PARAMETER CHOICE STRATEGIES FOR LEAST-SQUARES APPROXIMATION OF NOISY SMOOTH FUNCTIONS ON THE SPHERE

S. V. PEREVERZYEV*, I. H. SLOAN†, AND P. TKACHENKO*

Abstract. We consider a polynomial reconstruction of smooth functions from their noisy values at discrete nodes on the unit sphere by a variant of the regularized least-squares method of An et al., SIAM J. Numer. Anal. 50 (2012), 1513–1534. As nodes we use the points of a positive-weight cubature formula that is exact for all spherical polynomials of degree up to $2M$, where M is the degree of the reconstructing polynomial. We first obtain a reconstruction error bound in terms of the regularization parameter and the penalization parameters in the regularization operator. Then we discuss *a priori* and *a posteriori* strategies for choosing these parameters. Finally, we give numerical examples illustrating the theoretical results.

Key words. spherical polynomial, parameter choice strategy, regularization, penalization, continuous function on the sphere, *a posteriori* rules

AMS subject classifications. 65D32, 65H10

1. Introduction. In recent decades methods for approximation of a continuous function y on the sphere $\mathbb{S}^2 := \left\{\mathbf{x} = (x_1, x_2, x_3)^T \in \mathbb{R}^3 : \; x_1^2 + x_2^2 + x_3^2 = 1\right\}$ by means of polynomials have been discussed by many authors (see, for example, [9, 31, 39, 40]). Often the underlying motivation has been the need to approximate geophysical quantities. For example, such a task appears in the satellite gravity gradiometry problem (SGG-problem) [7], p. 120, 262, [28], in which the task is to find a spherical harmonic representation of Earth's gravitational potential from satellite observations. The present study was motivated by this example. We shall return to it several times throughout the paper.

The mathematical problem considered in this paper is to find a polynomial approximation to $y \in C(\mathbb{S}^2)$, given noisy data values $y^\epsilon(\mathbf{x}_i)$ at points $\mathbf{x}_i \in \mathbb{S}^2$, $i = 1, \dots N$, using a least-squares strategy developed in [1]. (In the SGG application the sphere in question is determined by the satellite orbits. The gravitational potential at the satellite height is smoother than at earth's surface, a complicating feature for the inverse problem.) We shall assume, in a slight generalization of [1], that the point set $X_N := \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ consists of the points of a cubature rule which is exact for all polynomials $p \in \mathbb{P}_{2M}$, where $\mathbb{P}_M$ is the set of all spherical polynomials of degree less than or equal to M, or in other words the restriction to $\mathbb{S}^2$ of the set of all polynomials in $\mathbb{R}^3$ of degree less than or equal to M. Thus the point set must satisfy

$$\forall p \in \mathbb{P}_{2M}, \quad \sum_{i=1}^{N} w_i p(\mathbf{x}_i) = \int_{\mathbb{S}^2} p(\mathbf{x}) d\omega(\mathbf{x}), \tag{1.1}$$

where $d\omega(\mathbf{x})$ denotes area measure on $\mathbb{S}^2$, and $w_i, i = 1, \dots, N$ are positive cubature weights associated with the pointset X_N. For sufficiently large N one can find in the

*Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Altenbergerstrasse 69, 4040 Linz, Austria (sergei.pereverzyev@oeaw.ac.at, pavlo.tkachenko@oeaw.ac.at).

†School of Mathematics and Statistics, University of New South Wales, Sydney NSW 2052, Australia (i.sloan@unsw.edu.au).

literature a variety of suitable cubature formulas (see, e.g., [21, 11, 14, 42]). Moreover, in principle the point sets for such a rule can be generated by selecting from any sufficiently dense set of points on the sphere, see [25, 15, 10].

The strategy is to take the approximant $y_M \in \mathbb{P}_M$ to be the minimizer of the regularized discrete least-squares problem

$$y_M = \arg\min \left\{ \sum_{i=1}^{N} w_i (p(\mathbf{x}_i) - y^\epsilon(\mathbf{x}_i))^2 + \alpha \sum_{i=1}^{N} w_i (R_M p(\mathbf{x}_i))^2,\ p \in \mathbb{P}_M \right\}, \tag{1.2}$$

where $y^\epsilon(\mathbf{x}_i) := y(\mathbf{x}_i) + \epsilon_i$ represent noisy values of a perturbed version y^ϵ of the original function y calculated at the points of X_N, α is a regularization parameter, and $R_M : \mathbb{P}_M \to \mathbb{P}_M$ is a linear "penalization" operator given by

$$\begin{aligned} R_M p(\mathbf{x}) &:= \sum_{k=0}^{M} \beta_k \frac{2k+1}{4\pi} \int_{\mathbb{S}^2} P_k(\mathbf{x}\cdot\mathbf{z}) p(\mathbf{z}) d\omega(\mathbf{z}) \\ &= \sum_{k=0}^{M} \beta_k \frac{2k+1}{4\pi} \sum_{i=1}^{N} w_i P_k(\mathbf{x}\cdot\mathbf{x}_i) p(\mathbf{x}_i),\ \mathbf{x} \in \mathbb{S}^2,\ p \in \mathbb{P}_M, \end{aligned} \tag{1.3}$$

where P_k is the Legendre polynomial of degree k, and in the last step we used (1.1). Here the numbers $\beta_k, k = 1, \dots, M$ are a non-decreasing sequence of positive parameters. With β_0 fixed in some appropriate way, the important feature of the parameters β_k is their rate of growth. The central task in this paper will be to assign appropriate values for the β_k.

As pointed out in [1], the expression in (1.3) is the most general rotationally invariant expression for a linear operator on the space $\mathbb{P}_M$.

In [1] the point set X_N was taken to be a spherical $2M$-design, which simply means that (1.1) must hold with equal weights $w_i = 4\pi/N$. We gain considerable freedom in this paper by allowing general positive weights w_i in (1.1). The only effective difference in the present approximation scheme is that the least-squares problem (1.2) is slightly non-standard because of the appearance of the cubature weights w_i.

It was observed in numerical experiments in [1] that a proper choice of the penalization operator R_M together with the regularization parameter α can significantly improve the approximation. However, the choice of the model parameters in (1.3) was not settled, and still remains an open issue. In our paper we will tackle this crucial question by proposing parameter choice strategies (strategies for choosing β_k and α) that allow good approximation of noisy smooth functions on the sphere.

The paper is organized as follows: in the next section we present necessary preliminaries, and give an explicit solution of the regularized least-squares problem.. In Section 3 we derive theoretical error bounds for the resulting approximation. Sections 4 and 5 discuss error bounds and parameter choice strategies. Finally, in the last section we present some numerical experiments that test the theoretical results from previous sections.

2. Preliminaries. We introduce a real spherical harmonic basis for $\mathbb{P}_M$, see [23]

$$\{Y_{k,j} :\ k = 0, 1, ..., M,\ j = 1, ..., 2k+1\},$$

assumed to be orthonormal with respect to the standard L^2 inner product,

$$\langle f, g \rangle_{L^2(\mathbb{S}^2)} := \int_{\mathbb{S}^2} f(\mathbf{x}) g(\mathbf{x}) d\omega(\mathbf{x}).$$

Then for $p \in \mathbb{P}_M$ an arbitrary spherical polynomial of degree $\leq M$ there exists a unique vector $\gamma = (\gamma_{k,j}) \in \mathbb{R}^{(M+1)^2}$ such that

$$p(\mathbf{x}) = \sum_{k=0}^{M} \sum_{j=1}^{2k+1} \gamma_{k,j} Y_{k,j}(\mathbf{x}), \quad \mathbf{x} \in \mathbb{S}^2. \tag{2.1}$$

The addition theorem for spherical harmonics (see [23]), which asserts

$$\sum_{j=1}^{2k+1} Y_{k,j}(\mathbf{x}) Y_{k,j}(\mathbf{z}) = \frac{2k+1}{4\pi} P_k(\mathbf{x} \cdot \mathbf{z}), \quad \mathbf{x}, \mathbf{z} \in \mathbb{S}^2, \tag{2.2}$$

will play an important role.

The assumption that a function y on the unit sphere is continuous implies that $y \in L^2(\mathbb{S}^2)$, and hence that its Fourier coefficients $\langle Y_{k,j}, y \rangle_{L^2(\mathbb{S}^2)}$ with respect to the basis of spherical harmonics are square-summable, i.e.

$$\sum_{k=0}^{\infty} \sum_{j=1}^{2k+1} \left| \langle Y_{k,j}, y \rangle_{L^2(\mathbb{S}^2)} \right|^2 < \infty.$$

To measure any additional smoothness of y it is convenient to introduce a Hilbert space $W^{\phi,\beta}$ that is especially tailored to the particular problem, namely

$$y \in W^{\phi,\beta} := \left\{ g : \|g\|^2_{W^{\phi,\beta}} := \sum_{k=0}^{\infty} \sum_{j=1}^{2k+1} \frac{\left| \langle Y_{k,j}, g \rangle_{L^2(\mathbb{S}^2)} \right|^2}{\phi^2(\beta_k^{-2})} < \infty \right\},$$

where ϕ is an non-decreasing function such that $\phi(0) = 0$ and $\beta = \{\beta_0, \beta_1, ..., \beta_M, \ldots\}$ is the sequence of coefficients appearing in the regularizer (1.3). In the literature, see, e.g., [17], the function ϕ goes under the name of *smoothness index function*.

In this context the smoothness of y is encoded in ϕ and β. For example, if the smoothness index function $\phi(t)$ and the sequence $\beta = \{\beta_k\}$ increase polynomially with t and k such that $\phi(t) = O\left(t^{\nu_1}\right), \beta_k = O\left(k^{\nu_2}\right), \nu_1\nu_2 > 1/2$, then the space $W^{\phi,\beta}$ becomes a spherical Sobolev space $H_{2\nu_1\nu_2}$ (see, e.g., [7], p. 64), and a spherical analog of the fundamental lemma due to Sobolev (see [7], Lemma 2.1.5) says that $H_{2\nu_1\nu_2}$ is embedded in the space $C^{(\nu)}(\mathbb{S}^2)$ of functions, which have ν continuous derivatives on $\mathbb{S}^2, \nu < 2\nu_1\nu_2 - 1$, and are the restrictions to $\mathbb{S}^2$ of functions satisfying the Laplace equation in the outer space of $\mathbb{S}^2$ and being regular at infinity. Then Jackson's theorem on the sphere (see [30], Theorem 3.3) tells us that for $y \in W^{\phi,\beta}$, there holds

$$\inf_{p \in \mathbb{P}_M} \|y - p\|_{C(\mathbb{S}^2)} = O\left(M^{-\nu}\right), \ \nu < 2\nu_1\nu_2 - 1. \tag{2.3}$$

On the other hand, if the sequence $\beta = \{\beta_k\}$ increases exponentially then for polynomially increasing ϕ and $y \in W^{\phi,\beta}$ we have

$$\inf_{p \in \mathbb{P}_M} \|y - p\|_{C(\mathbb{S}^2)} = O\left(e^{-qM}\right),$$

where q is some positive number that does not depend on M.

In the error analysis later in the paper we make use of a linear polynomial approximation that in a certain precise sense mimics best approximation in the space of spherical polynomials of half the degree. The approximation, studied in [20, 6, 35], approximates a function $y \in C(\mathbb{S}^2)$ by $V_M y \in \mathbb{P}_M$ defined by

$$V_M y(\mathbf{x}) := \sum_{k=0}^{M} h\left(\frac{k}{M}\right) \sum_{j=1}^{2k+1} Y_{k,j}(\mathbf{x}) \left\langle Y_{k,j}, y\right\rangle_{L^2(\mathbb{S}^2)} = \sum_{k=0}^{M} h\left(\frac{k}{M}\right) \frac{2k+1}{4\pi} \int_{\mathbb{S}^2} P_k(\mathbf{x}\cdot\mathbf{z}) y(\mathbf{z}) d\omega(\mathbf{z}), \tag{2.4}$$

where h is a real-valued function on $\mathbb{R}^+$, called a filter function, which satisfies

$$h(t) \in [0,1]\, \forall\, t \in \mathbb{R}^+, \quad h(t) = \begin{cases} 1, & t \in [0, 1/2]\,, \\ 0, & t \in (1, \infty)\,. \end{cases} \tag{2.5}$$

It is shown in [35] that for suitable choices of the filter h (including any filter in $C^3(\mathbb{R}^+)$, or the unique C^1 quadratic spline with breakpoints at 1/2, 3/4 and 1 that satisfies (2.5)), the norm of the operator V_M as an operator from $\mathbb{P}_M$ to $C(\mathbb{S}^2)$ is bounded independently of M. Since, as is easily seen, V_M reproduces polynomials of degree less than or equal to $M/2$, it follows in the usual way that

$$\|y - V_M y\|_{C(\mathbb{S}^2)} \le c \inf_{p \in \mathbb{P}_{[M/2]}} \|y - p\|_{C(\mathbb{S}^2)}\,,$$

where $[\cdot]$ denotes the floor function. (In this paper c is a generic constant, which may take different values at different occurrences.) In view of (2.3), for polynomially increasing ϕ, β and $y \in W^{\phi,\beta}$ we have

$$\|y - V_M y\|_{C(\mathbb{S}^2)} \le c\,[M/2]^{-\nu} \le c M^{-\nu}.$$

On the other hand, for exponentially increasing β and polynomially increasing ϕ the theory [32] suggests taking $h(t) = 1$ for $t \in [0,1]$ (in which case $V_M y$ is just the Mth-degree partial sum of the Fourier-Laplace series). Then for $y \in W^{\phi,\beta}$ there holds

$$\|y - V_M y\|_{C(\mathbb{S}^2)} \le c\sqrt{M} \inf_{p \in \mathbb{P}_M} \|y - p\|_{C(\mathbb{S}^2)} \le c\sqrt{M} e^{-qM}.$$

3. Weighted regularized least-squares problem and its solution. The penalization operator (1.3) can equivalently be written, using the addition theorem (2.2) and (2.1), as

$$R_M p(\mathbf{x}) = \sum_{k=0}^{M} \beta_k \sum_{j=1}^{2k+1} Y_{k,j}(\mathbf{x}) \left\langle Y_{k,j}, p\right\rangle_{L^2(\mathbb{S}^2)} = \sum_{k=0}^{M} \beta_k \sum_{j=1}^{2k+1} \gamma_{k,j} Y_{k,j}(\mathbf{x}), \tag{3.1}$$

allowing us to write the minimization problem as one of linear algebra. For the noisy function y^ϵ defined on $\mathbb{S}^2$, let $\mathbf{y}^\epsilon := \mathbf{y}^\epsilon(X_N)$ be the column vector

$$\mathbf{y}^\epsilon = \left[y^\epsilon(\mathbf{x}_1), ..., y^\epsilon(\mathbf{x}_N)\right]^T \in \mathbb{R}^N,$$

and let $\mathbf{Y_M} := \mathbf{Y_M}(X_N) \in \mathbb{R}^{(M+1)^2 \times N}$ be the matrix of spherical harmonics evaluated at the points of X_N. Using this notation we can reduce the minimization problem in (1.2) to the following discrete minimization problem:

$$\min_{\boldsymbol{\gamma} \in \mathbb{R}^{(M+1)^2}} \left\| \mathbf{W}^{1/2} \mathbf{Y_M}^T \boldsymbol{\gamma} - \mathbf{W}^{1/2} \mathbf{y}^\epsilon \right\|_2^2 + \alpha \left\| \mathbf{W}^{1/2} \mathbf{R_M}^T \boldsymbol{\gamma} \right\|_2^2, \ \alpha > 0, \tag{3.2}$$

where $\|\cdot\|_2$ is the standard Euclidean vector norm, $\mathbf{R_M} := \mathbf{R_M}(X_N) = \mathbf{B_M Y_M} \in \mathbb{R}^{(M+1)^2 \times N}$, $\mathbf{B_M}$ is a positive diagonal matrix defined by

$$\mathbf{B_M} := \operatorname{diag}(\beta_0, \underbrace{\beta_1, \beta_1, \beta_1}_{3}, ..., \underbrace{\beta_M, \beta_M, ..., \beta_M}_{2M+1}) \in \mathbb{R}^{(M+1)^2 \times (M+1)^2}, \tag{3.3}$$

and $\mathbf{W}$ is a diagonal matrix of cubature weights

$$\mathbf{W} := \operatorname{diag}(w_1, ..., w_N) \in \mathbb{R}^{N \times N}.$$

The solution of (1.2) can be found from the following system of linear equations

$$(\mathbf{Y_M W Y_M}^T + \alpha \mathbf{B_M Y_M W Y_M}^T \mathbf{B_M})\gamma = \mathbf{Y_M W y}^\epsilon. \tag{3.4}$$

THEOREM 3.1. *Assume $y^\epsilon \in C(\mathbb{S}^2)$. Let $M > 0$ be given, and let (1.1) hold true for the set of points X_N. Then (3.4) has the unique solution $\boldsymbol{\gamma} = (\gamma_{k,j}) \in \mathbb{R}^{(M+1)^2}$,*

$$\gamma_{k,j} = \frac{1}{1 + \alpha\beta_k^2} \sum_{i=1}^{N} w_i Y_{k,j}(\mathbf{x}_i) y^\epsilon(\mathbf{x}_i), \tag{3.5}$$

and the minimizer of (1.2) is given by

$$\begin{aligned} y_M(\mathbf{x}) = T^\beta_{\alpha,M} y^\epsilon(\mathbf{x}) : &= \sum_{k=0}^{M} \sum_{j=1}^{2k+1} \frac{Y_{k,j}(\mathbf{x})}{1 + \alpha\beta_k^2} \sum_{i=1}^{N} w_i Y_{k,j}(\mathbf{x}_i) y^\epsilon(\mathbf{x}_i) \\ &= \sum_{k=0}^{M} \frac{2k+1}{4\pi} \frac{1}{1 + \alpha\beta_k^2} \sum_{i=1}^{N} w_i P_k(\mathbf{x} \cdot \mathbf{x}_i) y^\epsilon(\mathbf{x}_i). \end{aligned} \tag{3.6}$$

Proof. On using (1.1) we have

$$\sum_{i=1}^{N} w_i Y_{k,j}(\mathbf{x}_i) Y_{\kappa,\iota}(\mathbf{x}_i) = \langle Y_{k,j}, Y_{\kappa,\iota} \rangle_{L^2(\mathbb{S}^2)} = \delta_{k,\kappa}\delta_{j,\iota},$$

where $k, \kappa = 0, ..., M, \ j = 1, ..., 2k+1, \ \iota = 1, ..., 2\kappa + 1$. Thus $\mathbf{Y_M W Y_M}^T$ is the identity matrix. Since $\mathbf{B_M}$ and $\mathbf{W}$ are diagonal matrices, the solution of (3.4) is given by (3.5) and from (2.1) we obtain (3.6). ∎

Remark 3.1. Note that one can also employ fast iterative algorithms for scattered least squares [12] to find the minimizer (1.2). Moreover, the evaluation of the coefficients (3.5) can be realized with fast spherical Fourier transform presented in [13].

4. Error bounds. In this section we estimate the uniform error of approximation of y by y_M, see (3.6). It is convenient here to regard y^ϵ as a continuous function on $\mathbb{S}^2$, constructed by some interpolation process from its values on the discrete set X_N. The operator $T^\beta_{\alpha,M}$ defined in (3.6) can then be considered as an operator on the space $C(\mathbb{S}^2)$. Since $y_M = T^\beta_{\alpha,M} y^\epsilon$ it is clear that

$$y - y_M = y - T^\beta_{\alpha,M} V_M y + T^\beta_{\alpha,M}(V_M y - y + y - y^\epsilon),$$

and hence

$$\begin{aligned}\|y - y_M\|_{C(\mathbb{S}^2)} \le \left\| y - T^\beta_{\alpha,M} V_M y \right\|_{C(\mathbb{S}^2)} \\ + \left\| T^\beta_{\alpha,M} \right\|_{C(\mathbb{S}^2)} \left(\|y - V_M y\|_{C(\mathbb{S}^2)} + \|y - y^\epsilon\|_{C(\mathbb{S}^2)} \right),\end{aligned} \tag{4.1}$$

where $\left\| T^\beta_{\alpha,M} \right\|_{C(\mathbb{S}^2)}$ is the norm of the operator $T^\beta_{\alpha,M} : C(\mathbb{S}^2) \to C(\mathbb{S}^2)$.

Let $\epsilon = [\epsilon_1, \epsilon_2, ..., \epsilon_N] \in \mathbb{R}^N$, and $\|\epsilon\|_\infty = \max |\epsilon_i|$. It is natural to assume, and from now on we shall do so, that $\|y - y^\epsilon\|_{C(\mathbb{S}^2)} = \|\epsilon\|_\infty$. This means that we adopt the deterministic noise model, which allows the worst noise level at any point of X_N. Then it is also natural to assume that M is large enough to ensure $\|y - V_M y\|_{C(\mathbb{S}^2)} \le \|\epsilon\|_\infty$, since otherwise data noise is dominated by the approximation error and no regularization is required. Then the bound (4.1) can be reduced to

$$\|y - y_M\|_{C(\mathbb{S}^2)} \le \left\| y - T^\beta_{\alpha,M} V_M y \right\|_{C(\mathbb{S}^2)} + 2 \left\| T^\beta_{\alpha,M} \right\|_{C(\mathbb{S}^2)} \|\epsilon\|_\infty . \tag{4.2}$$

We will call the first term of the right-hand side in (4.2) the *regularization error* and the second the *noise propagation error*.

The noise propagation error can be quantified by the following result for the norm of $T^\beta_{\alpha,M}$, which is a consequence of (3.6).

THEOREM 4.1. *Under the conditions of Theorem 3.1*

$$\begin{aligned}\left\| T^\beta_{\alpha,M} \right\|_{C(\mathbb{S}^2)} &= \max_{\mathbf{x} \in \mathbb{S}^2} \sum_{i=1}^N w_i \left| \sum_{k=0}^M \frac{2k+1}{4\pi(1+\alpha\beta_k^2)} P_k(\mathbf{x} \cdot \mathbf{x}_i) \right| \\ &\le \max_{\mathbf{x} \in \mathbb{S}^2} \sum_{i=1}^N w_i \sum_{k=0}^M \frac{2k+1}{4\pi(1+\alpha\beta_k^2)} \left| P_k(\mathbf{x} \cdot \mathbf{x}_i) \right| .\end{aligned} \tag{4.3}$$

Theorem 4.1 reduces to Proposition 5.1 in [1] on setting $w_i = 4\pi/N$, but note that the result as stated in [1] corresponds to the upper bound in (4.3), and so is not correctly stated.

Now we are going to bound the regularization error $\left\| y - T^\beta_{\alpha,M} V_M y \right\|_{C(\mathbb{S}^2)}$. We start with the following decomposition

$$y - T^\beta_{\alpha,M} V_M y = y - T_{0,M} V_M y + (T_{0,M} - T^\beta_{\alpha,M}) V_M y, \tag{4.4}$$

where $T_{0,M}$ is the so-called hyperinterpolation operator [34],

$$T_{0,M}g(\mathbf{x}) = \sum_{k=0}^{M}\sum_{j=1}^{2k+1} Y_{k,j}(\mathbf{x})\sum_{i=1}^{N} w_i Y_{k,j}(\mathbf{x}_i)g(\mathbf{x}_i) = \sum_{k=0}^{M}\frac{2k+1}{4\pi}\sum_{i=1}^{N} w_i P_k(\mathbf{x}\cdot\mathbf{x}_i)g(\mathbf{x}_i). \tag{4.5}$$

From (4.5) and (1.1) it immediately follows that for any $p \in \mathbb{P}_M$ we have $T_{0,M}p = p$. Therefore, $T_{0,M}V_M y = V_M y$. In view of this property and the decomposition (4.4) we can derive a bound for the regularization error

$$\left\| y - T^{\beta}_{\alpha,M}V_M y\right\|_{C(\mathbb{S}^2)} \le \|y - V_M y\|_{C(\mathbb{S}^2)} + \left\|(T_{0,M} - T^{\beta}_{\alpha,M})V_M y\right\|_{C(\mathbb{S}^2)} \le \|\epsilon\|_\infty + \left\|(T_{0,M} - T^{\beta}_{\alpha,M})V_M y\right\|_{C(\mathbb{S}^2)}. \tag{4.6}$$

An estimate of the term $\left\|(T_{0,M} - T^{\beta}_{\alpha,M})V_M\right\|_{C(\mathbb{S}^2)}$ in (4.6) is given by the following theorem.

THEOREM 4.2. *Assume that the smoothness index function ϕ is such that the function $t \to t/\phi(t)$ is monotone. Then for $y \in W^{\phi,\beta}$ there holds*

$$\left\|(T_{0,M} - T^{\beta}_{\alpha,M})V_M y\right\|_{C(\mathbb{S}^2)} \le cM\hat{\phi}(\alpha)\,\|y\|_{W^{\phi,\beta}}, \tag{4.7}$$

where $\hat{\phi}(\alpha) = \phi(\alpha)$ if $t/\phi(t)$ is non-decreasing, and $\hat{\phi}(\alpha) = \alpha$ if $t/\phi(t)$ is non-increasing.

Proof. In view of (3.6), (4.5) and (2.4), together with the fact that the cubature formula in (1.1) is exact for $p \in \mathbb{P}_{2M}$, we may write

$$\begin{aligned}
&\left\|(T_{0,M} - T^{\beta}_{\alpha,M})V_M y\right\|_{C(\mathbb{S}^2)}\\
&= \left\|\sum_{k=0}^{M}\sum_{j=1}^{2k+1} Y_{k,j}(\cdot)\frac{\alpha\beta_k^2}{1+\alpha\beta_k^2}\langle Y_{k,j}, V_M y\rangle_{L^2(\mathbb{S}^2)}\right\|_{C(\mathbb{S}^2)}\\
&= \left\|\sum_{k=0}^{M}\sum_{j=1}^{2k+1} h\left(\frac{k}{M}\right) Y_{k,j}(\cdot)\frac{\alpha\beta_k^2}{1+\alpha\beta_k^2}\langle Y_{k,j}, y\rangle_{L^2(\mathbb{S}^2)}\right\|_{C(\mathbb{S}^2)},
\end{aligned}$$

where in the last step we used $\langle Y_{k,j}, V_M y\rangle_{L^2(\mathbb{S}^2)} = h(k/M)\langle Y_{k,j}, y\rangle_{L^2(\mathbb{S}^2)}$. Now using the Nikolskii inequality (see, e.g., [24], Proposition 2.5) and also $h(k/M) \le 1$, we obtain

$$\begin{aligned}
\|(T_{0,M}-T^{\beta}_{\alpha,M})V_M y\|_{C(\mathbb{S}^2)} &\le cM \left\| \sum_{k=0}^{M}\sum_{j=1}^{2k+1} h\left(\frac{k}{M}\right) Y_{k,j}\frac{\alpha\beta_k^2}{1+\alpha\beta_k^2}\langle Y_{k,j},y\rangle_{L^2(\mathbb{S}^2)}\right\|_{L^2(\mathbb{S}^2)} \\
&= cM\left(\sum_{k=0}^{M}\sum_{j=1}^{2k+1} h\left(\frac{k}{M}\right)^2\left(\frac{\alpha\beta_k^2}{1+\alpha\beta_k^2}\right)^2\left|\langle Y_{k,j},y\rangle_{L^2(\mathbb{S}^2)}\right|^2\right)^{1/2} \\
&\le cM\left(\sum_{k=0}^{M}\sum_{j=1}^{2k+1}\left(\frac{\alpha\beta_k^2}{1+\alpha\beta_k^2}\right)^2\phi^2(\beta_k^{-2})\frac{\left|\langle Y_{k,j},y\rangle_{L^2(\mathbb{S}^2)}\right|^2}{\phi^2(\beta_k^{-2})}\right)^{1/2} \\
&\le cM\sup_{u\in[0,\beta_0^{-2}]}\left|\frac{\alpha}{\alpha+u}\phi(u)\right|\|y\|_{W^{\phi,\beta}} \le cM\hat{\phi}(\alpha)\,\|y\|_{W^{\phi,\beta}},
\end{aligned}$$

where the last inequality follows from [17], Proposition 2.7. □

It is instructive to note that if, for example, $\phi(t)=t^{\nu}$, then the function $\hat{\phi}$ defined in Theorem 4.2 is given by

$$\hat{\phi}(\alpha)=\begin{cases}\alpha, & \nu\ge 1,\\ \alpha^{\nu}, & 0<\nu<1.\end{cases}$$

Thus the error bound in the theorem does not improve if $\phi(t)$ grows faster than t.

5. Parameter choice strategies. In this section we will be concerned with the choice of the design parameters for the least-squares approximation y_M, namely the regularization parameter α and the penalization parameters β_k. In the first subsection we discuss an *a priori* choice for the penalization parameters β_k. In the next subsection we consider an adaptive strategy for choosing the regularization parameter α. In the third subsection we present an *a posteriori* choice for the penalization parameters β_k.

The choice of parameters is motivated by the error bound (4.2) for $y-y_M$. From (4.6) and (4.7) it follows that the bound (4.2) can be reduced to the following:

$$\begin{aligned}
\|y-y_M\| &\le \|\epsilon\|_\infty + cM\hat{\phi}(\alpha)\,\|y\|_{W^{\phi,\beta}} + 2\,\|\epsilon\|_\infty \left\|T^{\beta}_{\alpha,M}\right\|_{C(\mathbb{S}^2)} \\
&\le cM\hat{\phi}(\alpha)\,\|y\|_{W^{\phi,\beta}} + c\,\|\epsilon\|_\infty\left\|T^{\beta}_{\alpha,M}\right\|_{C(\mathbb{S}^2)}.
\end{aligned} \tag{5.1}$$

5.1. A priori choice of the penalization parameters. For definiteness, we assume in this subsection that $\phi(t)=t$, which means that $\hat{\phi}$ has the highest order in α, namely $\hat{\phi}(\alpha)=\alpha$. The error bound (5.1) now provides useful guidance in the choice of the regularization parameters β_k. If β_0 is considered to be fixed, and we increase the rate of growth of the β_k, then the first term on the right-hand side of the last line of (5.1) will increase, while from (4.3) the second term has an upper bound that decreases with increasing rate of growth of the β_k. Even more can be said: for the first term to be finite the $W^{\phi,\beta}$ norm of y must be finite, which imposes the constraint

$$\sum_{k=0}^{\infty}\sum_{j=1}^{2k+1}\beta_k^4\,\langle Y_{k,j},y\rangle^2_{L^2(\mathbb{S}^2)}<\infty. \tag{5.2}$$

To see what this condition means in a particular application, we consider the SGG-problem mentioned in the Introduction. In this problem y is the second order radial derivative of the gravitational potential measured pointwise at the orbital sphere of a satellite. It can be shown [8, 16, 37] that after a proper normalization of this sphere to $\mathbb{S}^2$ we have

$$\langle Y_{k,j}, y\rangle_{L^2(\mathbb{S}^2)} = a_k g_{k,j}, \tag{5.3}$$

where $a_k = \left(\frac{R}{\rho}\right)^k \frac{(k+1)(k+2)}{\rho^2}$, ρ is the radius of the orbital sphere, R is the radius of the surface of the Earth considered as a sphere, and $\{g_{k,j}\}$ is some (unknown) sequence of scaled Fourier coefficients of the gravitational potential g measured at the surface of the Earth. It is well-known (see, e.g., [38, 8]) that in the scale of the spherical Sobolev spaces $\{H_s\}$ mentioned above the Earth's gravitational potential has a smoothness index $s = 3/2$ at least, which means that the sequence $\{g_{k,j}\}$ should satisfy the requirement

$$\sum_{k=0}^{\infty}\sum_{j=1}^{2k+1}(k+1/2)^3 g_{k,j}^2 < \infty.$$

In view of the last requirement, the condition (5.2) is satisfied by the choice

$$\beta_k = a_k^{-1/2}(k+1/2)^{3/4}, \quad k = 0, 1, ... \tag{5.4}$$

Of course the condition (5.2) will also be satisfied if the β_k increase more slowly, but at the likely expense of a larger second term in the error bound (5.1).

Since $R < \rho$, it is clear that the β_k given by (5.4) increase exponentially. This is natural in view of the exponential decrease of the Fourier coefficients (5.3) of the approximated function, which implies that the exact function as measured at the satellite height is very smooth, even analytic. The regularization scheme (1.2) with weights (5.4) will penalize the presence of oscillating coefficients with large indexes in the approximant $T^{\beta}_{\alpha,M} y^{\epsilon}$. In the last section we illustrate a good performance of the scheme (1.2) with these penalization weights.

5.2. Regularization parameter choice strategy. For regularization of our problem we will implement an adaptive regularization parameter choice strategy known as the balancing principle (see, e.g., [17, 18, 29] and references therein). In this method the regularization parameter α is selected from some finite set, say $\Delta_L := \{\alpha_i = q^i\alpha_0, i = 1, 2, ..., L\}$, with $q \in (0, 1)$ and L large enough.

Applying the balancing principle to our problem we start with the smallest parameter α_L and increase stepwise $\alpha_{i-1} = \alpha_i/q, i = L, L-1, ...,$ until $\alpha_* := \alpha_z$ is the parameter for which

$$\left\|T^{\beta}_{\alpha_z,M} y^{\epsilon} - T^{\beta}_{\alpha_{z+1},M} y^{\epsilon}\right\|_{C(\mathbb{S}^2)} > æ \, \|\epsilon\|_\infty \left\|T^{\beta}_{\alpha_{z+1},M}\right\|_{C(\mathbb{S}^2)}$$

for the first time. Here æ is a design parameter. In all our numerical tests with the balancing principle (BP) reported below in Section 6, the value of æ is fixed as æ = 0.002 and is data independent, while the value of the regularization parameter α_* chosen according to BP varies with data.

Note that for choosing α_* we need only the knowledge of (3.6) and an upper bound of $\left\|T^{\beta}_{\alpha,M}\right\|_{C(\mathbb{S}^2)}$ given by (4.3).

In the Section 6 we will present a numerical test showing a good reconstruction of the function on the sphere from noisy observations with the above *a posteriori* regularization parameter. It is instructive to see that in all tests BP performs at the level of the ideal parameter choice $\alpha \in \Delta_L$.

5.3. A posteriori choice of the penalization weights. We start with the observation that the space $\mathbb{P}_M$ of spherical polynomials p is a reproducing kernel Hilbert space (RKHS) $\mathcal{H}$. By the Riesz representation theorem, to every RHKS $\mathcal{H}$ there corresponds a unique symmetric positive definite function $K : \mathbb{S}^2 \times \mathbb{S}^2 \to \mathbb{R}$, called the reproducing kernel of $\mathcal{H} = \mathcal{H}_K$, that has the following reproducing property: $p(\mathbf{x}) = \langle p(\cdot), K(\cdot, \mathbf{x})\rangle_{\mathcal{H}_K}$. A comprehensive theory of RKHSs can be found in [2].

It is easy to check that the kernel

$$K(\mathbf{x}, \mathbf{z}) = \sum_{k=0}^{M} \beta_k^{-2} \sum_{j=1}^{2k+1} Y_{k,j}(\mathbf{x}) Y_{k,j}(\mathbf{z}),\ \ \mathbf{x}, \mathbf{z} \in \mathbb{S}^2 \tag{5.5}$$

has the above mentioned reproducing property if the inner product in $\mathbb{P}_M$ is defined as follows

$$\langle f, g\rangle_{\mathcal{H}_K} = \sum_{k=0}^{M} \beta_k^{2} \sum_{j=1}^{2k+1} \langle Y_{k,j}, f\rangle_{L^2(\mathbb{S}^2)} \langle Y_{k,j}, g\rangle_{L^2(\mathbb{S}^2)}.$$

Indeed, for $p \in \mathbb{P}_M$ we find

$$\begin{aligned}\langle p(\cdot), K(\mathbf{x}, \cdot)\rangle_{\mathcal{H}_K} &= \sum_{k=0}^{M} \beta_k^{2} \sum_{j=1}^{2k+1} \langle Y_{k,j}, p\rangle_{L^2(\mathbb{S}^2)} \langle Y_{k,j}, K(\mathbf{x}, \cdot)\rangle_{L^2(\mathbb{S}^2)} \\ &= \sum_{k=0}^{M} \beta_k^{2} \sum_{j=1}^{2k+1} \langle Y_{k,j}, p\rangle_{L^2(\mathbb{S}^2)} \beta_k^{-2} Y_{k,j}(\mathbf{x}) = p(\mathbf{x}).\end{aligned}$$

In this RKHS setting the spherical polynomial $y_M = y_M(N, K, \alpha)$ defined by (1.2) also can be seen, using the addition theorem and (1.1), as the minimizer of the following quadratic functional

$$T_\alpha(N, K; p) = \sum_{i=1}^{N} w_i (p(\mathbf{x}_i) - y^{\epsilon}(\mathbf{x}_i))^2 + \alpha \|p\|^2_{\mathcal{H}_K},\ \ p \in \mathbb{P}_M, \tag{5.6}$$

which makes (5.5) a natural way of defining the reproducing kernel in this context.

At this point the problem of the choice of the penalization weights $\{\beta_k\}$ is transformed into that of selecting a kernel K from the set $\mathcal{K}$ of kernels of the form (5.5).

In the literature there are several methods for choosing a kernel from the available set of kernels (see, e.g., [22, 27], and references therein). For example, in [22] the authors suggest selecting a kernel by minimizing the value of the functional (5.6) evaluated at its minimizer y_M. In the present context, according to [22], the kernel $K = K_*$ of choice is given as

$$K_* = \arg\min \{T_\alpha(N, K; y_M(N, K, \alpha)),\ K \in \mathcal{K}\}.$$

Note that such K_* depends on the value of the regularization parameter α. Therefore, the approach [22] can be realized only for an *a priori* known α. However, in practice we are not provided with this knowledge, and have to use *a posteriori* regularization parameter choice strategies (for example, the balancing principle described in Subsection 4.1). Thus, in practice we are dealing with kernel dependent regularization parameter $\alpha = \alpha(K)$.

This situation has been discussed in [27]. In the present context the kernel choice suggested in [27] can be written as follows

$$K_+ = \arg\min \left\{T_{\alpha(K)}(N, K; y_M(N, K, \alpha(K))),\ K \in \mathcal{K}\right\}. \tag{5.7}$$

The existence of such K_+ has been proved in [27] under rather general assumptions on the set of admissible kernels $\mathcal{K}$ and regularization parameter choice strategy $\alpha = \alpha(K)$.

From a practical point of view, it is a challenging issue to use the strategy (5.7) in our case because one has to minimize a function depending on $M + 1$ unknown penalization weights β_k. Therefore, it is natural to reduce the complexity of the model before applying the strategy from [27].

For example, one may parametrize $\{\beta_k\}$ as follows: $\beta_k^2 = e^{\lambda_1(k+1)}(k+1)^{\lambda_2}, \lambda_1, \lambda_2 \geq 0$. In other words, in (5.7) the set of kernels $\mathcal{K}$ consists of the functions

$$K(t,\tau) = \sum_{k=0}^{M} e^{-\lambda_1(k+1)}(k+1)^{-\lambda_2} \sum_{j=1}^{2k+1} Y_{k,j}(t)Y_{k,j}(\tau),\ t,\tau \in \mathbb{S}^2. \tag{5.8}$$

Then the kernel K_+ can be found by minimizing a function of two variables λ_1, λ_2. In the last section we will illustrate such a reduced approach by a numerical test showing good performance of the scheme (1.2) with *a posteriori* chosen penalization weights.

6. Numerical examples. In this section we present some numerical experiments to verify the analysis from the previous sections. Note that we work not with real data but with artificially generated ones. In all our experiments we follow [10, 26] and assume that the set of points X_N is the set of Gauss-Legendre points, for which the positive quadrature weights are known analytically. The number of points in this case is $N = 2(M + 1)^2$, and the corresponding cubature formula (1.1) is indeed exact for all spherical polynomials of degree $2M$. In all our experiments $M = 30$.

Note that in real applications the spherical polynomials of much higher degree are used [28]. Moreover, the Gauss-Legendre points are known to have the drawback of having too many points concentrated at the poles, making it not suitable for real satellite data. In our experiments below we use the Gauss-Legendre points and polynomials of modest degree only for illustration purposes and as a proof of concept. At the same time, we note that even for the case $M = 30$ the corresponding discrete problem is rather ill-conditioned and, thus, should be treated with a regularization (see Figure 6.1 and the discussion below).

We start with an experiment illustrating that a proper choice of the penalization weights $\beta_0, ..., \beta_M$ is crucial for the approximation of functions on the sphere. Consider again the SGG-problem corresponding to (5.3). Note that for $k = 1, 2, ..., 30$ the values $a_k = \left(\frac{R}{\rho}\right)^k \frac{(k+1)(k+2)}{\rho^2}$ in (5.3) are increasing, and so, they do not exhibit a typical behavior of the singular values of the compact operators. This effect is well-known (see, e.g, [7], Fig. 4.2.3, p. 280).

Therefore, to mimic the SGG-problem for $M = 30$ one usually omits the factor $\frac{(k+1)(k+2)}{\rho^2}$ (see, e.g., [4]). In this case the decay character of the coefficients a_k in (5.3) can be modeled, for example, as

$$a_k = (1.2)^{-k},\ k = 0, 1, ..., M.$$

We conduct our first experiment in the following way. First we generate a spherical function

$$y = y(\mathbf{x}) = \sum_{k=0}^{M} (1.2)^{-k} \sum_{j=1}^{2k+1} g_{k,j} \frac{1}{\rho} Y_{k,j}(\mathbf{x}), \quad \mathbf{x} \in \mathbb{S}^2,$$

where $g_{k,j} = (k + 1/2)^{-3/2} x_{k,j},\ k = 0, ..., M,\ j = 1, ...2k + 1$, and $x_{k,j}$ are random numbers uniformly distributed on $[0, 1]$. The blurred spherical function y^ϵ is simulated by adding a random point-wise noise to the values of the initial function y at the point set X_N. The simulated noise values are given as the components of a random vector $0.05\epsilon/ \|\epsilon\|_\infty$, where $\epsilon = [\epsilon_1, \epsilon_2, ..., \epsilon_N]$, and ϵ_i are uniformly distributed on $[-1, 1]$. To mimic the SGG-problem we reconstruct the vector $g = (g_{k,j})$ by $g^{\alpha,M} = (g^{\alpha,M}_{k,j})$, where $g^{\alpha,M}_{k,j} = a_k^{-1} \gamma_{k,j}$, and $\gamma_{k,j}$ are given by (3.5).

To assess the obtained results and compare the performance of the considered schemes we measure the relative error

$$\frac{\left\|g - g^{\alpha,M}\right\|_2}{\|g\|_2}.$$

The results are displayed in Figure 6.1,where along the vertical axis we plot the relative errors in solving the problem with one of 50 simulated data. The relative errors are plotted in ascending order for each of four methods: a straightforward least-squares fit to noisy data without any regularization, the regularization with the penalization weights (5.4) and α chosen according to the balancing principle (BP) from $\Delta_{60} = \left\{\alpha_i = 8 \cdot q^i\ i = 1, 2, \ldots, 60\right\}, q = 0.8$, the regularization with default penalization weights $\beta_k = 1, k = 0, 1, \ldots, M$, and the best $\alpha \in \Delta_{60}$, the regularization with the penalization weights (5.4) and the best $\alpha \in \Delta_{60}$. Thus, in the latter two cases the choice of the regularization parameter α for both schemes was made to achieve the best possible performance of each method. As it can be seen from Figure 6.1 the balancing principle (BP) performs at the level of the ideal parameter choice strategy.

From Figure 6.1 one can also conclude that the proper choice of the penalization weights according to the proposed *a priori* recipe can significantly improve the accuracy of the reconstruction. Moreover, Figure 6.1 shows that a straightforward least-squares fit to noisy data without regularization leads to the relative error that

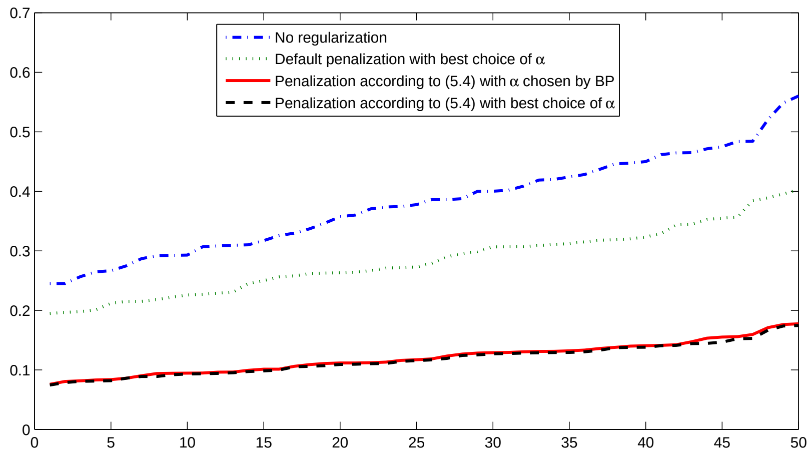


Fig. 6.1: Numerical illustration. The figure presents relative errors for 50 simulations of the data. The errors are plotted in ascending order for each of the discussed methods. Note that two bottom curves corresponding to penalization according to (5.4) nearly overlap.

is about 2-3 times larger than that after a regularization. This confirms that in the considered experiment we are really dealing with a rather ill-conditioned problem.

In our second experiment we again confirm that the balancing principle gives a value of the regularization parameter α_* that is competitive with the best value manually found in [1]. We choose the regularization parameter from the same geometric sequence Δ_{60} and use the same value of the design parameter æ $= 0.002$ in BP.

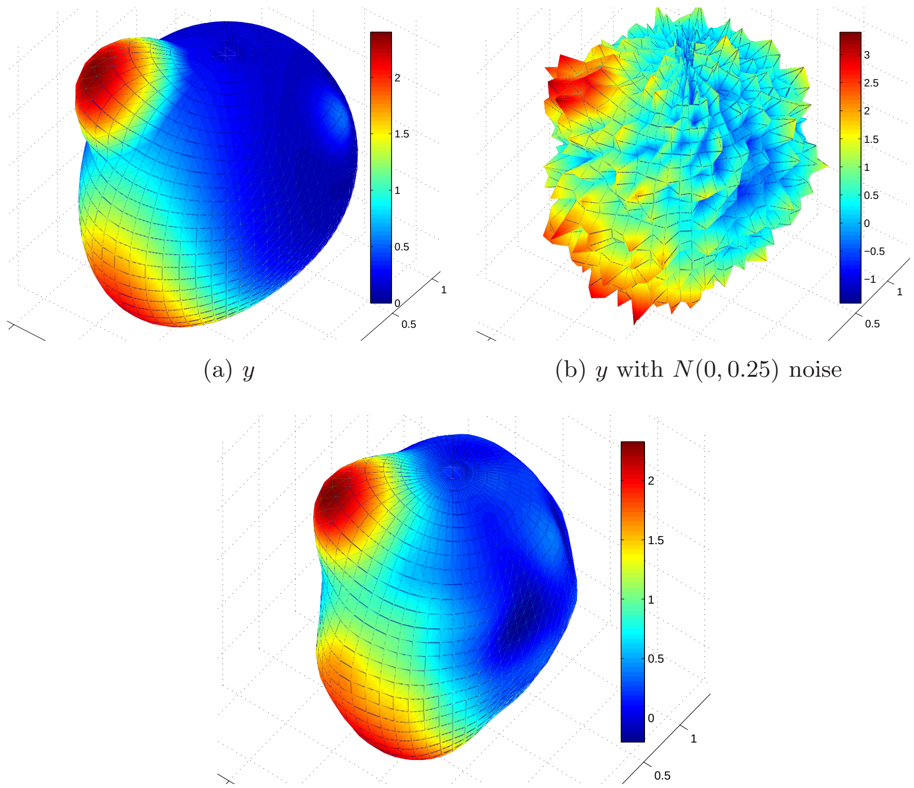
Similarly to [1], as a test function y we take the sum of the Franke function y_1 modified by Renka [33] (p.146) and a function y_{cap} [41], namely $y = y_1 + y_{\text{cap}}$ with

$$
\begin{aligned}
y_1(x_1,x_2,x_3) &= 0.75e^{-(9x_1-2)^2/4-(9x_2-2)^2/4-(9x_3-2)^2/4} \\
&+ 0.75e^{-(9x_1+1)^2/49-(9x_2+1)/49-(9x_3+1)/10} \\
&+ 0.5e^{-(9x_1-7)^2/4-(9x_2-3)^2/4-(9x_3-5)^2/4} \\
&- 0.2e^{-(9x_1-4)^2-(9x_2-7)^2-(9x_3-5)^2},\ (x_1,x_2,x_3)\in\mathbb{S}^2,
\end{aligned}
\tag{6.1}
$$

and

$$
y_{\text{cap}}(\mathbf{x}) = \begin{cases} 2\cos\left(\pi \arccos(\mathbf{x}_c\cdot\mathbf{x})\right), & \mathbf{x}_c\cdot\mathbf{x}\geq\cos(0.5), \\ 0, & \text{otherwise}, \end{cases}
\tag{6.2}
$$

where $\mathbf{x}_c = \left(-\frac{1}{2}, -\frac{1}{2}, \frac{1}{\sqrt{2}}\right)^T$ and $(\cdot)$ defines the dot product of two vectors. The function y was then contaminated by noise, taking for the noise $\epsilon(\mathbf{x})$ at each $\mathbf{x}\in X_N$ an independent sample of a normal random variable with mean 0 and standard deviation $\sigma = 0.5$.



(a) y

(b) y with $N(0, 0.25)$ noise

(c) $T_{\alpha_*,M} y^\epsilon$

Fig. 6.2: Franke function recovery

Figure 6.2a illustrates the function y, while Figure 6.2b shows the blurred function $y^\epsilon(\mathbf{x}) = y(\mathbf{x}) + \epsilon(\mathbf{x})$.

For the reconstruction, following [1] we choose a Laplace-Beltrami penalization operator that corresponds to the matrix

$$\mathbf{B_M} := \mathrm{diag}(0, \underbrace{4, 4, 4}_{3}, ..., \underbrace{(M(M+1))^2, ..., (M(M+1))^2}_{2M+1}) \in \mathbb{R}^{(M+1)^2 \times (M+1)^2}.$$

Figure 6.2c illustrates the reconstructed function $T^{\beta}_{\alpha_*,M} y^\epsilon$. The regularization parameter α_* was obtained according to the balancing principle described above. We found automatically the regularization parameter $\alpha_* = 1.42 \cdot 10^{-4}$ which agrees well with the value 10^{-4} from [1] obtained manually.

In our last experiment we will illustrate an application of the *a posteriori* rule (5.7) for choosing the penalization weights. As a test function y^ϵ we again consider the blurred function from the previous example, where we used the *a priori* chosen penalization weights $\beta_k = k(k+1)$ corresponding to Laplace-Beltrami operator. Now we are going to estimate the penalization weights using the *a posteriori* strategy described in Subsection 4.3.

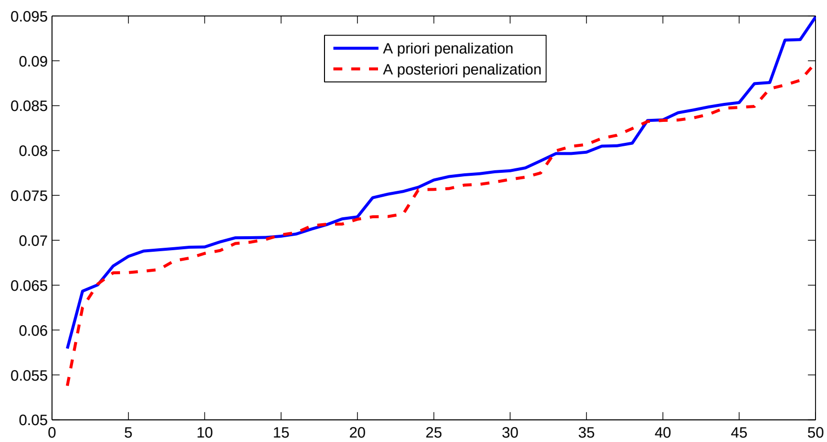


Fig. 6.3: Numerical illustration. The figure presents relative errors for 50 simulations of the data. The errors are plotted in ascending order for each of the discussed methods.

Recall that we are looking for the minimizer (5.7) among the set of admissible kernels $\mathcal{K}$ consisting of the functions (5.8). This approach allows us to take into account an exponential, as well as a polynomial growth of β_k.

To find an approximate minimizer of (5.7) we have implemented the Random Search method [19] over the set of parameters $(\lambda_1, \lambda_2) \in [0,5] \times [0,5]$. The method was implemented 10 times, and in each implementation 10 random steps have been performed. Then the mean values of the parameters λ_1, λ_2 appearing after each implementation of the Random Search method have been taken as an approximate minimum point. As the result, the values $\lambda_1 = 0.32, \lambda_2 = 1.9$ have been obtained.

Figure 6.3 displays the relative errors in solving the problem (6.1), (6.2) with one of 50 simulated noisy data, for each of two methods: regularization with the penalization weights $\beta_k = k(k+1)$, and regularization with *a posteriori* chosen weights.

From Figure 6.3 we see that the choice of the penalization weights according to the proposed *a posteriori* choice rule can improve the accuracy of the reconstruction.

Acknowledgments. The first and the third authors are supported by the Austrian Fonds Zur Forderung der Wissenschaftlichen Forschung (FWF), grant P25424. The work was initiated when the second author visited Johann Radon Institute for Computational and Applied Mathematics (RICAM) within the Special Semester on Applications of Algebra and Number Theory. The second author acknowledges the support of the Australian Research Council.

REFERENCES

[1] C. An, X. Chen, I. H. Sloan and R. S. Womersley, *Regularized least squares approximations on the sphere using spherical designs*, SIAM J. Numer. Anal., 50:3 (2012), pp. 1513–1534.
[2] N. Aronszajn, *Theory of reproducing kernels*, Trans. Amer. Math. Soc., 68 (1950), pp. 337–404.

[3] E. Bannai and E. Bannai, *A survey on spherical designs and algebraic combinatorics on spheres*, European J. Combin., 30 (2009), pp. 1392–1425.
[4] F. Bauer, P. Mathé and S. V. Pereverzev, *Local solutions to inverse problems in geodesy. The impact of the noise covariance structure upon the accuracy of estimation*, J. Geod., 81 (2007), pp. 39–51.
[5] D. Delsarte, J. M. Goethals and J. J. Seidel, *Spherical codes and designs*, Geom. Dedicata, 6 (1977), pp. 363–388.
[6] F. Filbir and W. Themistoclakis, *Polynomial approximation on the sphere using scattered data*, Math. Nachr., 281 (2008), pp. 650–668.
[7] W. Freeden, *Multiscale Modeling of Spaceborne Geodata*, Teubner, Stuttgart, Leipzig, 1999.
[8] W. Freeden and S. V. Pereverzev, *Spherical Tikhonov regularization wavelets in satellite gravity gradiometry with random noise*, J. Geod., 74 (2001), pp. 730–736.
[9] A. Gelb, *The resolution of the Gibbs phenomenon for Spherical harmonics*, Math. Comput., 66:218 (1997), pp. 699–717.
[10] M. Graf, S. Kunis and D. Potts, *On the computation of nonnegative quadrature weights on the sphere*, Appl. Comput. Harm. Anal., 27 (2009), pp. 124–132.
[11] K. Hesse, I. H. Sloan and R. S. Womersley, *Numerical integration on the sphere*, Handbook of Geomathematics, Volume 2, Springer-Verlag Berlin Heidelberg (2010), pp. 1185–1219.
[12] J. Keiner, S. Kunis, and D. Potts, *Efficient reconstruction of functions on the sphere from scattered data*, J. Fourier Anal. Appl., 13 (2007), pp. 435–458.
[13] J. Keiner and D. Potts, *Fast evaluation of quadrature formulae on the sphere*, Math. Comput., 77 (2008), pp. 397–419.
[14] Q. T. Le Gia and H. M. Mhaskar, *Polynomial operators and local approximations of solutions of pseudo-differential equations on the sphere*, Numer. Math., 103 (2006), pp. 299–322.
[15] Q. T. Le Gia, H. M. Mhaskar and Q. Thong, *Localized linear polynomial operators and quadrature formulas on the sphere*, SIAM J. Numer. Anal., 47 (2008/09), pp. 440–466.
[16] S. Lu and S. V. Pereverzev, *Multiparameter regularization in Downward Continuation of Satellite Data* , Handbook of Geomathematics, Volume 2, Springer-Verlag Berlin Heidelberg (2010), pp. 813–832.
[17] S. Lu and S. V. Pereverzev, *Regularization theory for ill-posed problems. Selected topics*, Walter de Gruyter GmbH, Berlin/Boston, 2013.
[18] P. Mathé and S. V. Pereverzev, *Geometry of linear ill-posed problems in variable Hilbert scales*, Inverse Problems, 19 (2003), pp. 789–803.
[19] J. Matyas, *Random optimiyation*, Automat. Remote Contr., 26 (1965), pp. 244–251.
[20] H. M. Mhaskar, *On the representation of smooth functions on the sphere using finitely many bits*, Appl. Comput. Harmon. Anal., 18 (2005), pp. 215–233.
[21] H. M. Mhaskar, F. J. Narcowich, and J. D. Ward, *Spherical Marcinkiewicz-Zygmund inequalities and positive quadrature*, Math. Comp., 235 (2001), pp. 1113–1130.
[22] C. A. Michelli and M. Pontil, *Learning the kernel function via regularization*, J. Machine Learning Res., 6 (2005), pp. 1099–1125.
[23] C. Muller, *Spherical Harmonics*, Lecture notes in Math. 17, Springer-Verlag, Berlin, 1966.
[24] F. Narcowich, P. Petrushev and J. Ward, *Decomposition of Besov and Triebel-Lizorkin spaces on the sphere*, J. Funct. Anal., 238 (2006), pp. 530–564.
[25] F. Narcowich, P. Petrushev and J. Ward, *Localized tight frames on spheres*, SIAM J. Math. Anal., 38 (2006), pp. 574–594.
[26] V. Naumova, S. V. Pereverzev and P. Tkachenko, *Regularized collocation for Spherical harmonics Gravitational Field Modeling*, Int. J. Geomath., 5 (2014), pp. 81–98.
[27] V. Naumova, S. V. Pereverzev and S. Sivananthan, *Extrapolation in variable RKHSs with application to the blood glucose reading*, Inverse Problems, 27(7):075010 (2011), 13 pp.
[28] N. K. Pavlis, S. A. Holmes, S. C. Kenyon and J. K. Factor, *An earth gravitational model to degree 2160: EGM2008*, EGU General Assembly, (2008), pp. 13–18.
[29] S. V. Pereverzev and E. Schock, *On the adaptive selection of the parameter in regularization of ill-posed problems*, SIAM J. Numer. Anal., 43 (2005), pp. 2060–2076.
[30] D. L. Ragozin, *Constructive polynomial approximation on spheres and projective spaces*, Trans. Amer. Math. Soc., 162 (1971), pp. 157–170.
[31] M. Reimer, *Multivariate polynomial approximation*, International Series of Numerical Mathematics, 144, Birkhauser Verlag, Basel, 2003.
[32] M. Reimer, *Hyperinterpolation on the sphere at the minimum projection order*, J. Approx. Theory, 104 (2000), pp. 272–286.
[33] R. J. Renka, *Multivariate interpolation of large sets of scattered data*, ACM Trans. Math. Software, 14 (1988), pp. 139–148.
[34] I. H. Sloan, *Polynomial interpolation and hyperinterpolation over general regions*, J. Approx.

Theory, 83 (1995), pp. 238–254.
[35] I. H. Sloan, *Polynomial approximation on spheres – generalizing de la Vallée-Poussin*, Comput. Methods Appl. Math., 11 (2011), pp. 540–552.
[36] I. H. Sloan and R. S. Womersley, *Filtered hyperinterpolation. A constructive polynomial approximation on the sphere*, Int. J. Geomath., 3 (2012), pp. 95–117.
[37] S. L. Svensson, *Pseudodifferential operators – a new approach to the boundary value problems of physical geodesy*, Manusc. Geod., 8 (1983), pp. 1–40.
[38] S. L. Svensson, *Solution of the altimetry-gravimetry problem*, Bull. Geod., 57 (1983), pp. 332–353.
[39] P. N. Swarztrauber, *On the Spectral Approximation of Discrete Scalar and Vector Functions on a Sphere*, SIAM J. Numer. Anal., 16 (1979), pp. 934–949.
[40] G. Wahba, *Spline interpolation and smoothing on the sphere*, SIAM J. Sci. Statist. Comput., 2 (1981), pp. 5–16.
[41] D. L. Williamson, J. B. Branke, J. J. Hack, R. Jakob and P. N. Swarztrauber, *A standard test set for numerical approximations to the shallow water equations in spherical geometry*, J. Comput. Phys., 102 (1992), pp. 211–224.
[42] Y. Xu, *Polynomial interpolation on the unit sphere*, SIAM J. Numer. Anal., 41 (2003), pp. 751–766.