

## Two-parameter regularization of ill-posed spherical pseudo-differential equations in the space of continuous functions

Hui Cao · Sergei V. Pereverzyev ·  
Ian H. Sloan · Pavlo Tkachenko

Received: date / Accepted: date

**Abstract** In this paper, a two-step regularization method is used to solve an ill-posed spherical pseudo-differential equation in the presence of noisy data. For the first step of regularization we approximate the data by means of a spherical polynomial that minimizes a functional with a penalty term consisting of the squared norm in a Sobolev space. The second step is a regularized collocation method. An error bound is obtained in the uniform norm, which is potentially smaller than that for either the noise reduction alone or the regularized collocation alone. We discuss an *a posteriori* parameter choice, and present some numerical experiments, which support the claimed superiority of the two-step method.

**Keywords** Two-parameter regularization · spherical pseudo-differential equations · quasi-optimality criterion

---

H. Cao  
Guangdong Province Key Laboratory of Computational Science, Sun Yat-sen University,  
Guangzhou, China  
E-mail: caohui6@mail.sysu.edu.cn

S. V. Pereverzyev  
Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy  
of Sciences, Altenbergerstrasse 69, 4040 Linz, Austria  
E-mail: sergei.pereverzyev@oeaw.ac.at

I. H. Sloan  
School of Mathematics and Statistics, University of New South Wales, Sydney NSW 2052,  
Australia  
E-mail: I.Sloan@unsw.edu.au

P. Tkachenko  
Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy  
of Sciences, Altenbergerstrasse 69, 4040 Linz, Austria  
E-mail: pavlo.tkachenko@oeaw.ac.at

## 1 Introduction

Mathematical models appearing in geoscience commonly have the form of an ill-posed spherical pseudo-differential equation

$$Ax = y, \quad (1)$$

where  $A$  is a pseudo-differential operator that relates continuous functions  $x \in C(\Omega_R)$  and  $y \in C(\Omega_\rho)$  defined on concentric spheres  $\Omega_R, \Omega_\rho \in \mathbb{R}^3$  of radii  $R \leq \rho$ . For example, in satellite geodesy, this approach has been introduced in [5, 25] where the spheres  $\Omega_R, \Omega_\rho$  are models of respectively the surface of the Earth and the surface traversed by the satellite.

Recall that a spherical pseudo-differential operator  $A : C(\Omega_R) \rightarrow C(\Omega_\rho)$  is a linear operator that assigns to any  $x \in C(\Omega_R)$  a function

$$Ax := \sum_{k=0}^{\infty} a_k \sum_{j=1}^{2k+1} \hat{x}_{k,j} \frac{1}{\rho} Y_{k,j} \left( \frac{\cdot}{\rho} \right) \in C(\Omega_\rho), \quad (2)$$

where

$$\hat{x}_{k,j} = \left\langle \frac{1}{R} Y_{k,j} \left( \frac{\cdot}{R} \right), x(\cdot) \right\rangle_{L_2(\Omega_R)} := \frac{1}{R} \int_{\Omega_R} x(\tau) Y_{k,j} \left( \frac{\tau}{R} \right) d\Omega_R(\tau)$$

are the spherical Fourier coefficients, and  $Y_{k,j}(\cdot), j = 1, 2, \dots, 2k + 1$ , are the spherical harmonics [16] of degree  $k$  which are  $L_2$ -orthonormal with respect to the unit sphere  $\Omega_1 \in \mathbb{R}^3$ , as a result of which

$$\left\langle \frac{1}{R} Y_{k,j} \left( \frac{\cdot}{R} \right), \frac{1}{R} Y_{k',j'} \left( \frac{\cdot}{R} \right) \right\rangle_{L_2(\Omega_R)} = \langle Y_{k,j}, Y_{k',j'} \rangle_{L_2(\Omega_1)} = \delta_{kk'} \delta_{jj'}. \quad (3)$$

The sequence of real numbers  $(a_k)_{k=0}^{\infty}$  is referred to as the spherical symbol of  $A$ . We shall assume that  $a_k$  is positive, and converges monotonically to zero.

In the case when the symbol sequence  $a_k$  tends to zero fast enough the operator  $A$  is compact. Therefore, its inverse  $A^{-1}$  is unbounded, and keeping in mind Hadamard's definition of a well-posed problem (existence, uniqueness, and continuity of inverse), we conclude that for this case the first and the third conditions are violated, and the problem (1), (2) with  $y \in C(\Omega_R)$  becomes ill-posed. Therefore, a regularization technique should be employed for solving it [4].

As examples of the ill-posed problem (1), (2) we can mention the satellite-to-satellite tracking problem (SST-problem) with  $a_k = \frac{k+1}{\rho} \left( \frac{R}{\rho} \right)^k$ , the satellite gravity gradiometry problem (SGG-problem) with  $a_k = \frac{(k+1)(k+2)}{\rho^2} \left( \frac{R}{\rho} \right)^k$ , etc. (for more details on these and other examples we can refer an interested reader to [5]). These problems are severely ill-posed because of the occurrence of the geometric factor  $(R/\rho)^k$ .

It is worth mentioning that many practical applications use a finite dimensional approximation of the solution of (1), (2). For example, Earth Gravity Models, such as EGM96 or EGM2008 [20] are parametrized by the spherical Fourier coefficients up to some prescribed degree  $M$ .

Note that in applications the function  $y$  is assumed to be continuous. However, in practice one is provided just with a finite number of points  $\{t_i\}_{i=1}^N \subset \Omega_\rho$  at which information about the values of  $y$  is collected. It should be noted also that the pointwise data  $y^\epsilon(t_i)$  contain measurement errors, which can be modeled, for example, in the following way:

$$|y^\epsilon(t_i) - y(t_i)| \leq \epsilon_i, \quad i = 1, \dots, N.$$

We assume for convenience that there exists a function  $y^\epsilon \in C(\Omega_\rho)$  standing for the noisy version of the original function  $y$ , such that

$$\|y^\epsilon - y\|_{C(\Omega_\rho)} \leq \epsilon := \max_{1 \leq i \leq N} \{\epsilon_i\},$$

where  $\epsilon_i$  are measurement errors.

In such a setup the problem (1), (2) is reduced to the following spherical pseudo-differential operator equation

$$A_M x := \sum_{k=0}^M a_k \sum_{j=1}^{2k+1} \hat{x}_{k,j} \frac{1}{\rho} Y_{k,j} \left( \frac{\cdot}{\rho} \right) = y^\epsilon. \quad (4)$$

Several regularization techniques can be used for treating (4). For example, the method where the discretization level  $M$  plays the role of regularization parameter was discussed in [2, 3, 9]. However, in our case we assume that the value of  $M$  is prescribed, thus this approach cannot be used.

The problem (4), but without noise in the right-hand side, was studied in [13]. In this noise-free case the solution  $x$  can be approximated by solving the equation

$$A_M x = \widehat{V}_M y,$$

where  $\widehat{V}_M : C(\Omega_\rho) \rightarrow \mathbb{P}_M(\Omega_\rho)$  is the so-called quasi-interpolatory operator. Here by  $\mathbb{P}_M(\Omega_\rho)$  we denote the set of all spherical polynomials of degree less than or equal to  $M$ , or in other words the restriction to  $\Omega_\rho$  of the polynomials in  $\mathbb{R}^3$  of degree less than or equal to  $M$ . Thus, in [13] the authors suggest constructing a polynomial approximation of  $y$  from the original pointwise data  $y(t_i), i = 1, 2, \dots, N$ , and then formally inverting the operator  $A_M$ .

In principle, this idea can be used also for the ill-posed case (4). However, in contrast to the approximation of a noise-free continuous function  $y$  on the sphere by means of polynomials, which has been discussed by many authors (see, for example, [7, 24, 26, 28, 30]), the approximation of noisy functions  $y^\epsilon$  has been studied only recently in [1, 21], to the best of our knowledge. Applying the method from [1, 21] we first perform so-called data-smoothing (or noise reduction). After this data preprocessing the formal inversion of  $A_M$  should

be safer. However, for performing the noise reduction step one needs *a priori* information about the smoothness of the function  $y$ , which is usually not available, a fact that makes the direct application of the scheme [21] not always appropriate.

In a different direction, for estimating the Fourier coefficients  $\widehat{x}_{k,j}$  directly from noisy measurements  $y^\epsilon(t_i)$  a regularized collocation method has been recently presented in [18]. This method is based on the standard and widely used Tikhonov-Phillips regularization. However, it is well-known that the Tikhonov-Phillips method suffers from saturation [15, 19], meaning that the accuracy of reconstruction cannot be improved regardless of the smoothness of the solution  $x$ . Another point is that while the Tikhonov-Phillips method has been well studied in the Hilbert space  $L_2$ , to the best of our knowledge, no analysis has been done in the space of continuous functions, a natural choice for our noise model.

In the present study we combine these two approaches. Moreover, in contrast to the previous results, we will analyze the approximation in the space of continuous functions. A combination of two regularization methods into one can be seen as a two-step regularization, in which we use the composition  $R_{\alpha,M} \circ T_{\lambda,M}$  of data smoothing operator  $T_{\lambda,M}$ , and a regularized collocation operator  $R_{\alpha,M}$ . In the literature there are not so many studies on two-step regularization, and we can refer only to [11, 12]. However, the analysis in those papers does not correspond to the setting of our problem (1)–(4).

The paper is organized as follows. In the next section we present the regularization method for noise reduction and define the data smoothing operator  $T_{\lambda,M}$ . In Section 3 we will give a short overview of the regularized collocation method, and define the operator  $R_{\alpha,M}$ . Section 4 is devoted to theoretical error bounds for the constructed two-parameter regularization. We will show that the approximation has the potential to perform at least as well as the better of the one-parameter regularizations which are involved in the composition. Finally, in the last section we discuss an *a posteriori* parameter choice rule, and present some numerical experiments supporting the claimed superiority of our method.

## 2 Data noise reduction

At the first step of our scheme we approximate the noisy continuous function  $y^\epsilon \in C(\Omega_\rho)$  by means of a spherical polynomial  $p_M \in \mathbb{P}_M(\Omega_\rho)$ . As discussed in the Introduction, instead of  $y$  we are provided only with pointwise measurements  $y^\epsilon(t_i), i = 1, 2, \dots, N$ . Therefore we introduce the sampling operator  $S_N : C(\Omega_\rho) \rightarrow \mathbb{R}_\omega^N$  for which

$$S_N y^\epsilon := (y^\epsilon(t_1), y^\epsilon(t_2), \dots, y^\epsilon(t_N)).$$

By  $\mathbb{R}_\omega^N$  we denote the vector space  $\mathbb{R}^N$  equipped with the inner product

$$\langle \eta, \gamma \rangle_\omega := \sum_{i=1}^N \omega_i \eta_i \gamma_i, \quad \eta, \gamma \in \mathbb{R}^N;$$

the corresponding norm is  $\|\eta\|_\omega = \langle \eta, \eta \rangle_\omega^{1/2}$ . Here  $\omega_1, \omega_2, \dots, \omega_N$  are positive weights in a cubature formula, and  $t_1, t_2, \dots, t_N \in \Omega_\rho$  are the corresponding cubature points, with the cubature rule having the property of being exact for all polynomials of degree up to  $2M$ ,

$$\forall p \in \mathbb{P}_{2M}(\Omega_\rho), \quad \sum_{i=1}^N \omega_i p(t_i) = \int_{\Omega_\rho} p(\zeta) d\Omega_\rho(\zeta). \quad (5)$$

A method for approximating  $y$  from  $S_N y^\epsilon \in \mathbb{R}_\omega^N$  based on such a cubature rule (but with all weights equal) was proposed recently in [1], while its generalization to other positive weights was presented in [21].

We briefly outline the method from [21] and its extension to pseudo-differential operator equations. We start with the observation that the space  $\mathbb{P}_M(\Omega_\rho)$  of spherical polynomials  $p$  of degree at most  $M$  can be considered as the reproducing kernel Hilbert space (RKHS)  $\mathcal{H}_K$  generated by the kernel

$$K(t, \tau) = \sum_{k=0}^M \beta_k^{-2} \sum_{j=1}^{2k+1} \frac{1}{\rho} Y_{k,j} \left( \frac{t}{\rho} \right) \frac{1}{\rho} Y_{k,j} \left( \frac{\tau}{\rho} \right), \quad t, \tau \in \Omega_\rho, \quad (6)$$

where  $\beta = (\beta_1, \beta_2, \dots, \beta_M, \dots)$  is a non-decreasing sequence of positive parameters. Note that the inner product of  $\mathcal{H}_K$  associated with this kernel can be written as

$$\langle f, g \rangle_{\mathcal{H}_K} = \sum_{k=0}^M \frac{\beta_k^2}{\rho^2} \sum_{j=1}^{2k+1} \left\langle Y_{k,j} \left( \frac{\cdot}{\rho} \right), f \right\rangle_{L^2(\Omega_\rho)} \left\langle Y_{k,j} \left( \frac{\cdot}{\rho} \right), g \right\rangle_{L^2(\Omega_\rho)}. \quad (7)$$

The reproducing property  $\langle f, K(\cdot, \tau) \rangle_{\mathcal{H}_K} = f(\tau)$  for  $\tau \in \Omega_R$  can be verified easily. In this paper the numbers  $\beta_1, \beta_2, \dots, \beta_N$  are assumed to be given *a priori*. We shall see their role later.

The approximation  $p_{\min}$  studied in [1, 21] appears as the minimizer of the following functional

$$p_{\min} = \arg \min \left\{ \|S_N p - S_N y^\epsilon\|_\omega^2 + \lambda \|p\|_{\mathcal{H}_K}^2, \quad p \in \mathbb{P}_M(\Omega_\rho) \right\}, \quad (8)$$

where  $\lambda \geq 0$  is a regularization parameter. Note that (8) can be seen as the definition of the data smoothing operator  $T_{\lambda, M} : C(\Omega_\rho) \rightarrow \mathbb{P}_M(\Omega_\rho)$  such that  $p_{\min} = T_{\lambda, M} y^\epsilon$ . Note that the regularization term involves the  $\mathcal{H}_K$ -norm of  $p$ , and hence depends on the rate of growth of the parameters  $\beta_k$ . The solution of (8) is given by the following theorem.

**Theorem 1** Assume that the points  $\{t_i\}$  and weights  $\{\omega_i\}$  are such that (5) holds. Then the minimizer  $T_{\lambda, M}y^\epsilon = p_{\min}$  in (8) has the form

$$(T_{\lambda, M}y^\epsilon)(\cdot) = \sum_{k=0}^M \frac{1}{1 + \lambda\beta_k^2} \sum_{j=1}^{2k+1} \frac{1}{\rho} Y_{k,j} \left( \frac{\cdot}{\rho} \right) \sum_{i=1}^N \omega_i \frac{1}{\rho} Y_{k,j} \left( \frac{t_i}{\rho} \right) y^\epsilon(t_i). \quad (9)$$

*Proof* It is known that the minimizer of the functional in (8) can be written as

$$T_{\lambda, M}y^\epsilon = (\lambda I + S_N^* S_N)^{-1} S_N^* S_N y^\epsilon, \quad (10)$$

where  $S_N^* : \mathbb{R}_\omega^N \rightarrow \mathcal{H}_K$  is the adjoint of  $S_N$ , and  $I$  is the identity operator in  $\mathcal{H}_K$ . By definition, for  $(\eta_i)_{i=1}^N \in \mathbb{R}_\omega^N$  we have

$$\begin{aligned} \langle \eta, S_N f \rangle_\omega &= \sum_{i=1}^N \eta_i f(t_i) \omega_i = \sum_{i=1}^N \eta_i \langle K(t_i, \cdot), f(\cdot) \rangle_{\mathcal{H}_K} \omega_i \\ &= \left\langle \sum_{i=1}^N \omega_i K(t_i, \cdot) \eta_i, f(\cdot) \right\rangle_{\mathcal{H}_K}, \end{aligned}$$

thus there holds

$$(S_N^* \eta)(\cdot) = \sum_{i=1}^N \omega_i K(t_i, \cdot) \eta_i, \quad \forall \eta \in \mathbb{R}_\omega^N,$$

and thereby

$$(S_N^* S_N f)(\cdot) = \sum_{i=1}^N \omega_i K(t_i, \cdot) f(t_i).$$

Inserting this expression into (10), we observe that  $p_{\min}$  solves

$$\lambda p_{\min}(\cdot) + \sum_{i=1}^N \omega_i K(t_i, \cdot) p_{\min}(t_i) = \sum_{i=1}^N \omega_i K(t_i, \cdot) y^\epsilon(t_i) \quad (11)$$

The spherical harmonic expansion of the polynomial  $p_{\min} = T_{\lambda, M}y^\epsilon$  is

$$p_{\min}(t) = \sum_{k=0}^M \sum_{j=1}^{2k+1} \left\langle p_{\min}, \frac{1}{\rho} Y_{k,j} \left( \frac{\cdot}{\rho} \right) \right\rangle_{L^2(\Omega_\rho)} \frac{1}{\rho} Y_{k,j} \left( \frac{t}{\rho} \right), \quad t \in \Omega_\rho. \quad (12)$$

To find the coefficients in this expansion, we write the second term on the left-hand side of (11), on using (5) and then (6), as

$$\begin{aligned} \sum_{i=1}^N \omega_i K(t_i, t) p_{\min}(t_i) &= \langle p_{\min}, K(\cdot, t) \rangle_{L^2(\Omega_\rho)} \\ &= \sum_{k=0}^M \beta_k^{-2} \sum_{j=1}^{2k+1} \left\langle p_{\min}, \frac{1}{\rho} Y_{k,j} \left( \frac{\cdot}{\rho} \right) \right\rangle_{L^2(\Omega_\rho)} \frac{1}{\rho} Y_{k,j} \left( \frac{t}{\rho} \right). \end{aligned}$$

After expanding the right-hand side of (11) using (6), and then equating the coefficients of  $Y_{k,j}\left(\frac{t}{\rho}\right)$ , we obtain

$$(\lambda + \beta_k^{-2}) \left\langle p_{\min}, Y_{k,j}\left(\frac{t}{\rho}\right) \right\rangle_{L^2(\Omega_\rho)} = \beta_k^{-2} \sum_{i=1}^N \omega_i Y_{k,j}\left(\frac{t}{\rho}\right) y^\epsilon(t_i), \quad t \in \Omega_\rho,$$

which on substituting into (12) yields the desired result.  $\square$

At this point the solution of the first step depends on the regularization parameter  $\lambda$ , and the penalization weights  $\beta_k$ . As mentioned in the Introduction, the choice of the regularization parameter  $\lambda$  will be addressed in the last section. A data-driven choice of the penalization weights  $\beta_k$  has been recently discussed in [21].

The assumption that the function  $x$  on  $\Omega_R$  is continuous implies that  $x \in L^2(\Omega_R)$ , and hence that its Fourier coefficients  $\left\langle \frac{1}{R} Y_{k,j}\left(\frac{\cdot}{R}\right), x \right\rangle_{L^2(\Omega_R)}$  with respect to the basis of spherical harmonics are square-summable, i.e.,

$$\sum_{k=0}^{\infty} \sum_{j=1}^{2k+1} \left| \left\langle \frac{1}{R} Y_{k,j}\left(\frac{\cdot}{R}\right), x \right\rangle_{L^2(\Omega_R)} \right|^2 < \infty.$$

Any additional smoothness of  $x$  can be measured in terms of the summability of Fourier coefficients with some increasing weights depending on the sequence  $(\beta_k)$  or, as it is usual for the regularization theory (see, e. g., [14]), on the symbol  $(a_k)$ . Therefore, it is convenient to introduce two Sobolev spaces  $W^{\phi,\beta}$ , and  $W^{\psi,a}$ , the first depending on the sequence  $(\beta_k)$  and the second on the symbol  $(a_k)$ , and defined by

$$W^{\phi,\beta} := \left\{ g \in L^2(\Omega_R) : \sum_{k=0}^{\infty} \sum_{j=1}^{2k+1} \frac{\left| \left\langle \frac{1}{R} Y_{k,j}\left(\frac{\cdot}{R}\right), g \right\rangle_{L^2(\Omega_R)} \right|^2}{\phi^2(\beta_k^{-2})} =: \|g\|_{W^{\phi,\beta}}^2 < \infty \right\}, \quad (13)$$

$$W^{\psi,a} := \left\{ g \in L^2(\Omega_R) : \sum_{k=0}^{\infty} \sum_{j=1}^{2k+1} \frac{\left| \left\langle \frac{1}{R} Y_{k,j}\left(\frac{\cdot}{R}\right), g \right\rangle_{L^2(\Omega_R)} \right|^2}{\psi^2(a_k^2)} =: \|g\|_{W^{\psi,a}}^2 < \infty \right\}, \quad (14)$$

where  $\phi, \psi$  are non-decreasing functions such that  $\phi(0) = 0$  and  $\psi(0) = 0$ . In the literature, see, e.g., [14], the functions  $\phi, \psi$  go under the name of “smoothness index functions”, and are usually unknown.

### 3 Regularized collocation method after noise reduction

After the first step of our method the reduced original equation (4) is transformed into the equation

$$A_M x = T_{\lambda, M} y^\epsilon. \quad (15)$$

The regularized collocation method (see [18]) is now applied to this equation, yielding an approximate solution  $x_{\alpha, \lambda}^\epsilon$  of the equation (15), defined as the minimizer of the functional

$$\|B_{N, M} x - S_N T_{\lambda, M} y^\epsilon\|_\omega^2 + \alpha \|x\|_{L_2(\Omega_R)}^2, \quad (16)$$

where  $B_{N, M} := S_N A_M : L_2(\Omega_R) \rightarrow \mathbb{R}_\omega^N$ , and  $\alpha \geq 0$  is the regularization parameter for the second step. The minimizer of (16) can be written in the form

$$x_{\alpha, \lambda}^\epsilon = (\alpha I + B_{N, M}^* B_{N, M})^{-1} B_{N, M}^* S_N T_{\lambda, M} y^\epsilon, \quad (17)$$

where  $B_{N, M}^* : \mathbb{R}_\omega^N \rightarrow L_2(\Omega_R)$  is the adjoint of  $B_{N, M}$ , given by

$$(B_{N, M}^* \eta)(\cdot) = \sum_{k=0}^M a_k \sum_{j=1}^{2k+1} \frac{1}{R} Y_{k, j} \left( \frac{\cdot}{R} \right) \sum_{i=1}^N \omega_i \frac{1}{\rho} Y_{k, j} \left( \frac{t_i}{\rho} \right) \eta_i,$$

from which it follows, using (4), that

$$B_{N, M}^* B_{N, M} x = \sum_{k=0}^M a_k^2 \sum_{j=1}^{2k+1} \frac{1}{R} Y_{k, j} \left( \frac{\cdot}{R} \right) \hat{x}_{k, j}.$$

Using (17) and (9) we then obtain explicitly

$$x_{\alpha, \lambda}^\epsilon := \sum_{k=0}^M \frac{a_k}{\alpha + a_k^2} \frac{1}{1 + \lambda \beta_k^2} \sum_{j=1}^{2k+1} \frac{1}{R} Y_{k, j} \left( \frac{\cdot}{R} \right) \sum_{i=1}^N \omega_i \frac{1}{\rho} Y_{k, j} \left( \frac{t_i}{\rho} \right) y^\epsilon(t_i). \quad (18)$$

If we now define  $R_{\alpha, M} : \mathbb{P}_M(\Omega_\rho) \rightarrow \mathbb{P}_M(\Omega_R)$  as

$$\begin{aligned} (R_{\alpha, M} p)(\cdot) &:= \left( (\alpha I + B_{N, M}^* B_{N, M})^{-1} B_{N, M}^* S_N p \right)(\cdot) \\ &= \sum_{k=0}^M \frac{a_k}{\alpha + a_k^2} \sum_{j=1}^{2k+1} \frac{1}{R} Y_{k, j} \left( \frac{\cdot}{R} \right) \sum_{i=1}^N \omega_i \frac{1}{\rho} Y_{k, j} \left( \frac{t_i}{\rho} \right) p(t_i) \end{aligned} \quad (19)$$

$$= \sum_{k=0}^M \frac{a_k}{\alpha + a_k^2} \sum_{j=1}^{2k+1} \frac{1}{R} Y_{k, j} \left( \frac{\cdot}{R} \right) \left\langle \frac{1}{\rho} Y_{k, j}, p \right\rangle_{L_2(\Omega_\rho)}. \quad (20)$$

then we can write  $x_{\alpha, \lambda}^\epsilon = R_{\alpha, M} T_{\lambda, M} y^\epsilon$ . Thus  $R_{\alpha, M}$  reflects the regularized collocation second step of the method, whereas  $T_{\lambda, M}$  corresponds to the noise reduction first step.

As one can see from (18), in the end we approximate the unknown solution  $x$  on  $\Omega_R$  by means of a spherical polynomial of degree  $M$ . In this context a question arises about the best approximation of a continuous function  $x$  by means of a polynomial of degree  $M$ . In turn the quality of best polynomial

approximation is determined by the smoothness of  $x$ . We shall assume that  $x$  lies in the intersection of  $W^{\phi,\beta}$  and  $W^{\psi,a}$ , see (13) or (14), thus the smoothness of  $x$  is encoded in  $\phi$  and  $\beta_k$ , or  $\psi$  and  $a_k$ . For example, if the smoothness index function  $\phi(t)$  and the sequence  $\beta = \{\beta_k\}$  increase polynomially with  $t$  and  $k$ , then Jackson's theorem on the sphere (see [22], Theorem 3.3) tells us that for  $x \in W^{\phi,\beta}$  there is  $\nu > 0$  such that

$$\inf_{p \in \mathbb{P}_M} \|x - p\|_{C(\Omega_R)} = O(M^{-\nu}). \quad (21)$$

At the same time, if the sequence  $\beta = \{\beta_k\}$  increases exponentially then for polynomially increasing  $\phi$  and  $x \in W^{\phi,\beta}$  we have

$$\inf_{p \in \mathbb{P}_M} \|x - p\|_{C(\Omega_R)} = O(e^{-qM}),$$

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

In the error analysis of the next section we shall make use of a constructive polynomial approximation introduced in [26], in which  $x \in C(\Omega_R)$  is approximated by  $V_M x \in \mathbb{P}_M(\Omega_R)$ , given by

$$(V_M x)(t) = \sum_{k=0}^M h\left(\frac{k}{M}\right) \sum_{j=1}^{2k+1} \frac{1}{R} Y_{k,j}\left(\frac{t}{R}\right) \left\langle \frac{1}{R} Y_{k,j}\left(\frac{\cdot}{R}\right), x \right\rangle_{L^2(\Omega_R)}, \quad t \in \Omega_R, \quad (22)$$

where  $h$  (a ‘‘filter function’’) is a continuously differentiable function on  $\mathbb{R}^+$  satisfying

$$h(t) = \begin{cases} 1, & t \in [0, 1/2], \\ 0, & t \in [1, \infty), \end{cases} \quad 0 \leq h(t) \leq 1 \text{ for } t \in \mathbb{R}^+$$

Explicit examples of suitable filter functions  $h$  can be found in [27]. It is important for our later analysis that, as shown in [26],  $V_M x$  is (up to a constant) an optimal approximation in the uniform norm, in the sense that,

$$\|x - V_M x\|_{C(\Omega_R)} \leq c \inf_{p \in \mathbb{P}_{[M/2]}} \|x - p\|_{C(\Omega_R)}, \quad (23)$$

where  $c$  is a generic constant, which may take different values at different occurrences, and  $[\cdot]$  denotes the floor function.

In view of (21), for polynomially increasing  $\phi, \beta$  and for  $x \in W^{\phi,\beta}$  we have

$$\|x - V_M x\|_{C(\Omega_R)} \leq c [M/2]^{-\nu} \leq c M^{-\nu}.$$

On the other hand, for exponentially increasing  $\beta$  and polynomially increasing  $\phi$  the theory [23] suggests taking  $h(t) \equiv 1, t \in [0, 1]$ . In this case the right-hand side of (23) has to be modified by multiplying by  $\sqrt{M}$ , and by replacing  $[M/2]$  by  $M$ , thus for  $x \in W^{\phi,\beta}$  there holds

$$\|x - V_M x\|_{C(\Omega_R)} \leq c\sqrt{M} \inf_{p \in \mathbb{P}_M} \|x - p\|_{C(\Omega_R)} \leq c\sqrt{M} e^{-qM}.$$

#### 4 Error estimation in uniform norm

In this section we will estimate the uniform error of approximation of  $x$  by the polynomial  $x_{\alpha,\lambda}^\epsilon$  given by (18). It is clear that

$$\begin{aligned} \|x - x_{\alpha,\lambda}^\epsilon\|_{C(\Omega_R)} &= \|x - R_{\alpha,M}T_{\lambda,M}y^\epsilon\|_{C(\Omega_R)} \\ &\leq \|x - V_Mx\|_{C(\Omega_R)} + \|V_Mx - R_{\alpha,M}U_My\|_{C(\Omega_R)} \\ &\quad + \|(R_{\alpha,M} - R_{\alpha,M}T_{\lambda,M})U_My\|_{C(\Omega_R)} \\ &\quad + \|R_{\alpha,M}T_{\lambda,M}\|_{C(\Omega_\rho) \rightarrow C(\Omega_R)} (\|U_My - y\|_{C(\Omega_\rho)} + \|y - y^\epsilon\|_{C(\Omega_\rho)}), \end{aligned} \quad (24)$$

where  $U_My$  is the same as  $V_Mx$  in (22) except that it is an approximation on the larger sphere  $\Omega_\rho$  instead of on  $\Omega_R$ ,

$$(U_My)(t) = \sum_{k=0}^M h \left( \frac{k}{M} \right) \sum_{j=1}^{2k+1} \frac{1}{\rho^2} Y_{k,j} \left( \frac{t}{\rho} \right) \left\langle Y_{k,j} \left( \frac{\cdot}{\rho} \right), y(\cdot) \right\rangle_{L^2(\Omega_\rho)}. \quad (25)$$

It is natural to assume that  $\|U_My - y\|_{C(\Omega_\rho)} < \epsilon$ , since otherwise data noise is dominated by the approximation error and no regularization is required. We also restrict ourselves to the case when  $\|x - V_Mx\|_{C(\Omega_R)} < c\epsilon$ , otherwise the term  $\|x - V_Mx\|_{C(\Omega_R)}$ , representing the error of almost best polynomial approximation, will dominate the error bound.

Then the bound (24) can be reduced to the following one

$$\begin{aligned} \|x - x_{\alpha,\lambda}^\epsilon\|_{C(\Omega_R)} &\leq \|V_Mx - R_{\alpha,M}U_My\|_{C(\Omega_R)} \\ &\quad + \|(R_{\alpha,M} - R_{\alpha,M}T_{\lambda,M})U_My\|_{C(\Omega_R)} + c(1 + \|R_{\alpha,M}T_{\lambda,M}\|_{C(\Omega_\rho) \rightarrow C(\Omega_R)})\epsilon. \end{aligned} \quad (26)$$

An estimate of the coefficient in the last term is given by the following theorem.

**Theorem 2** *Under the conditions of Theorem 1*

$$\|R_{\alpha,M}T_{\lambda,M}\|_{C(\Omega_\rho) \rightarrow C(\Omega_R)} \leq \frac{1}{R\rho} \max_{t \in \Omega_R} \left| \sum_{i=1}^N \omega_i \sum_{k=0}^M \frac{(2k+1)a_k}{4\pi(\alpha + a_k^2)(1 + \lambda\beta_k^2)} P_k \left( \frac{t \cdot t_i}{R\rho} \right) \right|,$$

where  $P_k$  are the Legendre polynomials of degree  $k$ .

*Proof* In view of the definition (9) and (19), together with (3) with  $R$  replaced by  $\rho$  and the property (5) of the cubature rule, we can write

$$\begin{aligned} &(R_{\alpha,M}T_{\lambda,M}y)(t) \\ &= \sum_{k=0}^M \frac{a_k}{\alpha + a_k^2} \frac{1}{1 + \lambda\beta_k^2} \sum_{j=1}^{2k+1} \frac{1}{R} Y_{k,j} \left( \frac{t}{R} \right) \sum_{i=1}^N \omega_i \frac{1}{\rho} Y_{k,j} \left( \frac{t_i}{\rho} \right) y(t_i) \\ &= \frac{1}{R\rho} \sum_{k=0}^M \frac{2k+1}{4\pi} \frac{a_k}{\alpha + a_k^2} \frac{1}{1 + \lambda\beta_k^2} \sum_{i=1}^N \omega_i P_k \left( \frac{t \cdot t_i}{R\rho} \right) y(t_i), \quad \forall t \in \Omega_R, \end{aligned}$$

where in the last step we used the addition theorem for the spherical harmonics [16]

$$\sum_{j=1}^{2k+1} Y_{k,j}(x)Y_{k,j}(y) = \frac{2k+1}{4\pi} P_k(x \cdot y), \quad x, y \in \Omega_1.$$

□

Finally, we estimate the first two terms of the error bound (26).

**Theorem 3** *Assume that the smoothness index functions  $\phi, \psi$  are such that the functions  $t/\phi(t)$ , and  $t/\psi(t)$  are monotone. Then for  $x \in W^{\phi,\beta} \cap W^{\psi,a}$  and  $Ax = y$  we have*

$$\begin{aligned} \|V_M x - R_{\alpha,M} U_M y\|_{C(\Omega_R)} + \|(R_{\alpha,M} - R_{\alpha,M} T_{\lambda,M}) U_M y\|_{C(\Omega_R)} \\ \leq cM[\widehat{\psi}(\alpha) \|x\|_{W^{\psi,a}} + \widehat{\phi}(\lambda) \|x\|_{W^{\phi,\beta}}], \end{aligned}$$

where  $\widehat{\phi}(\lambda) = \phi(\lambda)$ ,  $\widehat{\psi}(\alpha) = \psi(\alpha)$ , if  $t/\phi(t)$  and  $t/\psi(t)$  are non-decreasing, and  $\widehat{\phi}(\lambda) = \lambda$ ,  $\widehat{\psi}(\alpha) = \alpha$  otherwise.

*Proof* It follows easily from (2) and (3) that

$$a_k \left\langle \frac{1}{R} Y_{k,j} \left( \frac{\cdot}{R} \right), x \right\rangle_{L^2(\Omega_R)} = \left\langle \frac{1}{\rho} Y_{k,j} \left( \frac{\cdot}{\rho} \right), y \right\rangle_{L^2(\Omega_\rho)}.$$

Hence from (22) and (25) we obtain

$$\begin{aligned} (V_M x)(t) &= \sum_{k=0}^M h \left( \frac{k}{M} \right) \frac{1}{a_k} \sum_{j=1}^{2k+1} \frac{1}{R} Y_{k,j} \left( \frac{t}{R} \right) \left\langle \frac{1}{\rho} Y_{k,j} \left( \frac{\cdot}{\rho} \right), y \right\rangle_{L^2(\Omega_\rho)} \\ &= \sum_{k=0}^M \frac{1}{a_k} \sum_{j=1}^{2k+1} \frac{1}{R} Y_{k,j} \left( \frac{t}{R} \right) \left\langle \frac{1}{\rho} Y_{k,j} \left( \frac{\cdot}{\rho} \right), (U_M y) \right\rangle_{L^2(\Omega_\rho)}. \end{aligned}$$

Moreover, in view of (26), (25) and (5) we also have the representations

$$\begin{aligned} (R_{\alpha,M} U_M y)(t) &= \sum_{k=0}^M \frac{a_k}{\alpha + a_k^2} \sum_{j=1}^{2k+1} \frac{1}{R} Y_{k,j} \left( \frac{t}{R} \right) \sum_{i=1}^N \omega_i \frac{1}{\rho} Y_{k,j} \left( \frac{t_i}{\rho} \right) (U_M y)(t_i) \\ &= \sum_{k=0}^M \frac{a_k}{\alpha + a_k^2} \sum_{j=1}^{2k+1} \frac{1}{R} Y_{k,j} \left( \frac{t}{R} \right) \left\langle \frac{1}{\rho} Y_{k,j} \left( \frac{\cdot}{\rho} \right), (U_M y)(\cdot) \right\rangle_{L^2(\Omega_\rho)}, \end{aligned}$$

while from (7) and (3) and two uses of (4) we obtain

$$\begin{aligned} (R_{\alpha,M} T_{\lambda,M} U_M y)(t) &= \sum_{k=0}^M \frac{a_k}{\alpha + a_k^2} \frac{1}{1 + \lambda\beta_k^2} \sum_{j=1}^{2k+1} \frac{1}{R} Y_{k,j} \left( \frac{t}{R} \right) \sum_{i=1}^N \omega_i \frac{1}{\rho} Y_{k,j} \left( \frac{t_i}{\rho} \right) (U_M y)(t_i) \\ &= \sum_{k=0}^M \frac{a_k}{\alpha + a_k^2} \frac{1}{1 + \lambda\beta_k^2} \sum_{j=1}^{2k+1} \frac{1}{R} Y_{k,j} \left( \frac{t}{R} \right) \left\langle \frac{1}{\rho} Y_{k,j} \left( \frac{\cdot}{\rho} \right), (U_M y)(\cdot) \right\rangle_{L^2(\Omega_\rho)}. \end{aligned}$$

Then we can write, using again (25),

$$\begin{aligned} & \|V_M x - R_{\alpha, M} U_M y\|_{C(\Omega_R)} \\ &= \left\| \sum_{k=0}^M \left( \frac{1}{a_k} - \frac{a_k}{\alpha + a_k^2} \right) \sum_{j=1}^{2k+1} \frac{1}{R} Y_{k,j} \left( \frac{t}{R} \right) \left\langle \frac{1}{\rho} Y_{k,j} \left( \frac{\cdot}{\rho} \right), U_M y \right\rangle_{L^2(\Omega_\rho)} \right\|_{C(\Omega_R)} \\ &= \left\| \sum_{k=0}^M h \left( \frac{k}{M} \right) \frac{\alpha}{a_k(\alpha + a_k^2)} \sum_{j=1}^{2k+1} \frac{1}{R} Y_{k,j} \left( \frac{t}{R} \right) \left\langle \frac{1}{\rho} Y_{k,j} \left( \frac{\cdot}{\rho} \right), y \right\rangle_{L^2(\Omega_\rho)} \right\|_{C(\Omega_R)}. \end{aligned}$$

Now using the property that  $0 \leq h \left( \frac{k}{M} \right) \leq 1$  and then the Nikolskii inequality (see, e.g., [17], Proposition 2.5), we obtain, for  $x \in W^{\psi, \alpha}$ ,

$$\begin{aligned} & \|V_M x - R_{\alpha, M} U_M y\|_{C(\Omega_R)} \\ &\leq \left\| \sum_{k=0}^M \frac{\alpha}{\alpha + a_k^2} \sum_{j=1}^{2k+1} \frac{1}{R} Y_{k,j} \left( \frac{t}{R} \right) \left\langle \frac{1}{R} Y_{k,j} \left( \frac{\cdot}{R} \right), x \right\rangle_{L^2(\Omega_R)} \right\|_{C(\Omega_R)} \\ &\leq cM \left\| \sum_{k=0}^M \frac{\alpha}{\alpha + a_k^2} \sum_{j=1}^{2k+1} \frac{1}{R} Y_{k,j} \left( \frac{t}{R} \right) \left\langle \frac{1}{R} Y_{k,j} \left( \frac{\cdot}{R} \right), x \right\rangle_{L^2(\Omega_R)} \right\|_{L^2(\Omega_R)} \\ &= cM \left\| \sum_{k=0}^M \frac{\alpha}{\alpha + a_k^2} \frac{\psi(a_k^2)}{\psi(a_k^2)} \sum_{j=1}^{2k+1} \frac{1}{R} Y_{k,j} \left( \frac{t}{R} \right) \left\langle \frac{1}{R} Y_{k,j} \left( \frac{\cdot}{R} \right), x(\cdot) \right\rangle_{L^2(\Omega_R)} \right\|_{L^2(\Omega_R)} \\ &\leq cM \sup_{u \leq a_1^2} \left| \frac{\alpha}{\alpha + u} \psi(u) \right| \|x\|_{W^{\psi, \alpha}}. \end{aligned}$$

The proof can be completed by using Proposition 2.7 from [14]. For convenience, we present the corresponding argument here.

Let us consider two cases. First, we assume that  $t/\psi(t)$  is a non-decreasing function. Then keeping in mind that the function  $\psi$  is increasing, for  $0 < u < \alpha$  we obtain

$$\sup_{u < \alpha} \left| \frac{\alpha}{\alpha + u} \psi(u) \right| \leq \sup_{u < \alpha} \psi(u) < \psi(\alpha).$$

If  $\alpha \leq u$  we can write

$$\begin{aligned} \sup_{\alpha \leq u} \left| \frac{\alpha}{\alpha + u} \psi(u) \right| &\leq \sup_{\alpha \leq u} \left| \frac{\alpha u}{\alpha + u} \frac{\psi(u)}{u} \right| \leq \sup_{\alpha \leq u} \left| \frac{\alpha u}{\alpha + u} \right| \sup_{\alpha \leq u} \left| \frac{\psi(u)}{u} \right| \\ &\leq \frac{\alpha}{\inf_{\alpha \leq u} \frac{u}{\psi(u)}} \leq \frac{\alpha}{\psi(\alpha)} = \psi(\alpha). \end{aligned}$$

Assume now that  $t/\psi(t)$  is a decreasing function. Then

$$\sup_{u \leq a_1^2} \left| \frac{\alpha}{\alpha + u} \psi(u) \right| \leq \sup_{u \leq a_1^2} \left| \frac{\alpha u}{\alpha + u} \frac{\psi(u)}{u} \right| \leq \alpha \frac{\psi(a_1^2)}{a_1^2} \leq c\alpha.$$

Combining everything together we obtain

$$\|V_M x - R_{\alpha, M} U_M y\|_{C(\Omega_R)} \leq cM \widehat{\psi}(\alpha) \|x\|_{W^{\psi, a}}.$$

The second part of the regularization error can be bounded by applying similar steps and using the assumption of the theorem that  $x \in W^{\phi, \beta}$ .  $\square$

From (26) and Theorems 2 and 3 we can estimate the error of the approximation of  $x$  with the use of the proposed two-step regularization scheme as follows:

$$\begin{aligned} \|x - x_{\alpha, \lambda}^\epsilon\|_{C(\Omega_R)} &\leq cM[\widehat{\phi}(\lambda) \|x\|_{W^{\phi, \beta}} + \widehat{\psi}(\alpha) \|x\|_{W^{\psi, a}}] \\ &+ c \max_{t \in \Omega_R} \sum_{k=0}^M \frac{(2k+1)a_k}{4\pi(a_k^2 + \alpha)(1 + \lambda\beta_k^2)} \frac{1}{R\rho} \left| \sum_{i=1}^N \omega_i P_k \left( \frac{t \cdot t_i}{R\rho} \right) \right| \epsilon. \end{aligned} \quad (27)$$

From the error decomposition (27) one can easily obtain the error bounds for the single parameter regularization schemes involved in the two-step combination. If  $\lambda = 0$ , which means that no noise reduction has been done, we derive the following error bound in the uniform norm:

$$\begin{aligned} \|x - x_{\alpha, \lambda}^\epsilon\|_{C(\Omega_R)} &\leq cM \widehat{\psi}(\alpha) \|x\|_{W^{\psi, a}} \\ &+ c \max_{t \in \Omega_R} \sum_{k=0}^M \frac{(2k+1)a_k}{(\alpha + a_k^2)} \frac{1}{R\rho} \left| \sum_{i=1}^N \omega_i P_k \left( \frac{t \cdot t_i}{R\rho} \right) \right| \epsilon. \end{aligned} \quad (28)$$

On the other hand, if  $\alpha = 0$ , which means that direct inversion after data pre-smoothing has been done, then

$$\begin{aligned} \|x - x_{\alpha, \lambda}^\epsilon\|_{C(\Omega_R)} &\leq cM \widehat{\phi}(\lambda) \|x\|_{W^{\phi, \beta}} \\ &+ c \max_{t \in \Omega_R} \sum_{k=0}^M \frac{2k+1}{4\pi a_k (1 + \lambda\beta_k^2)} \frac{1}{R\rho} \left| \sum_{i=1}^N \omega_i P_k \left( \frac{t \cdot t_i}{R\rho} \right) \right| \epsilon. \end{aligned} \quad (29)$$

Now we illustrate the potential advantage of the two-step approximation compared to the single-parameter schemes. Assume that  $\phi(t) = t^\mu$ , and  $\psi(t) = t^\nu$  for some  $\mu, \nu \in (0, 1)$ . In this case the error bound (27) is reduced to the following one

$$\begin{aligned} \|x - x_{\alpha, \lambda}^\epsilon\|_{C(\Omega_R)} &\leq cM(\lambda^\mu \|x\|_{W^{\phi, \beta}} + \alpha^\nu \|x\|_{W^{\psi, a}}) \\ &+ c \max_{t \in \Omega_R} \sum_{k=0}^M \frac{(2k+1)a_k}{(\alpha + a_k^2)(1 + \lambda\beta_k^2)} \frac{1}{R\rho} \left| \sum_{i=1}^N \omega_i P_k \left( \frac{t \cdot t_i}{R\rho} \right) \right| \epsilon. \end{aligned} \quad (30)$$

Consider for simplicity the situation in which for fixed  $\mu$  and  $\nu$ ,  $\alpha = \lambda^{\mu/\nu}$ . Then the first terms of (28), (29) and (30) are essentially the same, whereas the last term in (30) is obviously smaller than that in either (28) or (29).

## 5 Numerical illustrations

In all our experiments the regularization parameters  $\lambda, \alpha$  are chosen by means of the so-called quasi-optimality criterion. This heuristic approach was originally proposed in [29] and has been recently advocated in [10].

In the case of the noise reduction method (8), the quasi-optimality criterion chooses a regularization parameter  $\lambda = \lambda_l$  from a set

$$Q_L^\lambda = \{\lambda = \lambda_i = \lambda_0 q^i, i = 0, 1, 2, \dots, L\}, \quad q > 1, \quad (31)$$

such that

$$\|A_M^{-1}(T_{\lambda_l, My^\epsilon} - T_{\lambda_{l-1}, My^\epsilon})\| = \min \left\{ \|A_M^{-1}(T_{\lambda_i, My^\epsilon} - T_{\lambda_{i-1}, My^\epsilon})\|, i = 0, 1, \dots, L \right\},$$

where  $\|\cdot\|$  denotes the uniform norm in the space  $C = C(\Omega_R)$  of continuous functions.

In the similar way we can apply the quasi-optimality criterion in the regularized collocation method. Then a regularization parameter  $\alpha = \alpha_k$  is chosen from a set

$$Q_K^\alpha = \{\alpha = \alpha_j = \alpha_0 r^j, j = 0, 1, 2, \dots, K\}, \quad r > 1, \quad (32)$$

in such a way that

$$\|R_{\alpha_k, My^\epsilon} - R_{\alpha_{k-1}, My^\epsilon}\| = \min \left\{ \|R_{\alpha_j, My^\epsilon} - R_{\alpha_{j-1}, My^\epsilon}\|, j = 0, 1, \dots, K \right\}.$$

For the two-step regularization the quasi-optimality criterion can be implemented as follows. First, for every  $\alpha = \alpha_j \in Q_K^\alpha$  we choose  $\lambda = \lambda_l = \lambda(\alpha_j)$  from the set (31) such that

$$\|x_{\alpha_j, \lambda_l}^\epsilon - x_{\alpha_j, \lambda_{l-1}}^\epsilon\| = \min \left\{ \|x_{\alpha_j, \lambda_i}^\epsilon - x_{\alpha_j, \lambda_{i-1}}^\epsilon\|, i = 0, 1, \dots, L \right\}.$$

Next, we apply the quasi-optimality criterion to the sequence  $\left\{ x_{\lambda(\alpha_j), \alpha_j}^\epsilon \right\}$  parametrized by  $\alpha_j \in Q_K^\alpha$ . More specifically, we select  $\alpha_k \in Q_K^\alpha$  such that

$$\|x_{\alpha_k, \lambda(\alpha_k)}^\epsilon - x_{\alpha_{k-1}, \lambda(\alpha_{k-1})}^\epsilon\| = \min \left\{ \|x_{\alpha_j, \lambda(\alpha_j)}^\epsilon - x_{\alpha_{j-1}, \lambda(\alpha_{j-1})}^\epsilon\|, j = 0, 1, \dots, K \right\}.$$

Then, a regularized approximate solution  $x_{\alpha, \lambda}^\epsilon$  of our choice is defined by (18) with  $\lambda = \lambda(\alpha_k)$  and  $\alpha = \alpha_k$ .

In all our experiments we follow [6, 18] and assume that  $\{t_i\}_{i=1}^N$  is the set of Gauss-Legendre points, for which the positive quadrature weights are known analytically. In this case  $N = 2(M+1)^2$ , and we take  $M = 30$ .

The data are simulated in the following way. First we generate a spherical function

$$y(t) = \sum_{k=0}^M a_k \sum_{j=1}^{2k+1} \hat{x}_{k,j} \frac{1}{\rho} Y_{k,j} \left( \frac{t}{\rho} \right), \quad t \in \Omega_\rho,$$

where  $\hat{x}_{k,j} = (k+1/2)^{-v} g_{k,j}$ ,  $k = 0, \dots, M$ ,  $j = 1, \dots, 2k+1$ ,  $v > 0$ , and  $g_{k,j}$  are uniformly distributed random values from  $[-1, 1]$ . This means that the solutions of the simulated problems have Sobolev smoothness  $v$ , and  $y$  can be seen as a result of the action of the operator  $A$  on a function  $x$  from the spherical Sobolev space  $H_2^v$ . A noisy spherical function  $y^\epsilon$  is simulated by adding a Gaussian white noise of intensity  $\epsilon = 0.05$  to the values of the initial function  $y$  at the points  $\{t_i\}_{i=1}^N$ .

To assess the performance of the considered schemes we measure the relative error

$$\frac{\|x - x_{\alpha, \lambda}^\epsilon\|}{\|x\|},$$

where

$$x = \sum_{k=0}^M \sum_{j=1}^{2k+1} \hat{x}_{k,j} \frac{1}{R} Y_{k,j} \left( \frac{\cdot}{R} \right),$$

and the approximation  $x_{\alpha, \lambda}^\epsilon$  is given by (18).

The results are displayed in Figure 1, where each subfigure corresponds to different values of  $a_k, v$ , and penalization weights  $\beta_k$ . In Figure 1 each circle exhibits a value of the relative error in solving the problem with one of 10 simulated data, for each of three methods: regularized collocation method corresponding to the case when  $\lambda = 0$ , noise reduction method with direct inversion (the case when  $\alpha = 0$ ), and the two-step regularization method. Note that such a form of graphical representation of the performance of different regularization algorithms is rather common (see, e.g., [8]).

In our experiments we tried to cover different degrees of ill-posedness of (1), (2), i.e. the so-called moderate ill-posedness when the spherical symbol tends to zero polynomially (Figures 1c, 1d, 1e), and severely ill-posed case with exponential symbol decrease (Figures 1a, 1b). Note that the spherical pseudo-differential operators with polynomial symbol have been studied in [13], while

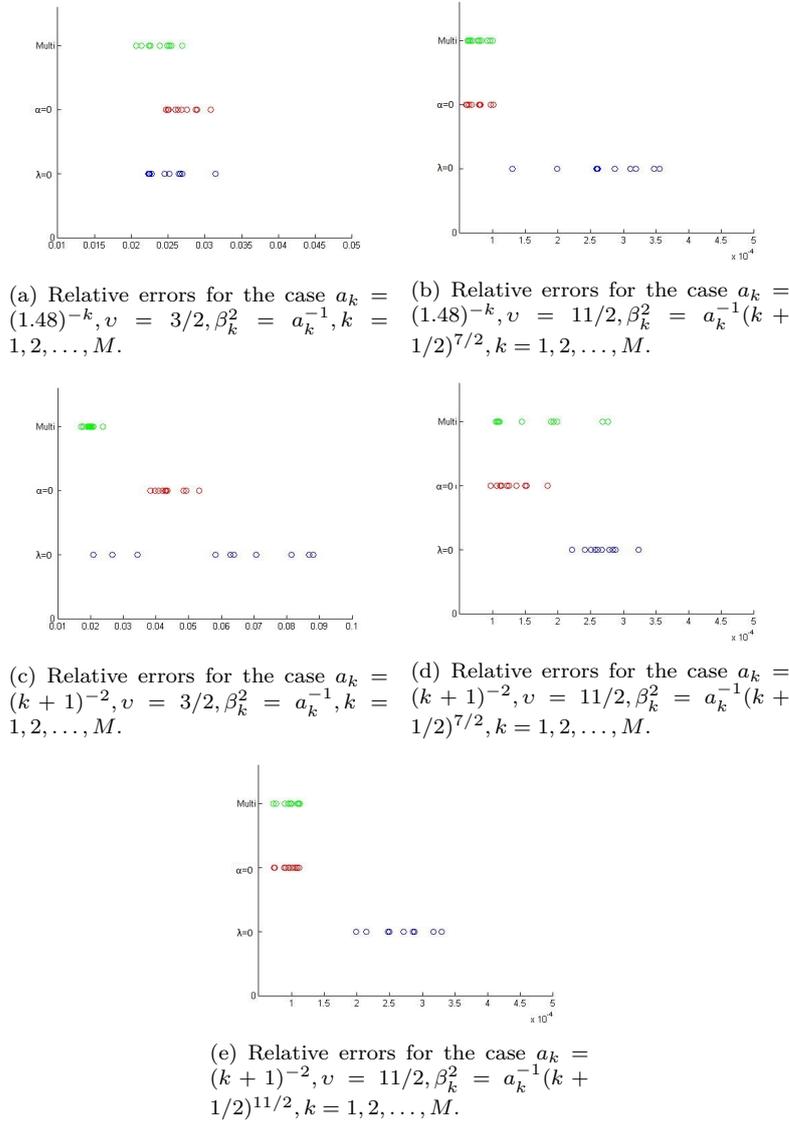


Fig. 1: Numerical illustrations. The comparison of the performance of single-parameter regularizations and their combination (Multi). In each of the figures the results corresponding to the two-step regularization are displayed at the top, the case of noise reduction with direct inversion is shown in the middle, and the case of regularized collocation corresponds to the bottom row. The horizontal axes show values of the relative errors (the vertical axes have no significance)

the operators with exponentially decreasing symbol, appearing in geoscience, have been discussed in the Introduction.

It is instructive to look at Figures 1a, 1b and 1c, 1d from the view point of [21], where *a priori* choice of the penalization weights  $\beta_k$  has been discussed. For the situations described in the above mentioned figures the choice of  $\beta_k$  suggested in [21] would be  $\beta_k^2 = a_k^{-1}(k+1)^v$ . But such a choice requires knowledge of the smoothness  $v$  of the unknown solutions. In practice the smoothness  $v$  is usually unknown, and the Figures 1a–1d correspond to the case of “underpenalization”.

As can be seen from Figures 1a–1d in such a case it is not clear *a priori* which of the considered single-parameter regularizations performs better. At the same time, these figures show that the proposed two-step method automatically follows the leader. The situation does not change when the penalization weights are chosen according to [21] (see Figure 1e).

In all our experiments the choice of the regularization parameters was done by using the quasi-optimality criterion described above with  $\alpha_0 = \lambda_0 = 1.78e - 5$ ,  $q = r = 1.25$ , and  $L = K = 50$ . Note also that we add the term  $\lambda = \alpha = 0$  to each of the sets (31), and (32).

These numerical results confirm the theoretical conclusion that the constructed two-step regularization methods can perform similarly to (or even better than) the best of single-parameter schemes.

## 6 Conclusion

We studied a two-parameter regularization scheme, which is a combination of the inversion after a data presmoothing and regularized collocation. In principle, each of the combined methods can be used alone as a regularization method. However, their performance very much depends on the interplay between the degree of ill-posedness of (1), (2), coded in the rate of increase of the spherical symbol  $\{a_k\}$ , and the smoothness of the solution.

For example, from Figure 1 one may conclude that for a severely ill-posed problem and moderately smooth solution (Figure 1a) the regularized collocation performs a bit better than the inversion after a data presmoothing, but for a moderately ill-posed problem (Figure 1c) this is not the case.

Since for ill-posed problems the smoothness of their solutions is usually unknown, it is not clear which of the single parameter regularizations is the better one. In this situation the combination of the single parameter regularizations, resulting in a two-parameter scheme, can be seen as a method of choice. The above figures illustrate this point.

At the end of Section 4 we have used Theorems 2 and 3 to demonstrate theoretically that the two-parameter scheme (18) is able to provide a better error bound than ones for the schemes involved in the combination. However, this demonstration is limited to an *a priori* choice of the regularization parameters, requiring knowledge of the solution smoothness, and therefore seldom usable in practice.

On the other hand, in Section 5 the above mentioned ability of the two-parameter scheme has been demonstrated for an *a posteriori* choice of the regularization parameters, where no knowledge of the solution smoothness has been used. At the same time, a *a posteriori* parameter choice employed in Section 5 is of a heuristic nature. We hope that this paper will stimulate research towards theoretical justification of *a posteriori* rules for multi-parameter regularization.

**Acknowledgements** The research of the first author was partially supported by the SRFPD fund 20120171120005 of China, Guangdong Provincial Government of China through the “Computational Science Innovative Research Team” program and Guangdong Province Key Laboratory of Computational Science at the Sun Yat-sen University. The second and the fourth authors are supported by the Austrian Fonds Zur Forderung der Wissenschaftlichen Forschung (FWF), grant P25424. The third author acknowledges the support of the Australian Research Council, project DP 120101816.

## 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. G. Bruckner, S. Prössdorf, G. Vainikko, Error bounds of discretization methods for boundary integral equations with noisy data, *Appl. Analysis* 63 (1996), pp. 25–37
3. G. Bruckner, S. V. Pereverzyev, Self-regularization of projection methods with a *a posteriori* discretization level choice for severely ill-posed problems, *Inverse Problems* 19 (2003), pp. 147–156
4. H. W. Engl, M. Hanke, A. Neubauer, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.
5. W. Freeden, V. Michel, *Multiscale potential theory. With Application to Geoscience*, Birkhäuser Boston Inc., Boston, MA, 2004.
6. 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.
7. A. Gelb, The resolution of the Gibbs phenomenon for Spherical harmonics, *Math. of Comput.*, 66:218 (1997), pp. 699–717.
8. P. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems. Numerical Aspects of Linear Inversion*, SIAM, Philadelphia, 1998.
9. H. Harbrecht, S. V. Pereverzyev, R. Schneider, Self-regularization by projection for noisy pseudodifferential equations of negative order, *Numer. Math.* 95 (2003) pp. 123–143.
10. S. Kindermann and A. Neubauer, On the convergence of the quasi-optimality criterion for (iterated) Tikhonov regularization, *Inverse Problems Imaging*, 2 (2008), pp. 291–299
11. E. Klann, P. Maass and R. Ramlau, Two-step regularization methods for linear inverse problems, *J. Inverse Ill-Posed Probl.*, 14 (2006), pp. 583–607
12. E. Klann, R. Ramlau, Regularization by fractional filter methods and data smoothing, *J. Inverse Ill-Posed Probl.*, 24 (2008), 26 pp.
13. Q. T. Le Gia, H. N. Mhaskar, Polynomial operators and local approximation of solutions of pseudo-differential equations on the sphere, *Numer. Math.*, 103 (2006), pp. 299–322
14. S. Lu, S. V. Pereverzev, *Regularization theory for ill-posed problems. Selected topics*, Walter de Gruyter GmbH, Berlin/Boston, 2013.
15. P. Mathé, Saturation of regularization methods for linear ill-posed problems in Hilbert spaces, *SIAM J. Numer. Anal.* 42 (2004), pp. 968–973
16. C. Muller, *Spherical Harmonics*, Lecture notes in Math. 17, Springer-Verlag, Berlin, 1966.
17. F. Narcowich, P. Petrushev, J. Ward, Decomposition of Besov and Triebel-Lizorkin spaces on the sphere, *J. Funct. Anal.*, 238 (2006), pp. 530–564.

18. V. Naumova, S. V. Pereverzev, P. Tkachenko, Regularized collocation for Spherical harmonics Gravitational Field Modeling, *Int. J. Geomath.*, DOI: 10.1007/s13137-013-0054-9.
19. A. Neubauer, On converse and saturation results for Tikhonov regularization of linear ill-posed problems, *SIAM J. Numer. Anal.* 34 (1997), pp. 517–527
20. 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), p. 1318
21. S. V. Pereverzyev, I. H. Sloan and P. Tkachenko, Parameter choice strategies for least-squares approximation of noisy smooth functions on the sphere, *RICAM-Report* 2014-04.
22. D. L. Ragozin, Constructive polynomial approximation on spheres and projective spaces, *Trans. Amer. Math. Soc.*, 162 (1971), pp. 157–170.
23. M. Reimer, Hyperinterpolation on the sphere at the minimum projection order, *J. Approx. Theory*, 104 (2000), pp. 272–286.
24. M. Reimer, Multivariate polynomial approximation, *International Series of Numerical Mathematics*, 144, Birkhauser Verlag, Basel, 2003.
25. S. L. Svensson, Pseudodifferential operators – a new approach to the boundary value problems of physical geodesy, *Manusc. Geod.*, 8 (1983), pp. 1–40.
26. I. H. Sloan, Polynomial approximation on spheres – generalizing de la Vallee-Poussin, *Comput. Methods Appl. Math.*, 11 (2011), pp. 540–552.
27. I. H. Sloan and R. S. Womersley, Filtered hyperinterpolation. A constructive polynomial approximation on the sphere, *Int. J. Geomath.*, 3 (2012), pp. 95–117.
28. P. N. Swartztrauber, On the Spectral Approximation of Discrete Scalar and Vector Functions on a Sphere, *SIAM J. Numer. Anal.*, 16:6 (1979), pp. 934–949.
29. A. N. Tikhonov and V. B. Glasko, Use of regularization methods in non-linear problems, *USSR Comput. Math. Phys.*, 5 (1965), pp. 93–107.
30. G. Wahba, Spline interpolation and smoothing on the sphere, *SIAM J. Sci. Statist. Comput.*, 2 (1981), pp. 5–16.