Baur, O., Gerrit, G., Kusche, J. (2008): Efficient GOCE satellite gravity field recovery based on least-squares using QR decomposition. - Journal of Geodesy, 82, 4-5, 207-221

DOI: 10.1007/s00190-007-0171-z

Efficient GOCE satellite gravity field recovery based on least-squares using QR decomposition

Oliver Baur[1]**, Gerrit Austen**[1]**, Jürgen Kusche**[2,3]

[1] Institute of Geodesy, Universität Stuttgart, Geschwister-Scholl-Str. 24D, 70174 Stuttgart, Germany; Phone: +49 (0)711 68584640, Fax: +49 (0)711 68583285

e-mail: baur@gis.uni-stuttgart.de, austen@gis.uni-stuttgart.de

[2] Delft Institute of Earth Observation and Space Systems (DEOS), Delft University of Technology, Kluyverweg 1, 2629 HS Delft, The Netherlands

[3] now at the Department of Geodesy and Remote Sensing, GFZ Potsdam, Telegrafenberg, Germany

e-mail: jkusche@gfz-potsdam.de

Abstract. We develop and apply an efficient strategy for Earth gravity field recovery from satellite gravity gradiometry data. Our approach is based upon the Paige-Saunders iterative least-squares method using QR decomposition (LSQR). We modify the original algorithm for space-geodetic applications: Firstly, we investigate how convergence can be accelerated by means of both subspace and block-diagonal preconditioning. The efficiency of the latter dominates if the design matrix exhibits block-dominant structure. Secondly, we address Tikhonov-Phillips regularization in general. Thirdly, we demonstrate an effective implementation of the algorithm in a high-performance computing environment. In this context, an important issue is to avoid the twofold computation of the design matrix in each iteration. The computational platform is a 64-processor shared-memory supercomputer. The runtime results prove the successful parallelization of the LSQR solver. The numerical examples are chosen in view of the forthcoming satellite mission GOCE (Gravity field and steady-state Ocean Circulation Explorer). The closed-loop scenario covers one month of simulated data with 5 s sampling. We focus exclusively on the analysis of radial components of satellite accelerations and gravity gradients. Our extensions to the basic algorithm enable the method to be competitive with well-established inversion strategies in satellite geodesy, such as conjugate gradient methods or the brute-force approach. In its current development stage, the LSQR method appears ready to deal with real-data applications.

Key words. Least-squares – Iterative solvers – QR decomposition – Preconditioning – Gravity field recovery – GOCE – Parallel computing

1 Introduction

Nowadays, geodetic scientists are in the favorable situation to perform gravity field recovery by near-globally distributed satellite-based observations. The CHAMP (Challenging Minisatellite Payload) and GRACE (Gravity Recovery and Climate Experiment) missions have been operational for a few years and still continuously provide gravity data. These missions will be complemented with the GOCE (Gravity field and steady-state Ocean Circulation Explorer) satellite to be launched towards the end of 2007.

Modelling of both the static terrestrial gravity field and its temporal variations facilitate advanced research in a multitude of geoscientific applications, such as the study of the dynamics of the Earth's lithosphere and upper mantle, global sea level variations, ocean circulation and ocean mass and heat transport and ice mass balance (ESA 1999). However, solving for the unknown parameters of the gravitational potential model, i.e., the coefficients of the corresponding harmonic series expansion, in a least-squares (LS) adjustment procedure is a challenging task. Millions of observations have to be processed to resolve up to 100,000 unknown gravity field parameters.

Paige and Saunders (1982a;b) published the LSQR algorithm. LSQR is an acronym for a special method to solve linear (ill-conditioned) LS problems using QR decomposition. Whereas the LSQR method is frequently applied in geophysics, in seismic tomography in particular (Van der Sluis and Van der Vorst 1987, Nolet 1993, Yao et al. 1999), it has found only little use in geodetic applications. Actually, large LS problems in geodesy are mostly treated by conjugate gradient (CG) methods (Hestenes and Stiefel 1952), leading to the CGLS procedure. Application of CGLS in terms of satellite gravity field recovery can be found in, e.g., Schuh (1996), Pail and Plank (2002) and Ditmar et al. (2003a). Geodetic problems are treated with LSQR in, e.g., Kusche and Mayer-Gürr (2001) and Baur and Austen (2005).

Both LSQR and CGLS are Krylov subspace-based iterative methods designed to solve linear systems of equations by means of successive approximations. They share approximately the same storage and work requirement demands and, mathematically, they generate the same sequence of approximations to the solution in exact arithmetic. Within the scope of this contribution, we focus exclusively on the LSQR algorithm, which

has been claimed to outperform CG methods with regard to stability of the iterative process on the grounds of theoretical arguments (Paige and Saunders 1982a, Björck 1996, Jacobsen et al. 2003).

Since memory requirements are small for these solvers, they can be implemented on an ordinary personal computer (PC) if runtime criteria are disregarded. However, to solve large-scale problems within a reasonable time-frame, parallel implementation is indispensable. This allows one to devote more computation time to tuning and optimization efforts, which generally involve multiple solutions of the individual LS problem. Here, we present the LSQR algorithm in its parallelized version using OpenMP (Chandra et al. 2001). The computational platform for numerical calculations is a 64-processor cc-NUMA (cache-coherent non-uniform memory access) system supported by the Center for Computing and Networking Services (SARA) in Amsterdam.

In terms of LSQR tuning, we perform preconditioning and regularization by tailored adoption to satellite-based gravity field recovery. Preconditioning can be realized in different ways to accelerate the convergence rate. In satellite geodesy, it is well known that an ideal satellite configuration, characterized by a circular and repeating orbit with fixed inclination, leads to a strictly block-diagonal normal matrix (Colombo 1984). In reality, deviations from the ideal conditions have to be taken into account. Nevertheless, the real normal matrix typically shows block-dominant structure. Thus, its block-diagonal approximation constitutes an appropriate preconditioner.

Normal matrix preconditioning with LSQR is investigated in Baur and Austen (2005). It turns out that the twofold computation of the design matrix in each iteration is unavoidable. This fact is of minor relevance for solving small linear systems of equations. However, in satellite gravity recovery, the number of observations, i.e., the dimension of the design matrix row-space, amounts to several millions, causing the assembly of the design matrix to become the most time consuming part of the algorithm. Here, we reformulate block-diagonal normal matrix preconditioning to the level of the design matrix. The new formulation circumvents the need to calculate the design matrix twice per iteration and thus considerably reduces the computational effort.

An alternative for LSQR preconditioning has been proposed by Jacobsen et al. (2003), denoted as subspace preconditioning. The basic idea is to split the solution space into two subspaces of different size. The splitting decomposes the original problem into two subproblems. The small subspace problem is solved by direct inversion, whereas the larger one is treated by LSQR with the small subspace solution projected to the larger space as a preconditioner. Finally, the partial estimates compose the overall solution vector.

Subspace preconditioning does not rely on a special block-dominant structure of the LS problem. Thus, it is generally applicable to any linear minimization problem. Because of the analogy to multi-level methods,

subspace preconditioning using two subspaces is denoted as two-level method. Xu (1997) showed that multigrid methods fit into the framework of multi-level techniques. Application of these methods in satellite gravity recovery have been investigated in, e.g., Kusche (2001) and Keller (2002).

Boxhammer (2006) and Boxhammer and Schuh (2006) present a very efficient method for the combination of high-resolution and low-resolution data sets in a common LS procedure. The numbering scheme developed, allows tailored preconditioning of systems of equations resulting from observation configurations providing different types of data.

Regularization is of fundamental importance when treating ill-posed LS problems. To stabilize the inversion procedure, Tikhonov-Phillips regularization (Phillips 1962, Tikhonov 1963) has been proven to perform very well for space-geodetic applications (Kusche and Klees 2002, Ditmar et al. 2003b, Schmidt et al. 2007). Paige and Saunders (1982a;b) formulated a regularized LSQR procedure. However, they only refer to regularization with the identity matrix, known as ordinary ridge regression. We adopt the method for general Tikhonov-Phillips regularization by transformation of the general regularization problem to its standard form (Björck 1996). Moreover, we investigate the combination of preconditioning and regularization.

The paper is organized as follows. The next section briefly reviews the original LSQR algorithm, including an outline concerning decorrelation and variance-covariance propagation issues. Section 3 treats the reformulation of the block-diagonal normal matrix preconditioner to the level of the design matrix. Subspace preconditioning is addressed in Sect. 4. Section 5 introduces Tikhonov-Phillips regularization to the algorithms presented earlier. The methodology for GOCE gravity field recovery is presented in Sect. 6. In Sect. 7, we demonstrate numerical examples in the context of a GOCE-like closed-loop simulation study. Moreover, we focus on aspects concerning implementation and parallelization. Finally, Sect. 8 summarizes the conclusions of this contribution.

2 The basic LSQR method

A linear (or linearized) inverse problem can be defined as that of solving the system

$$\mathbf{A}\mathbf{x} = \mathbf{y} + \mathbf{r}. \tag{1}$$

The design matrix $\mathbf{A}$ relates the vector of the unknown parameters $\mathbf{x}$ to the vector of observations $\mathbf{y}$ and $\mathbf{r}$ is a vector of data errors. Solving Eq. (1) for $\mathbf{x}$ by minimizing the L_2-norm of the residual vector results in the linear LS problem

$$\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|^2. \tag{2}$$

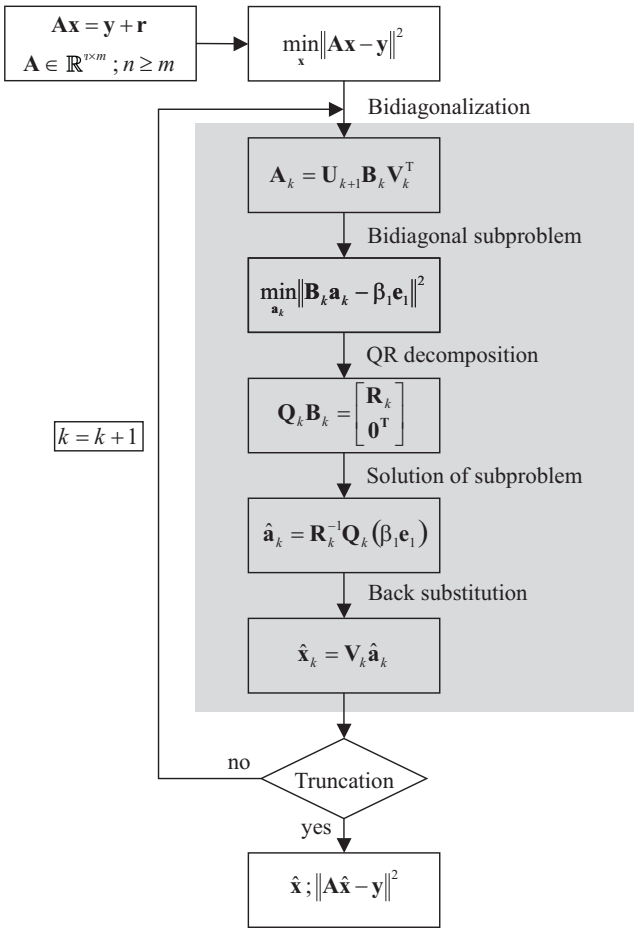

Fig. 1. Flowchart of the LSQR method

Equation (2) is the starting point for LSQR representation. According to Fig. 1, the algorithm works as follows:

(i) Lanczos lower bidiagonalization procedure (Golub et al. 1965, Paige and Saunders 1982a) transforms the original minimization problem to the simpler bidiagonal one

$$\min_{\mathbf{a}_k} \|\mathbf{B}_k\mathbf{a}_k - \beta_1\mathbf{e}_1\|^2. \tag{3}$$

This transformation is achieved by decomposing the design matrix $\mathbf{A}$ $(n \times m)$ in an iterative manner into two orthogonal matrices $\mathbf{U}_{k+1} = [\mathbf{u}_1, \ldots, \mathbf{u}_{k+1}]$ and $\mathbf{V}_k = [\mathbf{v}_1, \ldots, \mathbf{v}_k]$, of dimension $(n \times (k+1))$ and $(m \times k)$, as well as a lower bidiagonal matrix $\mathbf{B}_k$ of dimension $((k+1) \times k)$. Thus, the relationship $\mathbf{A} \approx \mathbf{A}_k = \mathbf{U}_{k+1}\mathbf{B}_k\mathbf{V}_k^T$ holds.

The bidiagonalization process reads

$$\beta_{k+1}\mathbf{u}_{k+1} = \mathbf{A}\mathbf{v}_k - \alpha_k\mathbf{u}_k, \tag{4}$$

$$\alpha_{k+1}\mathbf{v}_{k+1} = \mathbf{A}^T\mathbf{u}_{k+1} - \beta_{k+1}\mathbf{v}_k \tag{5}$$

with the initial conditions $\beta_1 \mathbf{u}_1 = \mathbf{y}$ and $\alpha_1 \mathbf{v}_1 = \mathbf{A}^T \mathbf{u}_1$. The scalars α_k, $\beta_k \geq 0$ are chosen subject to $\|\mathbf{u}_k\| = \|\mathbf{v}_k\| = 1$. The values α_k constitute the main diagonal of $\mathbf{B}_k$ and β_k the first lower sub-diagonal. k denotes the number of iterations and for $k = m$ the matrix $\mathbf{A}$ is completely decomposed ($\mathbf{A}_{k=m} = \mathbf{A}$), but usually $k \ll m$ will be sufficient. $\beta_1 = \|\mathbf{y}\|$ corresponds to the norm of the observation vector $\mathbf{y}$ and $\mathbf{e}_1$ is the first column of a unit matrix of appropriate dimension.

(ii) The solution of the bidiagonal subproblem in Eq. (3),

$$\mathbf{a}_k = \mathbf{R}_k^{-1} \mathbf{Q}_k (\beta_1 \mathbf{e}_1), \tag{6}$$

is generated by a series of Givens rotations (Björck 1996) through which the bidiagonal matrix $\mathbf{B}_k$ is decomposed in an orthogonal matrix $\mathbf{Q}_k$ and an upper bidiagonal matrix $\mathbf{R}_k$.

(iii) The k^{th} iterate of the original parameter vector is determined from Eq. (6), the solution of the subproblem, and the matrix $\mathbf{V}_k$ subject to

$$\mathbf{x}_k = \mathbf{V}_k \mathbf{a}_k. \tag{7}$$

Note that the column vectors of $\mathbf{V}_k$ span a *k*-dimensional Krylov space.

To evaluate Eq. (7), it would be necessary to store all vectors $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$. Paige and Saunders (1982a) present a simple recursion relation to compute $\mathbf{x}_k$ only from the last iterate $\mathbf{x}_{k-1}$ and the present column vector $\mathbf{v}_k$. The iterative process is terminated if the k^{th} iterate meets a suitably chosen truncation criterion as will be discussed in Sect. 7.2.

For LSQR, the twofold computation of the design matrix $\mathbf{A}$ according to Eqs. (4) and (5) can be avoided by introducing the temporary variable $\mathbf{h}_k$ with $\mathbf{h}_k = \sum_j \mathbf{A}(j,:) u_{k+1}^j$. After calling the j^{th} row of $\mathbf{A}$, denoted as $\mathbf{A}(j,:)$, the j^{th} loading to the product $\mathbf{A}^T \mathbf{u}_{k+1}$ consists of the j^{th} row itself as well as the j^{th} element of the (unnormalized) vector $\mathbf{u}_{k+1}$, i.e., u_{k+1}^j, which is already available from Eq. (4).

Pseudocode 1 presents a guideline for LSQR implementation. In addition, Pseudocode 2 shows an implementation of the Givens rotation.

2.1 Decorrelation

The LS problem formulation in Eq. (2) implies simplified stochastic properties of the observation vector $\mathbf{y}$, i.e., uncorrelated and equally accurate observations. In its general representation, the minimization problem reads

$$\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|^2_{\mathbf{\Sigma}^{-1}} \tag{8}$$

Pseudocode 1: The LSQR method to solve

$$\min \|\mathbf{r}\|^2 = \min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|^2$$

Initialization

1. $\beta_1 = \|\mathbf{y}\|,\ \mathbf{u}_1 = \frac{\mathbf{y}}{\beta_1}$
2. $\mathbf{v}_1 = \mathbf{A}^T\mathbf{u}_1,\ \alpha_1 = \|\mathbf{v}_1\|,\ \mathbf{v}_1 = \frac{\mathbf{v}_1}{\alpha_1}$
3. $\alpha_1 = \frac{\alpha_1}{\beta_1}$
4. $\bar{\phi}_1 = \beta_1$
5. $\bar{\rho}_1 = \alpha_1$

First Iteration: $k = 1$

6. $\mathbf{u}_2 = \mathbf{A}\mathbf{v}_1 - \alpha_1\mathbf{u}_1,\ \beta_2 = \|\mathbf{u}_2\|,\ \mathbf{u}_2 = \frac{\mathbf{u}_2}{\beta_2}$
7. $\mathbf{h}_1 = \mathbf{A}^T\beta_2\mathbf{u}_2$
8. $[c_1, s_1, \rho_1] = \text{givrot}(\bar{\rho}_1, \beta_2)$
9. $\phi_1 = c_1\bar{\phi}_1$
10. $\bar{\phi}_2 = -s_1\bar{\phi}_1$
11. $\mathbf{q}_1 = \frac{1}{\rho_1}\mathbf{v}_1$
12. $\mathbf{x}_1 = \phi_1\mathbf{q}_1$
13. $\mathbf{v}_2 = \mathbf{h}_1 - \beta_2^2\mathbf{v}_1,\ \alpha_2 = \|\mathbf{v}_2\|,\ \mathbf{v}_2 = \frac{\mathbf{v}_2}{\alpha_2}$
14. $\alpha_2 = \frac{\alpha_2}{\beta_2}$

Further iterations: for $k = 2 : m$

15. $\mathbf{u}_{k+1} = \mathbf{A}\mathbf{v}_k - \alpha_k\mathbf{u}_k,\ \beta_{k+1} = \|\mathbf{u}_{k+1}\|,\ \mathbf{u}_{k+1} = \frac{\mathbf{u}_{k+1}}{\beta_{k+1}}$
16. $\mathbf{h}_k = \mathbf{A}^T\beta_{k+1}\mathbf{u}_{k+1}$
17. $\theta_{k-1} = s_{k-1}\alpha_k$
18. $\bar{\rho}_k = c_{k-1}\alpha_k$
19. $[c_k, s_k, \rho_k] = \text{givrot}(\bar{\rho}_k, \beta_{k+1})$
20. $\phi_k = c_k\bar{\phi}_k$
21. $\bar{\phi}_{k+1} = -s_k\bar{\phi}_k$
22. $\mathbf{q}_k = \frac{1}{\rho_k}(\mathbf{v}_k - \theta_{k-1}\mathbf{q}_{k-1})$
23. $\mathbf{x}_k = \mathbf{x}_{k-1} + \phi_k\mathbf{q}_k$
24. $\mathbf{v}_{k+1} = \mathbf{h}_k - \beta_{k+1}^2\mathbf{v}_k,\ \alpha_{k+1} = \|\mathbf{v}_{k+1}\|,\ \mathbf{v}_{k+1} = \frac{\mathbf{v}_{k+1}}{\alpha_{k+1}}$
25. $\alpha_{k+1} = \frac{\alpha_{k+1}}{\beta_{k+1}}$

Pseudocode 2: Givens rotation $[c, s, \rho] = \text{givrot}(\bar{\rho}, \beta)$

if $\beta = 0.0$
  a. $c = 1.0,\ s = 0.0,\ \rho = \bar{\rho}$
else if $|\beta| > |\bar{\rho}|$
  a. $t = \bar{\rho}/\beta,\ q = \sqrt{1.0 + t^2}$
  b. $s = 1.0/q,\ c = ts,\ \rho = q\beta$
else
  a. $t = \beta/\bar{\rho},\ q = \sqrt{1.0 + t^2}$
  b. $c = 1.0/q,\ s = tc,\ \rho = q\bar{\rho}$

with the observation variance-covariance information $\mathbf{\Sigma} = \sigma^2\mathbf{P}^{-1}$. It is the product of the variance of unit weight σ^2 and the inverse positive definite weight matrix $\mathbf{P}$. In the presence of colored noise, $\mathbf{P}^{-1}$ is a full matrix or at least typically shows band structure of Toeplitz type. The width of the band depends on the correlation length.

In terms of a noise whitening process, Schuh (1996) presented a method to transfer the correlated LS problem in Eq. (8) to the uncorrelated formulation in Eq. (2). It is based on the decomposition of $\mathbf{\Sigma}^{-1}$ subject to

$$\mathbf{\Sigma}^{-1} \approx \mathbf{F}^T\mathbf{F}. \tag{9}$$

The matrix $\mathbf{F}$ applied to $\mathbf{y}$ yields the decorrelated vector of (pseudo-)observations $\mathbf{y}_\mathrm{D} = \mathbf{F}\mathbf{y}$. Thus, in signal processing terminology, $\mathbf{F}$ is a linear filter. Under the assumption of stationary noise and regularly distributed data, the filter $\mathbf{F}$ gets Toeplitz structure and it is implemented by means of an ARMA (auto-regressive moving average) process.

For consistency reasons, according to Eq. (1), the filter has to be applied to both the columns of the design matrix $\mathbf{A}$ and the residual vector $\mathbf{r}$ as well, yielding $\mathbf{A}_\mathrm{D} = \mathbf{F}\mathbf{A}$ and $\mathbf{r}_\mathrm{D} = \mathbf{F}\mathbf{r}$. This results in the decorrelated system of equations

$$\mathbf{A}_\mathrm{D}\mathbf{x} = \mathbf{y}_\mathrm{D} + \mathbf{r}_\mathrm{D} \tag{10}$$

and thus, the decorrelated minimization problem

$$\min_{\mathbf{x}} \|\mathbf{A}_\mathrm{D}\mathbf{x} - \mathbf{y}_\mathrm{D}\|^2. \tag{11}$$

The decorrelation process transforms the correlated LS problem (Eq. 11) to the uncorrelated formulation in Eq. (2). For this reason, throughout this contribution, we restrict ourselves to uncorrelated LS problem formulations.

In our numerical applications (Sect. 7), we implemented decorrelation as described in Schuh (1996). After the computation of a new row of $\mathbf{A}$, the filter is applied to each column of it as well as to the observations. An alternative approach has been proposed by Klees et al. (2003), based on a fast method to solve a Toeplitz system of linear equations.

2.2 Variance-covariance propagation

The variance-covariance matrix $\mathbf{\Sigma}(\hat{\mathbf{x}})$ of the parameter estimate is of utmost importance for the further use of the LS result in terms of error propagation. Unfortunately, in contrast to direct LS solvers, iterative methods do not provide this information. Here we give an outline of two possibilities for evaluating $\mathbf{\Sigma}(\hat{\mathbf{x}})$ with LSQR.

The first one has been derived by Yao et al. (1999) and is specifically related to the algorithm itself. It is based on the explicit representation of the general inverse of the design matrix yielding

$$\mathbf{\Sigma}_k(\hat{\mathbf{x}}) = \hat{\sigma}^2 \mathbf{V}_k (\mathbf{B}_k^T \mathbf{B}_k)^{-1} \mathbf{V}_k^T. \tag{12}$$

where $\mathbf{\Sigma}_k(\hat{\mathbf{x}})$ is the k^{th} approximation to the "true" variance-covariance matrix subject to $\mathbf{\Sigma}_{k=m}(\hat{\mathbf{x}}) = \mathbf{\Sigma}(\hat{\mathbf{x}})$. As the derivation of Eq. (12) is straightforward, we skip the details here. The interested reader is advised to consult the cited literature.

From the implementation point of view, the computation of $\mathbf{\Sigma}_k(\hat{\mathbf{x}})$ only requires marginal additional effort. The matrices $\mathbf{V}_k$ and $\mathbf{B}_k$ have to be stored and the product $\mathbf{V}_k(\mathbf{B}_k^T\mathbf{B}_k)^{-1}\mathbf{V}_k^T$ has to be evaluated. However, the extended memory requirement for matrix storage contradicts the philosophy of LSQR. Thus, dependent on the computation platform, $\mathbf{\Sigma}_k(\hat{\mathbf{x}})$ evaluation is restricted to some maximum value for k.

On the other hand, the goodness of the approximation predominantly depends on the number of iterations. Moreover, as shown in Yao et al. (1999), $\hat{\mathbf{x}}$ and $\mathbf{\Sigma}_k(\hat{\mathbf{x}})$ do not necessarily exhibit the same convergence behavior, but in general $\hat{\mathbf{x}}$ converges faster than the variance-covariance matrix. All in all, this demonstrates the necessity to investigate the method in detail for practical applications.

An alternative and very effective approach to estimate the variance-covariance matrix has been published recently by Alkhatib and Schuh (2007). As well as the Gibbs sampler method (Gundlich et al. 2003), it is based on Monte Carlo integration techniques. However, in contrast to the Gibbs sampler, it is particularly suited for iterative solvers.

3 Block-diagonal preconditioning

The convergence behavior of an iterative solver is predominantly determined by the spreading of the singular values of Eq. (1) or, equivalently, the condition number of the normal matrix. Preconditioning is applied to improve the condition of the normal equation system, and thus to increase the speed of convergence of the iterative solver.

In order to outline the methodology of preconditioning on the level of the design matrix, we briefly review preconditioning on the level of the normal matrix as presented in Baur and Austen (2005).

3.1 Normal matrix preconditioning: PCN-LSQR

The solution of Eq. (2) in the sense of the L_2-norm is explicitly given as

$$(\mathbf{A}^T\mathbf{A})\hat{\mathbf{x}} = \mathbf{A}^T\mathbf{y}. \tag{13}$$

Introducing the normal matrix $\mathbf{N} = (\mathbf{A}^T\mathbf{A})$ and the right-hand side vector $\mathbf{b} = \mathbf{A}^T\mathbf{y}$ leads to

$$\mathbf{N}\hat{\mathbf{x}} = \mathbf{b}, \tag{14}$$

the normal equation system of a linear Gauss-Markov model. Inversion of Eq. (14) yields the best linear unbiased estimate (BLUE) of the unknown parameter vector

$$\hat{\mathbf{x}} = \mathbf{N}^{-1}\mathbf{b}. \tag{15}$$

Let $\mathbf{N}_{\mathrm{bd}}$ denote the block-diagonal approximation of the true normal matrix. If $\mathbf{N}$ itself shows block-dominant structure, then $\mathbf{N}_{\mathrm{bd}}$ represents an appropriate preconditioner for $\mathbf{N}$ in terms of

$$\mathbf{N}\mathbf{N}_{\mathrm{bd}}^{-1} \cong \mathbf{I}. \tag{16}$$

Multiplying the unit matrix $\mathbf{I} = \mathbf{N}_{\mathrm{bd}}^{-1}\mathbf{N}_{\mathrm{bd}}$ in between the normal matrix and the unknown parameter vector on the left-hand side of Eq. (14) results in

$$\mathbf{N}\mathbf{N}_{\mathrm{bd}}^{-1}\mathbf{N}_{\mathrm{bd}}\hat{\mathbf{x}} = \mathbf{b}. \tag{17}$$

Let $\mathbf{N}^* = \mathbf{N}\mathbf{N}_{\mathrm{bd}}^{-1}$ denote the modified normal matrix and $\hat{\mathbf{x}}^* = \mathbf{N}_{\mathrm{bd}}\hat{\mathbf{x}}$ the new parameter vector, then Eq. (17) reads

$$\mathbf{N}^*\hat{\mathbf{x}}^* = \mathbf{b}. \tag{18}$$

Thus, inserting $\mathbf{N}^*$ and $\mathbf{b}$ for $\mathbf{A}$ and $\mathbf{y}$ into the LSQR procedure means solving the uniquely determined problem in Eq. (18) instead of the overdetermined one in Eq. (1). The original parameter vector is subsequently obtained by

$$\hat{\mathbf{x}} = \mathbf{N}_{\mathrm{bd}}^{-1}\hat{\mathbf{x}}^*. \tag{19}$$

The benefit of solving Eq. (18) instead of Eq. (14) is a better condition number of $\mathbf{N}^*$ opposed to $\mathbf{N}$, resulting in a higher convergence rate and a reduced number of necessary iterations k.

However, the procedure of preconditioning on the level of the normal matrix entails two significant drawbacks: (i) The design matrix needs to be accumulated twice per iteration and (ii) normal matrix preconditioning involves solving a uniquely determined problem, whereas LSQR is essentially designed to tackle overdetermined ones. Therefore, preconditioning on the level of the design matrix is presented next in order to overcome (i) and (ii).

3.2 Design matrix preconditioning: PCA-LSQR

The principle for design matrix preconditioning is similar to the procedure above. The preconditioner itself is derived from the Cholesky factorization of the block-diagonal approximation $\mathbf{N}_{\mathrm{bd}}$ of the true normal matrix $\mathbf{N}$ subject to

$$\mathbf{N}_{\mathrm{bd}} = \mathbf{L}_{\mathbf{N}}^{T}\mathbf{L}_{\mathbf{N}}. \tag{20}$$

Expanding the left-hand side of Eq. (1) with the unit matrix $\mathbf{I} = \mathbf{L}_{\mathbf{N}}^{-1}\mathbf{L}_{\mathbf{N}}$ yields

$$\mathbf{A}\mathbf{L}_{\mathbf{N}}^{-1}\mathbf{L}_{\mathbf{N}}\mathbf{x} = \mathbf{y} + \mathbf{r}. \tag{21}$$

Further, with the substitutions $\mathbf{A}_{\mathbf{N}} = \mathbf{A}\mathbf{L}_{\mathbf{N}}^{-1}$ and $\mathbf{z} = \mathbf{L}_{\mathbf{N}}\mathbf{x}$, Eq. (21) becomes

$$\mathbf{A}_{\mathbf{N}}\mathbf{z} = \mathbf{y} + \mathbf{r}. \tag{22}$$

Thus, $\mathbf{A}_{\mathbf{N}}$ and $\mathbf{y}$ enter the LSQR procedure to solve the overdetermined LS problem resulting from Eq. (22). The initial vector of unknowns is computed from

$$\hat{\mathbf{x}} = \mathbf{L}_{\mathbf{N}}^{-1}\hat{\mathbf{z}}. \tag{23}$$

Again, the twofold evaluation of the design matrix per iteration can be avoided. With $\mathbf{A}_{\mathbf{N}} = \mathbf{A}\mathbf{L}_{\mathbf{N}}^{-1}$, the bidiagonalization process (refer to Eqs. (4) and (5)) becomes

$$\beta_{k+1}\mathbf{u}_{k+1} = \mathbf{A}\mathbf{L}_{\mathbf{N}}^{-1}\mathbf{v}_k - \alpha_k\mathbf{u}_k, \tag{24}$$

$$\alpha_{k+1}\mathbf{v}_{k+1} = (\mathbf{L}_{\mathbf{N}}^{-1})^{T}\mathbf{A}^{T}\mathbf{u}_{k+1} - \beta_{k+1}\mathbf{v}_k. \tag{25}$$

For practical use, we apply $\mathbf{L}_{\mathbf{N}}^{-1}$ to $\mathbf{v}_k$ before calling $\mathbf{A}$. Then the temporary vector update is similar as described in Sect. 2. Finally, apply $(\mathbf{L}_{\mathbf{N}}^{-1})^{T}$ to $\mathbf{h}_k$. Note that the numerical costs for assembling $\mathbf{N}_{\mathrm{bd}}$ and $\mathbf{L}_{\mathbf{N}}$ have to be taken into account.

4 Subspace preconditioning: SP-LSQR

Preconditioning in terms of subspace splitting is treated in Hanke and Vogel (1999). Its special application to LSQR as outlined in this section is based as illustrated in Jacobsen et al. (2003).

The basic idea is to split the solution space $\mathbb{R}^m$ into two subspaces $\mathbb{W}_1^s \in \mathbb{R}^m$ and $\mathbb{W}_2^{m-s} \in \mathbb{R}^m$ with $s \ll m$ and $\mathbb{W}_1^s \cup \mathbb{W}_2^{m-s} = \mathbb{R}^m$. This kind of subspace preconditioning is also known as the two-level method (Hanke and Vogel 1999), i.e., a special case of multi-level methods.

When assuming the columns of the matrix $\mathbf{W}_1 \in \mathbb{R}^{m \times s}$ spanning the subspace $\mathbb{W}_1^s$ and analogously the columns of the matrix $\mathbf{W}_2 \in \mathbb{R}^{m \times (m-s)}$ spanning the subspace $\mathbb{W}_2^{m-s}$, then the overall solution $\mathbf{x} \in \mathbb{R}^m$ of the minimization problem in Eq. (2) can be formulated as the linear combination

$$\mathbf{x} = \mathbf{W}_1 \mathbf{w}_1 + \mathbf{W}_2 \mathbf{w}_2. \tag{26}$$

where the unknown parameter vectors $\mathbf{w}_1$ and $\mathbf{w}_2$ denote the partial solutions according to the solution space decomposition. Inserting Eq. (26) into Eq. (2) yields

$$\min_{\mathbf{w_1}, \mathbf{w_2}} \left\| \mathbf{A} \left[\mathbf{W}_1 \ \mathbf{W}_2 \right] \begin{bmatrix} \mathbf{w}_1 \\ \\ \mathbf{w}_2 \end{bmatrix} - \mathbf{y} \right\|^2. \tag{27}$$

Next, we perform QR factorization of the product $\mathbf{AW}_1$. The factorization can be written as

$$\mathbf{AW}_1 = \mathbf{Q}\hat{\mathbf{R}} = \left[\mathbf{Y} \ \mathbf{Z} \right] \begin{bmatrix} \mathbf{R} \\ \\ \mathbf{0} \end{bmatrix} = \mathbf{YR}. \tag{28}$$

For an overdetermined system $(n > m)$, $\hat{\mathbf{R}}$ $(n \times s)$ can be split in the upper triangular matrix $\mathbf{R}(s \times s)$ and the zero matrix $\mathbf{0}((n-s) \times s)$ subject to $\hat{\mathbf{R}} = \left[\mathbf{R}^T \ \mathbf{0}^T \right]^T$. Splitting the square orthogonal matrix $\mathbf{Q}(n \times n)$ accordingly, i.e., $\mathbf{Q} = [\mathbf{Y} \ \mathbf{Z}]$ with $\mathbf{Y}(n \times s)$ and $\mathbf{Z}(n \times (n-s))$, results in Eq. (28).

Pre-multiplying Eq. (27) by $\mathbf{Q}^T = [\mathbf{Y} \, \mathbf{Z}]^T$ yields

$$\min_{\mathbf{w}_1, \mathbf{w}_2} \left\| \begin{bmatrix} \mathbf{Y}^T\mathbf{AW}_1 & \mathbf{Y}^T\mathbf{AW}_2 \\ \\ \mathbf{Z}^T\mathbf{AW}_1 & \mathbf{Z}^T\mathbf{AW}_2 \end{bmatrix} \begin{bmatrix} \mathbf{w}_1 \\ \\ \mathbf{w}_2 \end{bmatrix} - \begin{bmatrix} \mathbf{Y}^T\mathbf{y} \\ \\ \mathbf{Z}^T\mathbf{y} \end{bmatrix} \right\|^2. \tag{29}$$

The lower minimization problem in Eq. (29),

$$\min_{\mathbf{w}_2} \|\mathbf{Z}^T\mathbf{AW}_2\mathbf{w}_2 - \mathbf{Z}^T\mathbf{y}\|^2 = \min_{\mathbf{p}} \|\mathbf{Z}^T\mathbf{Ap} - \mathbf{Z}^T\mathbf{y}\|^2, \tag{30}$$

is dependent on $\mathbf{w}_2$ only (or $\mathbf{p}$ with $\mathbf{p} = \mathbf{W}_2\mathbf{w}_2$), as $\mathbf{Z}^T\mathbf{AW}_1 = \mathbf{0}$ holds. The latter is a direct outcome of the QR decomposition in Eq. (28).

The upper minimization problem in Eq. (29),

$$\min_{\mathbf{w}_1, \mathbf{w}_2} \|\mathbf{Rw}_1 + \mathbf{Y}^T\mathbf{AW}_2\mathbf{w}_2 - \mathbf{Y}^T\mathbf{y}\|^2, \tag{31}$$

is dependent on both $\mathbf{w}_1$ and $\mathbf{w}_2$. In Eq. (31), $\mathbf{AW}_1 = \mathbf{YR}$ and thus $\mathbf{Y}^T\mathbf{AW}_1 = \mathbf{R}$ is considered which again results from Eq. (28). If $s \ll m$ holds, Eq. (31) denotes the small subspace problem, which can be solved easily by direct inversion applied to

$$\mathbf{Rw}_1 = \mathbf{Y}^T \left(\mathbf{y} - \mathbf{AW}_2\mathbf{w}_2 \right) = \mathbf{Y}^T \left(\mathbf{y} - \mathbf{Ap} \right). \tag{32}$$

Thus, the main computational effort occurs for solving the minimization problem in Eq. (30), which is performed here by LSQR. The critical part of the algorithm is the QR decomposition in Eq. (28). Even when storing and applying the orthogonal matrix $\mathbf{Q}$, or $\mathbf{Z}$, by means of successive Householder transformations, the effort for the decomposition itself as well as the application of $\mathbf{Q}$ in each iteration increases very fast when expanding the subspace dimension of $\mathbb{W}_1$. Thus, only if $s \ll m$ holds, the effort remains considerably small.

Subspace preconditioning modifies the original LSQR algorithm concerning the Lanzcos bidiagonalization process. Actually, the minimization problem in Eq. (2) is manipulated by replacing $\mathbf{A}$ by $\mathbf{Z}^T\mathbf{A}$ and $\mathbf{y}$ by $\mathbf{Z}^T\mathbf{y}$. The latter substitution is not problematic. The transformation of the observation vector can be performed in the initialization step of the algorithm. However, the manipulation of the design matrix induces the bidiagonalization process to be modified according to

$$\beta_{k+1}\mathbf{u}_{k+1} = \mathbf{Z}^T\mathbf{A}\mathbf{v}_k - \alpha_k\mathbf{u}_k, \tag{33}$$

$$\alpha_{k+1}\mathbf{v}_{k+1} = \mathbf{A}^T\mathbf{Z}\mathbf{u}_{k+1} - \beta_{k+1}\mathbf{v}_k. \tag{34}$$

Unfortunately, the principle to avoid the assembly of the design matrix twice per iteration can not be transferred to SP-LSQR (subspace preconditioned LSQR). For calculating the element u^j_{k+1}, the whole matrix $\mathbf{A}$ has to be processed due to the multiplication of $\mathbf{Z}^T$ from the left-hand side in Eq. (33). Thus, the temporary vector $\mathbf{h}_k$ can not be updated as outlined in Sect. 2 by calling the j^{th} row $\mathbf{A}(j,:)$ only once each iteration. This is the major drawback of the method assuming the setup of $\mathbf{A}$ is a costly process.

To demonstrate the application of subspace preconditioning in gravity field recovery, we consider the simplest case of subspace splitting, i.e., $\mathbf{W}_1 = \left[\mathbf{I}_{(s\times s)}\,\mathbf{0}_{(s\times(m-s))}\right]^T$ and $\mathbf{W}_2 = \left[\mathbf{0}_{(s\times s)}\,\mathbf{I}_{(s\times(m-s))}\right]^T$. Depending on the sequence of the unknowns assembled in the parameter vector $\mathbf{x}$ (by degree or by order), a low-degree, respectively low-order, solution is obtained by direct inversion of the small system.

The subspace solver is used to accelerate the iterative process of the large system. Note that the subspace method can be additionally accelerated by transformation of the unknown parameter vector in terms of block-diagonal preconditioning as illustrated in Sect. 3. This requires the algorithm to be adapted accordingly and is not addressed in this contribution.

For our application of subspace preconditioning in satellite geodesy, we have chosen the subspace $\mathbb{W}_1$ to constitute the solution space of the long-wavelength part of the gravitational spectrum. This is only one of other different possibilities, but in general it seems to be reasonable to shift a certain frequency domain of the gravitational spectrum to the subspace $\mathbb{W}_1$.

5 Regularization of LSQR, PCA-LSQR and SP-LSQR

Regularization is a commonly used tool to overcome the instability of ill-posed inversion problems. Ill-posedness means that the solution vector $\hat{\mathbf{x}}$ does not continuously depend on the given data $\mathbf{y}$ (Phillips 1962, Tikhonov 1963). Different types of regularization have been applied for various applications such as the method of Tikhonov-Phillips, truncated singular value decomposition (Hansen 1987) or iterative regularization methods. As already mentioned, here we exclusively focus on Tikhonov-Phillips regularization.

5.1 Regularization of LSQR: R-LSQR

Extending Eq. (2) in terms of the auxiliary condition of minimizing the functional $\min_{\mathbf{x}} \|\mathbf{x}\|^2$, the regularized LSQR minimization problem according to Paige and Saunders (1982b) reads

$$\min_{\mathbf{x}} \left\{ \|\mathbf{A}\mathbf{x} - \mathbf{y}\|^2 + \kappa \|\mathbf{x}\|^2 \right\} = \min_{\mathbf{x}} \left\| \begin{bmatrix} \mathbf{A} \\ \sqrt{\kappa}\mathbf{I} \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathbf{y} \\ \mathbf{0} \end{bmatrix} \right\|^2 \tag{35}$$

with the regularization parameter κ and the regularization matrix $\mathbf{K} = \mathbf{I}$.

Equation (35) is known as regularization in standard form. The relative balance of the terms $\|\mathbf{A}\mathbf{x} - \mathbf{y}\|$ and $\|\mathbf{x}\|$ depends on the size of κ. If the regularization parameter is chosen too small, Eq. (35) is close to the ill-conditioned minimization problem. On the other hand, if κ is too large, the solution of Eq. (35) might be far away from the problem to be solved originally. Thus, an appropriate balancing is of fundamental importance.

As the optimal regularization parameter κ_{opt} is usually not known a priori, it is advantageous to solve the linear LS problem for various regularization parameters κ_i; $i = 1, 2, ..., i_{\max}$. The optimal one is chosen according to some quality criterion of the estimate $\hat{\mathbf{x}}_{\kappa_i}$, cf., e.g., Hanke and Hansen (1993), Kusche and Klees (2002) or Ditmar et al. (2003b). In fact, the determination of κ_{opt} is a delicate topic.

Due to the extended minimization problem in Eq. (35), $\mathbf{B}_{k,\kappa_i}$ changes from a bidiagonal to a tridiagonal matrix. The secondary lower sub-diagonal throughout contains the value $\sqrt{\kappa_i}$. Only the QR decomposition is affected by regularization but not the Lanczos step, where $\mathbf{a}_{k,\kappa_i}$ and $\mathbf{x}_{k,\kappa_i}$ is derived. Thus, the additional computational effort constitutes $i_{\max}$ Givens rotations in each iteration, which is marginal even for a large number of regularization parameters.

For regularization with matrix $\mathbf{K} \neq \mathbf{I}$, referred to as regularization in general form, it is convenient to transform the extended minimization problem

$$\min_{\mathbf{x}} \left\{ \|\mathbf{A}\mathbf{x} - \mathbf{y}\|^2 + \kappa \|\mathbf{x}\|_{\mathbf{K}}^2 \right\} \tag{36}$$

to standard form. With Eq. (36), the extended normal equation system (cf. Eq. (13)) reads

$$(\mathbf{A}^T\mathbf{A} + \kappa\mathbf{K})\hat{\mathbf{x}}_\kappa = \mathbf{A}^T\mathbf{y}. \tag{37}$$

Introducing the Cholesky factorization of the quadratic matrix $\mathbf{K}$ with $\mathbf{K} = \mathbf{L}_\mathbf{K}^T\mathbf{L}_\mathbf{K}$, applying $(\mathbf{L}_\mathbf{K}^{-1})^T$ from the left-hand side to Eq. (37) and substituting the unknown parameter vector $\hat{\mathbf{x}}_\kappa$ according to $\hat{\mathbf{c}}_\kappa = \mathbf{L}_\mathbf{K}\hat{\mathbf{x}}_\kappa$ yields

$$((\mathbf{L}_\mathbf{K}^{-1})^T\mathbf{A}^T\mathbf{A}\mathbf{L}_\mathbf{K}^{-1} + \kappa\mathbf{I})\hat{\mathbf{c}}_\kappa = (\mathbf{L}_\mathbf{K}^{-1})^T\mathbf{A}^T\mathbf{y}. \tag{38}$$

Finally, the substitution $\mathbf{A}_\mathbf{K} = \mathbf{A}\mathbf{L}_\mathbf{K}^{-1}$ results in

$$\min_\mathbf{c} \left\| \begin{bmatrix} \mathbf{A}_\mathbf{K} \\ \sqrt{\kappa}\mathbf{I} \end{bmatrix} \mathbf{c} - \begin{bmatrix} \mathbf{y} \\ \mathbf{0} \end{bmatrix} \right\|^2. \tag{39}$$

Equation (38) transforms the general regularization problem (Eq. 36) to its standard form in Eq. (39). Thus, the properties of Eq. (35) solving for $\mathbf{x}$ hold for Eq. (39) when solving for $\mathbf{c}$. An estimate of the original parameter vector is obtained by back substitution subject to $\hat{\mathbf{x}}_\kappa = \mathbf{L}_\mathbf{K}^{-1}\hat{\mathbf{c}}_\kappa$. Note that for Tikhonov-Phillips regularization, $\mathbf{K}$ typically is a diagonal matrix. Thus, $\mathbf{L}_\mathbf{K} = \mathbf{L}_\mathbf{K}^T$ holds with $L_{ii} = \sqrt{K_{ii}}$.

5.2 Regularization of PCA-LSQR: PCAR-LSQR

Two possibilities for the combination of block-diagonal preconditioning as outlined in Sect. 3.2 and regularization seem to be straightforward, i.e., (i) applying preconditioning to the regularized problem, or (ii) applying regularization to the preconditioned problem. In fact, the two approaches are equivalent as will be shown next.

Extending Eq. (38) by preconditioning subject to $\mathbf{I} = \mathbf{L}_{\bar{\mathbf{N}}}^{-1}\mathbf{L}_{\bar{\mathbf{N}}}$ yields

$$((\mathbf{L}_\mathbf{K}^{-1})^T\mathbf{A}^T\mathbf{A}\mathbf{L}_\mathbf{K}^{-1} + \kappa\mathbf{I})\mathbf{L}_{\bar{\mathbf{N}}}^{-1}\mathbf{L}_{\bar{\mathbf{N}}}\hat{\mathbf{c}}_\kappa = (\mathbf{L}_\mathbf{K}^{-1})^T\mathbf{A}^T\mathbf{y} \tag{40}$$

with $\mathbf{L}_{\bar{\mathbf{N}}} = \mathbf{L}_\mathbf{N}\mathbf{L}_\mathbf{K}^{-1}$. Using $\hat{\mathbf{c}}_\kappa = \mathbf{L}_\mathbf{K}\hat{\mathbf{x}}_\kappa$, Eq. (40) can be rewritten as

$$((\mathbf{L}_\mathbf{K}^{-1})^T\mathbf{A}^T\mathbf{A}\mathbf{L}_\mathbf{K}^{-1} + \kappa\mathbf{I})\mathbf{L}_\mathbf{K}\mathbf{L}_\mathbf{N}^{-1}\mathbf{L}_\mathbf{N}\hat{\mathbf{x}}_\kappa = (\mathbf{L}_\mathbf{K}^{-1})^T\mathbf{A}^T\mathbf{y}. \tag{41}$$

Equation (41) represents the preconditioned and regularized normal equation system. In order to transform it to standard form, the substitution $\hat{\mathbf{d}}_\kappa = \mathbf{L}_\mathbf{K}\mathbf{L}_\mathbf{N}^{-1}\mathbf{L}_\mathbf{N}\hat{\mathbf{x}}_\kappa = \mathbf{L}_\mathbf{K}\hat{\mathbf{x}}_\kappa = \hat{\mathbf{c}}_\kappa$ seems to be straightforward. But then we return to Eq. (38), i.e., preconditioning vanishes.

The alternative option is to introduce the auxiliary parameter vector according to $\hat{\mathbf{d}}_\kappa^* = \mathbf{L}_\mathbf{N}\hat{\mathbf{x}}_\kappa$. To ensure the symmetry of the normal matrix, in this case Eq. (41) has to be pre-multiplied with $(\mathbf{L}_\mathbf{K}\mathbf{L}_\mathbf{N}^{-1})^T$. With $\mathbf{A}_\mathbf{N} = \mathbf{A}\mathbf{L}_\mathbf{N}^{-1}$, this option leads to

$$(\mathbf{A}_\mathbf{N}^T\mathbf{A}_\mathbf{N} + \kappa(\mathbf{L}_\mathbf{N}^{-1})^T\mathbf{K}\mathbf{L}_\mathbf{N}^{-1})\hat{\mathbf{d}}_\kappa^* = \mathbf{A}_\mathbf{N}^T\mathbf{y} \tag{42}$$

or equivalently to

$$\min_{\mathbf{d}^*} \left\| \begin{bmatrix} \mathbf{A_N} \\ \sqrt{\kappa}(\mathbf{L_N^{-1}})^T\mathbf{K}\mathbf{L_N^{-1}} \end{bmatrix} \mathbf{d}^* - \begin{bmatrix} \mathbf{y} \\ \mathbf{0} \end{bmatrix} \right\|^2. \tag{43}$$

Equation (43) coincides with the one that is directly obtained by applying regularization to the preconditioned problem. The drawback of this approach is that it results in a minimization problem with regularization in general form. Thus, in the current implementation, the combination of general-form regularization and block-diagonal preconditioning can not be realized. Further investigations are necessary to overcome the combination difficulties.

5.3 Regularization of SP-LSQR: SPR-LSQR

Regularization can be added without any problems to the SP-LSQR procedure. Expanding Eq. (35) in terms of regularization in general form yields

$$\min_{\mathbf{x}} \left\| \begin{bmatrix} \mathbf{A} \\ \sqrt{\kappa}\mathbf{K} \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathbf{y} \\ \mathbf{0} \end{bmatrix} \right\|^2. \tag{44}$$

With the substitutions $\tilde{\mathbf{A}} = \begin{bmatrix}\mathbf{A}^T & \sqrt{\kappa}\mathbf{K}^T\end{bmatrix}^T$ and $\tilde{\mathbf{y}} = \begin{bmatrix}\mathbf{y}^T & \mathbf{0}^T\end{bmatrix}^T$, Eq. (44) results in the stacked minimization problem (Jacobsen et al. 2003)

$$\min_{\mathbf{x}} \|\tilde{\mathbf{A}}\mathbf{x} - \tilde{\mathbf{y}}\|^2 \tag{45}$$

which can be solved according to the achievements in Sect. 4, in particular Eqs. (26) to (32). All that has to be considered is replacing $\mathbf{A}$ by $\tilde{\mathbf{A}}$ and $\mathbf{y}$ by $\tilde{\mathbf{y}}$.

In Eq. (45), the leading dimension of the stacked arrays is $(n + c)$ with n the row-space dimension of $\mathbf{A}$ and c the number of additional conditions due to regularization. Thus, the leading dimension of the design matrix is enlarged compared to the original formulation. If $c = m$ holds, i.e., c equal to the leading dimension of $\mathbf{x}$, regularization is performed for the whole unknown parameter vector.

6 Methodology for gravity field recovery

In the context of gravity field recovery, we focus on the GOCE mission principle (ESA 1999). The positions of the spacecraft are deduced from GPS (Global Positioning System) measurements between the low-orbiting GOCE spacecraft ($\approx 250\,$km) and the high-flying GPS satellites ($\approx 20,000\,$km). This observation configuration is referred to as high-low satellite-to-satellite tracking (hl-SST). It is restricted to recover long-wavelength

features only. For high-resolution geopotential determination, in the framework of the GOCE mission, satellite gravity gradiometry (SGG) will be realized for the first time.

The combination of both techniques leads to the best possible GOCE-only gravity field solution (ESA 1999). It can be be performed by superposition of the observation systems of equations in a common LS adjustment. For proper weighting of different observation groups, variance component estimation using Monte-Carlo techniques has been proven to perform very well in a simulation environment (Kusche 2003). We investigate the hl-SST and SGG observation scenario separately from each other, to provide several examples for the use of the LSQR method in satellite geodesy. The achievements made within the scope of these case studies can be transferred to comparable space-geodetic experiments.

Gravitational field modeling in terms of a spherical harmonic series expansion of the potential function $V(\lambda, \varphi, r)$ is given with Eq. (46). Both the attenuation factor $(R/r)^{l+1}$ and the 4π-normalized surface spherical harmonics $\bar{Y}_{l,m}(\lambda, \varphi)$, cf. Eq. (47), depend on the spherical coordinates (λ, φ, r). Therein, λ denotes East longitude, φ latitude and r the radial distance from the origin. Further, $\bar{P}_{l,m}(\sin\varphi)$ are the fully normalized associated Legendre functions of the first kind.

$$V(\lambda, \varphi, r) = \frac{GM}{R} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left(\frac{R}{r}\right)^{l+1} \bar{Y}_{l,m}(\lambda, \varphi)\, \bar{v}_{l,m} \tag{46}$$

$$\bar{Y}_{l,m}(\lambda, \varphi) = \begin{cases} \bar{P}_{l,m}(\sin\varphi)\cos m\lambda & 0 \leq m \leq l \\ \bar{P}_{l,|m|}(\sin\varphi)\sin|m|\,\lambda & -l \leq m < 0 \end{cases} \tag{47}$$

In theory, an infinite series is necessary to fully characterize the geopotential, but practically the double sum in Eq. (46) is truncated at a maximum degree L determined by parameters of the satellite mission and the desired resolution of the field. Both the geocentric constant GM and the major semi-axis R of a reference ellipsoid are fixed. The series coefficients $\bar{v}_{l,m}$ are unknown parameters. Their estimation can be performed best by globally distributed observation data such as provided by GOCE.

The orbit of the spacecraft is tracked continuously by GPS. Considering the satellite free-falling around the Earth, the hl-SST observation equation can be formulated in a simple way. According to Eq. (48), the acceleration of the satellite (reduced by all disturbing effects such as tidal forces) is equal to the terrestrial attraction, i.e., the gradient of the Earth's gravitational potential:

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathbf{x}\,(\lambda(t), \varphi(t), r(t)) = \ddot{\mathbf{x}}(t) = \mathbf{e}_i a_i = \nabla\, V(\lambda, \varphi, r). \tag{48}$$

where a_i denote the components of the gravitational acceleration vector. They refer to the orthonormal base vectors $\mathbf{e}_i$ with $\mathbf{e}_1$ and $\mathbf{e}_2$ in direction tangential to the parameter lines of the sphere and $\mathbf{e}_3$ directed radially outwards.

The functional model in Eq. (48) is referred to as the acceleration approach. It balances kinematic orbit information to gravitational features. In particular, numerical differentiation techniques provide satellite accelerations $\ddot{\mathbf{x}}(t)$ derived from the observed positions $\mathbf{x}(t)$. The approach has been successfully applied to real CHAMP data analysis in Reubelt et al. (2005). Moreover, Baur and Grafarend (2006) investigate hl-SST analysis of simulated GOCE data.

SGG observations correspond to second-order derivatives of the geopotential in Eq. (46). They are by far more sensitive to short-wavelength features than gravitational accelerations (first-order derivatives). Application of the gradient operator to Eq. (48) yields

$$\nabla\nabla V(\lambda, \varphi, r) = \mathbf{e}_i \otimes \mathbf{e}_j V_{ij}. \tag{49}$$

Equation (49) results in an analytical expression of the gravitational tensor in terms of the gravity field model parameters $\bar{v}_{l,m}$. Each single component V_{ij}, referred to as gravity gradient (GG), constitutes one type of observation. Exemplary for the radial component V_{33}

$$V_{33} = \frac{GM}{R^3} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left(\frac{R}{r}\right)^{l+3} (l+2)(l+1) \bar{Y}_{l,m}(\lambda, \varphi) \bar{v}_{l,m} \tag{50}$$

holds.

Commonly, SGG analysis is performed by analyzing individual GGs, in particular the main diagonal elements of the gravitational tensor. Details can be found in, e.g., Schuh (1996), Rummel et al. (1993), Klees et al. (2000), Sneeuw (2000) and Pail and Plank (2002). A completely different approach is based on the rotational invariants of the gravitational tensor (Rummel 1986). The benefit of this method arises from its independence of the gradiometer instrument orientation in space. The invariants approach has been adopted to the GOCE mission scenario in Baur and Grafarend (2006) and Baur et al. (submitted).

From the numerical point of view, Eqs. (48) and (49) constitute, in each case, a linear system of equations with the unknown parameter vector $\mathbf{x} = [\bar{v}_{2,0}, ..., \bar{v}_{L,0}, \bar{v}_{2,1}, ..., \bar{v}_{L,L}, \bar{v}_{2,-1}, ..., \bar{v}_{L,-L}]^T$, the vector of observations $\mathbf{y} = [a_1(t_1), a_2(t_1), a_3(t_1), ..., a_3(t_n)]^T$, or $\mathbf{y} = [V_{11}(t_1), V_{12}(t_1), ..., V_{33}(t_1), ..., V_{33}(t_n)]^T$, and the design matrix $\mathbf{A}$ describing the functional relationship between them. Note that in satellite geodesy, typically the number of observations n is by far larger than the number of unknowns m to be resolved, i.e., $n \gg m$ holds yielding an overdetermined system.

7 Numerical examples

To demonstrate the feasibility and benefit of LSQR preconditioning and regularization, numerical simulation studies based on a 30-days GOCE-like synthetic test data set with a sampling rate of $\Delta t = 5\,\mathrm{s}$ is investigated. The simulation is performed up to degree and order $L = 300$ using the EGM96 gravity field model (Lemoine et al. 1998). The synthetic data set is a result of the IAG Special Commission 7 (SC7) activities (Ilk et al. 2003). Additionally, for hl-SST analysis, the satellite position coordinates $x_i;\ i = 1, 2, 3$ have been contaminated with a normal distributed random noise sequence $\varepsilon_{x_i} \sim N(0, \sigma_{x_i})$ with standard deviation $\sigma_{x_i} = 1\,\mathrm{cm}$.

Colored SGG noise is generated by a power spectral density model with a flat part of 3–4 mE $\mathrm{Hz}^{-\frac{1}{2}}$ in the measurement bandwidth (MBW) ranging from 5 mHz to 0.1 Hz (ESA 1999). Although the noise characteristic of the GOCE gradiometer has been updated due to several changes of the mission design (such as the abolition of the field emission electronic propulsion thrusters), we stick here to the original stochastic properties. For noisy SGG data analysis, decorrelation is applied as outlined in Sect. 2.1. Since the gradiometer reference frame axes of data simulation are defined according to “along-track”, “cross-track” and “quasi-radial”, we transform the functional model in Eq. (49) into the orbit frame.

Within this contribution, we restrict the attention on the radial hl-SST observation a_3 and the quasi-radial SGG measurement, denoted as V_{33}^q. Moreover, the reference gravity field model OSU86F (Rapp and Cruz 1986) complete to full analysis degree and order L has been reduced from the data in advance. Thus, we estimate corrections to a set of reference parameters rather than absolute values. The choice of the reference model is not critical at all in our closed-loop simulation study.

We perform the calculations on a 64-processor cc-NUMA supercomputer using OpenMP for parallelization. The platform is part of an SGI Altix 3700 system, consisting of 416 Intel Itanium2 $1.3\,\mathrm{GHz}$ CPUs, 832 GByte of memory and 2.8 TByte of scratch disk space. The total peak performance of the system is $2.2\,\mathrm{TFLOPS}$.

For LSQR implementation, the main effort occurs for design matrix assembly. Since each observation, i.e., each row of $\mathbf{A}$, can be treated separately, the parallelization of the algorithm is realized by distributing the number of observations to the individual CPUs. Thus, the parallel region covers the successive design matrix decomposition which is split (uniformly) among the processing elements used. Moreover, in case of PCA-LSQR, the computation of the design matrix preconditioner is performed in parallel.

Table 1. WRMS values (cm) of (quasi-)geoid height differences (computed on a global 1° grid) with respect to EGM96 neglecting low-order coefficients: $L = 100 : m_{\mathrm{thres}} = 10$; $L = 200 : m_{\mathrm{thres}} = 20$; $L = 300 : m_{\mathrm{thres}} = 30$

Type of observation	Resolution L	Considered resolution	Noise-free data	Noisy data
hl-SST	100	75	4.6	38.3
SGG	100	100	5.9	10.2
	200	200	2.0	6.1
	300	230	$1.1 \cdot 10^{-2}$	13.3

7.1 Analysis results

In a first study, we perform both hl-SST and SGG data analysis of noise-free observation data without using regularization. Thus, the model parameter estimates are affected by spectral leakage (and to a minor part by aliasing) effects only. Their influence increases as the maximum degree L decreases, since the unmodeled signal content of the data is mapped to the coefficients resolved. This behavior is proved by the third column in Table 1.

Table 1 presents latitude-weighted root mean square (WRMS) values of (quasi-)geoid height differences with respect to EGM96, approximated by

$$h_i = R \sum_{l=2}^{L} \sum_{m=0}^{l} \bar{Y}_{l,m}(\lambda, \varphi) \Delta\bar{v}_{l,m}. \tag{51}$$

The $\Delta\bar{v}_{l,m}$ indicate the residual geopotential coefficients subject to $\Delta\bar{v}_{l,m} = \hat{\bar{v}}_{l,m} + \bar{v}_{l,m}^{\mathrm{OSU86F}} - \bar{v}_{l,m}^{\mathrm{EGM96}}$ and $\hat{\bar{v}}_{l,m}$ the elements of $\hat{\mathbf{x}}$. Further

$$\mathrm{WRMS}_h^k = \sqrt{\frac{\sum_{i=1}^{N}((h_i^k)^2 \cos\varphi_i)}{\sum_{i=1}^{N} \cos\varphi_i}} \tag{52}$$

holds for the evaluation of the WRMS with N being the number of evaluation points considered, preferably distributed on a regular grid of, e.g., 1°.

Dependent on the satellite mission design, it is reasonable to leave out polar areas for the evaluation of Eq. (52). In particular, the GOCE satellite ground-track will have a coverage of $|\varphi_{\mathrm{GOCE}}| \approx 83^\circ$. Figure 2 presents the characteristic of (quasi-)geoid height errors (RMS values per circle of latitude) of the parameter estimate relative to the EGM96 model. The RMS values increase with increasing latitude. For $|\varphi| > \varphi_{\mathrm{GOCE}}$ the values are not representative at all.

The low-order coefficients reflect the properties in the polar regions. Thus, instead of evaluating Eq. (52) around a spherical belt only, alternatively the spherical harmonics with order $m < m_{\mathrm{thres}}$ can be neglected for WRMS_h^k computation. Van Gelderen and Koop (1997) deduced a degree-dependent rule of thumb for the

non-resolvable maximal order m_{thres} subject to $m_{\mathrm{thres}} \approx l \mid \frac{\pi}{2} - I \mid$ with I as the inclination of the satellite orbit. For the computation in Table 1, we neglect low orders $m_{\mathrm{thres}} = 10\,(L = 100)$, $m_{\mathrm{thres}} = 20\,(L = 200)$ and $m_{\mathrm{thres}} = 30\,(L = 300)$ respectively.

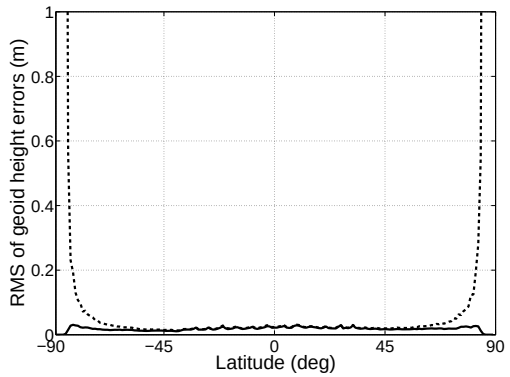

Fig. 2. RMS values (m) of (quasi-)geoid height errors per circle of latitude (SGG, $L = 200$). Solid line: $m_{\mathrm{thres}} = 20$, dashed line: $m_{\mathrm{thres}} = 0$

Next we analyze noisy data, again without incorporating regularization. Figures 3 and 4 illustrate degree error RMS (DE-RMS) values of the coefficient estimates. One month of observation data allows hl-SST analysis approximately up to degree and order 80, whereas ≈ 230 is reached for SGG. This agrees with what is currently expected for the GOCE mission (Pail 2005). Note that the DE-RMS curves are truncated at the points where they cross the EGM96 signal curve. These points we refer to as the "considered resolution". For WRMS computation in the fourth column of Table 1, only the coefficients up to the "considered resolution" have been taken into account. Higher degrees can not be resolved.

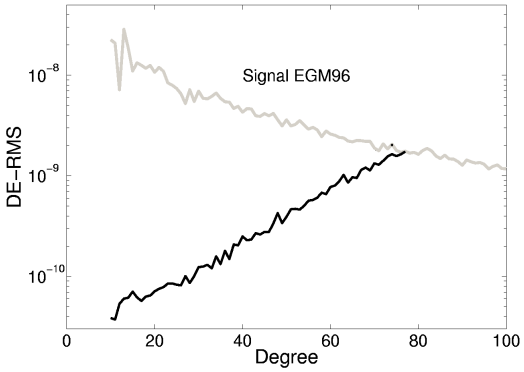

Fig. 3. Analysis results of noisy hl-SST data in terms of DE-RMS values, $L = 100$

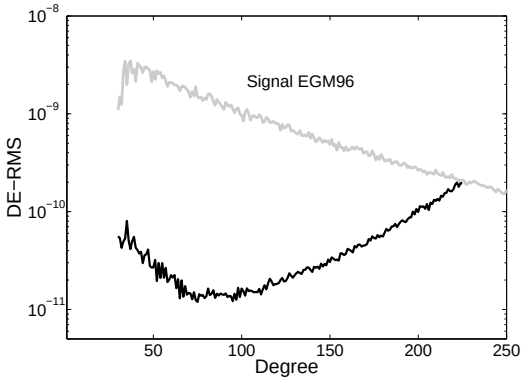

Fig. 4. Analysis results of noisy SGG data in terms of DE-RMS values, $L = 300$

7.2 Truncation criterion

With regard to an appropriate stopping rule for the iterative LSQR process, different truncation criteria can be applied, dependent on the philosophy for truncation. Mathematical stopping rules (Paige and Saunders 1982a) are mainly based on the magnitude of the residual vector $\|\hat{\mathbf{r}}_k\| = \|\mathbf{A}\hat{\mathbf{x}}_k - \mathbf{y}\|$, which is provided in each iteration k thus requiring no additional computational costs. The process is truncated if $\|\hat{\mathbf{r}}_k\|$, or a functional of it, falls below a pre-defined threshold. Thus, the termination is related to mathematical properties of the method itself, but can not serve any physical interpretation to the problem treated. A completely different idea offers the cross validation (CV) functional, an approach that is addressed in Björck (1996), Kusche and Mayer-Gürr (2001) and Ditmar et al. (2003b).

We have chosen a truncation criterion apart from the methods above. It is related to the physical correspondence of the estimate in terms of a gravity field functional. In particular, process termination is coupled to the difference $\Delta\mathrm{WRMS}_h^{k,k-1} = \mathrm{WRMS}_h^k - \mathrm{WRMS}_h^{k-1}$, which is evaluated at the end of each iteration k. If the value for $\Delta\mathrm{WRMS}_h^{k,k-1}$ falls below the pre-defined truncation threshold δ, i.e., if $\Delta\mathrm{WRMS}_h^{k,k-1} < \delta$ holds, the final estimate is found. However, due to the non-monotonic convergence of the LSQR method (cf. Fig. 5), we recommend to continue the process for a few iterations beyond to ensure having really reached the truncation threshold δ.

7.3 Block-diagonal preconditioning

Table 2 presents runtime results for the basic LSQR implementation using eight CPUs. The truncation threshold is fixed conservatively to $\delta = 0.25\,\mathrm{mm}$. The values displayed in Table 2 correspond to the number of iterations necessary to finally reach δ, i.e., the runtime for additional iterations beyond the threshold is not considered. For hl-SST analysis with $L = 100$, convergence is achieved after 53 iterations. Approximately the same num-

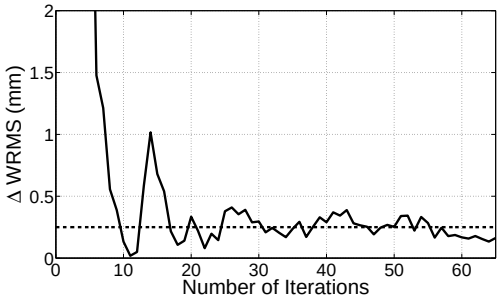

Fig. 5. Convergence behavior of LSQR (SGG, $L = 100$)

Table 2. Runtime results LSQR (8 CPUs)

Type of observation	Resolution L	Number of iterations	Wall time (h)
hl-SST	100	53	1.2
SGG	100	56	1.1
	200	180	13.4

Table 3. Runtime results PCA-LSQR (8 CPUs)

Type of observation	Resolution L	Number of iterations	Wall time total (h)	Wall time $\mathbf{L_N^{-1}}$ (h)
hl-SST	100	6	0.4	0.2
SGG	100	8	0.4	0.2
	200	14	2.6	1.4
	250	19	5.1	2.6
	300	32	14.4	4.8

ber holds for the equivalent SGG analysis. However, the number of iterations increases to 180 for a spectral resolution of $L = 200$.

Block-diagonal preconditioning according to Sect. 3 has been implemented and successfully parallelized. Runtime decreases significantly since the number of iterations drops dramatically compared to the basic algorithm. According to Table 3, for $L = 100$ convergence is achieved within less than 10 iterations. Runtime is approximately one third compared to Table 2. Regarding $L = 200$, the number of iterations drops from 180 to 14. Runtime is reduced to $19.4\,\%$ compared to basic LSQR.

Note that the computational costs within each iteration are comparable for LSQR and PCA-LSQR. The additional operations of applying $\mathbf{L_N^{-1}}$ and $(\mathbf{L_N^{-1}})^T$ to a vector in the PCA-LSQR case have no significant runtime effect. However, the setup of the (inverse) preconditioner itself is a quite costly process. Table 3 shows the part for $\mathbf{L_N^{-1}}$ calculation opposite to the total runtime. Roughly spoken, half the computing time is required for $\mathbf{L_N^{-1}}$

Table 4. Runtime results PCA-LSQR (SGG, $L = 200$)

Number CPUs	Iterations	Wall time (h)	Speed-up
1	14	19.68	1
8	14	2.59	7.6
16	14	1.38	14.3
32	14	0.72	27.3
64	14	0.41	48.0

Table 5. Runtime results SP-LSQR (hl-SST, $L = 100$, 1 CPU)

L_{SP}	Number iterations	Wall time total (h)	Wall time QR decomposition (h)
- (LSQR)	53	9.9	-
20	33	7.0	0.2
50	17	11.4	6.8

computation. This can be improved by the approximate calculation of the block-diagonal preconditioner, but is not considered here.

Moreover, exemplary for all scenarios, Table 4 displays runtime scaling dependent on the number of processing units. The speed-up S_p is defined as the ratio between serial runtime T_1 and the runtime achieved using p processors, denoted as T_p. Thus,

$$S_p = \frac{T_{1\,\mathrm{CPU}}}{T_{p\,\mathrm{CPUs}}} \tag{53}$$

holds. In the optimal case, the speed-up is equal to p.

This is not realizable in practical applications due to communication and synchronization reasons. However, the speed-up value is an appropriate quality measure of the parallel implementation. According to Table 4, for our implementation, the speed-up is largely close to the number of CPUs used. This demonstrates the efficient and powerful parallelization of the algorithm.

7.4 Subspace preconditioning

Table 5 summarizes runtime results for SP-LSQR. The subspace $\mathbb{W}_1$ has been chosen to constitute the solution space of the long-wavelength part of the gravitational spectrum with maximal resolution L_{SP}. With increasing subspace dimension, the number of iterations drops accordingly. Note that if the subspace dimension is zero, SP-LSQR equals the LSQR method from the methodological point of view.

On the other hand, the effort for preconditioning itself, in particular the QR decomposition according to Eq. (28), increases with increasing L_{SP}. For this reason, the wall time measurements in Table 5 do not scale

Table 6. Runtime results SP-LSQR (hl-SST, $L = 100$, $L_{\mathrm{SP}} = 50$)

Number CPUs	Wall time 1 iteration (s)	Speed-up 1 iteration	Wall time QR decomp. (h)
1	904	1	6.8
16	428	2.1	4.2

with the number of iterations. Even a deceleration can be observed for $L_{\mathrm{SP}} = 50$ compared to the LSQR version without preconditioning. For the former, more than half the runtime is required for the QR decomposition procedure.

From the computational point of view, subspace preconditioning turns out to be less effective than block-diagonal preconditioning for two reasons. Firstly, as already mentioned in Sect. 4, the design matrix $\mathbf{A}$ has to be assembled twice for each iteration. Secondly, the runtime effort for QR decomposition using LAPACK (linear algebra package) routines (Anderson et al. 1999) increases dramatically with increasing subspace dimension. Moreover, in the current implementation the problem concerning the QR decomposition also affects the speed-up of the method using different numbers of CPUs. This is illustrated in Table 6 comparing one to 16 CPUs. The speed-up of one iteration has the value of only 2.1.

Besides the computational drawbacks of SP-LSQR, there is also a methodological one for the scenarios investigated. Block-diagonal preconditioning performs much better than the subspace method. For hl-SST analysis, only six iterations are necessary using PCA-LSQR to reach the final estimate, whereas SP-LSQR does not fall below 17 iterations according to the scenarios in Table 5. This is mainly due to the special problem design considered, i.e., a satellite orbit configuration close to the ideal conditions as outlined in Sect. 1.

7.5 Regularization

Tikhonov-Phillips regularization has been added successfully to the implementations. For R-LSQR, multiple regularization parameter treatment has no significant effect on the overall runtime, cf. Table 7. The additional costs considering 100 regularization parameters opposite to the unregularized method are less than $1\,\%$.

However, to find the "best" regularization parameter amongst the a priori values, some quality criterion has to be evaluated for each of them (the corresponding runtime is neglected in Table 7). Depending on the quality measure, this is a more or less costly process and in general the procedure has to be repeated for each single regularization parameter. In the case studies presented, we performed global WRMS evaluation of successive estimates. This was done on a regular grid of 1° resolution in latitude and 10° resolution in longitude. For a

Table 7. Runtime results R-LSQR (SGG, $L = 200$, 8 CPUs, 1 iteration)

Number regularization param.	Wall time (s)	Wall time (%)
no regularization	242.7	100.0
1	242.7	100.0
10	243.6	100.4
100	244.9	100.9

spectral resolution of $L = 200$ and eight-CPUs parallel processing, truncation criterion evaluation requires 20 s wall time for each regularization parameter κ_i.

8 Conclusions

In this contribution, we investigated and applied the iterative LSQR algorithm for satellite-based gravity field recovery. In particular, we performed hl-SST and SGG data analysis of a GOCE-like simulation scenario covering one month of observation data with 5 s sampling.

The algorithm is characterized by transformation of the original LS problem to a bidiagonal subproblem, which is solved in terms of a QR decomposition. In general, iterative solvers have only small memory requirements since they avoid design and normal matrix storage. Moreover, they are largely independent of the processing platform, i.e., they are very well suited for parallel computing. However, opposite to brute-force LS solvers, they do not provide the variance-covariance matrix of the parameter estimate. This is the major drawback of iterative methods since the approximated computation of error estimates never yields exact information.

For space-geodetic applications, the basic LSQR algorithm according to Paige and Saunders (1982a;b) turns out to be not very effective. In terms of LSQR tuning, we addressed (i) preconditioning, (ii) regularization and (iii) parallel implementation.

Block-diagonal preconditioning has been proven to clearly outperform subspace preconditioning for two reasons. Firstly, for the subspace method the twofold assembly of the design matrix in each iteration is unavoidable. Secondly, the numerical costs for the QR decomposition increase dramatically with increasing subspace dimension, cf. Table 5. On the other hand, design matrix block-diagonal preconditioning performs very well. According to Tables 2 and 3, the number of iterations reduces by a factor of around 10. Due to the additional effort for the computation of the preconditioner, the wall time reduces by a factor of 3–5. Actually, preconditioner computation takes around half of the overall runtime.

Regularization in general form has been added to the original LSQR method by transformation of the extended minimization problem to standard form. Compared to the unregularized algorithm, runtime increases by

less than 1 % considering 100 regularization parameters at once. Thus, regularization requires only marginal additional computational effort. However, the combination of block-diagonal preconditioning and regularization in general form needs further investigation. In our current implementation, the transformation to standard form can not be realized warranting preconditioning.

The LSQR method is well suited for parallel implementation. Very good scaling results could be achieved up to 32 CPUs, cf. Table 4. For 64 CPUs, the speed-up decreases to 48. The parallel efficiency strongly depends on the problem dimension. Thus, we expect the scaling results to improve for larger problems.

In future, we aim to expand LSQR tuning for geopotential recovery. In particular, this involves variance component estimation for the proper weighting of different observation types and the approximate computation of error estimates.

Acknowledgements. This work was funded by the German Federal Ministry of Education and Research and the German Research Foundation (DFG) through the Geotechnologien II program, Grant No. 03F0329B. Moreover, it was carried out under the HPC-EUROPA project (RII3-CT-2003-506079), with the support of the European Community - Research Infrastructure Action (under the FP6 "Structuring the European Research Area" Program). The authors greatly acknowledge constructive comments by Pavel Ditmar and two anonymous reviewers of the manuscript.

References

Alkhatib H, Schuh W-D (2007) Integration of the Monte Carlo covariance estimation strategy into tailored solution procedures for large-scale least squares problems. J Geod 81: 53–66 DOI: 10.1007/s00190-006-0034-z

Anderson E, Bai Z, Bischof C, Blackford S, Demmel J, Dongarra J, Du Croz J, Greenbaum A, Hammarling S, McKenney A, Sorensen D (1999) LAPACK Users' Guide (third edition). SIAM, Philadelphia

Baur O, Austen G (2005) A parallel iterative algorithm for large-scale problems of type potential field recovery from satellite data. Proceedings Joint CHAMP/GRACE Science Meeting, Geoforschungszentrum Potsdam, online publication (www.gfz-potsdam.de/pb1/JCG))

Baur O, Grafarend EW (2006) High-Performance GOCE Gravity Field Recovery from Gravity Gradients Tensor Invariants and Kinematic Orbit Information. In: Flury J, Rummel R, Reigber C, Rothacher M, Boedecker G, Schreiber U (Eds) Observation of the Earth System from Space. Springer Berlin Heidelberg New York, pp 239–253

Baur O, Sneeuw N, Grafarend EW (submitted) Methodology and Use of Tensor Invariants for Satellite Gravity Gradiometry. J Geod (submitted)

Björck A (1996) Numerical methods for least squares problems. SIAM, Philadelphia

Boxhammer Ch (2006) Effiziente numerische Verfahren zur sphärischen harmonischen Analyse von Satellitendaten. Dissertation, University of Bonn, 95pp

Boxhammer Ch, Schuh W-D (2006) GOCE gravity field modeling: computational aspects - free kite numbering scheme. In: Flury J, Rummel R, Reigber C, Rothacher M, Boedecker G, Schreiber U (Eds): Observation of the Earth System from Space. Springer Berlin Heidelberg New York, pp 209–224

Chandra R, Menon R, Dagum L, Kohr D, Maydan D, McDonald J (2001) Parallel Programming in OpenMP. Academic Press

Colombo OL (1984) The global mapping of gravity with two satellites. Netherlands Geodetic Commission, New Series, 7(3)

Ditmar P, Klees R, Kostenko F (2003a) Fast and accurate computation of spherical harmonic coefficients from satellite gravity gradiometry data. J Geod 76: 690–705 DOI: 10.1007/s00190-002-0298-x

Ditmar P, Kusche J, Klees R (2003b) Computation of spherical harmonic coefficients from gravity gradiometry data to be acquired by the GOCE satellite: regularization issues. J Geod 77: 465–477 DOI: 10.1007/s00190-003-0349-1

ESA SP-1233 (1999) The four candidate Earth explorer core missions - gravity field and steady-state ocean circulation mission. European Space Agency Report SP-1233(1), Granada

Gelderen M van, Koop R (1997) The use of degree variances in satellite gradiometry. J Geod 71: 337–343 DOI: 10.1007/s001900050101

Golub GH, Kahan W (1965) Calculating the singular values and pseudoinverse of a matrix. SIAM J Numer Anal 2: 205–224

Gundlich B, Koch K-R, Kusche J (2003) Gibbs sampler for computing and propagating large covariance matrices. J Geod 77: 514–528 DOI: 10.1007/s00190-003-0350-5

Hanke M, Hansen PC (1993) Regularization methods for large-scale problems. Surv Math Ind 3: 253–315

Hanke M, Vogel CR (1999) Two-level preconditioners for regularized inverse problems I: Theory. Numer Math 83: 385–402

Hansen PC (1987) The truncated SVD as a method for regularization. Numer Math 27: 534–553

Hestenes MR, Stiefel E (1952) Methods of conjugate gradients for solving linear systems. J Res Nat Bur Stand 49: 409–436

Ilk KH, Visser P, Kusche J (2003) Satellite Gravity Field Missions. Final Report Special Commission 7, vol. 32, General and technical reports 1999–2003

Jacobsen M, Hansen PC, Saunders MA (2003) Subspace preconditioned LSQR for discrete ill-posed problems. Numer Math 43: 975–989

Keller W (2002) A Wavelet Approach for the Construction of Multi-Grid Solvers for Large Linear Systems. In: Adam J, Schwarz K-P (Eds) Vistas for geodesy in the new millennium. Springer Berlin, pp 265–270

Klees R, Koop R, Visser P, van den IJssel J (2000) Efficient gravity field recovery from GOCE gravity gradient observations. J Geod 74: 561–571 DOI: 10.1007/s001900000118

Klees R, Ditmar P, Broersen P (2003) How to handle coloured observation noise in large-scale least-squares problems. J Geod 76: 629–640 DOI: 10.1007/s00190-002-0291-4

Kusche J (2001) Implementation of multigrid solvers for satellite gravity anomaly recovery. J Geod 74: 773–782 DOI: 10.1007/s001900000140

Kusche J, Mayer-Gürr T (2001) Iterative Solution of Ill-Conditioned Normal Equations by Lanczos Methods. In: Adam J, Schwarz K-P (Eds) Vistas for geodesy in the new millennium. Springer Berlin, pp 248–252

Kusche J, Klees R (2002) Regularization of gravity field estimation from satellite gravity gradients. J Geod 76: 359–368 DOI: 10.1007/s00190-002-0257-6

Kusche J (2003) A Monte-Carlo technique for weight estimation in satellite geodesy. J Geod 76: 641–652 DOI: 10.1007/s00190-002-0302-5

Lemoine FG, Kenyon SC, Factor JK, Trimmer RG, Pavlis NK, Chinn DS, Cox CM, Klosko SM, Luthcke SB, Torrence MH, Wang YM, Williamson RG, Pavlis EC, Rapp RH, Olson TR (1998) The Development of the Joint NASA GSFC and NIMA Geopotential Model EGM96. NASA Goddard Space Flight Center, Greenbelt, 575pp

Nolet G (1993) Solving large linearized tomographic problems. In: Iyer HM, Hirahara K (Eds) Seismic Tomography: Theory and Practice. Chapman & Hall, London, pp 248–264

Paige CC, Saunders MA (1982a) LSQR: An algorithm for sparse linear equations and sparse least squares. ACM T Math Software 8: 43–71

Paige CC, Saunders MA (1982b) LSQR: Sparse linear equations and least squares problems. ACM T Math Software 8: 195–209

Pail R, Plank G (2002) Assessment of three numerical solution strategies for gravity field recovery from GOCE satellite gravity gradiometry implemented on a parallel platform. J Geod 76: 462–474 DOI: 10.1007/s00190-002-0277-2

Pail R (2005) A parametric study on the impact of satellite attitude errors on GOCE gravity field recovery. J Geod 79: 231–241 DOI: 10.1007/s00190-005-0464-z

Phillips DL (1962) A technique for the numerical solution of certain integral equations of the first kind. Journal of the Association for Computing Machinery 9: 54–97

Rapp RH, Cruz JY (1986) Spherical harmonic expansions of the Earth's gravitational potential to degree 360 using 30' mean anomalies. Technical Report 376, Department of Geodetic Sciences and Surveying, Ohio Sate University, Columbus

Reubelt T, Götzelmann M, Grafarend EW (2005) Harmonic Analysis of the Earth Gravitational Field from Kinematic CHAMP Orbits based on Numerically Derived Satellite Accelerations. In: Flury J, Rummel R, Reigber C, Rothacher M, Boedecker G, Schreiber U (Eds) Observation of the Earth System from Space. Springer Berlin Heidelberg New York, pp 27–42

Rummel R (1986) Satellite Gradiometry. In: Sünkel H (Ed) Mathematical and Numerical Techniques in Physical Geodesy. Lect Notes Earth Sci 7, Springer Berlin, pp 317–363

Rummel R, Sansò F, van Gelderen M, Brovelli M, Koop R, Miggliaccio F, Schrama E, Scerdote F (1993) Spherical harmonic analysis of satellite gradiometry. Netherlands Geodetic Commission, New Series, 39

Schmidt M, Fengler M, Mayer-Gürr T, Eicker A, Kusche J, Sánchez L, Han S-C (2007) Regional gravity modeling in terms of spherical base functions. J Geod 81: 17–38 DOI: 10.1007/s00190-006-0101-5

Schuh W-D (1996) Tailored numerical solutions strategies for the global determination of the Earth's gravity field. Mitteilungen der Universität Graz 81

Sneeuw N (2000) A semi-analytical approach to gravity field analysis from satellite observations. Deutsche Geodätische Kommission, Series C 527, Munich

Tikhonov AN (1963) Regularization of incorrectly posed problems. Sov. Mat. Dokl. 4: 1035–1038

Van der Sluis A, Van der Vorst HA (1987) Numerical solution of large, sparse linear algebraic systems arising from tomographic problems. In: Nolet G (Ed) Seismic Tomography. Reidel Publications, pp 49–84

Yao ZS, Roberts RG, Tryggvason A (1999) Calculating resolution and covariance matrices for seismic tomography with the LSQR method. Geophys J Int 138: 886–894

Xu J (1997) An Introduction to Multilevel Methods. In: Ainsworth M, Levesley K, Marietta M, Light W (Eds) Wavelets, Multilevel Methods and Elliptic PDEs. Numerical Mathematics and Scientific Computation, Clarendon Press, pp 213–302