Astron. Astrophys. 344, 696–708 (1999)

ASTRONOMY AND ASTROPHYSICS

Non linear regularization for helioseismic inversions

Application for the study of the solar tachocline

T. Corbard[1,⋆], **L. Blanc-Féraud**[2], **G. Berthomieu**[1], **and J. Provost**[1]

[1] Laboratoire G.-D. Cassini, CNRS UMR 6529, Observatoire de la Côte d'Azur, B.P. 4229, F-06304 Nice Cedex 4, France
[2] Projet Ariana, CNRS/INRIA/UNSA, 2004 route des Lucioles, B.P. 93, F-06902 Sophia Antipolis Cedex, France

Received 31 August 1998 / Accepted 8 January 1999

Abstract. Inversions of rotational splittings have shown that there exists at the base of the solar convection zone a region called the tachocline in which high radial gradients of the rotation rate occur. The usual linear regularization methods tend to smooth out any high gradients in the solution, and may not be appropriate for the study of this zone. In this paper we use, in the helioseismic context of rotation inversions, regularization methods that have been developed for edge-preserving regularization in computed imaging. It is shown from Monte-Carlo simulations that this approach can lead directly to results similar to those reached by linear inversions which however required some assumptions on the shape of the transition in order to be deconvolved. The application of this method to LOWL data leads to a very thin tachocline. From the discussions on the parameters entering the inversion and the Monte-Carlo simulations, our conclusion is that the tachocline width is very likely below $0.05R_\odot$ which lowers our previous estimate of $0.05 \pm 0.03R_\odot$ obtained from the same dataset (Corbard et al. 1998).

Key words: methods: numerical – Sun: interior – Sun: oscillations – Sun: rotation

1. Introduction

Helioseismic inverse problems consist in using the properties (namely frequencies or frequency splittings) of the oscillation pattern observed at the surface of the Sun in order to infer the internal variation of solar physical properties like the rotation rate, the sound speed or the density (see e.g. Gough & Thompson (1991) for a review). These problems can be expressed in terms of integral equations which represent ill-posed problems in the sense of Hadamard (1923).

The traditional approach used to override this difficulty consists in *regularizing the problem* by adding some a priori information on the solution (e.g. Kirsch 1996; Craig & Brown 1986). The well known Tikhonov regularization method (Phillips 1962; Twomey 1963; Tikhonov 1963) assumes a global smoothness of the solution by minimizing the norm of its derivative at a given order. Nevertheless, inversions for the solar rotation have shown (see e.g. Thompson et al. (1996) and Schou et al. (1998) for the latest results) that high gradients exist in the solar rotation profile near the surface and at the base of the convection zone in the so-called *solar tachocline* (Spiegel & Zahn 1992).

Therefore, enforcing global smoothness a priori may not be appropriate for the study of these zones, which are of particular interest for the study of solar dynamics. As a matter of fact, the tachocline represents a thin zone where the differential rotation of the convection zone becomes rigid in the radiative interior. It is thought to be the place where the solar dynamo originates and its precise structure is an important test for angular momentum transport theories. More precisely, the thickness of the tachocline can be related to the horizontal component of the turbulent viscosity and may be used as an important observational constraint on the dynamics properties of the turbulence (Spiegel & Zahn 1992; Gough & Sekii 1998; Elliot 1997).

Several works have already been performed to infer the fine structure of the tachocline (Kosovichev 1996, 1998; Basu 1997; Charbonneau et al. 1998; Antia et al. 1998; Corbard et al. 1998) using both forward analysis and inverse techniques. For the inverse approach, it may be interesting to change the global constraint which tends to smooth out every high gradients in the solution and to find a way to preserve these gradients in the inversion process. A first attempt in that direction has been carried out (Corbard et al. 1998) by using a nonlinear regularization term based on the Piecewise Polynomials-Truncated Singular Value Decomposition (PP-TSVD) method (Hansen & Mosegaard 1996) which uses a 1-norm. Nevertheless it has been shown that this method can produce solutions with sharp discontinuities even if the width of the tachocline is relatively large. This leads to very large error bars on the inferred width of high gradient zones and complicates the interpretation of the results obtained from real observations. More elaborate nonlinear techniques have been developed for edge-preserving regularization in computed imaging. In this paper, we investigate their use in the helioseismic context.

Sect. 2 briefly recalls how the solar internal rotation can be related to the frequency splittings determined from helioseismic

Send offprint requests to: T. Corbard
⋆ *Present address:* NCAR/High Altitude Observatory, P.O. Box 3000, Boulder, CO 80307, USA
Correspondence to: corbard@hao.ucar.edu

measurements, and introduces the corresponding discretized inverse problem. The non linear approach of regularization in inverse techniques is introduced in Sect. 3, and the computational aspects are discussed. For the particular case of solar rotation inversion and the determination of the tachocline width, Sects. 4 and 5 present how the regularizing parameters have been chosen and give the results of Monte-Carlo simulations for the estimation of the uncertainty on the tachocline width. Finally, the results obtained with LOWL data are discussed in Sect. 6, and we conclude in Sect. 7.

2. The astrophysical problem and its discretization

The internal rotation $\Omega(r,\theta)$ of the Sun expressed as a function of the solar radius r and colatitude θ can be related (Hansen et al. 1977), through a 2D integral equation, to the observed frequency splittings $\Delta\nu_{nlm}$ where each mode of solar acoustic oscillations is characterized by the degree l, the radial order n and the azimuthal order m ($-l \leq m \leq l$).

In this work, we focus on the application of non linear inversion to the description of the tachocline profile in the equatorial plane. The so-called 'tachocline parameters' are obtained by fitting, between $0.4R_\odot$ and $0.8R_\odot$, the solution $\bar{\Omega}(r)$ of the inversion by an error function (erf) of the form:

$$\Omega_{fit}(r) = \bar{\Omega}_0 + \frac{\bar{\Omega}_1 - \bar{\Omega}_0}{2}\left(1 + erf\left(\frac{r-\bar{r}_c}{0.5\bar{w}}\right)\right) + \bar{\alpha}(r-0.7). \quad (1)$$

This defines five tachocline parameters: $\bar{\Omega}_0, \bar{\Omega}_1, \bar{r}_c, \bar{w}$ and $\bar{\alpha}$. The coefficient $\bar{\alpha}$ has been introduced, following Antia et al. (1998), in order to take into account the linear behaviour sometimes found for the rotation rate in the convection zone just above the transition in the equatorial plane (see Sect. 6).

Thus, we consider the 1D problem of inferring the solar equatorial rotation profile $\Omega_{eq} = \Omega(r, 90^\circ)$ from the sum of odd-indexed a-coefficients defined by the expansion of the splittings on orthogonal polynomials (e.g. Schou et al. 1994)

$$\sum_{j=1,3,..} a_j^{nl} \simeq \int_0^{R_\odot} K_{nl}(r)\, \Omega_{eq}(r)\, dr, \quad (2)$$

where $K_{nl}(r)$ are the so-called rotational kernels, which have been calculated for each mode from a solar model taken from Morel et al. (1997). In the following, they are assumed to be known exactly.

This approximation of the 2D integral equation is valid only for high degree modes (e.g. Corbard 1997) but the influence of the low degree modes on the determination of the position and width of the tachocline and the rotation rate of the upper layers is thought to be small.

We discretize Eq. (2) by using a polynomial expansion method, which leads to the matrix equation:

$$\boldsymbol{W} = \boldsymbol{R}\boldsymbol{\Omega}, \quad (3)$$

where we have defined the vector $\boldsymbol{W} \equiv (W_i/\sigma_i)_{i=1,N}$, of N truncated sum of a-coefficients $W_i = \sum_{j=1}^{n_i} a_j$ weighted by the standard deviation σ_i for each mode $i \equiv (n,l)$. The number n_i of a-coefficients is fixed by the observations for each mode. The standard deviations have been computed straightforwardly from the uncertainties quoted on each a-coefficient. In this work, we assume no error correlation between the different modes.

We seek a solution $\bar{\Omega}(r)$ defined as a piecewise linear function of the radius:

$$\bar{\Omega}(r) = \sum_{p=1}^{N_p} \omega_p \psi_p(r) \qquad \boldsymbol{\Omega} \equiv (\omega_p)_{p=1,N_p} \quad (4)$$

where $\psi_p(r)$, $p = 1, N_p$ are piecewise straight lines ($N_p = 100$ in this work) between fixed break points distributed according to the density of turning points of the modes (cf. Corbard et al. 1997). The discretization matrix $\boldsymbol{R}$ is then defined by:

$$\boldsymbol{R} \equiv (R_{ip})_{\substack{i=1,N \\ p=1,N_p}} \qquad R_{ip} = \frac{1}{\sigma_i}\int_0^{R_\odot} K_i(r)\psi_p(r)dr \quad (5)$$

An inverse method should lead to a solution that is able to produce a good fit to the data. We define the goodness-of-fit by the χ^2 value obtained for any solution $\bar{\Omega}(r)$:

$$\chi^2(\bar{\Omega}(r)) = \sum_i \left[\frac{W_i - \int_0^{R_\odot} K_i(r)\bar{\Omega}(r)dr}{\sigma_i}\right]^2, \quad (6)$$

which can be written in the discretized form:

$$\chi^2(\boldsymbol{\Omega}) = \|\boldsymbol{R}\boldsymbol{\Omega} - \boldsymbol{W}\|_2^2. \quad (7)$$

3. Regularization: the non linear approach

3.1. Generalized regularization term and Euler equations

Unfortunately, the inverse integral problem is an ill-posed problem and the minimization of only the χ^2 value generally leads to oscillatory solutions that are not 'physically acceptable' in the sense that they do not correspond to our a priori knowledge on the shape of the solution. So, we have to use regularization techniques i.e. to introduce a priori information in the minimization process. A large class of these techniques, can be expressed in the general form of the minimization of a criterion J over the unknown solution $\bar{\Omega}(r)$:

$$J(\bar{\Omega}(r)) = \chi^2(\bar{\Omega}(r)) + \lambda^2 \int_0^{R_\odot} \varphi\left(\frac{1}{\delta}\left|\frac{d^q\bar{\Omega}(r)}{dr^q}\right|\right) dr, \quad (8)$$

where λ is the so-called trade-off parameter, chosen so that it establishes a balance between the goodness-of-fit to the data and the constraint introduced on the solution. The parameter δ allows to fix the threshold on the gradient modulus of the solution under which it is smoothed, and above which it is preserved (cf. Sect. 4). The order q of the derivative is usually taken equal to one or two. The two choices can lead to similar results with the appropriate choice of the regularizing parameter λ in the domains where the solution is well constrained by the data. However these two choices correspond to two different a priori constraints on the solution. As the rotation is known to be quasi-rigid in the radiative interior (at least down to $0.4R_\odot$), we have chosen in this work to use the first derivative. We note however

that the method described below can easily be generalized for any choice of the derivative order.

Two choices for the φ-function lead to well known regularization strategies:

- $\varphi(t) = t^2$ leads to the traditional Tikhonov approach with first derivative whereas
- $\varphi(t) = t$ is known as the *Total Variation* (TV) regularization method (e.g. Rudin et al. 1992; Acar & Vogel 1994). This approach uses the absolute value rather than the square modulus of the solution gradient. It has been shown that the solution is searched in a space composed of bounded variation functions which admit discontinuity points. This regularization method is therefore able to recover piecewise smooth solutions with steep gradients (see e.g. Dobson & Santosa 1994; Vogel & Oman 1996, 1997). The PP-TSVD method of Hansen & Mosegaard (1996), already used in helioseismic context in Corbard et al. (1998), can be seen as a 'truncated version' (in the sense that the regularizing parameter is a discrete truncation parameter) of the TV regularization in the same way as the MTSVD method introduced by Shibahashi & Sekii (1988) (see also: Hansen et al. 1992; Corbard et al. 1998) is a 'truncated version' of the Tikhonov regularization.

For a general φ-function, one can write the criterion Eq. (8) in a discretized form:

$$J(\mathbf{\Omega}) = \chi^2(\mathbf{\Omega}) + \lambda^2 J_2(\mathbf{\Omega}), \tag{9}$$

where $J_2(\mathbf{\Omega})$ represents the discretized regularization term defined by:

$$\int_0^{R_\odot} \varphi\left(\frac{1}{\delta}\left|\frac{d\bar{\Omega}(r)}{dr}\right|\right) dr = J_2(\mathbf{\Omega}) = \sum_{p=1}^{N_p-1} c_p \varphi\left(\frac{|\boldsymbol{L\Omega}|_p}{\delta}\right). \tag{10}$$

In this equation $(c_p)_{p=1,N_p-1}$ represent the weights used for the integration rule, and $\boldsymbol{L}$ is a discrete approximation of the first derivative operator. $|\boldsymbol{L\Omega}|_p$ is the absolute value of the p^{th} component of the vector $\boldsymbol{L\Omega}$. The expression of c_p and $\boldsymbol{L}$ are given in Appendix for the simple case of the polynomial expansion Eq. (4)

The minimization of the criterion $J(\mathbf{\Omega})$ over each component ω_p of $\mathbf{\Omega}$ leads to the following Euler equations (discretized form):

$$\nabla J(\mathbf{\Omega}) = 0 \Longleftrightarrow (\boldsymbol{R}^\top \boldsymbol{R} + \bar{\lambda}^2 \boldsymbol{L}^\top \boldsymbol{B}(\mathbf{\Omega}) \boldsymbol{L}) \mathbf{\Omega} = \boldsymbol{R}^\top \boldsymbol{W} \tag{11}$$

where $\bar{\lambda} = \frac{\lambda}{\delta}$ and $\boldsymbol{B}$ is a diagonal matrix with elements that depend on the gradient of the solution at each grid point:

$$\boldsymbol{B} = diag(b_p) \quad \text{with} \quad b_p = c_p \frac{\varphi^{'}(t)}{2t} \quad \text{and} \quad t = \frac{|\boldsymbol{L\Omega}|_p}{\delta} \tag{12}$$

For $\varphi(t) = t^2$, $\boldsymbol{B}$ is independent of $\mathbf{\Omega}$ and these normal equations reduce to a linear system which corresponds to the usual Tikhonov regularization with first derivative. On the other hand, for a general φ-function, this leads to a nonlinear problem which requires an appropriate iterative method of solution.

Now, with this general expression for the normal equations, the question for our particular problem becomes: what properties must the φ-functions satisfy to ensure the preservation of high gradients in the solution? The next section shows how the theoretical works developed in the field of computed imaging gives an answer to this question, and leads to an algorithm for solving the non linear Euler equation that is easy to implement.

3.2. Properties of the weighting function $\frac{\varphi^{'}(t)}{2t}$

From the Euler equation (Eqs. (11) and (12)) we can see that the function $\varphi^{'}(t)/2t$ acts as a weighting function in the smoothing process: at each grid point the gradient of the solution is used as an argument of this function in order to set locally the magnitude of regularization. This suggests an iterative process where the gradient of the solution at a given step is used for the computation of the regularization term at the next step. We show in the following that we can derive some basic properties of the weighting function so that high gradients can be preserved, and such that an iterative algorithm for solving the Euler equation is possible. First let us look at the behaviour of the weighting function at the limits of low and high gradients:

- For low gradients, we want to keep a Tikhonov regularization. From Eq. (11) this is the case if $\varphi^{'}(t)/2t$ is a non-null constant function.
- For high gradients, we want to remove smoothing. This happens when $\varphi^{'}(t)/2t$ is close to zero.
- Another property that sounds reasonable is to choose a decreasing function of the gradient between these two limits. Furthermore, in order to avoid numerical instabilities we will choose only strictly decreasing weighting functions.

Therefore the choice of the φ function must be made taking into account the following three properties needed on the weighting function $\varphi^{'}(t)/2t$ (Charbonnier et al. 1994, 1997):

1. Tikhonov smoothing for low gradients:
$$0 < \lim_{t\to 0} \frac{\varphi^{'}(t)}{2t} = M < \infty \tag{13}$$
2. no smoothing for high gradients:
$$\lim_{t\to \infty} \frac{\varphi^{'}(t)}{2t} = 0 \tag{14}$$
3. $\frac{\varphi^{'}(t)}{2t}$ strictly decreasing on$[0, +\infty[$ to avoid instabilities. (15)

Within these conditions, the φ-function may be chosen either convex (Green 1990; Charbonnier et al. 1994) or non-convex (Perona & Malik 1990; Geman & McClure 1985; Hebert & Leahy 1989) (see also Charbonnier et al. (1997) and Teboul et al. (1998) for examples in both cases). A non-convex function may be better suited for the search of high gradients. Nevertheless this choice leads to some numerical difficulties and instabilities related to the existence of local minima. This may

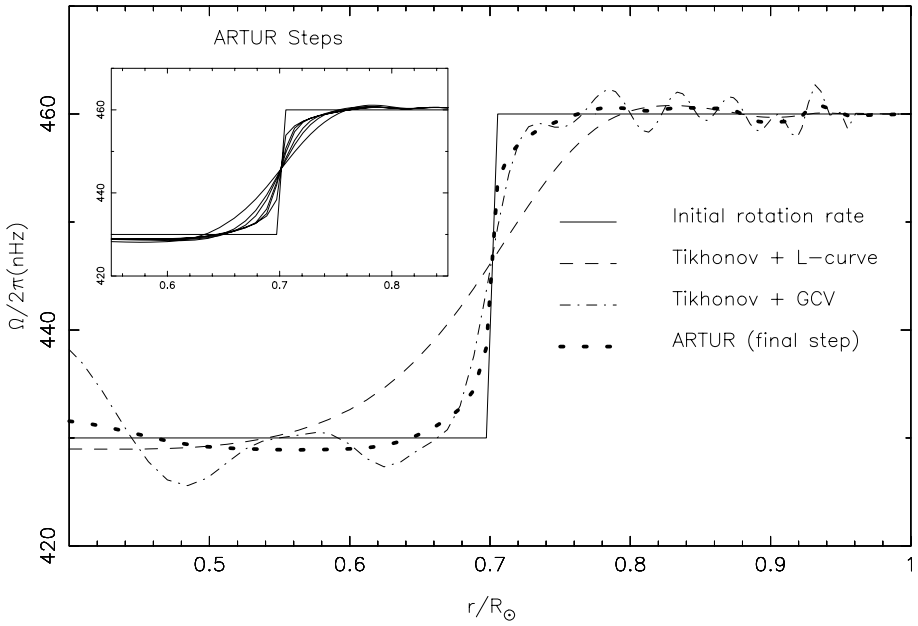

Fig. 1. Solutions obtained by inverting splittings computed from a discontinuous one dimension rotation profile (solid line) for the same modeset as in the LOWL data and including some 'realistic' noise (see text). The standard Tikhonov solution is given for two different automatic choices of the regularizing parameter. The successive steps of ARTUR algorithm are shown in the upper left insert, whereas the final step is shown on the main plot. The choice of the regularizing function and parameters for ARTUR algorithm are those discussed in Sect. 4. The solutions are plotted without error bars for clarity.

induce a high sensitivity of the inverse process to the choice of the regularization parameters (Blanc-Féraud et al. 1995). On the other hand, the choice of a convex function may avoid these numerical problems and is more suitable for relatively smooth transition (Blanc-Féraud 1998).

We note that in the case of the TV regularization (or equivalently the PP-TSVD method) the function $\varphi(t) = t$ does not satisfy the first property (Eq. 13). This may explain the difficulties encountered in using this method (Corbard et al. 1998). For smooth transitions, the dispersion of the results for different realizations of the noise became large, indicating some instabilities in the inversion process. In light of the above formalism, this may be related to the non differentiability of the corresponding weighting function $\varphi'(t)/2t$ at $t = 0$, which can lead to numerical instabilities.

3.3. The iterative algorithm: ARTUR

Under the three conditions (13)(14)(15), it has been shown by Charbonnier et al. (1994, 1997) that the non linear criterion can be solved by using an iterative scheme named ARTUR (Algebraic Reconstruction Technique Using Regularization) that is easy to implement: at each step k we calculate the regularization term using the derivative of the previous estimate $\mathbf{\Omega}^{k-1}$ and we simply compute the new estimate $\mathbf{\Omega}^k$ by solving the linear system:

$$\Big(\boldsymbol{R}^\top \boldsymbol{R} + \bar{\lambda}^2 \boldsymbol{L}^\top \boldsymbol{B}(\mathbf{\Omega}^{k-1})\boldsymbol{L}\Big)\mathbf{\Omega}^k = \boldsymbol{R}^\top \boldsymbol{W}. \tag{16}$$

When we use a convex φ-function, the convergence of this so-called *half quadratic algorithm* (minimization of a quadratic criterion at each step (Geman & Reynolds 1992)) to the minimum of the criterion given by Eq. (8) has been established (Charbonnier et al. 1997). This is also an adaptive regularization method which uses the information on the derivative of the solution obtained at each step in order to improve the regularization at the next step. This requires an initial guess for the solution, but we will show in the next section that a constant solution can always be used as the starting guess. Figs. 1 and 2 show examples of ARTUR steps in the case of a discontinuous rotation rate and for two different levels of noise. At each step of the ARTUR algorithm the linear system Eq. (16) has been solved using an iterative conjugate gradient method with Jacobi preconditioning (see e.g. Golub & Van-Loan 1989; Barrett et al. 1994) using $\mathbf{\Omega}^{k-1}$ as starting point. This leads to a very fast algorithm where the number of conjugate gradient iterations needed to solve the linear system decreases at each ARTUR step. The algorithm is stopped when the 2-norm of the relative difference between two solutions at two successive steps is below 10^{-6} i.e.:

$$\frac{\|\mathbf{\Omega}^k - \mathbf{\Omega}^{k-1}\|_2}{\|\mathbf{\Omega}^k\|_2} \leq 10^{-6} \tag{17}$$

4. The choice of φ-function and regularizing parameters for rotation inversion

4.1. The φ-function

In the particular case of the determination of the solar tachocline profile, the uncertainty on the width of the transition zone is still large (see Table 2 of Corbard et al. (1998) for a summary of some previous works). Therefore, according to the previous discussion, we have chosen to consider a convex regularizing φ-function, specifically the one defined in Charbonnier et al. (1994, 1997):

$$\varphi(t) = 2\sqrt{t^2 + 1} - 2 \tag{18}$$

This function is close to the absolute value function used in TV regularization and PP-TSVD method, but, unlike the absolute value, has a quadratic behaviour near 0 that satisfies the requirement of Eq. (13) on the contrary of the absolute value.

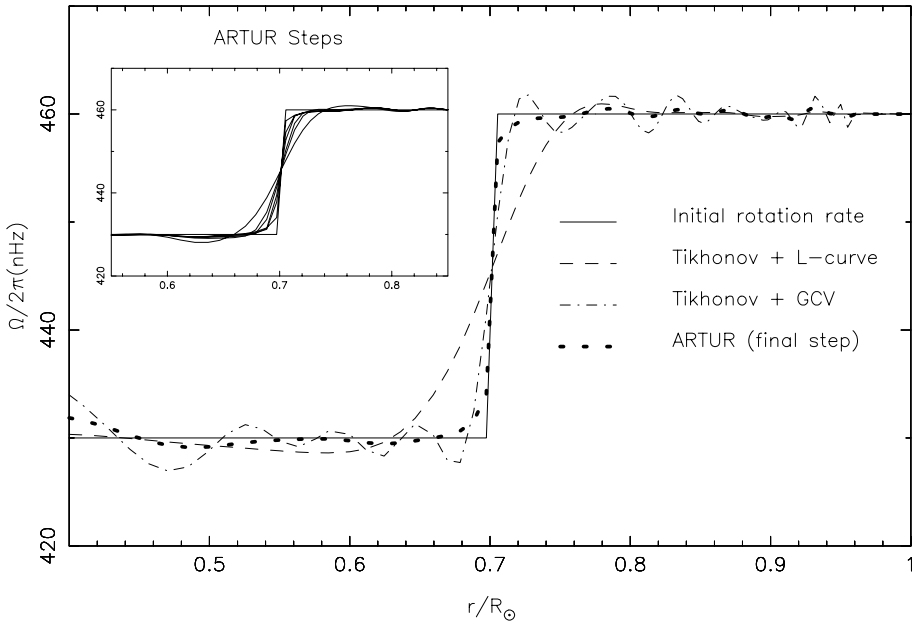

Fig. 2. The same as Fig. 1 but computed for a lower level of noise (standard deviations divided by $\sqrt{10}$). Comparison between Figs. 1 and 2 show the smoothing effect of the data noise level for the three methods.

The weighting function, shown as dashed lines in Figs. 3 and 6, is given by:

$$\varphi^{'}(t)/2t = \left(1+t^2\right)^{-\frac{1}{2}} . \tag{19}$$

4.2. The regularizing parameters

The choice of the parameters (λ, δ) is an important point, as in every regularization methods. For example, using a Generalized Cross Validation (GCV) strategy (Wahba 1977) for the regularizing parameter of a Tikhonov model leads to an oscillating solution while using the L-curve strategy produces a smoother solution. This choice influences evidently the estimation of the width of the tachocline. It is intuitively not surprising as the model represents a priori information on the solution (so the solution depends on this information).

Parameter estimation for regularizing problem is a deep question and represents an active research area of its own (eg. Lakshmanan & Derin 1989; Thompson et al. 1991; Galatsanos & Katsaggelos 1991). Some results on (λ, δ) parameters have been obtained for the proposed model in the field of image processing (Chan & Gray 1996; Jalobeanu et al. 1998; Zerubia et al. 1998). However, it is still an open problem, and studies and results depend sensitively on the considered inversion problem.

In the remainder of this paper, a choice of the parameters (λ, δ) is proposed, based on heuristic considerations (the following two sub-sections) and simulation results (Sect. 5). The application to real data is then discussed in Sect. 6.

4.2.1. The choice of $\overline{\lambda}$: using automatic strategies

If the initial guess is a constant function, then according to property 1 (Eq. 13) and Eq. (19), $M = 1$ and the solution at the first ARTUR step corresponds to a Tikhonov solution with $\overline{\lambda}$ as regularizing parameter. It has been shown in Corbard et al. (1998) that the GCV strategy for the choice of the regularizing parameter in Tikhonov inversions leads systematically to better results concerning the evaluation of the tachocline parameters, as compared to the L-curve strategy (Hansen 1992). In fact, the GCV strategy leads systematically to less smoothing than the L-curve approach ($\lambda_{Lcurve} \simeq 100 * \lambda_{GCV}$ in that work) and therefore is more suited to the study of a region with high gradients. Nevertheless, because of the low global regularization, this choice may lead to spurious oscillations below and above the tachocline (see Figs. 1, 2). The ARTUR algorithm will tend to enhance the high gradients found at the first step. Therefore it is important to start with a solution smooth enough to avoid spurious oscillations with high gradients. The GCV choice for $\overline{\lambda}$ may therefore not be well adapted for ARTUR initial step, and some experiments have shown that, on the other hand, the L-curve choice leads often to a solution which is too smooth and does not allow the buildup of the high gradients expected during the iterations. Nevertheless, the optimal choice of this parameter strongly depends on the level of noise included in the data, and therefore it is important to define an automatic choice of this parameter so that we use the same strategy for different realizations of the noise in Monte-Carlo simulations (cf. Sect. 5). The results of these simulations are shown in Fig. 4 with both $\bar{\lambda} = \lambda_{Lcurve}/10$ and $\bar{\lambda} = \lambda_{GCV}$. The second choice has been retained for the application to real data (cf. Sect. 6).

4.2.2. The choice of δ: using a priori knowledge on the searched gradients

The parameter δ has been introduced in order to adapt the shape of the weighting function to the gradient that we seek to detect in a particular application. We have chosen for simplicity to keep this parameter constant during the iterations. As we start the iterative process with a constant guess rotation, the first step is independent of δ and thus we can use the solution obtained

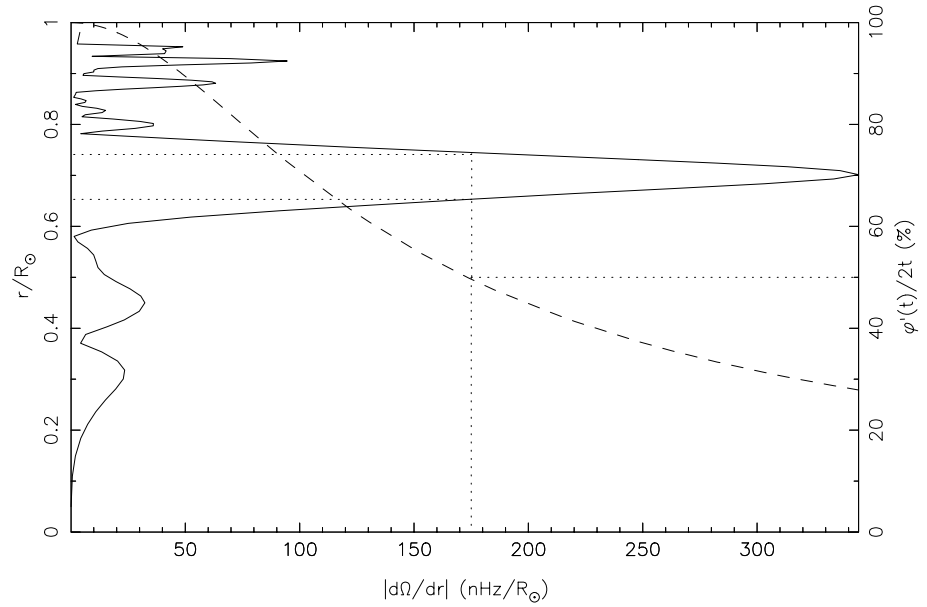

Fig. 3. The solid line shows the first derivative of a first step solution (cf. Fig. 1) in the ARTUR algorithm as a function of the fractional solar radius. For each value of the gradient, the dashed line gives the weight that will be given locally to the regularizing term at the following step of the algorithm. The weighting function is given by Eq. (19) with $t = |d\Omega/dr|/\delta$ and $\delta = 100R_\odot/$nHz. The dotted line indicates the gradient for which the local regularization will be reduced by 50% at the second step, as compared to the first step. Here, this occur for radius between $0.65R_\odot$ and $0.75R_\odot$.

after this initial step in order to adapt the parameter δ to the gradients found in the first step solution. As we have shown that the first step of ARTUR algorithm correspond to a classic Tikhonov inversion, this choice of δ can be viewed as a way to use our a priori knowledge (as given by Tikhonov inversion) of the searched gradients.

It is generally admitted that the width of the tachocline does not exceed 0.1 solar radius, which is also the resolution typically reached near the tachocline localization ($\simeq 0.69R_\odot$) with a Tikhonov method applied to the current datasets. Furthermore we have a good estimate of the difference between the rotation rate above and below the transition ($\simeq 30$nHz in Corbard et al. (1998)). Therefore we can estimate a level of $300\text{nHz}/R_\odot$ for the maximum gradient obtained at the first iteration of ARTUR process.

Fig. 3 shows (solid line) the first derivative of solution obtained at the first step by inverting artificial splittings which have been computed for a discontinuous rotation law (cf. Fig. 1) with a step of 30nHz and by adding some Gaussian noise with a standard deviation taken from the formal error given in LOWL data for each mode (cf. Corbard et al. 1998). The weighting function Eq. (19) is shown in dashed line for $\delta = 100R_\odot/nHz$. At the second step we want to preserve only high gradients i.e. to regularize less in these zones where high gradients have already been found at the first step. For example, according to Fig. 3, the choice of $\delta = 100R_\odot/nHz$ leads to regularize 50% less at the second step in that zones where the gradient of the first step solution is above ~ 175 nHz$/R_\odot$. For the particular realization of the noise introduced in artificial data this choice of $\delta = 100R_\odot/nHz$ looks reasonable, in the sense that it will tend to decrease the regularization especially in the transition zone. A smaller value would enhance the secondary peaks that are induced by the data noise.

The maximum gradient obtained, at the first step, with the artificial dataset ($350\text{nHz}/R_\odot$ in Fig. 3) corresponds approximately to our previous estimate of $300\text{nHz}/R_\odot$ expected for real data. For the Monte-Carlo simulations done in order to estimate the errors (cf. Sect. 5) the parameter δ has been fixed to $\delta = 100R_\odot/nHz$. Nevertheless, the shape of the solution derivative after the first step is a function of the dataset through the intrinsic resolution of the modeset and the level of noise. Therefore, with real data the choice of δ will always be made by looking at the derivative profile after the first step. We will see however in Sect. 6 that other indicators can help in the choice of δ, and that according to these indicators the choice $\delta = 100R_\odot/nHz$ seems to be a good compromise for LOWL data also.

5. On the error estimation on tachocline parameters using nonlinear methods

For nonlinear methods, we cannot compute straightforwardly the formal errors at each point of the solution as we can do for linear process. The ARTUR algorithm solves a linear system at each step, but the final result depends nonlinearly on the data since the coefficients of the matrix to be inverted at each step are functions of the data through the derivative of the previous estimate that is used as an argument of the weighting function.

We focus here only on having an estimate of the uncertainty on the width $\bar{w}$ and the location $\bar{r}_c$ of the tachocline. We cannot obtain directly error bars on the solution, and therefore the fit by an erf function does not give an estimate of the error on the inferred parameters. Instead, a first estimate of the uncertainty on the tachocline width inferred here has been computed using a Monte-Carlo method applied on given rotation profiles simulated by erf functions with widths lying between $0.01R_\odot$ and $0.08R_\odot$ (in steps of $0.01R_\odot$) and located at $r_c = 0.7R_\odot$. The LOWL set of modes splittings corresponding to these rotation profiles have been computed with addition of a Gaussian noise with a standard deviation taken for each mode from the formal error given in LOWL data. Since the ARTUR algorithm

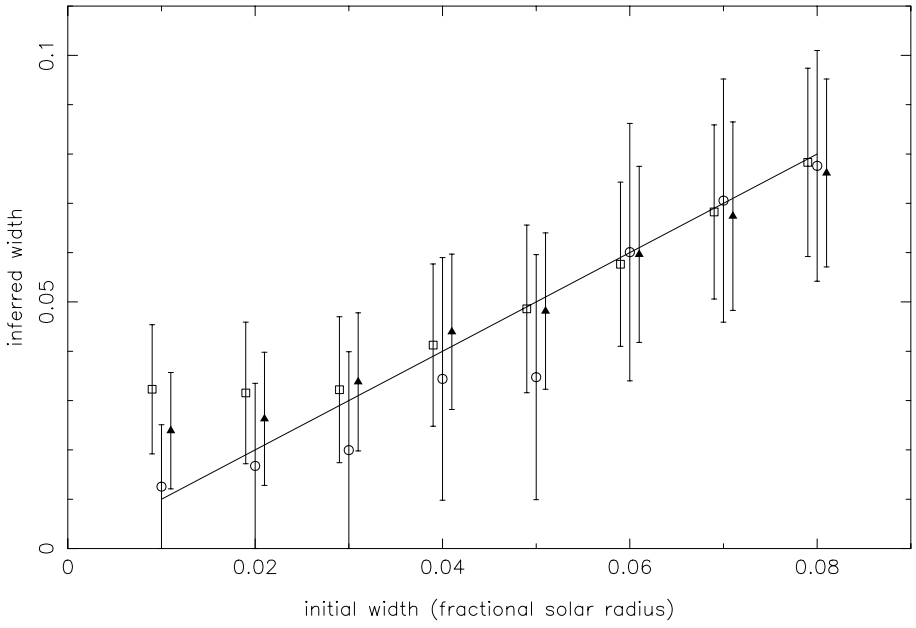

Fig. 4. Monte-carlo simulation for the estimation of the error on the width inferred with the ARTUR algorithm. Triangles are for $\bar{\lambda} = \lambda_{Lcurve}/10$ and circles for $\bar{\lambda} = \lambda_{GCV}$. In both cases $\delta = 100 R_\odot/nHz$. The rotation profile was taken as an erf function with different 'initial widths'. The 'inferred widths' are the mean value for 500 noise realizations of the results obtained by fitting directly the solutions by an erf function (cf. Eq. 1 where $\bar{\alpha}$ is assumed to be zero). Error bars represent a 68.3% confidence interval on the width. For comparison, the squares show the result obtained from Tikhonov method with GCV choice of the parameter after a 'local deconvolution' using the averaging kernels computed at the center of the transition (see Corbard et al. 1998).

is non linear, we can not define averaging kernels but Fig. 4 shows that the final step of the algorithm with $\bar{\lambda} = \lambda_{Lcurve}/10$ and $\delta = 100 R_\odot/nHz$ leads directly to results similar to those reached by Tikhonov inversion after a 'local deconvolution' using the averaging kernels computed at the center of the transition (see Corbard et al. 1998). The 1σ error interval on the width is found to be around $\pm 0.02 R_\odot$ for widths in the range 0.01–$0.08 R_\odot$. This uncertainty is probably related to the intrinsic resolution of the modeset and the level of noise contained in LOWL data.

A large bias is found with deconvolved Tikhonov method for $w \leq 0.02 R_\odot$. This bias is less important for the ARTUR method with $\bar{\lambda} = \lambda_{Lcurve}/10$ and vanishes if we set $\bar{\lambda} = \lambda_{GCV}$. Nevertheless, in this latter case error bars are bigger (around $\pm 0.025 R_\odot$ in the range 0.02–$0.08 R_\odot$ of initial widths) and the inferred widths tend to be underestimated for initials widths between $0.03 R_\odot$ and $0.05 R_\odot$. This is due to the fact that, in mean for the simulated data, the GCV criterion leads to a too strongly oscillating solution that is not a good starting point for ARTUR algorithm, as it tends to produce very sharp transitions in a large number of realizations for all initial widths below $0.05 R_\odot$.

The Monte-Carlo simulations have been performed for 500 realizations of the input errors for each initial width and the mean value over all the realizations (8×500) of the inferred center $\bar{r}_c$ is $0.701 \pm 0.004 R_\odot$. The center of the erf function seems therefore to be very well recovered by the inversion. Nevertheless, we must keep in mind that the center of the tachocline is defined as the center of the erf function that gives the best fit of the solution. This may not give an appropriate view of the tachocline profile if, for example, the solution is found to lack such a symmetry in the lower and upper parts of the tachocline when inverting real data.

Another important point that Fig. 4 demonstrates is the ability of the ARTUR algorithm to recover not only rotation with a discontinuity (as shown in the examples of Fig. 1 and 2), but also rotation with a relatively smooth transition. This property does not characterize the PP-TSVD method, and this was the reason for the difficulties encountered by Corbard et al. (1998) in interpreting the results obtained with this first non linear approach.

Even if both methods give similar results in the mean for some choices of the regularizing parameters, it sometimes happen that the two solutions differ strongly for a particular realization of the noise. Furthermore, the two approaches are very different in their underlying principles and therefore it is very interesting to compare the two results with observed data.

6. Application to observed data

6.1. LOWL data

The LOWL instrument is a Doppler imager based on a Potassium Magneto-Optical Filter that has been operating on Mauna Loa, Hawaii since 1994 (see Tomczyk et al. (1995) for a detailed description). The dataset includes the frequency splittings of 1102 modes (n, l) with degrees up to $l = 99$ and frequencies lower than $\nu = 3500$ μHz deduced from a two year period of observation (2/26/94–2/25/96). For each mode, individual splittings are given by, at best, $n_i = 5$ a-coefficients of the expansion on orthogonal polynomials defined by Schou et al. (1994).

6.2. The choice of inversion parameters and ARTUR solution for LOWL data

Since the first step of the ARTUR algorithm is standard Tikhonov inversion, we first present the results obtained from Tikhonov method and GCV choice of the regularization parameter (called T-GCV method hereafter). The equatorial profile obtained from T-GCV method on LOWL data is shown in Fig. 5. The fit of the solution with an erf function of the form

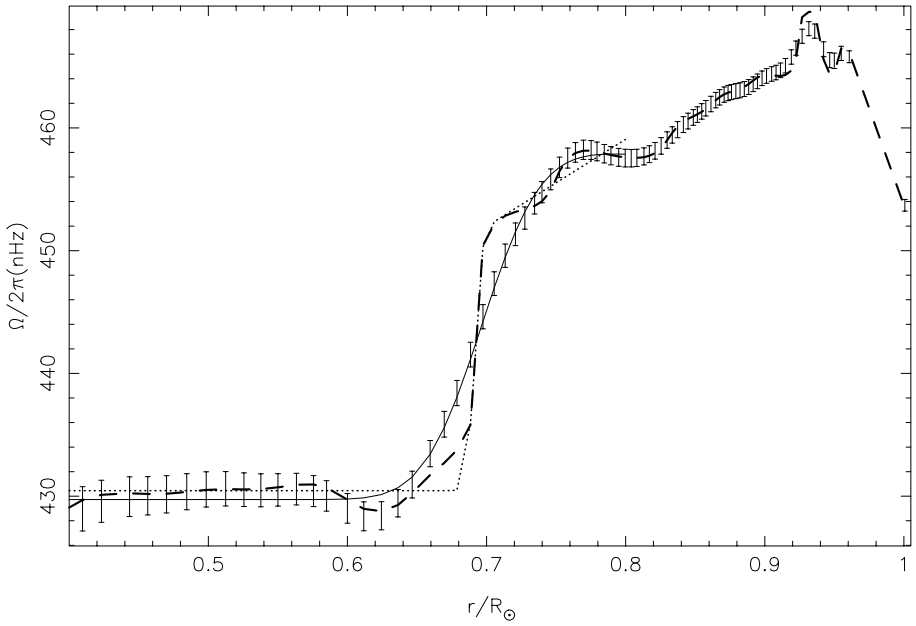

Fig. 5. Solar equatorial rotation rate estimated from LOWL data. The vertical error bars given at each grid points are the 1σ confidence intervals estimated for the T-GCV solution. The dashed line shows ARTUR solution obtained with $\bar{\lambda} = \lambda_{GCV}$ and $\delta = 100 R_\odot/nHz$. The fits by an erf function (Eq. (1)) of these two solutions are shown respectively by the solid and dotted lines. The fits have been computed only between $0.4R_\odot$ and $0.8R_\odot$ and the five parameters deduced from each fit are shown in Fig. 8

Table 1. Comparison between our previous work on inferring the equatorial tachocline profile from LOWL data and the results obtained with non linear regularization applied to the same data. The previous estimates of the parameters and their errors have been deduced from a comparison of three inversion methods, including the T-GCV method. See Sect. 6.3 for a more detailed discussion on the tachocline width using the ARTUR algorithm

	$\bar{\Omega}_0$(nHz)	$\bar{\Omega}_1$(nHz)	$\bar{\alpha}$(nHz/$R_\odot$)	$\bar{r_c}/R_\odot$	$\bar{w}/R_\odot$
Corbard et al. (1998)	431.0 ± 3.5	459.0 ± 1.5	0	0.695 ± 0.005	0.05 ± 0.03
ARTUR	430.5	452.0	70.0	0.691	0.01

Eq. (1) leads to a width of $w \simeq 0.09R_\odot$ (cf. Fig. 8d). By taking into account the width of the averaging kernel computed at $0.7R_\odot$, the corrected inferred width obtained from this 'local deconvolution' is $w \simeq 0.06R_\odot$.

Contrary to the example shown in Fig. 1 for simulated data, the GCV choice of the regularization parameter does not lead, with LOWL data, to an oscillating solution. This may indicate that this particular realization of the noise introduced in simulated data is rather different from the noise contained in LOWL data. The formal errors quoted on each a-coefficient are perhaps overestimated. Furthermore our model assumes that the errors are uncorrelated, which is probably not strictly the case. Therefore, with real data, the T-GCV solution may be a good starting point for the ARTUR algorithm and we choose to take $\bar{\lambda} = \lambda_{GCV}$ so that the first ARTUR iteration leads to the T-GCV solution.

Fig. 6 shows the absolute value of the first derivative of the T-GCV solution. High gradients are found not only in the tachocline, but also near the surface (above $0.09R_\odot$). However LOWL data include relatively few high degree modes ($l \leq 99$ in these data), so we will focus only on the tachocline in this work. As discussed in Sect. 4.2.2, Fig. 6 can help for the choice of the parameter δ. The weighting function is shown for two values of δ. The choice $\delta = 100 R_\odot/nHz$ lowers the regularization by more than a factor two at the second step in the tachocline and in the upper layers, whereas the choice $\delta = 200 R_\odot/nHz$ will never decrease the regularization by more than 50% after the first step. From this figure, we can guess that a choice of $\delta < 100 R_\odot/nHz$ will tend to enhance spurious oscillations due to the noise, whereas a choice $\delta > 200 R_\odot/nHz$ will lead to a result very similar to the T-GCV solution because only few points will be affected by the local change of regularization during the ARTUR steps. Between these two limits however, it is not clear which is the best choice for δ.

However, some indicators may help in the choice of δ:

- First, an important test for any global inversion is its capability of providing a good fit to the data. Fig. 7c shows, as a function of δ, the normalized χ^2 value for the modes which have their turning points between $0.6R_\odot$ and $0.8R_\odot$. As expected, this value is always lower for all ARTUR solutions than for the T-GCV solution, because we tend to regularize less. The gain in the χ^2 value is however small because the regularization is decreased only locally in a small region. The χ^2 value reaches a minimum for $\delta = 50 R_\odot/nHz$ but, as expected from Fig 6, for such low δ, the ARTUR solution becomes very oscillating and a discontinuity ($w \simeq 0$) is found near the tachocline but in fact the solution is found to be piecewise constant with many discontinuities.
- There is another indicator that we can use showing that this value of $\delta = 50 R_\odot/nHz$ is not appropriate, despite the

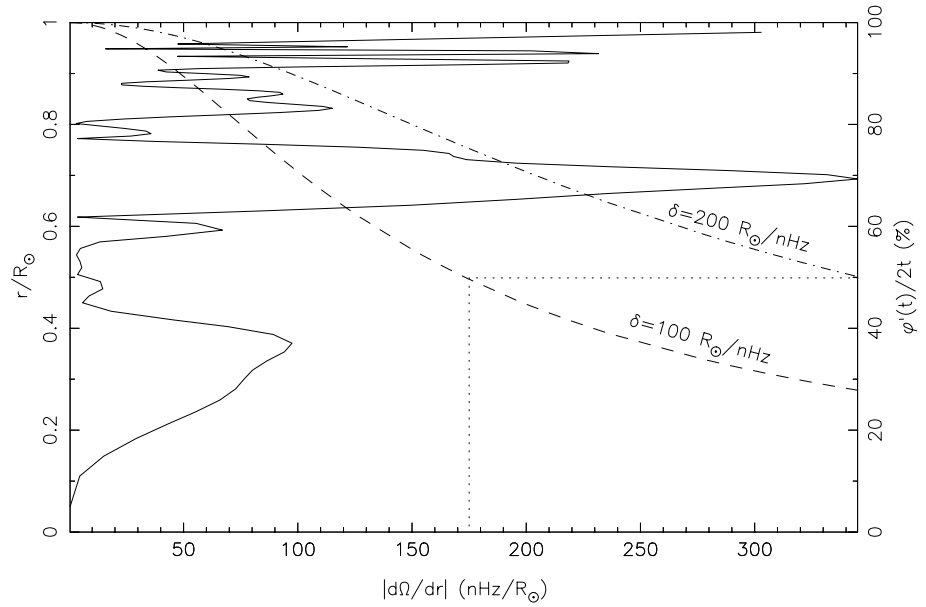

Fig. 6. Same as Fig. 3 but for LOWL data. The weighting function is shown for two choices of the parameter δ.

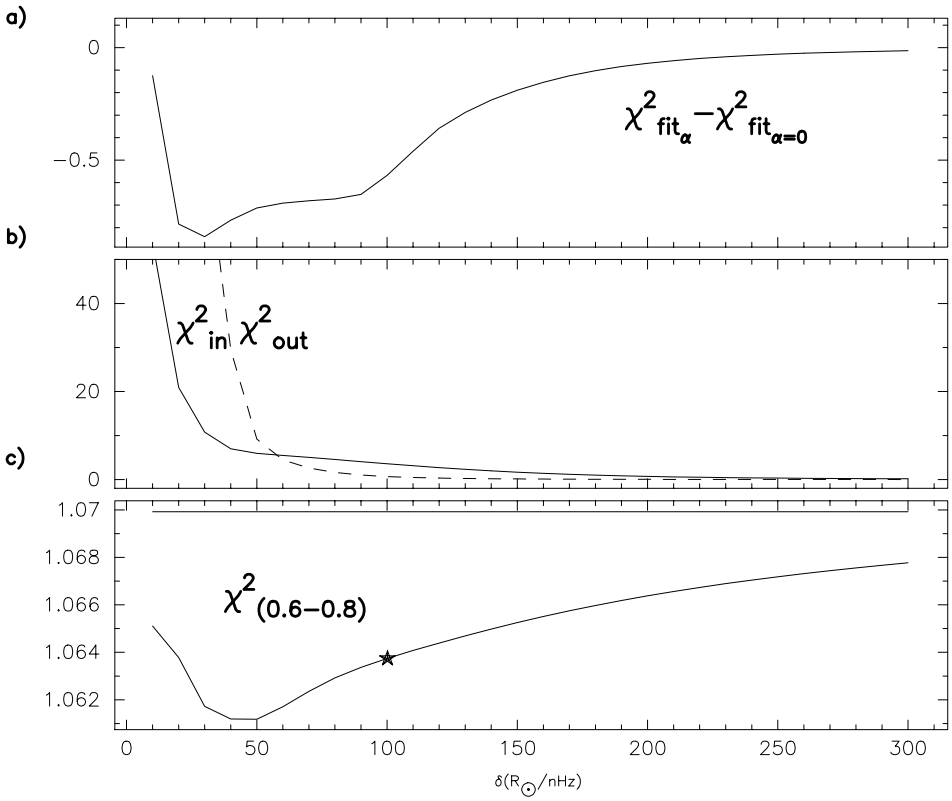

Fig. 7a–c. Variation of various χ^2 indicators with δ. **a** Difference between the goodness of the fits of ARTUR solutions by an erf function with or without a linear part (cf. Eq. (21)). **b** Difference between T-GCV and ARTUR solutions in and out of the zones where high gradients are expected (cf. Eq. (20)). **c** The normalized χ^2 value of ARTUR inversions for modes which have their turning points between $0.6R_\odot$ and $0.8R_\odot$. The horizontal line shows the value (1.07) obtained from T-GCV solution. The star symbol indicate the χ^2 value reached for the choice $\delta = 100R_\odot/nHz$ retained for the ARTUR algorithm.

good χ^2 value reached with this parameter. One of the objectives of the ARTUR method is to keep the same regularization as in Tikhonov method in zones without high gradients. Therefore one expects that the amount of change (compared to T-GCV solution) will be more important in the tachocline (and possibly near the surface) than in other zones. We denote by $\bar{\Omega}_A^\delta(r)$ the solution at the final step of ARTUR algorithm for a given δ, and by $\bar{\Omega}_T(r)$ the T-GCV solution; we then define the two quantities:

$$\chi^2_{\left|\begin{smallmatrix} in \\ out \end{smallmatrix}\right.} = \frac{1}{N_{\left|\begin{smallmatrix} in \\ out \end{smallmatrix}\right.}} \sum_{p \in I_{\left|\begin{smallmatrix} in \\ out \end{smallmatrix}\right.}} \left(\bar{\Omega}_A^\delta(r_p) - \bar{\Omega}_T(r_p)\right)^2 \qquad (20)$$

where $I_{in} = \{p \ / \ 0.6 < \frac{r_p}{R_\odot} \leq 0.8 \text{ or } r_p > 0.9R_\odot\}$, $I_{out} = [1, N_p] - I_{in}$ and N_{in}, N_{out} are the sizes of the two sets. Fig. 7b is a plot of these two quantities. It shows that for δ lower than $60R_\odot/nHz$ the ARTUR algorithm tends to alter the solution even in regions where no high gradients

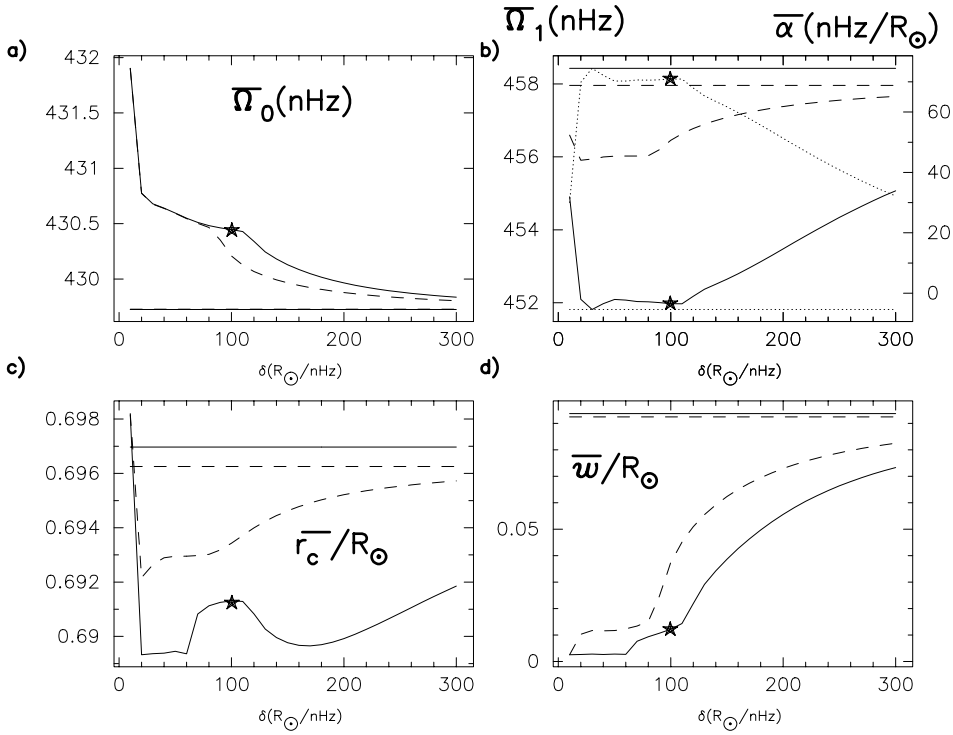

Fig. 8a–d. Variation with δ of the tachocline parameters as deduced by fitting ARTUR solutions by an erf function (cf. Eq. (1)). The dashed lines show the results obtained by searching only four parameters ($\bar{\Omega}_0$, $\bar{\Omega}_1$, $\bar{r}_c$, $\bar{w}$) after setting $\bar{\alpha} = 0$, whereas the solid lines show the results obtained when $\bar{\alpha}$ is a free parameter of the fit. In this latter case, the inferred value of $\bar{\alpha}$ is shown by the dotted line on panel b and the star symbols show the results obtained for $\delta = 100R_\odot/nHz$. The horizontal lines indicate the values (independent of δ) of the tachocline parameters obtained by fitting the T-GCV solution. These horizontal lines represent limits for high δ of the parameters inferred from ARTUR solutions. On panel d the horizontal lines indicate the width as inferred directly by the fit of T-GCV solution. The width corrected by a 'local deconvolution' using averaging kernels, is $w \simeq 0.06R_\odot$.

are expected, but that for values above $\delta \simeq 100R_\odot/nHz$ there is no longer any changes in these regions, as expected.

From these two criteria it seems that a choice of $\delta = 100R_\odot/nHz$ is a good compromise between minimizing the χ^2 value for modes which have their turning points near the tachocline, and operating changes essentially in zones where high gradients are expected. This was precisely the objective of the non linear regularization approach. The ARTUR solution is shown in Fig. 5 for this choice δ. The inferred tachocline parameters are summarize in Table 1 and compared to the results obtained by Corbard et al. (1998) for the same dataset. The estimates of the center of the tachocline and the rotation rate in the radiative interior obtained from ARTUR algorithm are in good agreement with our previous work. The inferred value of $\bar{\Omega}_1$ and $\bar{\alpha}$ are such that $\bar{\Omega}_1 + 0.1\bar{\alpha} \simeq 459$nHz, which correspond to the rotation rate inferred at $0.8R_\odot$ and to the value of the $\bar{\Omega}_1$ parameter obtained by the fit of the T-GCV solution. The tachocline width inferred from the ARTUR inversion ($w = 0.01R_\odot$) is smaller than our previous estimate but still compatible if we take into account an error of $0.02R_\odot$ as indicated by the Monte-Carlo simulations. The next section will give a more detailed discussion for the interpretation of this result by showing how the estimates of the width vary with the δ parameter, and during the ARTUR iterations.

6.3. On the inferred tachocline width

Fig. 8 shows how the parameters, inferred from a fit of the final step of the ARTUR algorithm are sensitive to the choice of δ. It shows that the relation $\bar{\Omega}_1 + 0.1\bar{\alpha} \simeq 459nHz$ is still valid for other choices of δ, and that $\bar{\Omega}_0$ and $\bar{r}_c$ vary only little with δ whereas the inferred width increases rapidly for $\delta > 100R_\odot/nHz$. It reaches $0.055R_\odot$ for $\delta = 200R_\odot/nHz$ which is the limit for the choice of δ above which we think that the ARTUR algorithm is no longer effective. For large values of δ, the number of step in the ARTUR process becomes very low, ARTUR solutions tend to T-GCV solution, and thus, tachocline parameters inferred from ARTUR solutions tend toward those inferred from T-GCV solution (cf. Fig. 8).

We have tried to fit the solutions with or without the linear part after the transition (i.e. by searching for the best $\bar{\alpha}$ coefficient or by setting $\bar{\alpha} = 0$ in Eq. (1)). The goodness of the fit is defined by:

$$\chi^2_{fit_{\bar{\alpha}}}(\bar{\Omega}) = \frac{1}{N_{fit} - 5} \sum_{p \in I_{fit}} \left(\bar{\Omega}(r_p) - \Omega_{fit}(r_p)\right)^2 \qquad (21)$$

where I_{fit} is the set of indices given by $I_{fit} \equiv \{p \ / \ 0.4 \leq \frac{r_p}{R_\odot} < 0.8\}$, and N_{fit} is the size of I_{fit}. We denote by $\chi^2_{fit_{\bar{\alpha}=0}}$ the goodness of a fit obtained with only four parameters ($\bar{\alpha} = 0$ in Eq. (1)). In this case the denominator of Eq. (21) becomes $N_{fit} - 4$. The two fits are almost equivalent when applied to the T-GCV solution ($\bar{\alpha} = -5$nHz when it is searched, cf. Fig. 8b), but the fit that allows a linear part has been found to be better suited for describing ARTUR solutions for all the choices of δ (cf. Fig. 7a). We note however that if one chooses to fit the solutions with $\bar{\alpha} = 0$ (i.e. to describe the tachocline by a simple erf function), then the inferred width would be systematically increased by a value up to $0.02R_\odot$ (cf. Fig. 8d).

As the increase of the χ^2 value between $\delta = 100R_\odot/nHz$ and $\delta = 200R_\odot/nHz$ is low (cf. Fig 7a), and because of our previous estimate of $\pm 0.02R_\odot$ for the uncertainty on the width, we cannot exclude a width larger than $0.01R_\odot$. Furthermore

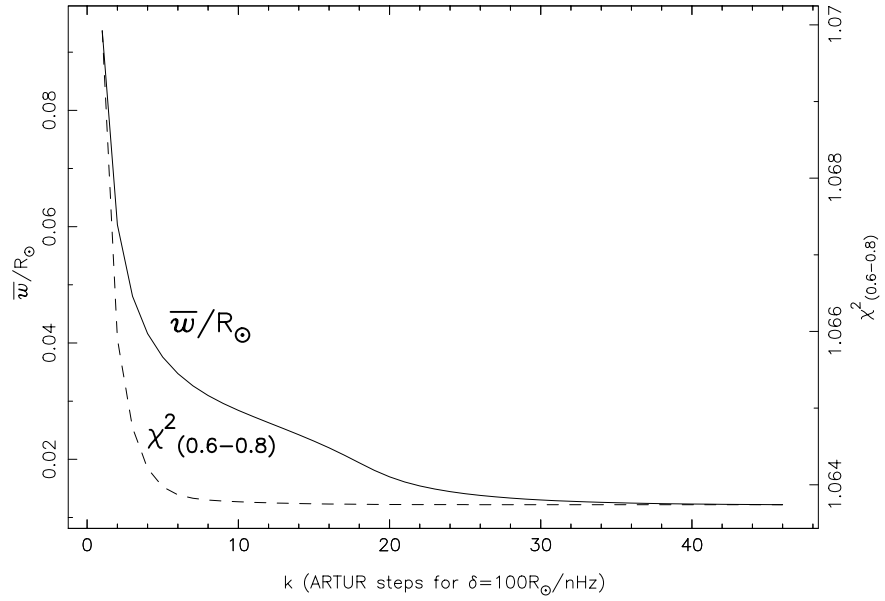

Fig. 9. Variation of the inferred width (solid line) and the χ^2 value for modes which have their turning points between $0.6R_\odot$ and $0.8R_\odot$ (dashed line) as a function of the iteration number in ARTUR process for $\delta = 100R_\odot/nHz$). The values found at the first step are equal to those obtained from the fit of the T-GCV solution shown as horizontal lines in Figs. 8d and 7c whereas the values obtained at the final step are shown by star graph markers on these plots.

Fig. 9 shows that the minimum value of the χ^2 is reached after only six ARTUR iterations, whereas the inferred width still decreases from $0.035R_\odot$ down to $0.01R_\odot$. This indicates that the data themselves do not allow us to choose between widths in that range. During the iterations as well as when we vary the value of δ, it is the amount of regularization introduced in high gradients zones that is changed. In that sense, stopping ARTUR iteration before its convergence according to the criterion Eq. (17) would be equivalent to increase the value of δ. As the needed amount of regularization is related to the level of noise contained in the data, the reliability of our result is strongly related to our knowledge of the data noise. Since we have shown that the GCV criterion may reveal some discrepancies between the data noise and the simulated noise, the only way to gain more confidence on the appropriate choice of δ, and thus on the result concerning the tachocline width inferred from this kind of non linear inversions, will be to increase our knowledge of the statistical properties of the data noise and to compare with other datasets. However, even the values of the width found at the final step for $\delta = 200R_\odot/nHz$ or at the sixth iteration with $\delta = 100R_\odot/nHz$ are still below the value of $0.06R_\odot$ inferred from T-GCV solution after a 'local deconvolution'. The use of nonlinear regularization argue in favor of a very sharp tachocline and even a discontinuity can not be excluded. Therefore our conclusion on the tachocline width is that it is very likely that it is less than $0.05R_\odot$. This is reinforce by the simulations (circles in Fig. 4) showing that initial widths up to $0.05R_\odot$ can lead to inferred widths of $0.01R_\odot$ whereas it is excluded (within 1σ error bars) for initial widths larger than $0.05R_\odot$.

7. Conclusions

This work introduces in helioseismic context an approach of the inverse problem that use an adaptive regularization which is then used toward on the estimation of the width of the tachocline. We do not claim that this paper gives a definitive answer to this question. Instead, this paper presents a new tool which allows to accurately reconstruct rotation profiles with possibly both smooth and abrupt variations with depth. This approach leads to a non linear problem that can be solved easily by an iterative process named ARTUR, initially developed in the field of image processing. It is shown that this allows to recover high gradients in the solution and avoids the spurious oscillations (known as 'Gibbs phenomena') that may be found when we try to recover such sharp transition zones with the usual Tikhonov approach.

The proposed procedure for choosing the regularizing parameters and the Monte-Carlo simulations represent a first step showing the feasibility and the capability of the method to retrieve both small and large tachocline widths from noisy observations. They have shown that this method, as well as the Tikhonov inversion with 'local deconvolution' is able to recover the width of erf functions with widths from to $0.03R_\odot$ up to $0.08R_\odot$ with the same error estimation of $\pm 0.02R_\odot$. For lower values of the width, the results are more sensitive to the choice of the regularizing parameters. However, some improvements may be brought by the studies on the optimal choice of the regularizing parameters that are underway and their future application to the rotation inverse problem.

The inversion of LOWL data using the ARTUR method gives an equatorial tachocline profile which differs from our previous work. These new results favor a sharp transition (down to a width of $0.01R_\odot$ with the adapted regularization parameters). From our study of the inferred width as a function of the parameter δ and from our estimate of the uncertainty on this parameter, we conclude that the width of the tachocline should be less than $0.05R_\odot$. This estimate is somewhat different from our previous estimate of $0.05 \pm 0.03R_\odot$ which allows relatively smooth transitions up to $0.08R_\odot$.

The change in our estimate of the width is partly due to the fact that we have changed the fitting function that defined the tachocline parameters. It is shown here that adding a lin-

ear behaviour in the upper part of the erf function allows a better fit of the solution. Whereas the T-GCV solution (before any deconvolution) can be well approximated by a simple erf function between $0.4R_\odot$ and $0.8R_\odot$, the ARTUR solution is better approximated, in the upper part of the tachocline, by a linear function with a slope around $70\mathrm{nHz}/R_\odot$ that goes from $452\mathrm{nHz}$ at $0.7R_\odot$ up to $459\mathrm{nHz}$ at $0.8R_\odot$. It was not possible to reach this conclusion from our previous approach using the T-GCV solution because the 'local deconvolution' used in this approach supposed explicitly that the rotation profile can be well approximated by a simple erf function (without linear behaviour). It will be therefore very interesting to study in future works and with other datasets the rotation profile just above the tachocline, in order to become more confident of our result.

An important contribution of the non linear regularization approach is that it allows to find directly a solution with sharp transitions without fixing a priori the shape of the transitions, as necessary in forward modeling or when we deconvolve a solution obtained from linear methods. In order to describe the tachocline with only few parameters, we can afterwards choose a shape for the fitting function that is adapted to the solution found.

Finally, we note that this approach to inversion with non linear regularization may find other applications in the helioseismic context, for any the problems where high gradients are expected in the solutions. We have shown that this is the case for the rotation of the surface layers, which can be studied more accurately with instruments such as MDI on board SoHO, or the GONG network which observe high degree modes. This algorithm may also be applied to the sound speed anomaly found between the real Sun and solar models by structural inversions. The width of this peak, which is located in the tachocline, can be related to the width of the mixed zone which is supposed to exist just below the convection zone (e.g. Morel et al. 1998). Some recent work (Elliott et al. 1998) have used a linear inversion which has been deconvolved to give an estimate of $0.02R_\odot$ for this width. It is interesting to note that this is of the same order as our estimate of the tachocline width deduced from the rotation profile. As the sound speed anomaly profile is also a zone with high gradients in the solution of an inverse problem, the use of non linear regularization may also be an alternate approach to address this problem in future work.

Acknowledgements. We gratefully thank S. Tomczyk and J. Schou for providing the LOWL data, and the referee for constructive remarks. This work has been performed using the computing facilities provided by the program "Simulations Interactives et Visualisation en Astronomie et Mécanique" (SIVAM, OCA, Nice) and by the "Institut du Développement et des Ressources en Informatique Scientifique" (IDRIS, Orsay).

Appendix A: terms of the discretization

In our application, the polynomial expansion Eq. (4) is such that:

$$\psi_p(r) = \begin{cases} \frac{r-r_{p-1}}{r_p-r_{p-1}} & \text{if } r_{p-1} < r \leq r_p \\ \frac{r_{p+1}-r}{r_{p+1}-r_p} & \text{if } r_p < r \leq r_{p+1} \\ 0 & \text{otherwise} \end{cases} \quad \text{(A1)}$$

where $(r_p)_{p=0,N_p+1}$ are the fixed break points distributed according to the density of turning points of modes. We have:

$$0 = r_0 = r_1 < r_2 < ..r_{N_p-1} < r_{N_p} = r_{N_p+1} = R_\odot, \quad \text{(A2)}$$

therefore, each coefficient ω_p of the expansion Eq. (4) simply represents the solution at the radius r_p:

$$\forall\, p = 1, ..N_p \;\; \bar{\Omega}(r_p) = \omega_p. \quad \text{(A3)}$$

Furthermore, with this expansion of the solution, the first derivative of $\bar{\Omega}(r)$ is represented by a piecewise constant function. Therefore, in this trivial case, one can take in Eq. (10):

$$c_p = r_{p+1} - r_p, \quad p = 1, ..N_p - 1 \qquad \boldsymbol{C} \equiv diag(c_p) \quad \text{(A4)}$$

and the first derivative operator is defined by the bi-diagonal matrix:

$$\boldsymbol{L} = \boldsymbol{C}^{-1}\boldsymbol{L}^{(1)} \quad \text{(A5)}$$

with:

$$\boldsymbol{L}^{(1)}_{i,j} = -\delta_{i,j} + \delta_{i,j-1} \begin{cases} i = 1, N_p - 1 \\ j = 1, N_p \end{cases}, \quad \text{(A6)}$$

References

Acar R., Vogel C.R., 1994, Inverse Problems 10(6), 1217
Antia H.M., Basu S., Chitre S.M., 1998, MNRAS 298, 543
Barrett R., Berry M., Chan T.F., et al., 1994, Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods. SIAM, Philadelphia, PA
Basu S., 1997, MNRAS 288, 572
Blanc-Féraud L., 1998, Problèmes inverses mal posés. In: Schatzman E. (ed.) Cours de Structure Interne: Génération et Interprétation des Oscillations Sollaires. Obs. de Paris, Meudon, France
Blanc-Féraud L., Charbonnier P., Aubert G., Barlaud M., 1995, Nonlinear image processing: Modeling and fast algorithm for regularization with edge detection. In: IEEE Proceedings of the 2nd International Conference of Image Processing, Washington DC, USA, p. 474
Chan K., Gray A., 1996, Statistics and Computing 6, 367
Charbonneau P., Christensen-Dalsgaard J., Henning R., et al., 1998, Observational constraints on the dynamical properties of the shear layer at the base of the solar convection zone. In: Provost J., Schmider F.X. (eds.) Sounding Solar and Stellar Interior. OCA & UNSA, Nice, p. 161
Charbonnier P., Aubert G., Blanc-Féraud L., Barlaud M., 1994, Two deterministic half-quadratic regularization algorithms for computed imaging. In: IEEE Proceedings of the 1st International Conference of Image Processing. Austin, USA, p. 168
Charbonnier P., Blanc-Féraud L., Aubert G., Barlaud M., 1997, IEEE Trans. on Image Processing 6(2), 298
Corbard T., 1997, Inferring the internal rotation of the sun from low degree helioseismology. In: Cacciani A. (ed.) Proceedings of IX IRIS meeting
Corbard T., Berthomieu G., Morel P., et al., 1997, A&A 324, 298
Corbard T., Berthomieu G., Provost J., Morel P., 1998, A&A 330, 1149
Craig I.J.D., Brown J.C., 1986, Inverse Problems in Astronomy: a guide to inversion strategies for remotely sensed data. Adam Hilger Ltd, Bristol
Dobson D., Santosa F., 1994, Inverse Problems 10, 317

Elliott J.R., 1997, A&A 327, 1222
Elliott J.R., Gough D.O., Sekii T., 1998, Helioseismic determination of the solar tachocline thickness. In: Korzennik S.G., Wilson A. (eds.) Structure and Dynamics of the Interior of the Sun and Sun-like Stars. ESA Publications Division, Noordwijk, The Netherlands, p. 763
Galatsanos N., Katsaggelos A., 1991, Cross-Validation and other criteria for estimating the regularizing parameter. In: IEEE Proceedings of ICASSP, p. 3021
Geman S., McClure D.E., 1985, Baysian image analysis: an application to single photon emission tomography. In: Assoc. A.S. (ed.) Proc. atist. Comput. Sect. , Washington DC, USA, p. 12
Geman S., Reynolds G., 1992, IEEE Trans. Patern Analysis and Machine Intelligence 14, 367
Golub G.H., Van-Loan C.F., 1989, Matrix Computations (2nd edition), Chapt. 10, The Johns Hopkins University Press, Baltimore, USA
Gough D.O., Sekii T., 1998, On the solar tachocline. In: Provost J., Schmider F.X. (eds.) Sounding Solar and Stellar interior. OCA & UNSA, Nice, p. 93
Gough D.O., Thompson M.J., 1991, The inverse problem. In: Cox A.N., Livingston W.C., Matthews M. (eds.) Solar Interior and Atmosphere. The University of Arizona Press, Tucson, p. 519
Green P.J., 1990, IEEE Trans. on Medical Imaging MI-9, 1, 84
Hadamard J., 1923, Lectures on the Cauchy Problem in linear Partial Differential Equation. Yale University Press, New Haven
Hansen C.J., Cox J.P., Van-Horn H.M., 1977, ApJ 217, 151
Hansen P.C., 1992, SIAM Review 34, 561
Hansen P.C., Mosegaard K., 1996, Numerical Linear Algebra with Applications 3(6), 513
Hansen P.C., Sekii T., Sibahashi H., 1992, SIAM J. Sci. Stat. Comput. 13, 1142
Hebert T., Leahy R., 1989, IEEE Trans. on Medical Imaging MI-8, 2, 194
Jalobeanu A., Blanc-Féraud L., Zerubia J., 1998, Estimation d'hyperparametres pour la restauration d'images satellitaires par une méthode MCMCML. Technical Report 3469, Inria
Kirsch A., 1996, An Introduction to the Mathematical Theory of Inverse Problems. Springer-Verlag, New York
Kosovichev A.G., 1996, ApJ 469, L61
Kosovichev A.G., 1998, Seismic obseervation of solar tachocline. In: Provost J., Schmider F.X. (eds.) Sounding Solar and Stellar Interior. OCA & UNSA, Nice, p. 97
Lakshmanan S., Derin H., 1989, IEEE Transations on Pattern Analysis and Machine Intelligence 11(8), 322
Morel P., Provost J., Berthomieu G., 1997, A&A 327, 349
Morel P., Provost J., Berthomieu G., 1998, How solar models fit the SoHO observations. In: Korzennik S.G., Wilson A. (eds.) Structure and Dynamics of the Interior of the Sun and Sun-like Stars. ESA Publications Division, Noordwijk, The Netherlands
Perona P., Malik J., 1990, IEEE Trans. on Pattern Analysis and Machine Intelligence PAMI-12, 7, 629
Phillips D.L., 1962, J. Assoc. Comput. Mach. 9(1), 84
Rudin L.I., Osher S., Fatemi E., 1992, Physica D 60, 259
Schou J., Christensen-Dalsgaard J., Thompson M.J., 1994, ApJ 433, 389
Schou J., Antia H.M., Basu S., et al., 1998, ApJ 505, 390
Shibahashi H., Sekii T., 1988, An inversion method based on the moore-penrose generalized inverse matrix. In: Rolfe E.J. (ed.) Seismology of the Sun and Sun-Like Stars. ESA Publication Division, Noordwijk, the Netherlands, p. 521
Spiegel E.A., Zahn J.P., 1992, A&A 265, 106
Teboul S., Blanc-Féraud L., Aubert G., Barlaud M., 1998, IEEE Trans. on Image Processing 7(3)
Thompson A., Brown J., J.W.Kay, Titterington D., 1991, IEEE Transations on Pattern Recognition and Machine Intelligence 13(4), 326
Thompson M.J., Toomre J., Anderson E.R., et al., 1996, Science 272, 1300
Tikhonov A.N., 1963, Sov. Maths. Doklady 4, 1035
Tomczyk S., Streander K., Card G., et al., 1995, sp 159, 1
Twomey S., 1963, J. Assoc. Comput. Mach. 10(1), 97
Vogel C.R., Oman M.E., 1996, SIAM Journal of Scientific Computing 17(1)
Vogel C.R., Oman M.E., 1997, IEEE Trans. on Image Processing, submitted
Wahba G., 1977, SIAM J. Numer. Anal. 14, 651
Zerubia J., Khoumri M., Blanc-Féraud L., 1998, Hyperparameter estimation of a variational model using a stochastic gradient method. In: SPIE's international Symposium on Optical Science, Engineering, and instrumentation - in the part dedicated to Bayesian Inference for Inverse Problem. San Diego, USA