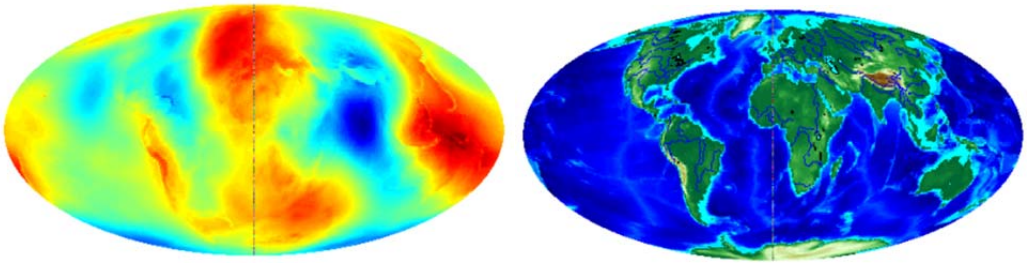

Siegen Preprints on Geomathematics

How to combine spherical harmonics and localized bases for regional gravity modelling and inversion

D. Fischer and V. Michel

Geomathematics Group
Department of Mathematics
University of Siegen, Germany

www.geomathematics-siegen.de

8

UNIVERSITÄT SIEGEN

How to combine spherical harmonics and localized bases for regional gravity modelling and inversion

D. Fischer and V. Michel
Geomathematics Group, University of Siegen, Germany
www.geomathematics-siegen.de
Email: michel@mathematik.uni-siegen.de

May 15, 2012

Abstract

In geophysics, one is oftentimes confronted with the choice of the best system of trial functions to solve a problem. This choice is crucial for the quality of the approximation and the computing time, which are very important topics with respect to today's achievements in data accuracy and abundance. However, mostly this choice is not an obvious one. Here, we present a new method – called the Regularized Functional Matching Pursuit (RFMP) – that constructs the best basis out of an arbitrary collection of different systems of trial functions. We show the potential of the RFMP on inverting the gravitational potential EGM2008 for the density distribution in the area of the Himalayas and India. To compute and represent the solution in this case, the RFMP chooses autonomously the best basis out of a collection of four different types of trial functions, where one has a global character and the other three are localized ones with different degrees of localization.

keywords: best basis; spherical harmonics; spline; wavelet; inverse gravimetry; regional gravity field; sparse regularization; Himalayas; EGM2008

1 Introduction

A common approach for the approximation on the sphere is to use an expansion in terms of spherical harmonics, which form an orthogonal basis system, to represent the data continuously on the whole sphere. However, because of the global character of these functions, small local changes of the data lead to changes in all spherical harmonic coefficients. Furthermore, spherical harmonics are strongly limited by heterogeneous data grids. An analogous problem occurs on the ball, where orthogonal polynomials are known, too, but are connected to the same limitations.
Different groups proposed localized basis systems to remedy the disadvantages of spherical harmonics. Slepian functions, for example, are a locally and globally orthogonal system of functions on the sphere that is optimally suited for a local reconstruction in areas of interest to minimize the effects of data gaps (see [1, 2, 25, 28, 29, 30] for theoretical results and applications). However, this concept is, up to now, limited to the sphere.
The Geomathematics Groups at the University of Kaiserslautern and the University of Siegen developed space localizing kernel functions on the sphere as well as the ball for spline and wavelet methods. These trial functions also allow us to minimize the effects of data gaps or differences in the data density. In this work, we will exploit the localizing character of these functions for a regional reconstruction of the detail structure of the density distribution of the Earth. For further

theoretical aspects of these trial functions and their applications mostly to geophysical problems, we refer to [4, 5, 6, 7, 13, 16, 18, 20, 21, 22, 23, 24].
The choice of the right type of trial functions is very important in regard to the quality of the solution as well as the computing effort. However, mostly it is not an obvious one. Here, we present a method which constructs the best basis out of an arbitrary collection of different systems of trial functions. The main idea is as follows: In every step, the best basis is supplemented by another trial function which is chosen to minimize the remaining data misfit. This yields a more accurate approximation in comparison to the previous iteration step. As a consequence, a solution is generated that is adapted to the local detail structure of the target function as well as the data structure. Note that we do not require initial values for, e.g., the position of a localized basis function.
The ideas for such a method stem from the field of sparse regularization where one is concerned with solving under-determined or ill-conditioned systems of linear equations with respect to the sparsity of the solution (see, e.g., [9, 11, 17, 26]). However, the new method is additionally applicable to inverse problems and also allows the treatment of the ball as a domain. Furthermore, we enhanced this algorithm by including a regularization term to treat ill-posed problems such as the inverse gravimetric problem.
We call this new method the Regularized Functional Matching Pursuit (RFMP). It can be divided into a preprocessing part, which has already been parallelized in our implementation, and the main part where, in every step, we just need to search for the optimal trial function to best match the data structure. This main part can be parallelized as well.

Preliminaries:
The Euclidean space $\mathbb{R}^l$ is equipped with the usual dot product

$$\langle x, y\rangle_{\mathbb{R}^l} := \sum_{j=1}^{l} x_j y_j, \quad x, y \in \mathbb{R}^l,$$

and its induced norm

$$||x||_{\mathbb{R}^l} := \sqrt{\langle x, x\rangle_{\mathbb{R}^l}}.$$

The closed ball with radius $a > 0$ is denoted by $\mathcal{B} := \{x \in \mathbb{R}^3 |\ |x| \le a\}$. $\mathrm{L}^2(\mathcal{B})$ denotes the space of all square-integrable scalar functions on $\mathcal{B}$, i.e. all $F : \mathcal{B} \to \mathbb{R}$ with

$$||F||_{\mathrm{L}^2(\mathcal{B})} := \left(\int_{\mathcal{B}} [F(x)]^2 \mathrm{d}x\right)^{1/2} < \infty.$$

The inner product for $\mathrm{L}^2(\mathcal{B})$ is defined by

$$\langle F, G\rangle_{\mathrm{L}^2(\mathcal{B})} := \int_{\mathcal{B}} F(x)G(x)\mathrm{d}x, \quad F, G \in \mathrm{L}^2(\mathcal{B}).$$

2 Regularized Functional Matching Pursuit (RFMP)

The idea to develop a solution adaptively and iteratively is not a new one. A corresponding algorithm was introduced as a Matching Pursuit in [17]. An enhanced version was constructed in [32], where kernel functions were introduced into the setting. The hitherto existing Matching Pursuits intrinsically require that the unknown function F is given in terms of grid-based data $y_i = F(x_i)$, $i = 1, \ldots, l$. For instance, in [17], the projection of F on every single trial function is

calculated out of the data. However, this is not applicable if an inverse problem is to be solved, i.e. if the data is given in terms of a (linear) functional $\mathcal{F}^i$ applied to the target function $y_i = \mathcal{F}^i F$, $i = 1, \ldots, l$. For instance, $\mathcal{F}^i F$ could be the gravitational potential which corresponds to the mass density function F and is measured at a particular point x_i outside the Earth. We enhance the present concept for our purposes in order to include the resolution of inverse problems.
Note that an essentially different approach to the construction of wavelets (based on a cubed sphere) for tomographic problems, which also includes aspects of sparsity, is presented in [31]. We propose here an alternative technique. The disadvantages and advantages of both have to be investigated further in the future.
Let us denote the collection of all available trial functions as the dictionary $\mathcal{D}$. Furthermore, we collect all functionals in the operator $\mathcal{F}F := (\mathcal{F}^1 F, \ldots, \mathcal{F}^l F)$. Now in every step, the iterative method chooses that trial function d out of $\mathcal{D}$ and the corresponding coefficient $\alpha \in \mathbb{R}$ that fits the data best, where we measure the fit by the norm of the residual, i.e. the difference between approximation and data. Thus, we get as a first step $F_1 = \alpha_1 d_1$ consisting of a trial function $d_1 \in \mathcal{D}$ and a coefficient $\alpha_1 \in \mathbb{R}$ which are chosen such that the data misfit $\sum_{i=1}^{l} (y_i - \mathcal{F}^i(\alpha_1 d_1))^2$ is minimal. Then d_2 and α_2 are selected such that the residual is further minimized, i.e.

$$\sum_{i=1}^{l} [(y_i - \mathcal{F}^i(\alpha_1 d_1)) - \mathcal{F}^i(\alpha_2 d_2)]^2 = \min_{\alpha \in \mathbb{R}, d \in \mathcal{D}} \sum_{i=1}^{l} [(y_i - \mathcal{F}^i(\alpha_1 d_1)) - \mathcal{F}^i(\alpha d)]^2.$$

Generally, in step $n+1$ the algorithm chooses d_{n+1} and α_{n+1} such that the norm of the residual

$$||R^{n+1}||_{\mathbb{R}^l}^2 \quad = \quad ||y - \mathcal{F}F_{n+1}||_{\mathbb{R}^l}^2 = \sum_{i=1}^{l} (y_i - \mathcal{F}^i F_{n+1})^2 = ||R^n - \alpha_{n+1} \mathcal{F} d_{n+1}||_{\mathbb{R}^l}^2$$

is minimized, where $F_{n+1} = \sum_{k=1}^{n+1} \alpha_k d_k$.
The ill-posedness of the inverse gravimetric problem requires the use of a regularization technique. In this work, we use a Tikhonov regularization, i.e. we try to achieve a trade-off between fitting the data and reducing the norm of the solution. The regularization parameter λ correlates both terms. Note that other regularization terms such as an l^1-norm have been used for Matching Pursuits as well. However, e.g., in [8] was shown that such a choice is not appropriate for ill-posed problems, whereas a Tikhonov regularization is more successful in this case.
For this reason, we introduce a regularized version called the Regularized Functional Matching Pursuit (RFMP) where the penalty term is concerned with the (non-)smoothness of the solution, i.e. its $\mathrm{L}^2(\mathcal{B})$-norm. In addition to the known advantages of this regularization, this choice is also more practicable in the implementation and yields, in contrast to the L^1-version, an analytical expression for the coefficient α.
Using a Tikhonov regularization, in step $n+1$, we need to find the trial function d_{n+1} and the corresponding coefficient α_{n+1} such that they minimize

$$||R^n - \alpha_{n+1} \mathcal{F} d_{n+1}||_{\mathbb{R}^l}^2 + \lambda ||F_n + \alpha_{n+1} d_{n+1}||_{\mathrm{L}^2(\mathcal{B})}^2,$$

where $\lambda > 0$ is the regularization parameter.
The Regularized Functional Matching Pursuit (RFMP) starts with $F_0 = 0$ (or some model), where the algorithm iteratively appends trial functions to the initially empty set while trying to reduce the residual combined with some penalty term at each stage.

Algorithm 2.1 (RFMP)
Start with $F_0 := 0$ *(or some model) and* $R^0 := y$.
Given F_n.
Build $F_{n+1} := F_n + \alpha_{n+1} d_{n+1}$ *such that*

$$d_{n+1} \text{ maximizes } \left| \frac{\langle R^n, \mathcal{F}d \rangle_{\mathbb{R}^l} - \lambda \langle F_n, d \rangle_{\mathrm{L}^2(\mathcal{B})}}{\sqrt{||\mathcal{F}d||^2_{\mathbb{R}^l} + \lambda ||d||^2_{\mathrm{L}^2(\mathcal{B})}}} \right| \text{ and}$$

$$\alpha_{n+1} := \frac{\langle R^n, \mathcal{F}d_{n+1} \rangle_{\mathbb{R}^l} - \lambda \langle F_n, d_{n+1} \rangle_{\mathrm{L}^2(\mathcal{B})}}{||\mathcal{F}d_{n+1}||^2_{\mathbb{R}^l} + \lambda ||d_{n+1}||^2_{\mathrm{L}^2(\mathcal{B})}}.$$

Set $R^{n+1} := R^n - \alpha_{n+1} \mathcal{F} d_{n+1}$.

This algorithm can be derived from the ansatz described above. For further details, see [14, 15]. Note that this algorithm does not necessarily provide us with the theoretically best match to the target function F. Since we determine the expansion functions and the corresponding coefficients stepwise, the expansion with n elements is possibly not optimal at step $n+1$. To remedy this inaccuracy, we can do a back-projection in analogy to [10]. That means we choose the function d_{n+1} as in the original algorithm but recompute the optimal set of coefficients $\alpha_1, \ldots, \alpha_{n+1}$ in each step. This extension of the algorithm gives us a better approximation while the computation time is increased. To get an even more accurate result, we may use pre-fitting, again in analogy to [10], where we directly optimize for the function d_{n+1} and the optimal coefficients $\alpha_1, \ldots, \alpha_{n+1}$ jointly. Although this is the most time-consuming version of the three, it will give the best-fitting solution as well. Nonetheless, in this work we will only use the original version of algorithm 2.1 (RFMP) to reduce the computing effort.
Code optimization and parallelization now allow a fast and competitive inversion of the data (see the numerical part for details). Moreover, if the data includes coherent structures with respect to $\mathcal{F}d$, where d is a trial function from the dictionary, then we get a faster decay of the norm of the residuals. Thus, it is an important step in the preprocessing to choose a well-matched dictionary with respect to the structure of the solution. If we have some idea about the structure of the target function, we may impose this information on the choice of the dictionary. Otherwise, we recommend to use a dictionary with more general functions of different kinds to get a faster convergence of the algorithm.
We refer to [14] and [15] for further details on the theoretical properties of the method, i.e. the existence of the solution, the stability of the method and the convergence of the regularization.

3 Numerical Application: The Inverse Gravimetric Problem

As a numerical example to demonstrate the power of the novel method, we invert the gravitational potential given by the Earth Gravitational Model 2008 (EGM2008, [27]) developed by the National Geospatial Intelligence Agency (NGA) in the area of the Himalayas and India. Further numerical tests including the advantages of the RFMP over spline and wavelet methods are shown in [15].
It is well-known that Newton's Law of Gravitation

$$(TF)(x) := \gamma \int_{\mathcal{B}} \frac{F(y)}{|x-y|} \, \mathrm{d}y, \; x \in \mathbb{R}^3 \setminus \mathcal{B},$$

where γ is the gravitational constant, represents the relation between the gravitational potential $TF = V$, where V is given, and the mass density distribution F. However, the determination of a (harmonic) solution F from given TF is instable such that this inverse problem is ill-posed.
Note that we use here a harmonicity constraint for the solution F to guarantee its uniqueness, though there is only a mathematical but not a physical justification for this constraint. However, there is no known constraint with an appropriate physical interpretation that yields a unique solution of the inverse gravimetric problem (see the survey article [23]).
Let us consider the functionals that map the density $F \in \mathrm{L}^2(\mathcal{B})$ to the gravitational potential

$$\mathcal{F}_{\mathrm{G}}^k \; F := \gamma \int_{\mathcal{B}} \frac{F(y)}{|x_k - y|} \; \mathrm{d}y,$$

where $x_k \in \mathbb{R}^3 \setminus \mathcal{B}$ is a point outside the Earth where the potential is given (see [19, 20] for a series representation of $\mathcal{F}_{\mathrm{G}}^k \; F$).
There already exist approximate models F_{M} of the mass density distribution of the Earth - for example PREM (see [12]). Thus, it is useful to apply the functionals to the deviation $\delta F := F - F_{\mathrm{M}}$ which is the difference of the mass density and the density of a reference model. For PREM as a reference model, we get $V_\delta(x) := V(x) - \gamma \; \frac{4\pi a^3}{3|x|} \; 5.5134 \; \frac{\mathrm{g}}{\mathrm{cm}^3}$ (see e.g. [18]).
Since it is well-known that the harmonicity constraint in particular and gravitational data in general are only appropriate for the determination of mass anomalies in the uppermost layer of the Earth (see [23]), we only reconstruct the density close to the surface. Here we recover the mass density variation of the Himalayas and India out of EGM2008 from degree 3 up to degree 1 500.
The algorithm may choose between four different types of trial functions: It may reconstruct global trends with the polynomials $G_{0,n,j}^{\mathrm{I}}$ while the localized trial functions (wavelet-based scaling functions) $K_h^{\mathrm{I}}(\cdot, x)$ for three different parameters h are a very good choice to recover detail structures of the target function.
There are two known systems of global orthonormal basis functions on $\mathcal{B}$, namely $G_{m,n,j}^{\mathrm{I}}$ and $G_{m,n,j}^{\mathrm{II}}$ (see, e.g., [21] and the references therein). Since we only aim to recover the harmonic part of the density distribution, we decided to use the inner harmonics

$$G_{0,n,j}^{\mathrm{I}}(x) \quad := \quad \sqrt{\frac{2n+3}{a^3}} \; \left(\frac{|x|}{a}\right)^n \; Y_{n,j}\left(\frac{x}{|x|}\right), \quad x \in \mathcal{B},$$

as global trial functions, where $n \in \mathbb{N}_0, \; j = 1, \ldots, 2n+1$ and $Y_{n,j}, n \in \mathbb{N}_0, j = 1, \ldots, 2n+1$, are the (fully normalized) real spherical harmonics.
Furthermore, we use the normalized and localized kernel functions $K_h^{\mathrm{I}} \in \mathrm{L}^2(\mathcal{B} \times \mathcal{B})$ based on [3, 5, 7, 13, 21] which are spline basis functions or scaling functions on the ball. For our numerical applications, we use certain parameter-dependent kernels, i.e.

$$\tilde{K}_h^{\mathrm{I}}(x, y) := \sum_{n=0}^{E_n} \sum_{j=1}^{2n+1} h^n G_{0,n,j}^{\mathrm{I}}(x) G_{0,n,j}^{\mathrm{I}}(y), \quad x, y \in \mathcal{B},$$

where every fixed $h \in]0, 1[$ yields one particular kernel. Note that h is a parameter to influence the localizing character of the kernel function. The hat-width decreases for h getting closer to 1 (see figure 1). The peak of $y \mapsto \tilde{K}_h(y, x)$ is centred at x.
In the following, we will always consider normalized kernel functions and denote them with $K_h^{\mathrm{I}}(\cdot, x)$, i.e.

$$K_h^{\mathrm{I}}(\cdot, x) := \frac{\tilde{K}_h^{\mathrm{I}}(\cdot, x)}{||\tilde{K}_h^{\mathrm{I}}(\cdot, x)||_{\mathrm{L}^2(\mathcal{B})}}.$$

Figure 1: Kernel function $K_h^{\mathrm{I}}(\cdot, x)$ on the surface of the ball $\mathcal{B}$ with $h = 0.5$ (left-hand), 0.8 (right-hand)

Explicitly, the dictionary is now given as

$$\begin{aligned} \mathcal{D} := \quad & \left\{ K_h^{\mathrm{I}}(\cdot, x) \;\middle|\; h \in \{0.97, 0.98, 0.99\}, x \in \mathrm{grid}(\mathcal{B}) \right\} \\ & \cup \left\{ G_{0,n,j}^{\mathrm{I}} \;\middle|\; n = 3, \ldots, 50, j = 1, \ldots, 2n+1 \right\}, \end{aligned} \tag{1}$$

where $\mathrm{grid}(\mathcal{B})$ is a nearly quadratic grid, which is equiangular each in longitude and latitude. After the restriction to a spherical rectangle covering the target area we are left with 39 800 grid points. Furthermore, we will stop the summation in the kernel functions at the degree of the spherical harmonics coefficients used to compute the data, i.e. we set $E_n := 1\,500$. This dictionary now contains approximately 120 000 elements of four different types.
We can use well-known formulae for the data corresponding to the trial functions:

$$\mathcal{F}_{\mathrm{G}}^k G_{m,n,j}^{\mathrm{I}} = \delta_{m0}\, \gamma\, \tfrac{4\pi}{2n+1} \sqrt{\tfrac{a^3}{2n+3}} \left(\tfrac{a}{|x_k|}\right)^n \tfrac{1}{|x_k|}\, Y_{n,j}\left(\tfrac{x_k}{|x_k|}\right)$$

and

$$\begin{aligned} \mathcal{F}_{\mathrm{G}}^k K_h^{\mathrm{I}}(\cdot, x) &= \frac{1}{||\tilde{K}_h^{\mathrm{I}}(\cdot, x)||_{\mathrm{L}^2(\mathcal{B})}}\, \mathcal{F}_{\mathrm{G}}^k \tilde{K}_h^{\mathrm{I}}(\cdot, x) \\ \mathcal{F}_{\mathrm{G}}^k \tilde{K}_h^{\mathrm{I}}(\cdot, x) &= \gamma \sum_{n=0}^{E_n} h^n \left(\frac{|x|}{|x_k|}\right)^n \frac{1}{|x_k|}\, P_n\left(\frac{x}{|x|} \cdot \frac{x_k}{|x_k|}\right) \\ ||\tilde{K}_h^{\mathrm{I}}(\cdot, x)||_{\mathrm{L}^2(\mathcal{B})} &= \sum_{n=0}^{E_n} h^{2n}\, \frac{2n+3}{a^3}\, \frac{2n+1}{4\pi} \left(\frac{|x|}{a}\right)^{2n}. \end{aligned}$$

Note that δ_{m0} denotes the Kronecker delta, which equals 1 if $m = 0$ and vanishes otherwise.
If not stated differently, the data will be given at 25 440 points on an equiangular grid in analogy to $\mathrm{grid}(\mathcal{B})$ which is located slightly above the Earth's surface at 7 km height. The method will be stopped after 20 000 iterations, i.e. 20 000 (not necessarily pairwise distinct) dictionary functions will be chosen out of dictionary (1), which contains approximately 120 000 elements.
We use algorithm 2.1 (RFMP) with the regularization parameter $\lambda = 500$, where this choice of λ is based on some numerical experiments.
These parameter choices yield the solution $F_{20\,000}$ displayed in the left-hand side of figure 2. Note that the regularization parameter yields an acceptable balance between noise and smoothing, since the solution is smooth enough without a too strong attenuation of the signal. We clearly see the continental boundary between the Indian plate and the Eurasian plate and the dominating

structures of the Himalayas as well as the beginning of the Java trench in Indonesia. In the right-hand side of figure 2, we display the centre points x of the chosen kernel functions $K_h^{\mathrm{I}}(\cdot, x)$, where we artificially included the coast lines of Asia in blue as an orientation. Obviously, the localized dictionary functions are chosen mainly in areas where the detail structure of the solution is more complicated such as the Himalayas. Thus, algorithm 2.1 (RFMP) clearly recovers a solution that is adapted to the structure of the target function, i.e. the density of the basis functions is correlated to the data density in a region.

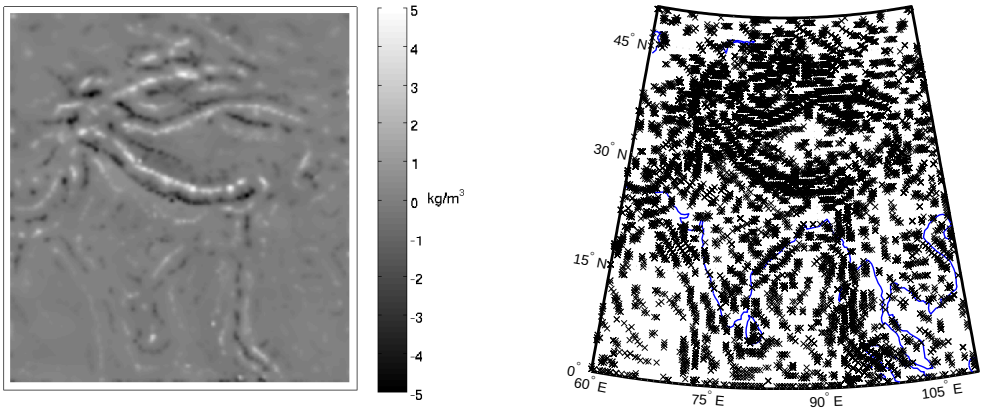

Figure 2: We used algorithm 2.1 (RFMP) with $\lambda = 500$ and dictionary (1) to reconstruct the density deviation (left-hand) out of the potential given at 25 440 data points, where we stopped after 20 000 iterations (i.e. $F_{20\,000}$ is shown). In addition, we display the centre points x of the (not necessarily pairwise distinct) chosen expansion functions $K_h^{\mathrm{I}}(\cdot, x)$ (right-hand).

We report the following computational performance of our algorithm: The computation of the scalar products and norms as well as all updates were parallelized with OpenMP. On a compute server with 16 cores and 48 GB RAM, the computing effort of this (relatively large) example is about 11 hours of time and about 15% of the RAM, where about 1 hour is used for preprocessing and the rest is used for the search of the best basis and the evaluation of the solution on an arbitrary grid. We plan to use a cluster for the computations such that the search can be parallelized as well with the help of MPI. Thus, there is still capacity to decrease the computing time much further.

In figure 3, we display the data in form of the potential given by EGM2008 at 25 440 data points (top left-hand), the approximation of the potential by the solution $F_{20\,000}$ (i.e. $\mathcal{F}F_{20\,000}$) at the same points (top right-hand) and the absolute values of the difference between both (i.e. the absolute values of the residual $|R^{20\,000}|$, bottom). Clearly, the main structures of the potential are approximated well enough while some of the detail structures remain as an error. However, the values of the error are mostly far below 10% of the original data.

These results are competitive to other methods such as splines and wavelets. However, in comparison, the novel method allows a larger data density on irregular grids as well. Note that the advantages of the novel method over previous approaches such as spline and wavelet methods were studied in [15] in further detail.

In the top of figure 4, we display the choice of trial functions with respect to the parameter h, that controls the localization of the trial function, for $F_{20\,000}$, where the choice of a $G_{0,n,j}^{\mathrm{I}}$ is denoted with 1 on the vertical axis. Out of the 20 000 expansion functions of $F_{20\,000}$, 766 were inner

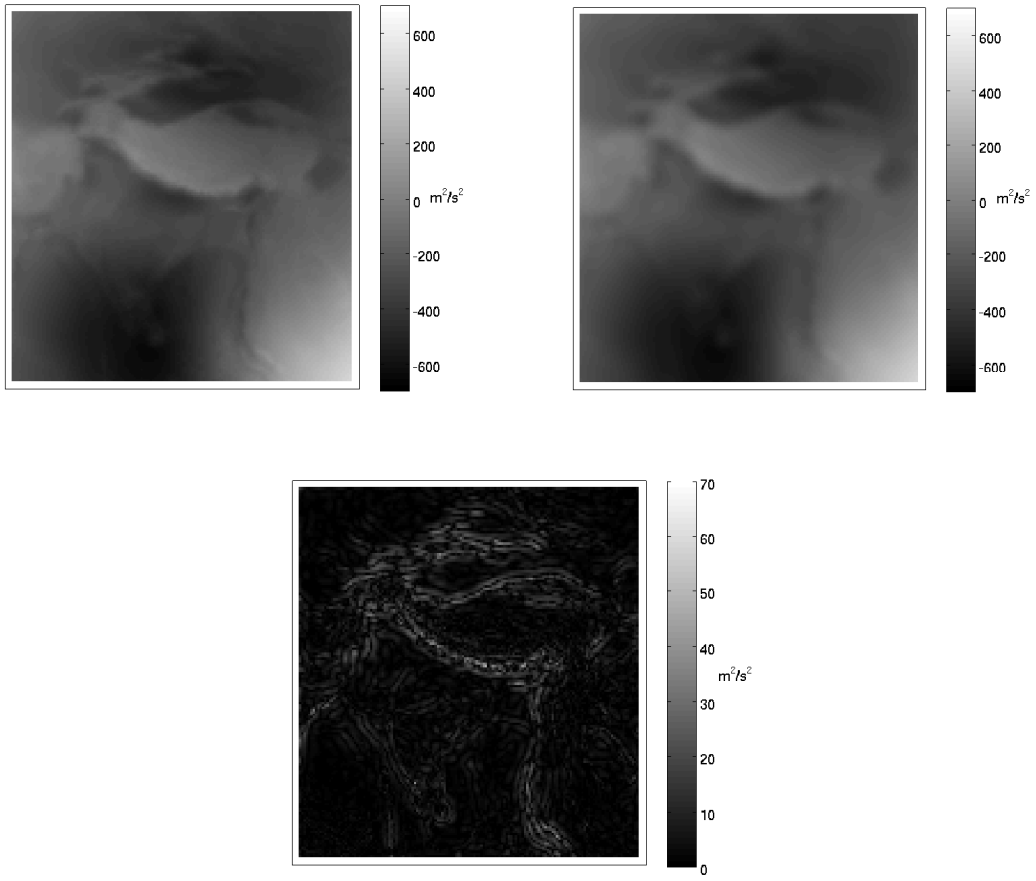

Figure 3: We compare the data input EGM2008 (top left-hand) and the potential generated by the solution $F_{20\,000}$ for $\lambda = 500$ (top right-hand) and display the absolute values of the difference of both (bottom) for 25 440 data points.

harmonics $G^{\mathrm{I}}_{0,n,j}$, i.e. trial functions with a global character. Moreover, there were 1 323, 2 831 and 15 263 functions corresponding to localized trial functions with the parameters $h = 0.97$, 0.98 and 0.99, respectively. They were chosen throughout the whole process to approximate the detail structures.
On the bottom of figure 4, we display only the first appearance of a dictionary element. Overall, only 4 522 different dictionary elements are used in this expansion of 20 000 elements. Thus, a back-projection or pre-fitting can be expected to increase the sparsity of the solution further.
Moreover, we observe that localized trial functions with lower localization (smaller h) are more often chosen in the beginning than later on, where smaller details are primarily added with kernels corresponding to $h = 0.99$ (where these trial functions are used throughout the process).

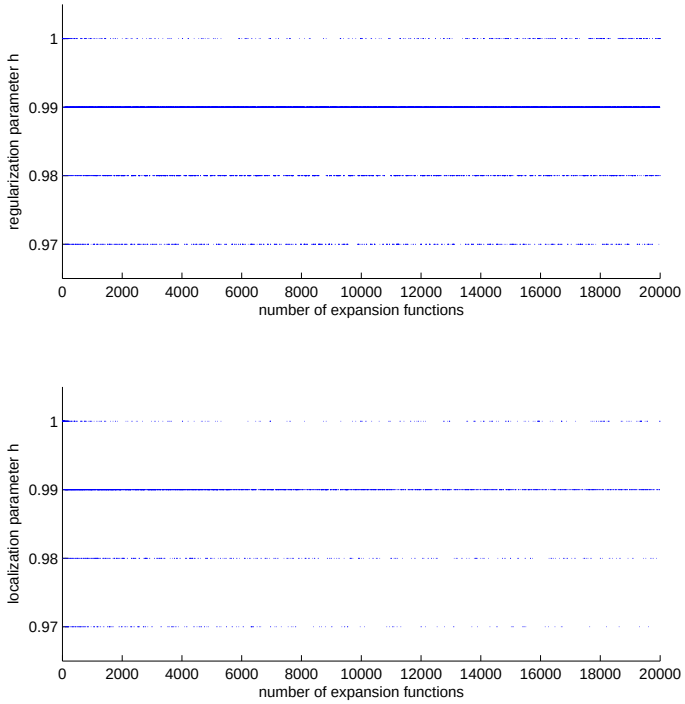

Figure 4: We compare the choice of trial functions for $F_{20\,000}$ with respect to the localization parameter $h = 0.97, 0.98, 0.99$, where the choice of a function $G^{\mathrm{I}}_{0,n,j}$ is denoted with 1 on the vertical axis. In the two diagrams, we display all expansion functions (top) and only the first appearance of a dictionary element (bottom), respectively.

The stability of the solution

As a stability test, we use the approximation $\mathcal{F}F_{20\,000}$ to the gravitational potential given by the solution $F_{20\,000}$ for $\lambda = 500$ as a data input (see the top left-hand side of figure 2 for the solution $F_{20\,000}$ and the top right-hand side of figure 3 for its approximation to the gravitational potential $\mathcal{F}F_{20\,000}$). Note that the data is now given in terms of inner harmonics as well as localized kernel functions. Moreover, it is given exactly such that we may use the (faster) unregularized algorithm with $\lambda = 0$. Apart from that, we use the same parameters as before for the dictionary and the data grid.

In the top left-hand side of figure 5 we display the solution of this stability test. Plotting the absolute values of the difference to $F_{20\,000}$ gives us some deviations from the original as displayed in the top right-hand side of figure 5. These deviations are mostly far below 10% of the original value and can be explained by the iterative character of the method. As we noted before, back-projection or pre-fitting would probably increase the stability and accuracy of the method. If we consider the quality of the approximation with respect to the gravitational potential, i.e. the absolute value of the difference of the potential generated by this solution to $\mathcal{F}F_{20\,000}$ as displayed in the bottom of figure 5, we observe that the error is even mostly far below 5% of the original values. The lower error on the data side underlines the ill-posedness of the inverse gravimetric problem, which certainly also contributes to the obtained error.
In total, however, the error in this stability test can be considered to be acceptably low.

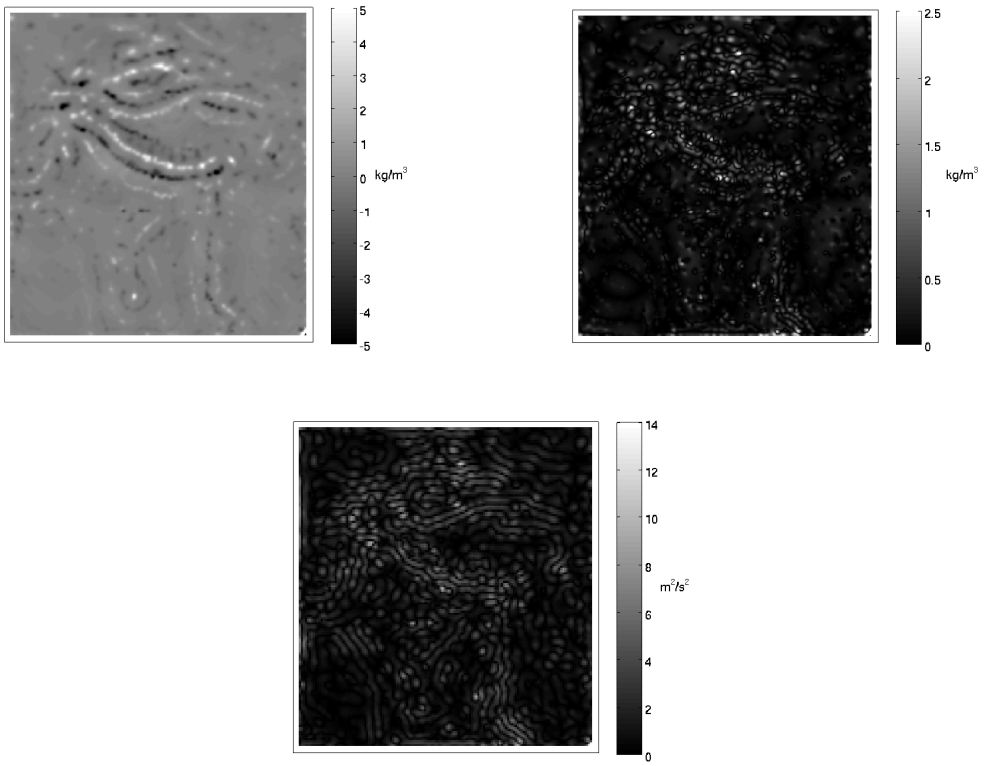

Figure 5: We used algorithm 2.1 (RFMP) and dictionary (1) to reconstruct the density deviation (top left-hand) out of the potential $\mathcal{F}F_{20\,000}$ (given at 25 440 data points) for $\lambda = 0$ and stopped the method after 20 000 iterations. We display the absolute values of the difference to $F_{20\,000}$ (top right-hand) and the absolute values of the difference of the here generated potential to $\mathcal{F}F_{20\,000}$ (bottom)

The localized character of the solution

Since we use localized trial functions in our dictionary, we can expect our solution to depend at each point only on data from the neighbourhood. This is, indeed, the case, as we will demonstrate here.

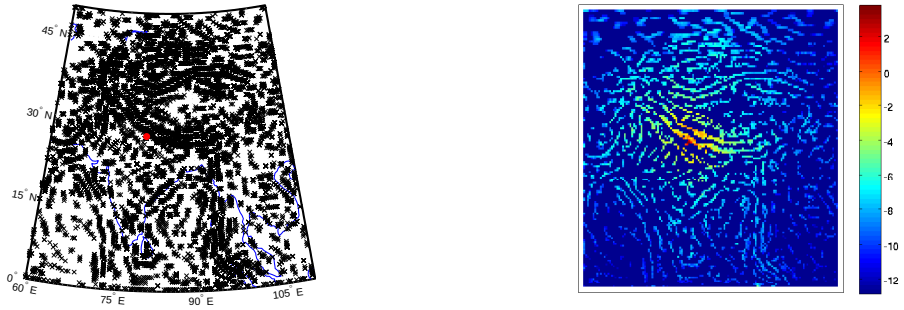

Figure 6: We display the influence of the choice of the expansion functions in solution $F_{20\,000}$ (right-hand) on the solution at point $x = (0.153525,\ 0.870648,\ 0.467296)^{\mathrm{T}}a$ (red dot) located in the Himalayas (left-hand).

On the bottom of figure 6, we examine the influence of the chosen trial functions of solution $F_{20\,000}$ on exactly one point (marked with a red dot in the left-hand plot of figure 6)

$$x = (0.153525,\ 0.870648,\ 0.467296)^{\mathrm{T}}a$$

situated in the Himalayas. Note that $F_{20\,000}(x)$ is a sum of 20 000 summands connected to the chosen trial functions $d_1, \ldots, d_{20\,000}$. For this reason, we display (the logarithm of) the contribution $\log|\alpha_k d_k(x)|$ of each expansion element d_k to $F_{20\,000}(x)$. We have to distinguish whether each d_k is a localized trial function or a global trial function: $\log|\alpha_k K^{\mathrm{I}}_{h_k}(x, x_k)|$ is given as its value at the point x_k while $\log|\alpha_k G^{\mathrm{I}}_{0,n_k,j_k}|$ is included as an additive at all points in the plot. As a consequence, the blue color denotes a very small impact while the red color denotes a large impact on the value at the point x. Obviously, the value at x is mostly controlled by localized expansion functions. Furthermore, we clearly observe that the influence of the localized expansion functions increases when the distance between x and the centre x_k decreases. Hence, local irregularities of all kinds only have a local influence on the computed approximation.

Dealing with noise

Let us examine the behaviour of algorithm 2.1 (RFMP) when applied to noisy data y^ε, where ε denotes the noise level. For example, a value $\varepsilon = 0.05$ corresponds to a data input y^ε that is disturbed with 5% uniformly distributed random noise relative to the exact data y, i.e.

$$y^\varepsilon_i = y_i + \varepsilon\ \mathrm{rand}_i\ y_i,\ \ i = 1, \ldots, l,$$

where rand_i is a random number in the interval $[-1, 1]$. In figure 7, we consider the reconstructed density deviations out of $l = 25\,440$ noisy data of the kind above, where we use dictionary (1) and stop algorithm 2.1 (RFMP) with regularization parameter $\lambda = 600$ after 20 000 iterations.
Obviously, the main structures are recognized in spite of the noise. Furthermore, in comparison to figure 2, where the same setup was considered for undisturbed data and a regularization parameter $\lambda = 500$, we clearly observe that the introduction of noise does not have an overly negative influence on the reconstruction quality of the algorithm. The absolute values of the difference between the solutions is mostly below 10 % (see the bottom of figure 7 as well). However, we hardly see structural differences between the original and the reconstruction out of noisy data. Moreover, the algorithm still positions the chosen dictionary elements according to the detail structures of

the solution as displayed in the top right-hand side of figure 7. Thus, the chosen regularization, indeed, works in the sense that noise only has a small influence on the inversion and the sparsity (i.e. the detail structure based choise of the localized trial functions) is not substantially influenced by noise.

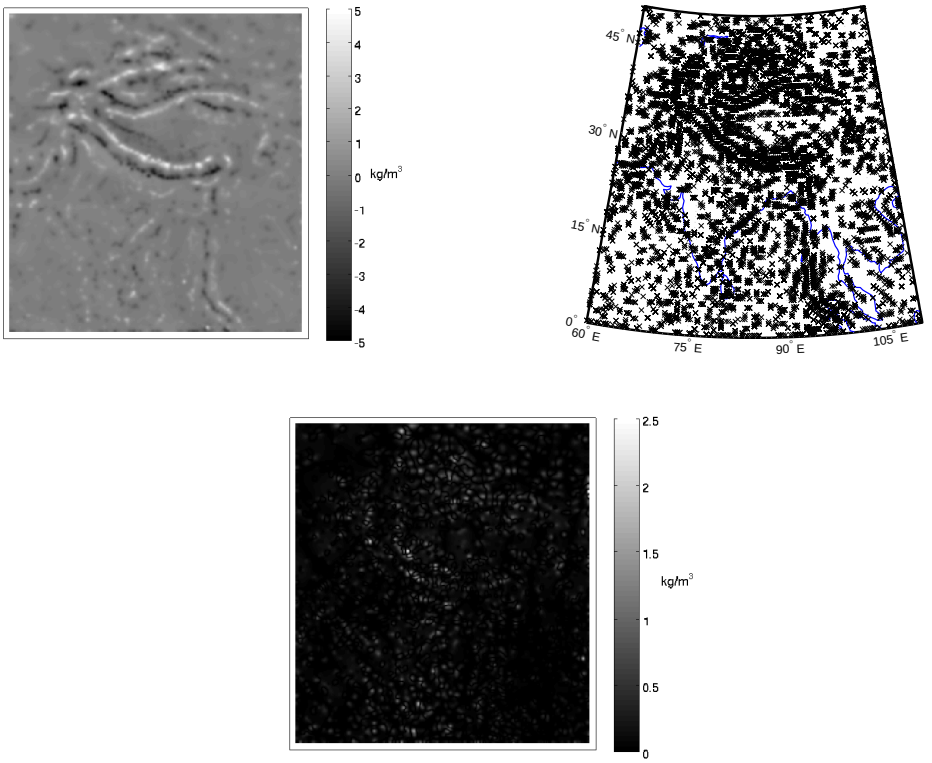

Figure 7: We used algorithm 2.1 (RFMP) to reconstruct the density deviation (top left-hand) out of the potential given at 25 440 data points with 5% noise ($\varepsilon = 0.05$) for $\lambda = 600$, where we stopped the iteration after 20 000 iterations. We display the centre points (top right-hand) of the (not necessarily pairwise distinct) chosen localized expansion functions and the absolute difference to $F_{20\,000}$ in figure 2 (bottom).

4 Conclusions and Outlook

We presented a novel method – called the Regularized Functional Matching Pursuit (RFMP) – that constructs the best basis out of an arbitrary collection of different systems of trial functions to approximate the solution of an inverse problem.
We emphasized the potential of the RFMP on inverting the gravitational potential EGM2008 for the density distribution in the area of the Himalayas and India, where we allowed the RFMP to choose a set of trial functions out of a collection of four different types - some of them with a global character, others localized ones. The less localized trial functions are chosen to reconstruct the main structures of the solution, while the more localized ones are used for the details.
Moreover, when we inverted the EGM2008 potential, which is based on spherical harmonics, the algorithm chose a large number of spherical harmonics (more precisely, inner harmonics) as well.

However, inverting the potential $\mathcal{F}F_{20\,000}$, which is generated out of a combination of spherical harmonics and localized functions, we gained a better approximation quality. For this reason, we expect our method to perform even better if data sets are used which have not been generated out of a model with global trial functions. For example, grid-based data could yield further improved results.
Note also that it is a particular feature of the presented method that it allows us to combine different data types, where much more data than previously may be used. Our main goal is to recover a model of the density distribution of the interior of the Earth as is done with a spline method in [6]. Gravitational data only gives information about the harmonic part of the density (and, consequently, about the uppermost layer of the Earth). The anharmonic part (and, in particular, deeper structures of the Earth) can be partially recovered from seismic data, e.g., normal mode splitting or travel times. However, the dimension of present data sets, e.g., in gravimetry, is beyond the numerical limitation of the spline method. We hope that our novel technique will overcome these problems, while the features of localized approaches are saved. Detailed studies of the combined inversion with this new method are currently being investigated and will be published in a forthcoming work. However, first results can be found in [14].
In addition, further improvements of the algorithm will be investigated in the near future. This includes a more accurate determination of the optimum, i.e. the inclusion of techniques such as a back-projection and a pre-fitting, and a more sophisticated choice of the regularization parameter.

Acknowledgments: We gratefully acknowledge the support by the German Research Foundation (DFG), projects MI 655/2-2 and MI 655/7-1. Moreover, the authors thank Frederik J. Simons (Princeton University) for inspiring discussions and competent advice concerning the geophysical topics. Furthermore, we are grateful that Abel Amirbekyan (Fraunhofer ITWM Kaiserslautern) helped us to optimize and parallelize our code.

References

[1] Albertella A, Sansò F and Sneeuw N 1999 Band-limited functions on a bounded spherical domain: the Slepian problem on the sphere *J. Geodesy* **73** 436–47

[2] Albertella A and Sneeuw N 1999 The analysis of radiometric data with Slepian functions *Phys. Chem. Earth* **25** 667–72

[3] Amirbekyan A 2007 *The Application of Reproducing Kernel Based Spline Approximation to Seismic Surface and Body Wave Tomography: Theoretical Aspects and Numerical Results* PhD thesis University of Kaiserslautern *http://kluedo.ub.uni-kl.de/volltexte/2007/2103/pdf/ThesisAbel.pdf*

[4] Amirbekyan A, Michel V and Simons F J 2008 Parameterizing surface–wave tomographic models with harmonic spherical splines *Geophys. J. Int.* **174** 617–28

[5] Berkel P 2009 *Multiscale Methods for the Combined Inversion of Normal Mode and Gravity Variations* PhD thesis University of Kaiserslautern (Aachen: Shaker Verlag)

[6] Berkel P, Fischer D and Michel V 2011 Spline multiresolution and numerical results for joint gravitation and normal mode inversion with an outlook on sparse regularisation *Int. J. Geomath.* **1** 167–204

[7] Berkel P and Michel V 2010 On mathematical aspects of a combined inversion of gravity and normal mode variations by a spline method *Math. Geosci.* **42** 795–816

[8] Candés E J, Eldar Y C, Needell D and Randall P 2011 Compressed sensing with coherent and redundant dictionaries *Appl. Comput. Harmon. Anal.* **31** 59–73

[9] Dai W and Milenkovic O 2009 Subspace pursuit for compressive sensing: closing the gap between performance and complexity *IEEE Trans. Inf. Theory* **55** 2230–48

[10] Davis G, Mallat S G and Zhang Z 1994 Adaptive time-frequency approximations with matching pursuits *SPIE Journal of Optical Engineering* **33** 2183–91

[11] Donoho D L, Tsaig Y, Drori I and Starck J-L 2006 Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit *Technical report at http://www-stat.stanford.edu/∼donoho/Reports/2006/*

[12] Dziewonski A and Anderson D L 1981 The preliminary reference Earth model *Phys. Earth Planet. Inter.* **25** 297–356

[13] Fengler M J, Michel D and Michel V 2006 Harmonic spline-wavelets on the 3-dimensional ball and their application to the reconstruction of the Earth's density distribution from gravitational data at arbitrarily shaped satellite orbits *ZAMM* **86** 856–73

[14] Fischer D 2011 *Sparse Regularization of a Joint Inversion of Gravitational Data and Normal Mode Anomalies* PhD thesis (München: Verlag Dr. Hut)

[15] Fischer D and Michel V 2012 Sparse regularization of inverse gravimetry – case study: spatial and temporal mass variations in South America *accepted for publication by Inv. Probl.*

[16] Freeden W and Michel V 2004 *Multiscale Potential Theory (with Applications to Geoscience)* (Boston: Birkhäuser)

[17] Mallat S G and Zhang Z 1993 Matching pursuits with time–frequency dictionaries *IEEE Trans. Signal Process.* **41** 3397–415

[18] Michel V 1999 *A Multiscale Method for the Gravimetry Problem: Theoretical and Numerical Aspects of Harmonic and Anharmonic Modelling* PhD thesis University of Kaiserslautern (Aachen: Shaker Verlag)

[19] Michel V 2002 *A Multiscale Approximation for Operator Equations in Separable Hilbert Spaces – Case Study: Reconstruction and Description of the Earth's Interior* Habilitation thesis University of Kaiserslautern (Aachen: Shaker Verlag)

[20] Michel V 2005 Regularized wavelet-based multiresolution recovery of the harmonic mass density distribution from data of the Earth's gravitational field at satellite height *Inv. Probl.* **21** 997–1025

[21] Michel V 2010 Tomography: problems and multiscale solutions *Handbook of Geomathematics* ed W Freeden et al (Heidelberg: Springer) pp 949–72

[22] Michel V 2011 *Lectures on Constructive Approximation - Fourier, Spline and Wavelet Methods on the Real Line, the Sphere and the Ball* (Boston: Birkhäuser), submitted

[23] Michel V and Fokas A S 2008 A unified approach to various techniques for the non-uniqueness of the inverse gravimetric problem and wavelet-based methods *Inv. Probl.* **24** 045091

[24] Michel V and Wolf K 2008 Numerical aspects of a spline-based multiresolution recovery of the harmonic mass density out of gravity functionals *Geophys. J. Int.* **173** 1–16

[25] Miranian L 2004 Slepian functions on the sphere, generalized Gaussian quadrature rule *Inv. Probl.* **20** 877–92

[26] Needell D and Tropp J A 2009 CoSaMP: iterative signal recovery from incomplete and inaccurate samples *Appl. Comput. Harmon. Anal.* **26** 301–21

[27] Pavlis N K, Holmes S A, Kenyon S C and Factor J K 2008 An Earth gravitational model to degree 2160: EGM2008 *presented at the 2008 General Assembly of the European Geosciences Union* Vienna

[28] Simons F J 2010 Slepian functions and their use in signal estimation and spectral analysis *Handbook of Geomathematics* ed W Freeden (Heidelberg: Springer) pp 891–923

[29] Simons F J and Dahlen F A 2006 Spherical Slepian functions and the polar gap in geodesy *Geophys. J. Int.* **166** 1039–61

[30] Simons F J, Dahlen F A and Wieczorek M A 2006 Spatiospectral concentration on a sphere *SIAM Review* **48** 504–36

[31] Simons F J, Loris I, Nolet G, Daubechies I C, Voronin S, Judd J S, Vetter P A, Charléty J and Vonesch C 2011 Solving or resolving global tomographic models with spherical wavelets, and the scale and sparsity of seismic heterogeneity *Geophys. J. Int.* **187** 969–88

[32] Vincent P and Bengio Y 2002 Kernel matching pursuit *Mach. Learn.* **48** 169–91

Siegen Preprints on Geomathematics

The preprint series “Siegen Preprints on Geomathematics” was established in 2010. See www.geomathematics-siegen.de for details and a contact address. At present, the following preprints are available:

1. P. Berkel, D. Fischer, V. Michel: *Spline multiresolution and numerical results for joint gravitation and normal mode inversion with an outlook on sparse regularisation*, 2010.
2. M. Akram, V. Michel: *Regularisation of the Helmholtz decomposition and its application to geomagnetic field modelling*, 2010.
3. V. Michel: *Optimally Localized Approximate Identities on the 2-Sphere*, 2010.
4. N. Akhtar, V. Michel: *Reproducing Kernel Based Splines for the Regularization of the Inverse Spheroidal Gravimetric Problem*, 2011.
5. D. Fischer, V. Michel: *Sparse Regularization of Inverse Gravimetry - Case Study: Spatial and Temporal Mass Variations in South America*, 2011.
6. A.S. Fokas, O. Hauk, V. Michel: *Electro-Magneto-Encephalography for the three-Shell Model: Numerical Implementation for Distributed Current in Spherical Geometry,* 2011.
7. M. Akram, I. Amina, V. Michel: *A Study of Differential Operators for Complete Orthonormal Systems on a 3D Ball*, 2011.
8. D. Fischer, V. Michel: *How to combine spherical harmonics and localized bases for regional gravity modelling and inversion*, 2012.

Geomathematics Group Siegen
Prof. Dr. Volker Michel

Contact at:
Geomathematics Group
Department of Mathematics
University of Siegen
Walter-Flex-Str. 3
57068 Siegen
www.geomathematics-siegen.de

UNIVERSITÄT SIEGEN