

# Local RBF Approximation for Scattered Data Fitting with Bivariate Splines

Oleg Davydov, Alessandra Sestini and Rossana Morandi

## Abstract

In this paper we continue our earlier research [4] aimed at developing efficient methods of local approximation suitable for the first stage of a spline based two-stage scattered data fitting algorithm. As an improvement to the pure polynomial local approximation method used in [5], a hybrid polynomial/radial basis scheme was considered in [4], where the local knot locations for the RBF terms were selected using a greedy knot insertion algorithm. In this paper standard radial local approximations based on interpolation or least squares are considered and a faster procedure is used for knot selection, significantly reducing the computational cost of the method. Error analysis of the method and numerical results illustrating its performance are given.

## 1 Introduction

Let  $\mathbf{X} \subset \Omega$ , be a set of scattered distinct sites and  $\{(\mathbf{x}, f_{\mathbf{x}}) : \mathbf{x} \in \mathbf{X}, f_{\mathbf{x}} \in \mathbb{R}\}$  the set of data points to be approximated,  $\Omega \subset \mathbb{R}^d$ . The idea of the two-stage method [16] is to compute in the first stage a large number of *local approximations* to the data and use them in the second stage as a source of information (e.g., function values and gradients at vertices of a triangulation) for building a *global spline approximation* of the full data set using a localized *quasi-interpolation* type operator. This helps to avoid solving large linear systems and large scale optimization problems arising if the interpolating, smoothing or minimal energy spline is directly computed from the data. For a long time it has been believed that two-stage methods cannot produce approximations of the same quality as the above mentioned global methods.

Recently, a promising two-stage bivariate spline algorithm has been developed and tested in [5, 9]. Convincing numerical evidence has been provided that the new method is efficient, robust and avoids the drawbacks usually associated with the two-stage methods. One of the goals of [4] and this paper is to improve the performance of this method at the first stage by achieving the approximation quality of the radial basis function (RBF) methods [2], in the same time also avoiding their well known computational difficulties, by applying them only to small subsets of the data.

In the original approach (see [5]) the local polynomial approximations are computed as discrete least squares, with the polynomial degree automatically adjusted to local data by taking into account the estimates of the approximation power of local least squares

[3]. In [4] and here we consider local approximation schemes defining non-polynomial approximations which are later converted into polynomials and then extended to a spline by the same method as in [5]. In [4] we have provided numerical evidence that better accuracy of the approximation may be achieved if local polynomials are augmented by linear combinations of radial basis functions, so defining *hybrid* approximations which are still computed by discrete least squares. The knot set used for each local hybrid approximation is chosen using an adaptive greedy algorithm based on successive knot insertion and estimates from [3].

In this paper we consider the standard radial approximations in the local stage that are computed by interpolation or by the least-squares method, with the local knots selected using a thinning algorithm similar to that suggested in [7] in the context of multiresolution. (Note that our motivation for thinning is entirely based on the computational considerations since the condition numbers of the matrices arising in the RBF method depend on the so called *separation distance* of the knots.)

The paper is organized as follows. In Section 2 we introduce the local approximation scheme. In Section 3 we provide an error analysis of this version of the two-stage method based on available estimates for the RBF interpolants. Sections 4 and 5 are devoted to extensive numerical tests with two goals: to verify the approximation order of the method, and to compare the performance of this new method with the method of [4] for some real world data sets.

## 2 Local RBF Approximation

At the first stage of a two-stage method, the local approximations are needed for each cell  $T$  of a partition of  $\Omega$  associated with the spline method used. (Such a cell is usually a  $d$ -dimensional simplex or cube.) The task of the first stage is to find a good approximation of the underlying function on  $T$ . To this end, the data from some domain  $\omega$ , where  $T \subset \omega \subset \Omega$ , are used. As in [4, 5, 9], we select the domain  $\omega$  initially as a ball with center at the barycenter of  $T$  and of radius equal to the diameter of  $T$ . If the number of data points located in this ball is smaller than a user specified number  $M_{min}$ , then the radius of the ball  $\omega$  is enlarged until this number is achieved. Another user specified parameter,  $M_{max}$ , controls the maximal number of points to be used, and a uniform type thinning is employed, if needed. Thus, this *data selection* procedure delivers a set of data sites  $\mathbf{X}_\omega = \{\mathbf{x}_1, \dots, \mathbf{x}_{N_\omega}\} \subset \mathbf{X} \cap \omega$ . where

$$N_\omega \leq M_{max}. \quad (1)$$

The local RBF approximation has the following form

$$\ell_\omega(\cdot) = \sum_{j=1}^m a_j p_j(\cdot) + \sum_{j=1}^{n_\omega} b_j \phi_\omega(\|\cdot - \mathbf{y}_j\|_2), \quad (2)$$

where the set of *knots*  $\mathbf{Y}_\omega = \{\mathbf{y}_j : j = 1, \dots, n_\omega\}$  is a subset of  $\mathbf{X}_\omega$ ,  $\{p_1, \dots, p_m\}$ ,  $m = \binom{d+q}{d}$ , is a suitable basis for the space  $\Pi_q^d$  of  $d$ -variate polynomials of total degree

$q \geq 0$ , and  $\phi_\omega : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$  is a *radial basis function*, i.e., a positive definite function or a conditionally positive definite function of order  $s \leq q + 1$  on  $\mathbb{R}^d$  [2], adjusted to the size of  $\omega$  by scaling. Thus, we take

$$\phi_\omega(r) = \phi\left(\frac{r}{\delta d_\omega}\right), \quad r \geq 0, \quad (3)$$

where  $\phi$  is a fixed radial basis function,  $d_\omega$  is the diameter of  $\omega$ , and  $\delta$  is a user specified parameter.

In this paper we consider only positive definite radial basis functions or conditionally positive definite radial basis functions of order 1. Therefore, it is sufficient to take

$$q = 0.$$

The function  $\ell_\omega$  of the form (2) is selected by using *interpolation* on the coarse set  $\mathbf{Y}_\omega$ , i.e. requiring

$$\ell_\omega(\mathbf{y}_j) = f_{\mathbf{y}_j}, \quad j = 1, \dots, n_\omega, \quad (4)$$

with additional *orthogonality condition*

$$\sum_{j=1}^{n_\omega} b_j = 0. \quad (5)$$

The existence and uniqueness of such a function is guaranteed for any  $\mathbf{Y}_\omega$  (see e.g. [2]). In particular, the matrix of the corresponding linear system,

$$\begin{bmatrix} e^T & A_{\mathbf{Y}_\omega} \\ 0 & e \end{bmatrix},$$

where  $e := (1, \dots, 1)$ ,

$$A_{\mathbf{Y}_\omega} := \begin{bmatrix} \phi_\omega(\|\mathbf{y}_1 - \mathbf{y}_1\|_2) & \dots & \phi_\omega(\|\mathbf{y}_1 - \mathbf{y}_{n_\omega}\|_2) \\ \vdots & & \vdots \\ \phi_\omega(\|\mathbf{y}_{n_\omega} - \mathbf{y}_1\|_2) & \dots & \phi_\omega(\|\mathbf{y}_{n_\omega} - \mathbf{y}_{n_\omega}\|_2) \end{bmatrix},$$

is nonsingular as soon as all knots  $\mathbf{y}_1, \dots, \mathbf{y}_{n_\omega}$  are distinct.

Since the linear system arising in this interpolation problem is of the size  $n_\omega + 1 \leq M_{max} + 1$ , its solution can be easily computed if  $M_{max}$  is not large, and if the matrix  $A_{\mathbf{Y}_\omega}$  is well conditioned.

To complete the description of the method we now explain how we choose  $\mathbf{Y}_\omega$ . It is known that the condition number of  $A_{\mathbf{Y}_\omega}$  can be bounded in terms of the reciprocal of the *separation distance*

$$s(\mathbf{Y}_\omega) = \frac{1}{2} \min_{1 \leq i < j \leq n_\omega} \|\mathbf{y}_i - \mathbf{y}_j\|_2.$$

Therefore, we choose  $\mathbf{Y}_\omega$  such that

$$d_\omega/s(\mathbf{Y}_\omega) \leq S, \quad (6)$$

where  $S$  is again a user specified number. To guarantee (6), the thinning algorithm from [7] is adapted.

As an alternative to interpolation, the *discrete least squares approach* [10, 12] can also be considered, i.e.  $\ell_\omega$  of the form (2) can be selected via the minimization of the *least-squares error* (the  $\ell_2$ -norm of the residual on  $\mathbf{X}_\omega$ ),

$$\left( \sum_{i=1}^{N_\omega} (f_i - \ell_\omega(\mathbf{x}_i))^2 \right)^{1/2}, \quad (7)$$

using the orthogonality condition (5) as a linear equality constraint. The existence and uniqueness of the least squares approximation follows from the theory of constrained least squares, see [1].

Regardless whether we use interpolation or least squares, and besides the choice of the radial basis function  $\phi$ , the scheme depends on the following *parameters* that are supposed to be specified by the user globally, i.e., the same values are used for all local approximations:

$$M_{min}, M_{max}, \delta, S. \quad (8)$$

In real world applications these parameters have to be adjusted to a particular type of data by some *calibration* procedure. The local error estimates discussed in the next section can also be useful for this.

### 3 Error Bounds

To facilitate a correct comparison to the approximation results for global methods, we mention that the approximation order of a two-stage method is the minimum of the order of the spline operator and that of the local scheme [16]. More precisely, let us assume for simplicity that the subdomains where local approximations are needed are the *cells*  $T$  of a *uniform partition*  $\Delta$  of  $\Omega$  associated with the spline space, which is the case for the splines used in [5]. Then the approximation error of the two-stage scheme in the uniform norm for a sufficiently smooth function can be estimated by

$$C_1 h^{p+1} + C_2 \max\{e_T : T \in \Delta\}, \quad (9)$$

where  $h$  is the diameter of the cells,  $p+1$  is the approximation order of the spline quasi-interpolation operator,  $e_T$  is the error of local approximation, and  $C_1, C_2$  are some positive constants.

To assess the approximation error of the first stage of the two-stage method we invoke some results from the approximation theory of radial basis functions on bounded domains.

Let  $\ell_\omega(f)$  be the sum (2) determined by the conditions (4) and (5) with  $f_{\mathbf{y}_j} = f(\mathbf{y}_j)$ ,  $j = 1, \dots, n_\omega$ , for a function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ . We assume that  $f$  is smooth enough to belong to the *native space*  $\mathcal{F}_{\phi_\omega}$  associated with the radial basis function  $\phi_\omega$ ,

$$\mathcal{F}_{\phi_\omega} = \{f \in L_2(\mathbb{R}^d) : |f|_{\phi_\omega} < \infty\},$$

where

$$|f|_{\phi_\omega} := \left( \int_{\mathbb{R}^d} \frac{|\hat{f}(\mathbf{x})|^2}{\hat{\Phi}_\omega(\mathbf{x})} d\mathbf{x} \right)^{1/2}, \quad \Phi_\omega(\cdot) := \phi_\omega(\|\cdot\|_2),$$

and  $\hat{f}$  denotes the generalized Fourier transform.

Well known error bounds for the interpolation with radial basis functions (see, e.g. [2, 11, 13, 17]) lead in the case  $q = 0$  to the estimate

$$|f(\mathbf{x}) - \ell_\omega(f, \mathbf{x})| \leq 2 \sqrt{E_0(\Phi_\omega)_{C(B_{h(\mathbf{x}, \mathbf{Y}_\omega)})}} |f|_{\phi_\omega}, \quad \mathbf{x} \in \mathbb{R}^d, \quad (10)$$

where  $h(\mathbf{x}, \mathbf{Y}_\omega)$  is the distance between  $\mathbf{x}$  and  $\mathbf{Y}_\omega$ ,

$$h(\mathbf{x}, \mathbf{Y}_\omega) := \inf_{\mathbf{y} \in \mathbf{Y}_\omega} \|\mathbf{x} - \mathbf{y}\|_2,$$

$E_0(\Phi_\omega)_{C(B_{h(\mathbf{x}, \mathbf{Y}_\omega)})}$  is the error of the best constant approximation of  $\Phi_\omega$ ,

$$E_0(\Phi_\omega)_{C(B_{h(\mathbf{x}, \mathbf{Y}_\omega)})} = \inf_{p \in \Pi_0^d} \|f - p\|_{C(B_{h(\mathbf{x}, \mathbf{Y}_\omega)})},$$

and  $B_r$  denotes the ball in  $\mathbb{R}^d$  with center  $\mathbf{0}$  and radius  $r$ .

Assuming that  $\phi$  is monotone (which is true for all available radial basis functions at least in a neighborhood of zero) and considering the *fill distance* of  $\mathbf{Y}_\omega$  with respect to  $T$ ,

$$h(T, \mathbf{Y}_\omega) = \sup_{\mathbf{x} \in T} h(\mathbf{x}, \mathbf{Y}_\omega),$$

we have for all  $x \in T$ ,

$$E_0(\Phi_\omega)_{C(B_{h(\mathbf{x}, \mathbf{Y}_\omega)})} \leq E_0(\Phi_\omega)_{C(B_{h(T, \mathbf{Y}_\omega)})} = \frac{1}{2} |\phi_\omega(h(T, \mathbf{Y}_\omega)) - \phi_\omega(0)|,$$

which leads to the estimate

$$\|f - \ell_\omega(f)\|_{C(T)} \leq \sqrt{2} |\phi_\omega(h(T, \mathbf{Y}_\omega)) - \phi_\omega(0)| \cdot |f|_{\phi_\omega}. \quad (11)$$

Obviously,  $h(T, \mathbf{Y}_\omega) \leq d_\omega$ , and taking into account (3) we have

$$|\phi_\omega(h(T, \mathbf{Y}_\omega)) - \phi_\omega(0)| = \left| \phi\left(\frac{h(T, \mathbf{Y}_\omega)}{\delta d_\omega}\right) - \phi(0) \right| \leq |\phi(1/\delta) - \phi(0)|, \quad (12)$$

which shows that increasing the value of the parameter  $\delta$  may have a positive effect on the error. It should, however, be taken into account that the seminorm  $|f|_{\phi_\omega}$  also depends on  $\delta$ , in view of (3).

Among the most commonly used radial basis functions are the *thin plate splines*

$$\phi^{\text{TP}, \beta}(r) = \begin{cases} (-1)^{\lceil \beta/2 \rceil} r^\beta, & \beta \in \mathbb{R}_{>0} \setminus 2\mathbb{N}, \\ (-1)^{\beta/2+1} r^\beta \log r, & \beta \in 2\mathbb{N}, \end{cases} \quad (13)$$

that are conditionally positive definite of order  $\lceil \beta/2 \rceil$  if  $\beta \in \mathbb{R}_{>0} \setminus 2\mathbb{N}$ , and  $\beta/2 + 1$  if  $\beta \in 2\mathbb{N}$ . (Here  $\lceil x \rceil$  denotes the smallest integer greater or equal to  $x \in \mathbb{R}$ .) Therefore they can be used in our scheme (where the polynomial degree is  $q = 0$ ) if  $0 < \beta < 2$ .

The approximation order of the thin plate splines is understood better than that of the other available RBFs because their Fourier transform is *homogeneous*,  $\hat{\Phi}^{\text{TP},\beta}(\mathbf{x}) = K\|\mathbf{x}\|_2^{-\beta-d}$ , with some constant  $K$  independent of  $\mathbf{x}$ . Therefore,

$$|f|_{\phi_\omega^{\text{TP},\beta}}^2 = (\delta d_\omega)^\beta |f|_{\phi^{\text{TP},\beta}}^2,$$

and we obtain from (11) and (12),

$$\|f - \ell_\omega(f)\|_{C(T)} \leq \sqrt{2h(T, \mathbf{Y}_\omega)^\beta} |f|_\phi, \quad 0 < \beta < 2. \quad (14)$$

By our algorithm (see Section 2), we always have  $d_\omega \geq 2h$ , where  $h$  is the diameter of the cell  $T$ , which is the same for all cells in our setting. On the other hand, we may assume that there is enough data so that  $d_\omega \leq ch$ , for a constant  $c$ . Taking into account (6) and the obvious inequality  $s(\mathbf{Y}_\omega) \leq h(T, \mathbf{Y}_\omega)$ , we have

$$2h/S \leq d_\omega/S \leq h(T, \mathbf{Y}_\omega) \leq d_\omega \leq ch.$$

Therefore, (9) and (14) suggest that the approximation order of the two-stage method with thin plate splines in the local stage should be  $\mathcal{O}(h^{\min\{\beta/2, p+1\}})$  or, assuming that  $p$  is high enough,  $\mathcal{O}(h^{\beta/2})$ . Note that since the cell  $T$  where we use the local approximations covers only the central part of  $\omega$ , the deterioration of the error near the boundary of  $\omega$  only affects the quality of the local approximations at the boundary of the entire domain  $\Omega$ .

Finally, we mention that the above estimates can be improved if  $f$  satisfies some more stringent requirements than  $f \in \mathcal{F}_{\phi_\omega}$ , see [2, 14, 15]. The improvement amounts basically (up to a constant factor) to removing the square root sign in (10), (11) and (14), and replacing the seminorm  $|f|_{\phi_\omega}$  with a stronger seminorm  $|f|$ , whose boundedness requires “higher smoothness” of  $f$ .

In particular, for the thin plate splines the approximation order becomes  $\mathcal{O}(h^\beta)$ . Moreover, in this case the order  $\mathcal{O}(h^{\beta+d/2})$  for scattered data and  $\mathcal{O}(h^{\beta+d})$  for grid data has been proved (see [2]).

## 4 Numerical Results: Approximation Order

In our numerical experiments we restrict ourselves to the two dimensional case  $d = 2$ . This section is devoted to numerical tests with randomly generated data for the well known Franke test function [8]. The goals of the tests are to measure the approximation order of the two-stage method, compare it with the theoretical error bounds, and get hints for the selection of good values of the parameters (8) for the local approximation.

More precisely, 40 different random data sets  $\mathbf{X}_i$ ,  $i = 1, \dots, 40$ , of cardinality  $\#\mathbf{X}_i = N$  were generated in the reference square  $[0, 1]^2$ , for  $N = 10^2, 10^3, 10^4, 10^5$ . For the second stage of the two-stage method we have chosen the method  $SQ_2^{av}$  of [5] which produces  $C^2$  piecewise sextic splines on the four directional mesh. Based on our experiments with the Franke test function in [4], we take the grid size for the spline

space to be  $n \times n$ , where  $n$  is the closest integer to  $\sqrt{N}/2$ . The experiments have been performed using the implementation of the spline operators in [6].

To measure the approximation error, we compute the maximum error  $\epsilon_i$  of the spline relative to the exact function values on a dense  $(10n + 1) \times (10n + 1)$  grid in a suitably reduced window  $([0.2, 0.8]^2)$  for every data set  $\mathbf{X}_i$  and take the geometric average  $max = \exp(\frac{1}{40} \sum_{i=1}^{40} \ln \epsilon_i)$  of these errors. We think that the geometric average is the most appropriate way of averaging for the approximation order tests. The motivation for using the reduced window is our desire to avoid boundary effects.

In the local stage we use the interpolation method described in Section 2 above and choose 1) the *thin plate spline*  $\phi(r) = -r^\beta$ ,  $\beta = 3/2$  or  $\beta = 7/4$ , and 2) the *multiquadrics*  $\phi(r) = -\sqrt{1 + r^2}$  for the experiments. We have chosen a high value  $M_{max} = 400$  to eliminate the influence of this parameter and tried to find nearly optimal values for  $M_{min}$ ,  $S$  and  $\delta$ . The results are presented in Tables 1 and 2.

N	spline grid	max ( $\beta = 3/2$ )	max ( $\beta = 7/4$ )
$10^2$	$5 \times 5$	$8.55 \cdot 10^{-2}$	$6.92 \cdot 10^{-2}$
$10^3$	$16 \times 16$	$5.22 \cdot 10^{-3}$	$3.37 \cdot 10^{-3}$
$10^4$	$50 \times 50$	$2.60 \cdot 10^{-4}$	$1.17 \cdot 10^{-4}$
$10^5$	$158 \times 158$	$2.37 \cdot 10^{-5}$	$7.09 \cdot 10^{-6}$

Table 1: Maximum error using the local RBF interpolation scheme based on  $\phi(r) = -r^\beta$ ,  $\beta = 3/2$  and  $7/4$ . Parameter values:  $M_{min} = 100$ ,  $S = 100$ ,  $\delta = 1$ .

For the thin plate spline (Table 1), the experiments confirm that the parameter  $\delta$  does not influence the error significantly. Therefore, we have chosen a nominal value  $\delta = 1$ . Note that the average number of RBF knots in the local approximations approaches 140 for the larger data sets, which makes these tests particularly slow. Although an increase of  $M_{min}$  was always profitable for  $\phi(r) = -r^\beta$ , fewer knots were sufficient to achieve nearly optimal errors for  $N < 10^5$ . In this sense nearly optimal values of the parameter  $M_{min}$  are:  $M_{min} = 20$  for  $N = 10^2$  (28 knots),  $M_{min} = 30$  for  $N = 10^3$  (45 knots), and  $M_{min} = 60$  for  $N = 10^4$  (65 knots). (We have taken  $S = M_{min}$  in these tests.)

Table 1 suggests the approximation order about  $h^{\beta+1}$ , which conforms nicely to the available theoretical results, see the comments at the end of Section 3. Note that the approximation error of the spline operator  $SQ_2^{av}$  is  $\mathcal{O}(h^7)$  [5] and hence it is negligible for this test.

It is clear from Table 2 that in the case of multiquadrics the correct choice of the parameter  $\delta$  (which clearly is related to the reciprocal of the classical multiquadric coefficient  $c$  in  $\sqrt{c^2 + r^2}$ ) is important. However, we had to choose the separation parameter  $S$  such that  $\delta S \leq 16$  since otherwise the computation with multiquadrics turned out numerically unstable. (Note that for the real world data, like those tested below in Section 5,  $\delta S$  must be even smaller.) The values of  $M_{min}$  lower than 100 were disadvantageous in our experiments for all  $N$  except  $N = 10^2$ . In the case  $N = 10^2$ , however,  $M_{min} = 100$  delivers relatively high errors:  $1.94 \cdot 10^{-2}$  for  $\delta = 0.2, S = 80$ ,  $3.75 \cdot 10^{-2}$  for  $\delta = 0.4, S = 40$ ,  $6.59 \cdot 10^{-2}$  for  $\delta = 0.8, S = 20$ ,  $1.08 \cdot 10^{-1}$  for  $\delta = 1.2, S = 40/3$ ,

N	$\delta = 0.4$ $S = 40$	$\delta = 0.8$ $S = 20$	$\delta = 1.2$ $S = 40/3$	$\delta = 1.6$ $S = 10$
$10^2$	$2.27 \cdot 10^{-2}$	$2.81 \cdot 10^{-2}$	$3.66 \cdot 10^{-2}$	$4.48 \cdot 10^{-2}$
$10^3$	$1.26 \cdot 10^{-5}$	$4.44 \cdot 10^{-6}$	$6.13 \cdot 10^{-5}$	$5.42 \cdot 10^{-4}$
$10^4$	$4.20 \cdot 10^{-6}$	$1.98 \cdot 10^{-7}$	$1.00 \cdot 10^{-7}$	$2.36 \cdot 10^{-7}$
$10^5$	$2.03 \cdot 10^{-6}$	$9.28 \cdot 10^{-8}$	$3.54 \cdot 10^{-8}$	$5.60 \cdot 10^{-8}$
$\#knots (N = 10^5)$	122.2	87.2	60.4	42.8

Table 2: Maximum error using the local RBF interpolation scheme based on  $\phi(r) = -\sqrt{1+r^2}$  for different values of  $\delta$ . The spline grid is the same as in Table 1. Other parameters:  $M_{min} = 20$  if  $N = 10^2$  and  $M_{min} = 100$  otherwise.

$1.65 \cdot 10^{-1}$  for  $\delta = 1.6, S = 10$ . (Note that  $N = M_{min} = 100$  means, in fact, that all local approximations are the same, and, hence, our spline does not differ much from the corresponding *global* multiquadric approximation.) Therefore, we use  $M_{min} = 20$  if  $N = 10^2$ , and  $M_{min} = 100$  for other  $N$ . The value  $M_{min} > 100$  may be advantageous for smaller  $\delta$ . For example, for  $N = 10^3$  and  $M_{min} = 200$  we have  $3.36 \cdot 10^{-6}$  if  $\delta = 0.4, S = 40$ , and  $1.26 \cdot 10^{-5}$  if  $\delta = 0.8, S = 20$ .

The results in Table 2 confirm that greater values of  $\delta$  tend to provide better errors. Indeed, for higher  $N$  we have to increase  $\delta$  in order to obtain the best errors, even though numerical stability considerations force us to take smaller  $S$ , which in turn leads to the reduction of the number of RBF knots (see the last row of the table). For any fixed  $\delta$ , however, Table 2 shows a substantial deterioration of the approximation order as  $N$  increases. The estimates of Section 2 do not provide a full theoretical explanation for this behavior. In particular, (11) includes the term  $|f|_{\phi_\omega}$ , whose behavior for  $d_\omega \rightarrow 0$  is not clear to us in the case of multiquadrics.

## 5 Numerical Results: Real World Data

The second group of our experiments is aimed at verifying the performance of the proposed scheme compared with the hybrid approach introduced in [4].

To this end we consider the same real-world data sets as in [4, 5], namely, the Glacier data (GL, 8345 points), the Black Forest data (BF, 15885 points) and the Rotterdam Port data (RP, 621624 points after cleaning, see [5]). Referring to [4, 5] for the description of these data sets, we only mention that GL is available from [8] and that RP has been provided by Quality Positioning Services (Zeist, The Netherlands), and it has been recorded using the QINSy software.

Note that in the first stage we use the *least-squares* method as described at the end of Section 2 since it consistently produced better results than interpolation for the real world data in our tests. To solve the constrained least squares problems we employ the routine DGESDD from LAPACK. (Note that the interpolation method in our implementation is also treated as special case of least squares.)

The results are reported in Table 3, where maximum (*max*), mean (*mean*) and



root mean square (*rms*) errors at the data points, the average number of RBF knots (*#knots*) used for the local approximations, and the computational time (*time*) are shown. Results obtained with the method suggested in this paper (R) are compared with the hybrid approach (H) of [4]. As in [4], we use *multiquadric* RBF in these tests. The degree  $q$  of the polynomial term is 0 in all tests except RP/H, where  $q = 1$ . In the second stage we use the spline methods  $RQ_2^{av}$  (piecewise sextic) for GL and BF and  $Q_1^{av}$  (piecewise cubic) for RP, as in the respective tests in [4, 5]. The computer used for these experiments is a Pentium 4m / 1.9 GHz / 768 MB RAM.

	GL/H	GL/R	BF/H	BF/R	RP/H	RP/R
<i>max</i>	15.6 <i>m</i>	17.8 <i>m</i>	32.0 <i>m</i>	30.0 <i>m</i>	92.5 <i>cm</i>	90.2 <i>cm</i>
<i>mean</i>	1.57 <i>m</i>	1.49 <i>m</i>	1.39 <i>m</i>	1.25 <i>m</i>	5.46 <i>cm</i>	5.23 <i>cm</i>
<i>rms</i>	2.26 <i>m</i>	2.19 <i>m</i>	2.17 <i>m</i>	2.00 <i>m</i>	7.46 <i>cm</i>	7.22 <i>cm</i>
<i>#knots</i>	14.9	20.8	12.2	8.2	5.4	11.6
<i>time</i>	33.0 <i>sec</i>	3.7 <i>sec</i>	134 <i>sec</i>	7.11 <i>sec</i>	316 <i>sec</i>	97.3 <i>sec</i>

Table 3: Results for the data sets GL, BF and RP. Parameters (see [4] for the meaning of  $\kappa_P$  and  $\kappa_H$ ): 1) GL. Spline grid  $20 \times 24$ ,  $M_{min} = 60$ ,  $M_{max} = 160$ ,  $\delta = 0.4$  for both H and R methods,  $\kappa_H = 10^5$  for H, and  $S = 8$  for R. 2) BF. Spline grid  $80 \times 80$ ,  $M_{max} = 100$  for both H und R methods,  $M_{min} = 12$ ,  $\delta = 0.3$ ,  $\kappa_H = 10^4$  for H, and  $M_{min} = 3$ ,  $\delta = 0.4$ ,  $S = 7$  for R. 3) RP. Spline grid  $100 \times 281$ ,  $M_{min} = 3$ ,  $M_{max} = 100$ ,  $\delta = 0.4$  for both H and R,  $\kappa_P = 100$ ,  $\kappa_H = 2 \cdot 10^4$  for H, and  $S = 5$  for R.

In addition to Table 3, we provide Figures 1–3 that present zooms into the same subareas of the surfaces as ones used in [4]. They are produced with MATLAB using dense grid evaluations of the spline surfaces (see [4]). The figures confirm the high visual quality of our approximations, as they show no artificial oscillation or other unnatural behaviour. Table 3 also shows that the errors for both H and R methods are comparable, whereas the computational time for the method of this paper is substantially lower.

## References

- [1] A. Björck, Numerical methods for least squares problems, SIAM, Philadelphia, 1996.
- [2] M. D. Buhmann, Radial basis functions, Cambridge University Press, 2003.
- [3] O. Davydov, On the approximation power of local least squares polynomials, in algorithms for approximation IV, J. Levesley, I. J. Anderson and J. C. Mason (eds), 2002, 346–353.
- [4] O. Davydov, R. Morandi and A. Sestini, Local hybrid approximation for scattered data fitting with bivariate splines, manuscript, 2003. <http://www.uni-giessen.de/~gcn5/davydov/>

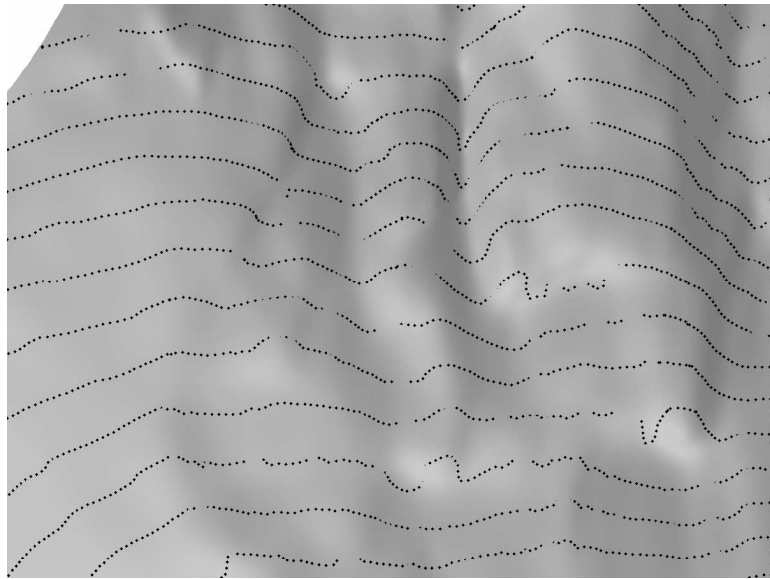


Figure 1: Glacier test.

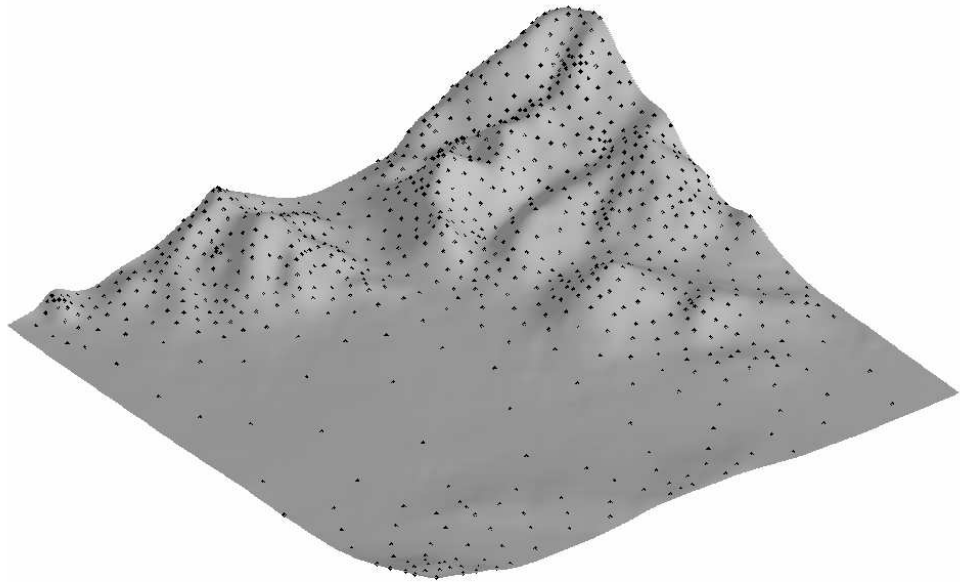


Figure 2: Black Forest test.



Figure 3: Rotterdam Port test.

- [5] O. Davydov and F. Zeilfelder, Scattered data fitting by direct extension of local polynomials to bivariate splines, *Adv. Comp. Math.* 21 (2004), 223–271.
- [6] O. Davydov and F. Zeilfelder, Toolbox for two-stage scattered data fitting, in preparation.
- [7] M. S. Floater and A. Iske, Thinning algorithms for scattered data interpolation, *BIT* 38 (1998), 705–720.
- [8] R. Franke, Homepage, <http://www.math.nps.navy.mil/~rfranke/>, Naval Post-graduate School.
- [9] J. Haber, F. Zeilfelder, O. Davydov, H.-P. Seidel, Smooth approximation and rendering of large scattered data sets, in: *Proceedings of IEEE Visualisation 2001* eds. T. Ertl, K. Joy and A. Varshney, 2001, pp. 341–347, 571.
- [10] A. Iske, Reconstruction of smooth signals from irregular samples by using radial basis function approximation, in: *Proceedings of the 1999 International Workshop on Sampling Theory and Applications*, The Norwegian University of Science and Technology, Trondheim, 1999, pp. 82-87.
- [11] K. Jetter, J. Stöckler and J. D. Ward, Error estimates for scattered data interpolation on spheres, *Math. Comp.* 68 (1999), 733–747.

- [12] J. R. McMahon and R. Franke, Knot selection for least squares thin plate splines, *SIAM J. Sci. Stat. Comput.* 13 (1992), 484–498.
- [13] R. Schaback, Reconstruction of multivariate functions from scattered data, manuscript, 1997. <http://www.num.math.uni-goettingen.de/schaback/>
- [14] R. Schaback, Improved error bounds for scattered data interpolation by radial basis functions, *Math. Comp.* 68 (1999), 201–216.
- [15] R. Schaback and H. Wendland, Inverse and saturation theorems for radial basis function interpolation, *Math. Comp.* 71 (2002), 669–681.
- [16] L. L. Schumaker, Fitting surfaces to scattered data, in: *Approximation Theory II* eds. G. G. Lorentz, C. K. Chui, and L. L. Schumaker, Academic Press, New York, 1976, pp. 203–268.
- [17] H. Wendland, Gaussian interpolation revisited, in: *Trends in Approximation Theory* eds. K. Kopotun, T. Lyche and M. Neamtu, Vanderbilt University Press, Nashville, TN, 2001, pp. 427–436.

Oleg Davydov  
Mathematisches Institut  
Justus-Liebig-Universität Giessen  
D-35392 Giessen  
GERMANY  
[oleg.davydov@math.uni-giessen.de](mailto:oleg.davydov@math.uni-giessen.de)

Alessandra Sestini and Rossana Morandi  
Dipartimento di Energetica  
Università di Firenze  
Via Lombroso 6/17, 50134 Firenze  
ITALY  
[morandi@de.unifi.it](mailto:morandi@de.unifi.it)  
[sestini@de.unifi.it](mailto:sestini@de.unifi.it)