Fast and sparse inversion of magnetic data with truncated generalized cross validation

Gholam Abbas Fanaee Kheirabad\textsuperscript{a,}\textsuperscript{∗}, Saeed Vatankhah\textsuperscript{b}, Rosemary A. Renaut\textsuperscript{c}

\textsuperscript{a}Birjand University of Technology, Birjand, Iran
\textsuperscript{b}Institute of Geophysics, University of Tehran, Tehran, Iran
\textsuperscript{c}School of Mathematical and Statistical Science, Arizona State University, Tempe, AZ, USA

Abstract

In this paper a fast method for sparse inversion of magnetic data is considered. The $L_1$-norm stabilizer is used to generate models with sharp and distinct interfaces. To deal with the non-linearity introduced by the $L_1$-norm, a model-space iteratively reweighted least squares algorithm is used. The original model matrix is factorized using the Golub-Kahan bidiagonalization and provides the projection of the problem onto a Krylov subspace of significantly reduced dimension. The model matrix of the projected system inherits the ill-conditioning of the original matrix but the spectrum of the projected system accurately captures only a portion of the full spectrum. Equipped with the singular value decomposition of the projected system matrix, the solution of the projected problem is expressed using a filtered singular value expansion. This expansion depends on a regularization parameter which is determined using the method of generalized cross validation for the truncated

\textsuperscript{∗}Corresponding author

Email addresses: afanaee@birjandut.ac.ir (Gholam Abbas Fanaee Kheirabad), svatan@ut.ac.ir (Saeed Vatankhah), renaut@asu.edu (Rosemary A. Renaut)

Preprint submitted to Elsevier July 22, 2016
spectrum. This new technique based on truncation is effective especially for high noise levels as compared with the standard method of generalized cross validation and a weighted version that has been previously suggested for solution of regularized problems at the subspace level. Numerical results for synthetic examples and real field data demonstrate the efficiency of the presented algorithm.

Keywords: magnetic survey, inverse problems, Golub-Kahan bidiagonalization, regularization parameter estimation, truncated generalized cross validation
1. Introduction

The inversion of magnetic field data presents many practical difficulties. Most importantly the solution of the problem is not unique. While the physics of the problem introduces inherent ambiguity in the solution, algebraic ambiguity arises because the number of data measurements is considerably smaller than the number of unknown model parameters for the subsurface discretization, (Li & Oldenburg, 1996; Pilkington, 1997). The determination of the solution is further complicated by the contamination of the measured data by noise which is amplified in the inversion process due to the ill-conditioning of the model matrix. Therefore, in order to estimate a stable and geologically meaningful solution it is necessary to both include additional constraints on the model and to impose regularization, (Portniaguine & Zhdanov, 1999). Finally, seeking a three dimensional model of the subsurface presents a computational challenge and powerful algorithms are needed to reduce both memory and CPU requirements of the process, (Li & Oldenburg, 2003; Pilkington, 2009; Portniaguine & Zhdanov, 2002; Boulanger & Chouteau, 2001). This paper addresses these issues through the use of an appropriately constrained edge-preserving regularization implemented through a computationally efficient iterative algorithm for finding the solution on a relevant subspace, and extends an approach introduced for the solution of the gravity inverse problem Vatankhah et al. (2016).

For the inversion of magnetic data a weighted $L_2$ error between the observed and predicted data due to the inverse solution is typically used to measure the misfit of the solution that is obtained from the minimization of an objective function consisting of this data misfit and a stabilizing term,
The general inversion methodology developed by Li & Oldenburg (1996) employs stabilization by $L_2$ regularization with respect to the low order finite difference approximation of the derivatives of the model parameters in each of the three orthogonal directions. The approach has been successfully applied in a wide range of the geophysical applications but yields relatively smooth models that are not always consistent with real geological structure (Farquharson, 2008). On the other hand, stabilization with a minimum support (MS) or minimum gradient support (MGS) function yields models that provide better contrast (Last & Kubic, 1983; Portniaguine & Zhdanov, 1999, 2002; Vatankhah et al., 2014a, 2015). Sharp and focused images of the subsurface are also achieved using the $L_1$-norm stabilizer (Loke et al., 2003; Farquharson, 2008; Vatankhah et al., 2016) and with the sparseness constraint introduced by the use of the Cauchy norm applied to the model parameters (Pilkington, 2009). Recently, a method for adaptively recovering both smooth and blocky models was developed by Sun & Li (2014). Here we use the approximate $L_1$-norm stabilization which is implemented using an iteratively reweighted least squares (IRLS) algorithm and was studied for gravity data inversion in Vatankhah et al. (2016). The results are contrasted with those obtained using the Cauchy norm for the real data studied in Pilkington (2009).

Due to the aim to provide an algorithm which is effective for large problems it is important to use a computationally efficient approach. The algorithm in Pilkington (2009) adopts a conjugate gradient algorithm for efficient solution of the underlying large scale systems of equations, as is adopted also
in (Li & Oldenburg, 2003; Pilkington, 1997). An alternative, chosen here, is to use the Golub-Kahan bidiagonalization (sometimes called Lanczos bidiagonalization) of the model matrix appearing in the data misfit term as discussed in (Vatankhah et al., 2016; Abedi et al., 2013), and which leads to a hybrid LSQR algorithm for the solution of the stabilized problem, Chung et al. (2008); Kilmer & O’Leary (2001). An extensive discussion of the multiplicity of techniques for finding the regularization parameter in the objective function, and which provides the trade off between the data misfit and the stabilizing terms, is provided in Vatankhah et al. (2016) and points to many of the key references. This discussion is therefore not reproduced here. Rather, we emphasize that the method developed here is based on the extension of the method of generalized cross validation (GCV) for use with the hybrid LSQR algorithm. Unlike in Abedi et al. (2013) we suggest the use of the GCV for the truncated spectrum of the projected system rather than the weighted GCV (WGCV) Chung et al. (2008). This further develops the approach in Vatankhah et al. (2016) which adopted a truncation for the unbiased predictive risk estimator (UPRE), but requires at the same time knowledge of the underlying noise distribution in the measured data that is not needed for the GCV related techniques. We use TGCV to denote the truncated GCV technique developed here. It is particularly appropriate for the inversion of magnetic data because the underlying model system is only mildly ill-posed, rather than severely ill-posed, as will be elaborated on in the discussion and results.

The remainder of the paper is organized as follows. In section 2 we describe the inversion methodology based on the $L_1$-norm stabilizer. The
solution of the inverse problem using the hybrid LSQR algorithm and methods for estimating the regularization parameter with respect to the Krylov subspace are discussed in sections 2.1 and 2.2, respectively. Furthermore, the application of the SVD at the subspace level for both the inverse solution and parameter-choice methods is also shown in section 2.2. In section 3 we show the results of the presented algorithm on two different synthetic models. The dipping dike discussed in section 3.1 is used in section 3.1.2 to demonstrate the impact of using the hybrid LSQR algorithm for a problem which is only mildly to moderately ill-posed. To further examine the approach we consider a model with multiple bodies in section 3.2. The aeromagnetic data over a portion of the Wuskwatim Lake region, Manitoba, Canada is inverted and contrasted with the results presented in Pilkington (2009) in section 4. Conclusions and future work are given in section 5.

2. Inversion methodology

We divide the subsurface volume into a large number of cubes of fixed size but unknown susceptibility. This allows the maximum flexibility for the model to represent the subsurface structures (Li & Oldenburg, 1996; Boulanger & Chouteau, 2001). For the inversion methodology presented here, we assume that there is no remanent magnetization and only the induced magnetization is considered. This magnetization is uniform within each cube and is given by the product of the susceptibility, $\kappa$, and inducing geomagnetic field (Li & Oldenburg, 1996). Suppose the susceptibilities of the cubes are collected in vector $\mathbf{m} = (\kappa_1, \kappa_2, \ldots, \kappa_n)^T$ where $n$ is the total number of cubes, and that $\mathbf{d}_{\text{obs}} \in \mathcal{R}^m$ contains the measured total field data. The
relationship between the observed data and model parameters is given by

\[ \mathbf{d}_{\text{obs}} = G \mathbf{m}, \]  

(1)

in which the sensitivity matrix \( G \in \mathbb{R}^{m \times n}, m \ll n \), has elements \( g_{ij} \) which represent the effects of unit susceptibility in the \( j \)th cell on data location \( i \). A simplified and computationally fast formula for evaluating the analytically given total field of a cube was developed by Rao & Babu (1991), and is used here to form the elements of the sensitivity matrix. Thus, given \( G \) and \( \mathbf{d}_{\text{obs}} \), the goal of inversion is to find a stable and geologically meaningful susceptibility model that reproduces the observed data at the noise level.

As noted in section 1, problem (1) is ill-posed and an acceptable solution may be found with a \( L_1 \) stabilizing regularization following the method used in Vatankhah et al. (2016). The solution of

\[ \Phi(\mathbf{m}, \alpha) = \| W_{d}(G \mathbf{m} - \mathbf{d}_{\text{obs}}) \|_2^2 + \alpha^2 \| (\mathbf{m} - \mathbf{m}_{\text{apr}}) \|_1 \]  

(2)

\[ \approx \| W_{d}(G \mathbf{m} - \mathbf{d}_{\text{obs}}) \|_2^2 + \alpha^2 \| W_{L_1}(\mathbf{m} - \mathbf{m}_{\text{apr}}) \|_2^2, \]

is determined. Here \( \mathbf{m}_{\text{apr}} \) is a vector of known reference information on the parameters, perhaps estimated from prior geological or geophysical investigation or can be taken to be 0, (Li & Oldenburg, 1996). Weighting matrix \( W_d \) is square and is diagonal with elements \( 1/\sigma_i, \ i = 1 : m, \) where \( \sigma_i \) is the standard deviation of the noise in the \( i \)th datum when the noise in \( \mathbf{d}_{\text{obs}} \) is uncorrelated and Gaussian. Following Farquharson & Oldenburg (2004), we assume that the absolute magnitudes of the errors are unknown, but that the relative magnitudes can be estimated. Then, for known \( \bar{\sigma}_i, \ i = 1 : m, \) \( \sigma_i = \sigma_0 \bar{\sigma}_i, \) for unknown constant \( \sigma_0 > 0. \) The \( L_p \)-norm of vector \( \mathbf{x} \) in (2) given by \( \| \mathbf{x} \|_p = (\sum_{i=1}^{n} |x_i|^p)^{1/p}, p \geq 1 \) can be approximated using the given
The $L_2$ norm with $W_{L_1} = \text{diag}(((m - m_{\text{spr}})^2 + \epsilon^2)^{-1/4})$ for small $\epsilon > 0$, (Wohlberg & Rodriguez, 2007; Voronin, 2012; Vatankhah et al., 2016). Then the sparse model that solves (2) is approximated by iteratively finding the solutions $m^{(k)}$ of the Tikhonov functions

$$
\Phi^{(k)}(m^{(k)}, \alpha^{(k)}) = \|W_d(Gm^{(k)} - d_{\text{obs}})\|_2^2 + (\alpha^{(k)})^2\|W_{L_1}^{(k)}(m^{(k)} - m^{(k-1)})\|_2^2,
$$

for $k = 1, \ldots$, initialized with $m^{(0)} = m_{\text{spr}}$, and $W_{L_1}^{(1)} = I_n$. Model weighting matrix is updated as $W_{L_1}^{(k)} = \text{diag}(((m^{(k-1)} - m^{(k-2)})^2 + \epsilon^2)^{-1/4})$, $k > 1$, and regularization parameter $\alpha^{(k)}$ is explicitly dependent on $k$. To counteract the natural decay of the kernel, the diagonal depth weighting matrix, $W_{\text{depth}} = 1/z_j^\beta$, is incorporated into (3) replacing $W_{L_1}^{(k)}$ by $W^{(k)} = W_{L_1}^{(k)}W_{\text{depth}}$, see Li & Oldenburg (1996); Pilkington (1997). Here, $z_j$ is the mean depth of the cell $j$ and $\beta$ is a weighting parameter. Theoretically, $m^{(k)}$ is obtained as the solution of the normal equations for (3), e.g. Golub & Van Loan (1996), which as detailed in Vatankhah et al. (2016) leads to the expression

$$
m^{(k)} = m^{(k-1)} + (W^{(k)})^{-1}(\tilde{G}^T\tilde{G} + (\alpha^{(k)})^2I_n)^{-1}\tilde{G}^T\tilde{r}^{(k)}
= m^{(k-1)} + (W^{(k)})^{-1}h^{(k)},
$$

for left and right preconditioned matrix $\tilde{G} = W_dG(W^{(k)})^{-1}$, updated weighted residual $\tilde{r}^{(k)} = W_d(d_{\text{obs}} - Gm^{(k-1)})$, and update $h^{(k)}$ solves the regularized problem

$$
\Phi^{(k)}(h^{(k)}, \alpha^{(k)}) = \|\tilde{G}h^{(k)} - \tilde{r}^{(k)}\|_2^2 + (\alpha^{(k)})^2\|h^{(k)}\|_2^2.
$$

Practically, at each iteration susceptibility limits $\kappa_{\text{min}}$ and $\kappa_{\text{max}}$ may be used to reduce the non-uniqueness of the problem by projecting values of
\(m^{(k)}\) that lie outside \([κ_{\min}, κ_{\max}]\) to the closest limit. The iteration terminates when either the solution satisfies the noise level, \(\|(d_{\text{obs}})_i - (d_{\text{pre}})_i\|_2^2 / σ_i^2 \leq m + \sqrt{2m},\) or a predefined maximum number of iterations, \(K_{\text{max}}\), is reached.

2.1. The LSQR algorithm

Computationally we use the LSQR algorithm to find the solution of (5), at each iteration \(k\), (Paige & Saunders, 1982a,b), in which the Golub-Kahan bidiagonalization process is applied to \(\tilde{G}\) yielding a projection of the problem to a space of much smaller dimension. Given \(\tilde{G}\) and vector \(\tilde{r}\), \(t\) steps of the process, initialized with \(H_{t+1}e_{t+1} = \tilde{r} / ∥\tilde{r}∥_2\), yields factorization

\[
\tilde{G}A_t = H_{t+1}B_t. \tag{6}
\]

Matrix \(B_t ∈ \mathcal{R}^{(t+1)×t}\) is bidiagonal, \(e_{t+1}\) is the unit vector of length \(t + 1\) and matrices \(H_{t+1} ∈ \mathcal{R}^{m×(t+1)}\) and \(A_t ∈ \mathcal{R}^{n×t}\) are column orthonormal, see e.g. Hansen (2007) and Kilmer & O’Leary (2001) for details. Defining the global solution \(h_t = A_tz_t\) with respect to the \(t\)–dimensional Krylov subspace spanned by the \(t\) columns of \(A_t\), where \(z_t\) is a vector of length \(t\), and noting the column orthonormality of \(H_{t+1}\), the data misfit term \(∥\tilde{G}h - \tilde{r}∥_2^2\) in (5) is replaced by a data misfit with respect to the projected solution

\[
∥B_tz_t - c∥_2^2, \tag{7}
\]

for \(c = ∥\tilde{r}∥_2e_{t+1}\), (Kilmer & O’Leary, 2001; Chung et al., 2008; Gazzola & Nagy, 2014). Furthermore, by the column orthonormality of \(A_t\), \(∥h_t∥_2^2 = ∥z_t∥^2\) and then (5) is replaced by the regularized projected problem, (Vatankhah et al., 2016),

\[
Φ(z_t, α^{(k)}) = ∥B_tz_t - c∥_2^2 + (α^{(k)})^2∥z_t∥_2^2. \tag{8}
\]
Under the assumption that \( t \ll \min(m,n) \), the cost of finding the solution of (8) is more efficient than solving (5), and update (4) is replaced by

\[
m_t^{(k)}(\alpha^{(k)}) = m_t^{(k-1)}(\alpha^{(k-1)}) + (W_t^{(k)})^{-1}A_t^{(k)}z_t^{(k)}(\alpha^{(k)}),
\]

for projected problem of size \( t \) at iteration \( k \). Note that both \( \tilde{G} \) and \( \tilde{r} \), and then the factorization (6), depend on \( k \).

It is clear for given \( k \) that the projected solution \( z_t(\alpha) \) depends on both subspace size \( t \) and regularization parameter \( \alpha \). As in Vatankhah et al. (2016) the focus here is the estimation of an optimal \( \alpha \) for given \( t \) and \( k \), and depends on examination of the ill-posedness of the specific problem (1). More details on finding an optimal \( t \) can be found in Renaut et al. (2015). The method described here allows the user to choose small \( t \) and hence is very effective in reducing the computational time.

2.2. Regularization parameter estimation

Whether solving (5), or (8) for fixed \( t \), an approach for determining the regularization parameter \( \alpha^{(k)} \) is required. Amongst the many possibilities, the method of generalized cross validation (GCV) has received much attention in conjunction with the solution of the projected problem, Kilmer & O’Leary (2001); Chung et al. (2008); Renaut et al. (2015). GCV is a statistical technique that does not depend on prior knowledge about the noise in data, rather it seeks to extract information from the observations. It is based on the leave one out principle that if a measurement is removed from the data set, then the corresponding regularized solution should provide a good estimate for the removed measurement (Golub et al., 1979).
As applied to (8), the GCV function to be minimized is given by

\[
G(\alpha, t, \omega) := \frac{\| (B_t B_t(\alpha) - I_{t+1})c \|_2^2}{\text{trace}(I_{t+1} - \omega B_t B_t(\alpha))^2}.
\]  

(9)

where \( B_t(\alpha) = (B_t^T B_t + \alpha^2 I_t)^{-1} B_t^T \) and \( \omega = 1 \). It has been demonstrated, however, that using \( \alpha_{\text{opt}} = \arg \min_{\alpha} \{G(\alpha)\} \) may lead to solutions which are over-smoothed when applied for the projected problem, Chung et al. (2008), and then the weight parameter \( 0 < \omega < 1 \) is required, denoted as the WGCV. Indeed, this approach was successfully used in Abedi et al. (2013) for the LSQR solution of the magnetic data inversion problem. Still, as already discussed in Chung et al. (2008), the method does require an argument for determining \( \omega \). Although an approach for estimating \( \omega \) was discussed in Renaut et al. (2015), particularly in relation to the connection between the problems (5) and (8), here we will further examine the GCV for the solution of (8) in relation to the degree of ill-posedness of the underlying problem. This discussion is motivated by recent results on the regularizing properties of the LSQR iteration presented in Huang & Jia (2016) and validates the suggestion to use the GCV for a truncated approximation to \( B_t \). In the following sections we will show, that although the WGCV may lead to acceptable results with low noise levels, the solutions deteriorate with increasing noise levels. In contrast the truncation-based GCV, Truncated GCV (TGCV), works very well especially for high noise levels.

To obtain the TGCV, we use the singular value decomposition of \( B_t \), given by \( B_t = \sum_{i=1}^{t'} \gamma_i u_i v_i^T \), (Golub & Van Loan, 1996) where the singular values are ordered \( \gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_t > 0 \) and \( u_i \) and \( v_i \) are the columns of orthonormal matrices \( U \in \mathbb{R}^{(t+1) \times (t+1)} \) and \( V \in \mathbb{R}^{t \times t} \), respectively. Introducing the filter factors \( f_i(\alpha) = \gamma_i^2 / (\alpha^2 + \gamma_i^2) \), the solution of (8) is given
by
\[ z_t(\alpha) = \sum_{i=1}^{t} f_i(\alpha) \frac{u_i^T c}{\gamma_i} v_i, \]  \hspace{1cm} (10)
and the WGCV by
\[ G(\alpha, t, \omega) = \sum_{i=1}^{t} (1 - f_i(\alpha))^2 \left( u_i^T c \right)^2 + \sum_{i=t+1}^{t+1} \left( u_i^T c \right)^2 \] \[ \frac{1 + t - \omega \sum_{i=1}^{t} f_i(\alpha))^2}{(1 + t - \omega \sum_{i=1}^{t} f_i(\alpha))^2}. \] \hspace{1cm} (11)

Now, replacing \( B_t \) in (9) by the best approximation of rank \( t' \), for \( t' < t \), \( B_t \approx \sum_{i=1}^{t'} \gamma_i u_i v_i^T \), as measured in the 2–norm, Golub & Van Loan (1996), yields the TGCV simply by replacing \( t \) in (11) by \( t' \), and taking \( \omega = 1 \). Given \( \alpha \) as the optimum for the TGCV, the solution is reconstructed using (10), specifically with \( t \) and not \( t' \). Thus the TGCV method does not correspond to taking the filtered TSVD solution with \( f_i(\alpha) = 0 \) for \( i > t' \). Instead \( \alpha \) is chosen to appropriately regularize the first \( t' \) terms, whereas choosing \( \alpha \) to regularize \( t \) terms will lead to a larger \( \alpha \) to handle the additional terms, so that the dominant terms are then over smoothed.

In the following we will show that in some cases the GCV and WGCV do not provide \( \alpha_{opt} \) which is as effective as that produced using TGCV. In contrast to WGCV which requires the choice of \( \omega \), for TGCV we have to find \( t' \). For magnetic inversion we will show by examination of the spectrum of \( B_t \) in relation to that of \( \tilde{G} \), that \( .7t \leq t' \leq t \) is suitable. This heuristic was used in Vatankhah et al. (2016) for the Truncated UPRE parameter-choice method. The steps to find the solution of the magnetic data problem are summarized in Algorithm 1, given in Appendix A, as also given in Vatankhah et al. (2016).
3. Synthetic examples

In this section we validate the strategies described in this paper using two different models, a simple dipping dike and a model of multiple bodies. The small-scale dipping dike model is used here to show the results of the inversion using the different GCV parameter-choice methods. This model allows us to show the singular values of both $\tilde{G}$ and $B_t$, for different $t$ and different noise levels, and hence demonstrate that the problem is only mildly to moderately ill-posed. Based on the heuristics that are developed through the examination of the small scale problem we find a comprehensive framework for the inversion of magnetic data using Algorithm 1 which we then verify for the solution of the large-scale multiple bodies model. All tests here were performed on a desktop computer, Intel Core i7-4790 CPU 3.6 GH.

3.1. Dipping dike

For the first model, we use a dipping dike with susceptibility 0.06 (in SI units) embedded in a nonsusceptible background, Fig. 1a, as used in (Li & Oldenburg, 1996, Figure 1). We generate the exact data, $d_{\text{exact}}$, on the surface over a $21 \times 21 = 441$ regular grid with 50 m grid spacing. The intensity of the magnetic field is 50000 nT, and inclination and declination are $I = 75^\circ$ and $D = 25^\circ$, respectively. Error contaminated data $d_{\text{obs}}$ are generated at three different noise levels, as given in Table 1. Fig. 1b shows noise-contaminated data for the second noise level $N2$. 
Table 1: Generation of contaminated data for testing the algorithms using \((\mathbf{d}_{\text{obs}})_i = (\mathbf{d}_{\text{exact}})_i + (\tau_1 (\mathbf{d}_{\text{exact}})_i + \tau_2 \|\mathbf{d}_{\text{exact}}\|)\) and the given parameter choices.

<table>
<thead>
<tr>
<th>Noise Level</th>
<th>N1</th>
<th>N2</th>
<th>N3</th>
</tr>
</thead>
<tbody>
<tr>
<td>((\tau_1, \tau_2))</td>
<td>(0.01, 0.001)</td>
<td>(0.02, 0.002)</td>
<td>(0.03, 0.01)</td>
</tr>
</tbody>
</table>

Figure 1: (a) Model of dipping dike with susceptibility equal to 0.06 (in SI units) embedded in a nonsusceptible background, cross section at Northing= 550 m. (b) The total field anomaly due to the model, contaminated with noise \(N_2\). The intensity of the magnetic field is 50000 nT, and inclination and declination are \(I = 75^\circ\) and \(D = 25^\circ\), respectively.

The parameters using Algorithm 1 for the inversion of the dipping dike model with the uniform subsurface discretization \(21 \times 21 \times 10\) are detailed in Table 2, where we use the notation \(\mu(\gamma_i)\) to denote the mean of the spectral values \(\gamma_i > 0\), and we note that picking a large initial value for \(\alpha^{(1)}\) is already justified in Farquharson & Oldenburg (2004) and Vatankhah et al. (2015). For comparison the parameters used for the synthetic multiple bodies model discussed in section 3.2 and the field data discussed in section 4 are also given.
Table 2: Parameters used for the Algorithm 1 for the solution of the dipping dike model, (DD), the multiple bodies model (MB), and the real field data (FD).

<table>
<thead>
<tr>
<th></th>
<th>$(m, n)$</th>
<th>$m_{apr}$</th>
<th>$(\kappa_{\text{min}}, \kappa_{\text{max}})$</th>
<th>$\epsilon^2$</th>
<th>$K_{\text{max}}$</th>
<th>$\alpha^{(1)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DD</td>
<td>(441, 4410)</td>
<td>0</td>
<td>$(0, 0.06)$</td>
<td>$1e^{-9}$</td>
<td>50</td>
<td>$(n/m)^5(\gamma_1/\mu(\gamma_i))$</td>
</tr>
<tr>
<td>MB</td>
<td>(5000, 75000)</td>
<td>0</td>
<td>$(0, 0.06)$</td>
<td>$1e^{-9}$</td>
<td>50</td>
<td>$(n/m)^5(\gamma_1/\mu(\gamma_i))$</td>
</tr>
<tr>
<td>FD</td>
<td>(4096, 81920)</td>
<td>0</td>
<td>$(0, 0.2)$</td>
<td>$1e^{-9}$</td>
<td>50</td>
<td>$(n/m)^5(\gamma_1/\mu(\gamma_i))$</td>
</tr>
</tbody>
</table>

3.1.1. Discussion of degree of ill-posedness: Dipping Dike

Before presenting the results we analysis the decay rate of the singular values $\tilde{G}$ in (5). Generally, if $\sigma_j \approx O(j^{-\nu})$, then the problem is called mildly ill-posed for $\nu < 1$ and moderately ill-posed for $\nu > 1$. If $\sigma_j \approx O(e^{-\nu j})$ for some $\nu > 1$ then the problem is said to be severely ill-posed, see e.g. Hoffman (1986); Aster et al. (2013); Huang & Jia (2016). From examination of the singular values of the $\tilde{G}$ we realize that the problem is either mildly or moderately ill-posed and then we do a nonlinear fit of the $\sigma_j$ to $Cj^{-\nu}$ to find $\nu$. The data fitting results for case $N3$ and over 4 iterations for a given sample are shown in Fig. 2. There is a very good fit to the data at each iteration, and $\nu$ is clearly iteration dependent. Equivalent results are obtained for both noise cases $N1$ and $N2$. We conclude that the magnetic inversion problem as used here is mildly ill-posed, with the parameter $\nu$ changing each iteration. Huang & Jia (2016) have shown that the LSQR iteration captures the underlying Krylov subspace of the system matrix better when that matrix is severely or moderately ill-posed, and therefore has better regularizing properties in these cases. When the problem is only mildly ill-posed, however, additional
regularization is needed.

Figure 2: The original spectrum of matrix $\tilde{G}$ at iterations 1 to 4 for case $N3$ and the data fit function in each case. In the title $\nu$ obtained at each iteration, .36163, .81138, .72383, and .65128, respectively.

It is the relatively mild ill-posedness of $\tilde{G}$ that now leads to the need to truncate the spectrum of the matrix $B_t$ when determining the regularization parameter. Specifically, we cannot expect that $B_t$ will accurately capture a right singular subspace of size $t$ for matrix $\tilde{G}$. In Fig. 3 we show the spectrum of the matrix $\tilde{G}$ as compared to that of $B_t$ with $t = 200$ and $t = 140$. We see immediately that we cannot use the right singular subspace for matrix $B_{140}$ because it, as with $B_{200}$, cannot accurately capture the subspace of size 140, resp. 200. On the other hand, over all iterations the truncation of the spectrum for $B_{200}$ at 140 gives a reasonable approximation to the spectrum.
of $\tilde{G}$ and we may expect to appropriately regularize the dominant $t' = 140$ terms in (10). In the these plots we also give the condition $\kappa(\tilde{G})$ at each iteration, and we can see that indeed $\tilde{G}$ is not severely ill-conditioned. At the same time, the fact that the condition is changing with the iteration supports the premise that the regularization parameter $\alpha$ does need to be estimated for each iteration $k$.

![Figure 3](image-url)

Figure 3: The original spectrum of matrix $\tilde{G}$ at iterations 1 to 4 for case $N3$ as compared to that for $B_t$ with $t = 200$ and $t = 140$, respectively. The vertical line cuts through at $t = 140$.

To further demonstrate that it is not sufficient to just find the solution of (7) without regularization we also illustrate the relative error of the reconstructed model, $RE^{(K)} = \frac{\|m_{\text{exact}} - m^{(K)}\|_2}{\|m_{\text{exact}}\|_2}$ obtained both by regularization, and by the projection only. For the projected case we
find the solution with minimum relative error over solutions obtained using $t = 10 : 10 : 250$. Results are shown for cases noise $N2$ and $N3$ in Fig. 4. As the noise level increases the benefit of the regularization is more apparent.

Figure 4: The relative error over 4 iterations using $t = 200$, as compared to the unregularized projected solution found using $t = 10 : 10 : 250$ with minimum relative error for cases noise $N2$, left, and noise $N3$, right, respectively.

Finally, in validating the new TGCV approach for finding the regularization parameter, we compare the GCV, WGCV and TGCV functions for iterations 2 to 4 for cases $N2$ and $N3$ in Figs. 5-6, respectively. It is immediate that the WGCV leads to a severe underestimate for the regularization parameter. While both the TGCV and GCV functions achieve a suitable minimum, the parameter achieved using GCV is larger and will lead to over smoothing in the solution because the GCV accounts for the smaller singular values from $B_t$. This is also evident from the same plots for $t = 50$, Figs. 7-8, respectively.
Figure 5: The GCV, WGCV and TGCV functions at iterations 2 to 4 for case \(N2\) using \(t = 200\).
Figure 6: The GCV, WGCV and TGCV functions at iterations 2 to 4 for case N3 using $t = 200$. 
Figure 7: The GCV, WGCV and TGCV functions at iterations 2 to 4 for case N2 using $t = 50$. 
Figure 8: The GCV, WGCV and TGCV functions at iterations 2 to 4 for case N3 using $t = 50$.

3.1.2. Dipping Dike Results

We now summarize the results for the inversion of the dipping dike problem. In each case the final iteration, $K$, the final regularization parameter, $\alpha^{(K)}$, and the relative error of the reconstructed model for three values of $t$, $t = 5, 50$ and 200, are given in Table 3. To illustrate the results, the reconstructed models are shown in Figs. 9, 10 and 11.

Algorithm 1 is generally able to yield non-smooth models with distinct boundaries. For $t = 5$, the reconstructed models using all three methods are in good agreement with the original problem. The issue, here, is that for very small $t$ the singular values of $B_t$ tend to approximate the largest
singular values of \( \tilde{G} \), see Fig. 12a, without capturing the spectral condition of \( \tilde{G} \), so that no regularization is needed. In fact, here, the regularization is achieved by Krylov filtering at the expense of losing some spectral properties of the original kernel. For \( t = 50 \) and \( 200 \), the matrix \( B_t \) captures the ill-conditioning of the full system matrix, see Figs. 12b and 12c, and as already noted in section 3.1.1 regularization is required. As anticipated from the discussion on the GCV, WGCV and TGCV parameter choice methods, the WGCV underestimates the regularization parameter and the solution is noisy, while both GCV and TGCV work very well, see Table 3. Although these results are for the case with weight parameter \( \omega = 0.8 \) in the GCV, further simulations, using \( \omega = 0.4 \) and \( \omega = 1.4 \), have shown that alternative choices for \( \omega \) are also not useful in this example. Indeed the solutions are quite comparable to those already presented in Table 3.

We now emphasize again that \( t \) should be selected as small as possible, \( t \ll \min(m, n) \), in order to provide an efficient algorithm, while simultaneously capturing the condition of the original kernel. This indicates that \( t = 5 \) and \( 200 \) are not good choices, one does not capture the condition and the other is relatively large as compared to \( m \). A compromise is needed, and while we suggest as in Vatankhah et al. (2016) that \( t \geq m/20 \), it will be the case that the size for \( t' \) depends on the spectrum of \( \tilde{G} \), and the degree of ill-posedness. We also note that contrary to the conclusion in Abedi et al. (2013) that the WGCV method is a good candidate for the determining the regularization parameter for the magnetic inverse problem, our results do not support the use of the WGCV. The standard GCV and TGCV both outperform the WGCV, especially for high noise levels. For \( t = 50 \) the TGCV
works better than even the GCV, indicating that the method can be used with high confidence. In all cases, it is a matter of examining the spectrum of $B_t$. While one may not want to estimate the spectrum of $\tilde{G}$, an approach to determine the extent to which $B_t$ does or does not capture a suitable right singular subspace can be ascertained by extending beyond $t$ by some small number and then contrasting the spectra that are obtained. Equivalently, one may reduce $t$, as indicated in Fig.2 to determine the extent by which the spectrum is preserved with respect to $t'$. When $B_t$ includes the small singular values from $\tilde{G}$, truncation is needed to assure that the solutions are not overly smooth due to larger estimates for $\alpha^{(k)}$ than needed.
Table 3: The results obtained by inverting the data from dipping dike for three noise levels \( N_1 \), \( N_2 \) and \( N_3 \) using Algorithm 1 and GCV, WGCV and TGCV parameter-choice methods. Here the results are presented for three values of parameter \( t \).

<table>
<thead>
<tr>
<th>GKB steps (t)</th>
<th>Parameter-choice method</th>
<th>Noise level</th>
<th>Iteration (K)</th>
<th>( \alpha^{(K)} )</th>
<th>( RE^{(K)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>GCV</td>
<td>( N_1 )</td>
<td>50</td>
<td>32039</td>
<td>0.4791</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( N_2 )</td>
<td>23</td>
<td>12338</td>
<td>0.5972</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( N_3 )</td>
<td>21</td>
<td>7730</td>
<td>0.7140</td>
</tr>
<tr>
<td></td>
<td>WGCV (( \omega = 0.8 ))</td>
<td>( N_1 )</td>
<td>39</td>
<td>31829</td>
<td>0.4531</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( N_2 )</td>
<td>17</td>
<td>12427</td>
<td>0.5907</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( N_3 )</td>
<td>13</td>
<td>8375</td>
<td>0.7082</td>
</tr>
<tr>
<td></td>
<td>TGCV (( t' = 3 ))</td>
<td>( N_1 )</td>
<td>50</td>
<td>31988</td>
<td>0.4973</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( N_2 )</td>
<td>22</td>
<td>12298</td>
<td>0.5957</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( N_3 )</td>
<td>21</td>
<td>7730</td>
<td>0.7140</td>
</tr>
<tr>
<td>50</td>
<td>GCV</td>
<td>( N_1 )</td>
<td>13</td>
<td>1583</td>
<td>0.4255</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( N_2 )</td>
<td>6</td>
<td>16625</td>
<td>0.6100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( N_3 )</td>
<td>8</td>
<td>9233</td>
<td>0.7041</td>
</tr>
<tr>
<td></td>
<td>WGCV (( \omega = 0.8 ))</td>
<td>( N_1 )</td>
<td>14</td>
<td>1547</td>
<td>0.4280</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( N_2 )</td>
<td>12</td>
<td>663</td>
<td>0.7589</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( N_3 )</td>
<td>13</td>
<td>414</td>
<td>0.9042</td>
</tr>
<tr>
<td></td>
<td>TGCV (( t' = 35 ))</td>
<td>( N_1 )</td>
<td>13</td>
<td>7432</td>
<td>0.3867</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( N_2 )</td>
<td>4</td>
<td>4204</td>
<td>0.6012</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( N_3 )</td>
<td>4</td>
<td>1237</td>
<td>0.7001</td>
</tr>
<tr>
<td>200</td>
<td>GCV</td>
<td>( N_1 )</td>
<td>14</td>
<td>9812</td>
<td>0.4054</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( N_2 )</td>
<td>5</td>
<td>10547</td>
<td>0.6377</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( N_3 )</td>
<td>4</td>
<td>3399</td>
<td>0.7263</td>
</tr>
<tr>
<td></td>
<td>WGCV (( \omega = 0.8 ))</td>
<td>( N_1 )</td>
<td>13</td>
<td>1469</td>
<td>0.4366</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( N_2 )</td>
<td>13</td>
<td>646</td>
<td>0.7581</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( N_3 )</td>
<td>12</td>
<td>359</td>
<td>0.9099</td>
</tr>
<tr>
<td></td>
<td>TGCV (( t' = 140 ))</td>
<td>( N_1 )</td>
<td>13</td>
<td>8315</td>
<td>0.3974</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( N_2 )</td>
<td>4</td>
<td>3584</td>
<td>0.6517</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( N_3 )</td>
<td>4</td>
<td>2517</td>
<td>0.7365</td>
</tr>
</tbody>
</table>
Figure 9: The reconstructed model using Algorithm 1 with $t = 5$. (a)-(c) using the GCV as parameter-choice method for noise cases $N_1, N_2$ and $N_3$, respectively. (d)-(f) using the WGCV as parameter-choice method for noise cases $N_1, N_2$ and $N_3$, respectively. (g)-(i) using the TGCV as parameter-choice method for noise cases $N_1, N_2$ and $N_3$, respectively.
Figure 10: The reconstructed model using Algorithm 1 with $t = 50$. (a)-(c) using the GCV as parameter-choice method for noise cases $N_1$, $N_2$ and $N_3$, respectively. (d)-(f) using the WGCV as parameter-choice method for noise cases $N_1$, $N_2$ and $N_3$, respectively. (g)-(i) using the TGCV as parameter-choice method for noise cases $N_1$, $N_2$ and $N_3$, respectively.
Figure 11: The reconstructed model using Algorithm 1 with $t = 200$. (a)-(c) using the GCV as parameter-choice method for noise cases $N1; N2$ and $N3$, respectively. (d)-(f) using the WGCV as parameter-choice method for noise cases $N1; N2$ and $N3$, respectively. (g)-(i) using the TGCV as parameter-choice method for noise cases $N1; N2$ and $N3$, respectively.

Figure 12: The singular values of $\tilde{\tilde{G}}$, blue $+$, and $B_t$, red $\ast$, at iteration 4 using Algorithm 1 for the noise $N2$. (a) $t = 5$; (b) $t = 50$ and (c) $t = 200$. 
3.2. Multiple bodies

To demonstrate the application of the method for a larger and more complex model, we now apply the algorithm on a model consisting of six different bodies. Fig. 13a shows a perspective view of the model. The dimensions and susceptibilities of each body are given in Table 4. Four plane-sections of the model are shown in Fig. 14. The surface magnetic data are generated on a $100 \times 50$ grid with 20 m spacing. Intensity of the magnetic field, inclination and declination are 47000 nT, 50° and 2°, respectively. Error contaminated data $d_{\text{obs}}$, illustrated in Fig. 13b, uses the noise model $N2$. The parameters of Algorithm 1 for the inversion of the model with the uniform subsurface discretization $100 \times 50 \times 15$ are detailed in Table 2. As already determined the projected problem should have size $t \geq m/20 = 250$. Here we take $t = 300$, yielding matrix $B_t$ of size $301 \times 300$.

Figure 13: (a) Model comprising 6 different bodies. Darker and brighter bodies have the susceptibilities 0.06 and 0.04, respectively; (b) The total noise contaminated anomaly due to the model. Intensity of the magnetic field, inclination and declination are 47000 nT, 50° and 2°, respectively.
Table 4: The susceptibility and dimension of each body for the model in Fig. 9a.

<table>
<thead>
<tr>
<th>Source number</th>
<th>Dimensions (m)</th>
<th>Depth to the surface (m)</th>
<th>Susceptibility (SI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100 × 100 × 20</td>
<td>20</td>
<td>0.06</td>
</tr>
<tr>
<td>2</td>
<td>200 × 400 × 120</td>
<td>20</td>
<td>0.06</td>
</tr>
<tr>
<td>3</td>
<td>600 × 100 × 80</td>
<td>40</td>
<td>0.04</td>
</tr>
<tr>
<td>4</td>
<td>200 × 200 × 140</td>
<td>20</td>
<td>0.04</td>
</tr>
<tr>
<td>5</td>
<td>100 × 400 × 100</td>
<td>20</td>
<td>0.04</td>
</tr>
<tr>
<td>6</td>
<td>600 × 100 × 60</td>
<td>40</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Figure 14: The model in Fig. 13a is displayed in four plane-sections. The depths of the sections are: (a) 20 m; (b) 60 m; (c) 100 m and (d) 140 m.

The inversion process is very fast, requiring less than 10 minutes to implement for each of the methods. The results of the algorithm are given in Table 5. For all cases the final iteration, $K$, $\alpha^{(K)}$, and $RE^{(K)}$, are reported.
The results indicate that TGCV outperforms both GCV and WGCV methods. To illustrate the results, we show the plane-sections of the reconstructed models in Figs. 15, 16 and 17 for GCV, WGCV and TGCV, respectively. The GCV, WGCV and TGCV functions at the final iterations and isosurfaces of the 3-D inversions are given in Figs. 18 and 19. The models reconstructed by GCV and TCGV are of higher quality than the model obtained using the WGCV which underestimates the regularization parameter. The solutions using GCV and TCGV have horizontal borders that are in good agreement with those of the original model, but as is typical for the inversion of magnetic data additional structures appear at depth due to the loss of resolution at depth. Still, unlike smoothing inversion algorithms, the models here are more focused even at depth.

Table 5: The inversion results for $K$, $\alpha^{(K)}$ and $RE^{(K)}$ obtained by inverting data from Fig. 13b using Algorithm 1 with $t = 300$ and GCV, WGCV and TGCV parameter-choice methods.

<table>
<thead>
<tr>
<th>Parameter-choice method</th>
<th>Iteration ($K$)</th>
<th>$\alpha^{(K)}$</th>
<th>$RE^{(K)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCV</td>
<td>5</td>
<td>774.8583</td>
<td></td>
</tr>
<tr>
<td>WGCV</td>
<td>3</td>
<td>14.8702</td>
<td></td>
</tr>
<tr>
<td>TGCV</td>
<td>4</td>
<td>190.8127</td>
<td></td>
</tr>
</tbody>
</table>
Figure 15: For the data in Fig. 13b; the reconstructed model using Algorithm 1 with $t = 300$ and GCV. The depths of the sections are: (a) 20 m; (b) 60 m; (c) 100 m and (d) 140 m.
Figure 16: For the data in Fig. 13b; the reconstructed model using Algorithm 1 with $t = 300$ and WGCV. The depths of the sections are: (a) 20 m; (b) 60 m; (c) 100 m and (d) 140 m.
Figure 17: For the data in Fig. 13b; the reconstructed model using Algorithm 1 with \( t = 300 \) and TGCV. The depths of the sections are: (a) 20 m; (b) 60 m; (c) 100 m and (d) 140 m.

Figure 18: (a) The GCV function at final iteration \( K = 5 \); (b) The WGCV function at final iteration \( K = 3 \) and (d) The TGCV function at final iteration \( K = 4 \).
4. Field data inversion

We applied the developed code for the inversion of aeromagnetic data obtained over the Wuskwatim Lake region in Manitoba, Canada, Fig. 20. The given area lies within a poorly exposed meta-sedimentary gneiss belt consisting of paragneiss, amphibolite, and migmatite derived from Proterozoic volcanic and sedimentary rocks (Pilkington, 2009). Pilkington (2009) applied a data-space sparse inversion methodology with a Cauchy norm sparsity constraint on the model parameters for the inversion of this data gridded onto a 64 × 64 grid. The results of the inversion using Li & Oldenburg (1996) algorithm are also presented in (Pilkington, 2009, Figure 5).

Consistent with the inversion detailed in (Pilkington, 2009) we use a 64 × 64 data grid with 100 m spacing and the uniform subsurface discretization 64 × 64 × 20. We assume that the data are contaminated with error using the noise model \((\tau_1, \tau_2) = (0.03, 0.005)\) in order to estimate the standard deviation of the error at measurement \(i\) as \((0.03d_{\text{obs}})_i + 0.005\|d_{\text{obs}}\|\). We again note the projected problem should have size \(t \geq m/20 \approx 205\) and take \(t = 250\). The parameters of Algorithm 1 for the inversion of the model are detailed.
in Table 2, and assume as in Pilkington (2009) $\kappa_{\text{max}} = 0.2$. For inversion we use the TGCV method for parameter choice. The convergence is achieved at $K = 11$ iterations and completes in less than 15 minutes. Two plane-sections of the recovered model at depths 400 m and 800 m are presented in Figs. 21a and 21b. The TGCV function at the final iteration is shown in Fig. 21c, and an isosurface of the 3-D recovered model with susceptibility greater than 0.05 (SI) is shown in fig. 21d.

Figure 20: Aeromagnetic data over a portion of the Wuskwatim Lake area, Monitoba, Canada.
Figure 21: The results of the inversion for the data in Fig. 20. (a) Plane-section at depth = 400 m; (b) Plane-section at depth = 800 m; (c) TGCV function at final iteration $K = 11$ and (d) The isosurface of the reconstructed model with susceptibility greater than 0.05 (SI).

The algorithm produces results that are competitive with, but not as sparse as, those presented in Pilkington (2009, Figure 5). The reduction in sparsity is a characteristic of the $L_1$-norm stabilizer. If greater sparsity is required, the minimum support $L_0$-norm stabilizer can be implemented in Algorithm 1 by replacing $W_{L_1}$ with $W_{MS} = \text{diag}\left(\left(\frac{(|m^{(k)} - m^{(k-1)})^2 + \epsilon^2}{\epsilon} \right)^{-1/2}\right)$ in step 13, see Vatankhah et al. (2016).
5. Conclusions

We developed an algorithm for sparse inversion of magnetic data using $L_1$-norm stabilization with projection of the solution to a smaller Krylov subspace using the LSQR algorithm. Generalized cross validation techniques were used to efficiently estimate the regularization parameter for the projected space solutions expressed using the singular value decomposition of the subspace system matrix. The results demonstrated that the weighted GCV is not as effective as the GCV for high noise levels. The regularization parameter is underestimated. In contrast, noting that the magnetic inversion problem is only mildly ill-posed, we introduced a truncated GCV to take account of the observation that the projected system only captures a portion of the full singular space accurately. The projected subspace is therefore truncated so that the components related to the inaccurate small singular values are ignored in calculating the regularization parameter. For relatively small subspaces acceptable solutions are obtained with a limited number of iterations for the IRLS algorithm. The algorithm was validated for two synthetic models, a dipping dike and a model of multiple bodies, and demonstrated that the algorithm permits inversion of large data sets and produces relatively focused images of the subsurface. The algorithm was applied for the inversion of real aeromagnetic data collected over Wuskwatim Lake in Manitoba, Canada, and compares favorably with results achieved using alternative algorithms presented in the literature.
Acknowledgements

We thank Dr. Mark Pilkington for providing aeromagnetic data from Wuskwatim Lake area.

Appendix A. Appendix

Algorithm 1: Iterative Projected $L_1$ Inversion Algorithm

**Require:** $d_{\text{obs}}$, $m_{\text{apr}}$, $G$, $W_d$, $\epsilon > 0$, $\kappa_{\text{min}}$, $\kappa_{\text{max}}$, $t$, $K_{\text{max}} = 50$, $\alpha^{(1)}$

1: Calculate $W_{\text{depth}}$, $\tilde{G} = W_dG$, and $\tilde{d}_{\text{obs}} = W_dd_{\text{obs}}$
2: Initialize $m^{(0)} = m_{\text{apr}}$, $W^{(1)}_{L_1} = I_n$, $W^{(1)} = W_{\text{depth}}$
3: Calculate $\tilde{r}^{(1)} = \tilde{d}_{\text{obs}} - \tilde{G}m^{(0)}$, $\tilde{G}^{(1)} = \tilde{G}(W^{(1)})^{-1}$, $k = 0$
4: **while** Not converged, noise level not satisfied, or $k < K_{\text{max}}$ **do**
5: $k = k + 1$
6: Apply GKB: $\tilde{G}^{(k)} A_t^{(k)} = H_t^{(k)} B_t^{(k)}$, $H_t^{(k)} e_{t+1} = \tilde{r}^{(k)}/||\tilde{r}^{(k)}||_2$
7: Find SVD: $B_t^{(k)} = UTV$. Calculate $c = ||\tilde{r}||_2 e_{t+1}$.
8: For $k > 1$ Estimate $\alpha^{(k)}$, using GCV, WGC or TGCV.
9: Set $z_t^{(k)} = \sum_{i=1}^{t} \frac{\gamma_i^2}{\gamma_i^2 + \alpha^{(t)}} \frac{u^T e}{\gamma_i} v_i$
10: Set $m^{(k)} = m^{(k-1)} + (W^{(k)})^{-1} A_t^{(k)} z_t^{(k)}$
11: Impose constraint conditions on $m^{(k)}$ to force $\kappa_{\text{min}} \leq m^{(k)} \leq \kappa_{\text{max}}$
12: Calculate residual $\tilde{r}^{(k+1)} = \tilde{d}_{\text{obs}} - \tilde{G}m^{(k)}$
13: Set $W^{(k+1)}_{L_1} = \text{diag} \left((m^{(k)} - m^{(k-1)})^2 + \epsilon^2 \right)^{-1/4}$ and $W^{(k+1)} = W^{(k+1)}_{L_1} W_{\text{depth}}$
14: Calculate $\tilde{G}^{(k+1)} = \tilde{G}(W^{(k+1)})^{-1}$
15: **end while**

**Ensure:** Solution $\kappa = m^{(k)}$. $K = k$. 39
References


Li, Y., Oldenburg, D. W., 1996. 3-D inversion of magnetic data, Geophysics, 61, 394-408.


cation of the method to the Safo manganese mine in northwest of Iran, Journal Of Geophysics and Engineering, 11, 045001.


