, *J. Glaciol.*, **48**(160), 95–108.

Parrenin, Frédéric and Richard Hindmarsh, 2007. Influence of a non-uniform velocity field on isochrone geometry along a steady flowline of an ice sheet, *J. Glaciol.*, **53**(183), 612–622.

Richardson, C. and P. Holmlund, 1999. Spatial variability at shallow snow-layer depths in central Dronning Maud Land, East Antarctica, *Ann. Glac.*, **29**, 10–16.

Richardson-Näslund, Cecilia, 2004. Spatial characteristics of snow accumulation in Dronning Maud Land, Antarctica, *Global and Planetary Change*, **42**, 31–43.

Rotschky, G., O. Eisen, F. Wilhelms, U. Nixdorf and H. Oerter, 2004. Spatial characteristics of accumulation patterns derived from combined data sets in Dronning Maud Land, Antarctica, *Ann. Glac.*, **39**, 265–270.

Scales, J. A., M. L. Smitch and S. Treitel, 2001. Introductory Geophysical Inverse Theory, Samizdat Press, Golden, White River Junction.

Schwerzmann, A., M. Funk, H. Blatter, M. Lüthi, M. Schwikowski and A. Palmer, 2006. A method to reconstruct past accumulation rates in alpine firn regions: A study on Fiescherhorn, Swiss Alps, *J. Geophys. Res.*, **111**(F01014).

Siegert, Martin J., Richard C. A. Hindmarsh and Gordon S. Hamilton, 2003. Evidence for a large surface ablation zone in central East Antarctica during the last Ice Age, *Quaternary Research*, **59**, 114–121.

Truffer, M., 2004. The basal speed of valley glaciers: an inverse approach, *J. Glaciol.*, **50**(169), 236–242.

Vieli, A. and A. J. Payne, 2003. Application of control methods for modelling the flow of Pine Island Glacier, West Antarctica, *Ann. Glac.*, **36**, 197–204.

Waddington, Edwin D., Thomas A. Neumann, Michelle R. Koutnik, Hans-Peter Marshall and David L. Morse, 2007. Inference of accumulation-rate patterns from deep layers in glaciers and ice sheets, *J. Glaciol.*, **53**(183), 694–712.

Wunsch, C., 1985. Can a Tracer Field Be Inverted for Velocity?, *J. Phys. Oceanogr.*, **15**(11), 1521–1531.

Wunsch, Carl, 1996. The ocean circulation inverse problem, Cambridge University Press, Cambridge, New York.

Xiaolan, W. U. and C. Jezek, 2004. Antarctic ice-sheet balance velocities from merged point and vector data, *J. Glaciol.*, **50**(169), 219–230.

APPENDIX

Notation

Convention of variables:
vectors: lower-case bold letters (e.g. $\mathbf{u}$)
matrices: upper-case bold letters (e.g. $\mathbf{M}$)

Section 2.1	
$A, A_{i,k}$	depositional age of particle
$\rho, \rho_k, \rho_{i,k}$	density
t	time
x, z	horizontal, vertical spatial coordinate
$\mathbf{r} = (x, z)$	coordinate vector
∂_i	partial derivative with respect to $i \in \{x, z, t\}$
$u, u_{i,k}$	horizontal velocity component
$w, w_{i,k}$	vertical velocity component
$\mathbf{u}, \mathbf{w}$	horizontal, vertical velocity field
$\mathbf{v}$	velocity (=model) vector $\in \mathcal{R}^N$ $= (\mathbf{u}^T, \mathbf{w}^T)^T$
Section 2.2	
$\dot{b}$	accumulation
ρ_0	density at surface
Section 2.3	
$\mathbf{u}^{ref}, \mathbf{w}^{ref}$	reference horizontal, vertical velocity field, ⌜**i.e. the correct solution**⌝
$\mathbf{u}^{est}, \mathbf{w}^{est}$	estimated model parameters: the SVD solution of horizontal and vertical velocity field
Section 2.4	
$\Delta x, \Delta z$	horizontal, vertical spatial increment
$\{c_{i,k}^{\alpha,\ldots,\nu}\}$	coefficients of linear system
I, K	number of horizontal, vertical nodes
i, k	horizontal, vertical index
n_u, n_u^x, n_u^z	number of nodes for u: total, x-, z-direction
n_w, n_w^x, n_w^z	number of nodes for w: total, x-, z-direction
M	dimension of data space (number of observations)
N	dimension of model space (number of unknowns)
$\mathbf{d}$	data vector $\in \mathcal{R}^M$
d_p	components of $\mathbf{d}$
v_q	components of $\mathbf{v}$
$\mathbf{M}$	model matrix $\in \mathcal{R}^{M\times N}$
$M_{p,q}$	components of $\mathbf{M}$
p, q	element indices
Section 3.1	
R	mathematical rank of $\mathbf{M}$
$\hat{R}$	effective/reduced rank of $\mathbf{M}$
$\mathbf{\Lambda}$	singular-value matrix $\in \mathcal{R}^{M\times N}$
$\Lambda_{p,q}$	components of $\mathbf{\Lambda}$
$\mathbf{\Lambda}_R$	submatrix of $\mathbf{\Lambda} \in \mathcal{R}^{R\times R}$
λ_p	singular value
$\mathbf{U}$	data/observation space $\in \mathcal{R}^{M\times M}$ $= \{\mathbf{U}_R\ \mathbf{U}_0\}$
$\mathbf{V}$	model/solution space $\in \mathcal{R}^{N\times N}$ $= \{\mathbf{V}_R\ \mathbf{V}_0\}$
$\mathbf{U}_R$	data range $\in \mathcal{R}^{M\times R}$
$\mathbf{V}_R$	model range $\in \mathcal{R}^{N\times R}$
$\mathbf{U}_0$	data nullspace $\in \mathcal{R}^{M\times M-R}$
$\mathbf{V}_0$	model nullspace $\in \mathcal{R}^{N\times N-R}$
$\boldsymbol{\alpha}$	coefficients of data nullspace
δ_{ij}	Kronecker symbol

Section 3.2	
$\mathbf{T_V}$	model/solution resolution matrix $= \mathbf{V}_R \mathbf{V}_R^T$
$\mathbf{T_U}$	data/observation resolution matrix $= \mathbf{U}_R \mathbf{U}_R^T$
$\mathbf{I}_N$	unit matrix $\in \mathcal{R}^{N \times N}$
Section 3.3	
$\mathbf{n}$	vector of residuals $\in \mathcal{R}^M$
$\mathbf{R_{nn}}$	residual covariance
$\mathbf{R}_{\alpha\alpha}$	covariance of nullspace coefficients
$\mathbf{C_{vv}}$	model covariance
$\mathbf{P_{vv}}$	model uncertainty
$\mathbf{P_{nn}}$	residual uncertainty
Section 3.4	
$\mathbf{S}, \mathbf{W}$	column-, row-scaling matrix
$\mathbf{M}^T$	transpose
$\mathbf{\Lambda}^{-1}$	inverse
$\mathbf{W}^{1/2}$	square root (Cholesky decomposition)
$\widetilde{\mathbf{M}}, \tilde{\mathbf{d}}, \tilde{\mathbf{v}}$	linear system in scaled space
Section 3.5	
$\bar{\mathbf{d}}, \bar{\mathbf{v}}$	vectors corresponding to flow-field mean
$\breve{\mathbf{d}}, \breve{\mathbf{v}}, \ldots$	vectors corresponding to flow-field variations
$\langle \mathbf{u} \rangle, \langle \mathbf{w} \rangle$	mean of vectors $\mathbf{u}, \mathbf{w}$
$\mathbf{i}_N$	diagonal of $\mathbf{I}_N$, vector with all ones
Section 4.1	
$x_{min}, x_{max}, z_{max}$	boundaries of x- and z-dimension
$\dot{b}_0, x_\sigma, x_\mu$	parameters of accumulation distribution
ρ_i, c_ρ	parameters of density distribution
c_u	parameter of horizontal velocity distribution
Section 4.2	
k_0	vertical index at surface
i_b	horizontal index of borehole position
$\Delta_x u, \Delta_z u$	horizontal, vertical difference of u over one spatial increment
Section 5	
$\Delta\breve{\mathbf{v}}, \Delta\breve{\mathbf{u}}, \Delta\breve{\mathbf{w}}$	residuals of velocity variation (reference ⌈**minus**⌉ solution)
$\|\|\Delta\breve{\mathbf{u}}\|\|, \|\|\Delta\breve{\mathbf{w}}\|\|$	norm of velocity residuals
$\|\|\breve{\mathbf{n}}\|\|, \|\|\breve{\mathbf{v}}\|\|$	norm of residual, solution/model vector

Staggered-grid differences and coefficients

⌈**Applying finite differences to (3) on the triplex-staggered grid yields the discrete equations**

$$\frac{1}{2\Delta x}\left[(A_{i+1,k} - A_{i,k})u_{i,k} + (A_{i,k} - A_{i-1,k})u_{i-1,k}\right] + \frac{1}{2\Delta z}\left[(A_{i,k+1} - A_{i,k})w_{i,k} + (A_{i,k} - A_{i,k-1})w_{i,k-1}\right] = 1 \quad \text{(A1a)}$$

$$\rho_k \frac{1}{\Delta x}(u_{i,k} - u_{i-1,k}) + \rho_k \frac{1}{\Delta z}(w_{i,k} - w_{i,k-1}) + \frac{1}{2\Delta z}\left[(\rho_{k+1} - \rho_k)w_{i,k} + (\rho_k - \rho_{k-1})w_{i,k-1}\right] = 0 \quad \text{(A1b)}$$

⌉where the i index for density $\rho_{i,k}$ has been dropped, as density is laterally homogeneous, and depends ⌈**only**⌉on depth index k. Rearranging and combining factors to the coefficients $\{c_{i,k}^{\alpha,\ldots,\nu}\}$ results in the expression for a unit cell (Figure 1),

$$c_{i-1,k}^{\alpha} u_{i-1,k} + c_{i,k}^{\beta} u_{i,k} + c_{i,k-1}^{\gamma} w_{i,k-1} + c_{i,k}^{\delta} w_{i,k} = 1 \quad \text{(A2a)}$$

$$c_k^{\kappa} u_{i-1,k} + c_k^{\lambda} u_{i,k} + c_{i,k-1}^{\mu} w_{i,k-1} + c_{i,k}^{\nu} w_{i,k} = 0, \quad \text{(A2b)}$$

which⌈**can be written in**⌉the matrix notation of equation (4). The coefficients are given by

$$\begin{aligned}
c_{i-1,k}^{\alpha} &= \frac{1}{2\Delta x}(A_{i,k} - A_{i-1,k}), \\
c_{i,k}^{\beta} &= \frac{1}{2\Delta x}(A_{i+1,k} - A_{i,k}), \\
c_{i,k-1}^{\gamma} &= \frac{1}{2\Delta z}(A_{i,k} - A_{i,k-1}), \\
c_{i,k}^{\delta} &= \frac{1}{2\Delta z}(A_{i,k+1} - A_{i,k}), \\
c_k^{\kappa} &= -\frac{\rho_k}{\Delta x}, \\
c_k^{\lambda} &= \frac{\rho_k}{\Delta x}, \\
c_{k-1}^{\mu} &= -\frac{1}{2\Delta z}(\rho_k + \rho_{k-1}), \\
c_k^{\nu} &= \frac{1}{2\Delta z}(\rho_{k+1} + \rho_k).
\end{aligned} \quad \text{(A3)}$$

Cases of determinacy and conditions for existence of nullspaces

⌈**Let us denoted by $\{\}$ empty sets of the model nullspace $\mathbf{V}_0$ or the data nullspace $\mathbf{U}_0$. If a data nullspace exists, $\mathbf{U}_0 \neq \{\}$, and the data vector has components in it, then it will be impossible to fit the data exactly (Scales and others, 2001). If a model nullspace exists, $\mathbf{V}_0 \neq \{\}$, and the true model vector has components in it, then it will be impossible to find the correct model. The following combinations are possible:**⌉

$M = N$	just determined $\mathbf{V}_0 = \mathbf{U}_0 = \{\}$
$M = N > R$	deficient rank just determined $\mathbf{V}_0 \neq \{\}$, $\mathbf{U}_0 \neq \{\}$
$M > N = R$	full-rank overdetermined $\mathbf{V}_0 = \{\}$, $\mathbf{U}_0 \neq \{\}$
$M > N > R$	deficient rank overdetermined $\mathbf{V}_0 \neq \{\}$, $\mathbf{U}_0 \neq \{\}$
$N > M = R$	full-rank underdetermined $\mathbf{V}_0 \neq \{\}$, $\mathbf{U}_0 = \{\}$
$N > M > R$	deficient rank underdetermined $\mathbf{V}_0 \neq \{\}$, $\mathbf{U}_0 \neq \{\}$

Definition of moments, norms, scaling, and weighting

Second-moment or covariance matrix

Let x be a random variable with samples $(x_1, x_2, \ldots, x_n)$ drawn from the population. The k-th sample moment of x is defined as

$$\frac{1}{n}\sum_{i=1}^{n} x_i^k. \quad \text{(A4)}$$

The sample mean $\langle x \rangle$ follows as the first moment of x. The k-th central moments are defined as

$$\frac{1}{n}\sum_{i=1}^{n} (x_i - \langle x \rangle)^k. \quad \text{(A5)}$$

The sample variance of x is the second central moment,

$$\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \langle x\rangle\right)^2 . \tag{A6}$$

Assuming that the true mean of x is zero, the second moment is equal to the second central moment or variance,

$$\frac{1}{n}\sum_{i=1}^{n} x_i^2 . \tag{A7}$$

Renaming x with ${}^1x$ with samples $\{{}^1x_i\}$, $(i = 1, \ldots, n)$ and considering a second random variable ${}^2x$ with samples $\{{}^2x_i\}$, and further assuming that ${}^1x$ and ${}^2x$ have zero mean, we can estimate the covariance of ${}^1x$ and ${}^2x$ as

$$r_{12} = \frac{1}{n}\sum_{i=1}^{n} {}^1x_i\,{}^2x_i . \tag{A8}$$

Extending this further to the random variable ${}^Nx$ with samples $\{{}^Nx_i\}$, we can define the random vector $\mathbf{x} = ({}^1x, {}^2x, \ldots, {}^Nx)$ with samples $\mathbf{x}_i = ({}^1x_i, {}^2x_i, \ldots, {}^Nx_i)$. The covariance (or second moments) for pairs of variables ${}^px$ and ${}^qx$ follows as

$$r_{pq} = \frac{1}{n}\sum_{i=1}^{n} {}^px_i\,{}^qx_i . \tag{A9}$$

The r_{pq} are the components of the covariance or second-moment matrix $\mathbf{R_{xx}}$, as introduced in section 3.3 for vectors $\mathbf{n}$ and $\boldsymbol{\alpha}$.

L_2-norm

The norm of a vector is a measure for its length. A general definition for the norm of a vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is given by

$$||\mathbf{x}||_p = \left(|x_1|^p + |x_2|^p + \cdots + |x_n|^p\right)^{1/p} , \tag{A10}$$

where $|x_i|$ denotes the absolute value of the component x_i, and $p \geq 1$ is a real number. For $p = 1$, (A10) is the so-called L_1-norm. For $p = 2$, we get the L_2-norm, usually referred to simply as the length of the vector $\mathbf{x}$ in Euclidean space, the **square root** of the sum of squares of its components. The L_2-norm is used throughout this paper.

Row and column scaling

Let $M_{p,q}$ **be** the components of the matrix $\mathbf{M}$, with $p = 1, \ldots, M$ denoting the row number, and $q = 1, \ldots, N$ denoting the column number. The L_2-norm of the i-th row is calculated by

$$||\mathbf{M}||_p^{row} = \left(|M_{p,1}|^2 + |M_{p,2}|^2 + \cdots + |M_{p,N}|^2\right)^{1/2} . \tag{A11}$$

For row scaling, each element $M_{p,q}$ of the p-th row is divided by the row norm $||\mathbf{M}||_p^{row}$. This leads to the row-scaling matrix $\mathbf{W}$, which has components $W_{p,q}$, defined as

$$W_{p,q} = \delta_{pq}||\mathbf{M}||_p^{row} , \tag{A12}$$

that is, the $||\mathbf{M}||_p^{row}$ on the main diagonal and zero elsewhere.

Now taking $M'_{p,q}$ as the components of the already row-scaled matrix $\mathbf{M}'$, the L_2-norm of the q-th column is determined from

$$||\mathbf{M}'||_q^{col} = \left(|M'_{1,q}|^2 + |M'_{2,q}|^2 + \cdots + |M'_{M,q}|^2\right)^{1/2} . \tag{A13}$$

For column scaling, each element $M'_{p,q}$ of q-th column is divided by the column norm $||\mathbf{M}'||_q^{col}$. The components of the column-scaling matrix $\mathbf{S}$ are defined as

$$S_{p,q} = \delta_{iq}/||\mathbf{M}'||_q^{col} , \tag{A14}$$

so that $\mathbf{S}$ has the $1/||\mathbf{M}'||_q^{col}$ on the main diagonal and zeros elsewhere.