Fast Evaluation of Radial Basis Functions: Methods Based on Partition of Unity

Holger Wendland

Abstract. We combine the theory of radial basis function interpolation with a partition of unity method to solve large-scale, scattered data problems. We analyze the computational complexity and pay special attention to the underlying data structure. Finally, we give a numerical example.

§1. Radial Basis Functions

Radial basis functions are nowadays a popular choice for solving interpolation problems, where the data sites $X = \{x_1, \dots, x_N\}$ are scattered points in $\mathbb{R}^d$ with $d \geq 2$. The idea is to choose a fixed conditionally positive definite function Φ of a certain order m and to form the interpolant to the data $f_1, \dots, f_N$ by

$$s_{f,X} = \sum_{j=1}^{N} \alpha_j \Phi(\cdot - x_j) + p, \tag{1}$$

where p is a polynomial of degree $m-1$. Additionally to the interpolation conditions $s_{f,X}(x_j) = f_j$, $1 \leq j \leq N$, the coefficients have to satisfy the relation $\sum \alpha_j q(x_j) = 0$ for all polynomials q of degree less than m. It is well known that under certain mild conditions on the location of the data sites, this interpolation problem is uniquely solvable for a wide range of basis functions, including thin plate splines, multiquadrics and Gaussians (see [7] for a recent overview). These globally supported basis functions have the drawback of being unfavourably expensive concerning computational complexity. To be more precise, direct methods for computing the interpolant $s_{f,X}$ need $\mathcal{O}(N^3)$ operations and each evaluation of (1) takes

Approximation Theory X: Wavelets, Splines, and Applications
Charles K. Chui, Larry L. Schumaker, and Joachim Stöckler (eds.), pp. 473–483.

ISBN 0-8265-1416-2.

another $\mathcal{O}(N)$ operations. This is obviously too much if problems with more than $N = 100,000$ points have to be dealt with and many people consider radial basis functions only to be useful for small problems. But this is definitely wrong. Much work in the direction of fast solvers for the interpolation problem and in the direction of fast evaluation of the interpolant (1) has been done in recent years (cf. [2,3,4,5,8]). All these methods retain the initial approximation space and allow to decrease the computational complexity dramatically. But since everything has its price, these methods need some insight into multipole expansions, which are rather delicate to implement.

Hence, we want to propose another method, which is actually not a new one. The idea is simply to solve a large number of small, local problems instead of one large-scale problem and to put the local solutions together by a partition of unity method. As a matter of fact, this is a very natural way to deal with radial basis functions, which are actually local methods. They are local in the following sense. If we investigate the approximation property of our interpolant (1) at a point x, we only need those data points x_j that are close to x. We will come back to this in a later section.

This paper is organized as follows. The next section is devoted to a general description of partition of unity approximation methods. In the third section we will use radial basis functions as the local spaces for the partition of unity method. This allows us to show that this method inherits the approximation properties of the global radial basis function interpolant. In the fourth section we discuss data structures and computational complexity. In the fifth section we shortly comment on stability of the procedure. In the final section we provide a numerical example.

§2. Partition of Unity

The idea of a partion of unity method is the following. We start with a mildly overlapping covering $\{\Omega_j\}_{j=1}^M$ of the region Ω. We will make the term "mildly overlapping" more precise in just a moment. Associated with this covering we choose a partition of unity, i.e. a family of compactly supported, continuous functions $\{w_j\}$ with $\sum_{j=1}^M w_j = 1$ and $\operatorname{supp}(w_j) \subseteq \overline{\Omega_j}$. Moreover, for every cell Ω_j we choose an approximation space V_j. Then, a function f is approximated on each cell by a local approximant $s_j \in V_j$ and the local approximants are put together by forming

$$s_f(x) = \sum_{j=1}^{M} s_j(x) w_j(x). \tag{2}$$

To be more precise, we introduce

Definition 1. *Let $\Omega \subseteq \mathbb{R}^d$ be a bounded set. Let $\{\Omega_j\}_{j=1}^M$ be an open and bounded covering of Ω. This means all Ω_j are open and bounded and Ω is contained in their union. Set $\delta_j = \operatorname{diam}(\Omega_j) = \sup_{x,y\in\Omega_j} \|x-y\|_2$. We call a family of functions $\{w_j\}_{j=1}^M$ with $w_j \in C^k(\mathbb{R}^d)$ a* k-stable partition of unity with respect to the covering *$\{\Omega_j\}$, if*

1) $\operatorname{supp}(w_j) \subseteq \overline{\Omega_j}$,
2) $\sum_{j=1}^{M} w_j \equiv 1$ *on* Ω,
3) *for every $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \le k$ there exists a constant $C_\alpha > 0$ such that*
$$\|D^\alpha w_j\|_{L_\infty(\Omega_j)} \le C_\alpha / \delta_j^{|\alpha|}$$
for all $1 \le j \le M$.

So far, we did not make any further assumptions on the covering $\{\Omega_j\}$, but for efficiency it is necessary that the cardinality of
$$I(x) := \{j : x \in \Omega_j\}$$
is uniformly bounded on Ω. Nonetheless, even without this assumption we can give a first convergence result. But without this assumption we will lose convergence orders.

Theorem 2. *Let $\Omega \subseteq \mathbb{R}^d$ be bounded. Suppose $\{\Omega_j\}_{j=1}^M$ is an open and bounded covering of Ω, and $\{w_j\}_{j=1}^M$ is a k-stable partition of unity. Let $f \in C^k(\Omega)$ be the function to be approximated. Let $V_j \subseteq C^k(\Omega_j)$ be given. Assume that the local approximation spaces V_j have the following approximation property: On each patch $\Omega_j \cap \Omega$, f can be approximated by a function $s_j \in V_j$ such that*
$$\|D^\alpha f - D^\alpha s_j\|_{L_\infty(\Omega\cap\Omega_j)} \le \varepsilon_j(\alpha).$$
Then the function (2) satisfies
$$|(D^\alpha f - D^\alpha s_f)(x)| \le \sum_{j\in I(x)} \sum_{\beta\le\alpha} \binom{\alpha}{\beta} C_{\alpha-\beta} \delta_j^{|\beta|-|\alpha|} \varepsilon_j(\beta) \tag{3}$$
for all $|\alpha| \le k$.

Proof: The proof is straightforward. We simply use Leibniz' rule and the fact that the $\{w_j\}$ form a partition of unity to derive
$$\begin{aligned}(D^\alpha f - D^\alpha s_f)(x) &= D^\alpha \sum_{j=1}^{M} w_j(x)(f(x) - s_j(x)) \\ &= \sum_{j\in I(x)} \sum_{\beta\le\alpha} \binom{\alpha}{\beta} D^{\alpha-\beta} w_j(x) D^\beta (f - s_j)(x).\end{aligned}$$

The assumed bounds on the derivatives of w_j and on $D^\beta(f - s_j)$ now yield the stated result. □

§3. Radial Basis Functions as the Local Method

It is now our goal to use radial basis functions to define the local approximation spaces. To this end and for the reader's convenience, we review some details on local error estimates for radial basis function interpolation, which can be found in [11].

To every conditionally positive definite function Φ and every region $\Omega \subseteq \mathbb{R}^d$, there exists a natural function space $\mathcal{N}_\Phi(\Omega)$, the native Hilbert space. This space can be defined in various ways, and we assume that the reader is familiar with the concept (see [10]). In many cases the native space turns out to be a classical smoothness space like Sobolev or Beppo Levi space. Here, we need only few properties. The first one is that if $\Phi \in C^k(\mathbb{R}^d)$, its smoothness is inherited by the native space via $\mathcal{N}_\Phi(\Omega) \subseteq C^{\lfloor k/2 \rfloor}(\Omega)$. Hence, we might naturally ask for bounds on $D^\alpha f - D^\alpha s_{f,X}$ for $f \in \mathcal{N}_\Phi(\Omega)$ and $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \leq k/2$. The second property we need is the extension property. Every function $f \in \mathcal{N}_\Phi(\Omega)$ has a natural, norm preserving extension to a function $Ef \in \mathcal{N}_\Phi(\mathbb{R}^d)$.

As usual we will measure the approximation error in terms of the fill distance

$$h_{X,\Omega} = \sup_{x \in \Omega} \min_{x_j \in X} \|x - x_j\|_2.$$

We will state our convergence results only in case of (conditionally) positive definite functions with a finite number of continuous derivatives. But it should be clear from the proofs that if Gaussians or multiquadrics are used, convergence orders are spectral again. To get the full approximation order, we use the idea of Hölder continuity, but we need only a weak form. Hence we define the space $C^k_\nu(\mathbb{R}^d)$ to be the space of all functions $f \in C^k(\mathbb{R}^d)$ such that their derivatives of order k satisfy $D^\alpha f(x) = \mathcal{O}(\|x\|_2^\nu)$ for $\|x\|_2 \to 0$.

The local version for error estimates on radial basis functions is taken from [11]. It can be formulated as

Theorem 3. *Suppose $\Phi \in C^k_\nu(\mathbb{R}^d)$ is conditionally positive definite of order m. Let $\tilde{\theta} \in (0, \pi/2)$ and $h > 0$ be given. Let $\tilde{\Omega} \subseteq \mathbb{R}^d$ satisfy an interior cone condition with angle $\tilde{\theta}$ and radius $\tilde{r} = C_s h$. Suppose further that $\tilde{X} = \{\tilde{x}_1, \ldots, \tilde{x}_M\} \subseteq \tilde{\Omega}$ satisfies $h_{\tilde{X},\tilde{\Omega}} \leq h$. Then there exist constants $C, C_0 > 0$ depending only on d, k, m, $\tilde{\theta}$, and Φ, but not on $h_{X,\Omega}$, such that*

$$\|D^\alpha f - D^\alpha s_{f,X}\|_{L_\infty(\tilde{\Omega})} \leq C h_{\tilde{X},\tilde{\Omega}}^{\frac{k+\nu}{2} - |\alpha|} |f|_{\mathcal{N}_\Phi(\Omega)}$$

for all $f \in \mathcal{N}_\Phi(\Omega)$ and all $|\alpha| \leq k/2$, provided that $C_s \geq C_0$ and $h \leq h_0$.

Now, let us come back to the partition of unity method. As mentioned before, we want to use the radial basis function interpolant as the local approximant. Hence, we set $X_j = X \cap \Omega_j$ and $s_j = s_{f,X_j}$, so that the global approximant becomes

$$s_f(x) = \sum_{j=1}^{M} s_{f,X_j}(x) w_j(x).$$

Since the partition of unity is given by compactly supported functions, and since we interpolate on each cell, it is clear that the resulting function s_f interpolates the data as well.

To use Theorem 3 in this context, we have to make some additional assumptions on the covering $\{\Omega_j\}$.

Definition 4. *Suppose $\Omega \subseteq \mathbb{R}^d$ is bounded and $X = \{x_1, \ldots, x_N\} \subseteq \Omega$ are given. An open and bounded covering $\{\Omega_j\}_{j=1}^M$ is called* regular for *(Ω, X) if the following properties are satisfied.*

1) *For every $x \in \Omega$, the number of cells Ω_j with $x \in \Omega_j$ is bounded by a global constant K.*
2) *There exists a constant $C_r > 0$ and an angle $\theta \in (0, \pi/2)$ such that every patch Ω_j satisfies an interior cone condition with angle θ and radius $r = C_r h_{X,\Omega}$.*
3) *There exists a constant $0 < C_{ae} < C_r$ such that $h_{X_j,\Omega_j} \leq C_{ae} h_{X,\Omega}$, where $X_j = X \cap \Omega_j$.*

This looks technical at first sight. But a closer look at each property shows that these requirements are more or less natural. For example, the first property is necessary for making sure that the outer sum in (3) is actually a sum over at most K summands. Since K is independent of N, while M is usually proportional to N, this is essential for not losing convergence orders. Moreover, it is crucial for an efficient evaluation of the global approximant that only a constant number of local approximants have to be evaluated. To this end, it also has to be possible to locate those K indices in constant time. The second and third property are important for employing our estimates on radial basis function interpolants, as we will see very soon. We will also need that C_r has to be reasonably larger than C_{ae}. This is in particular necessary to ensure that Ω_j contains enough points to allow a unique interpolant. Another consequence of these last two properties is that the union of all patches $\tilde{\Omega} = \cup_{j=1}^M \Omega_j$ cannot overlap Ω too much, i.e. there exists a constant $C_{ov} > 0$, such that

$$\operatorname{dist}(\tilde{\Omega}, \Omega) := \sup_{\tilde{x} \in \tilde{\Omega}} \inf_{x \in \Omega} \|\tilde{x} - x\|_2 \leq C_{ov} h_{X,\Omega}.$$

We are now able to state the main result concerning the approximation properties of a partition of unity method built on radial basis

functions. The bottom line is that it inherits the same approximation order as the global method.

Theorem 5. *Suppose $\Omega \subseteq \mathbb{R}^d$ is open and bounded, and that $X = \{x_1, \dots, x_N\} \subseteq \Omega$. Let $\Phi \in C_\nu^k(\mathbb{R}^d)$ be conditionally positive definite of order m. Let $\{\Omega_j\}$ be a regular covering for (Ω, X) and let $\{w_j\}$ be k-stable for $\{\Omega_j\}$. Let C_0 denote the constant from Theorem 3. If $C_r \geq C_0 C_{ae}$, then the error between $f \in \mathcal{N}_\Phi(\Omega)$ and its partition of unity interpolant $s_f(x) = \sum_j s_{f,X_j}(x) w_j(x)$ can be bounded by*

$$|D^\alpha f(x) - D^\alpha s_f(x)| \leq C h_{X,\Omega}^{\frac{k+\nu}{2}-|\alpha|} |f|_{\mathcal{N}_\Phi(\Omega)},$$

for all $x \in \Omega$ and all $|\alpha| \leq k/2$.

Proof: The function f has a norm preserving extension $Ef \in \mathcal{N}_\Phi(\mathbb{R}^d)$. Moreover, it is known that the restriction $f_j = Ef|\Omega_j$ satisfies $|f_j|_{\mathcal{N}_\Phi(\Omega_j)} \leq |Ef|_{\mathcal{N}_\Phi(\mathbb{R}^d)} = |f|_{\mathcal{N}_\Phi(\Omega)}$. Hence, if we denote all these functions by f again, we have $|f|_{\mathcal{N}_\Phi(\Omega_j)} \leq |f|_{\mathcal{N}_\Phi(\Omega)}$.

Furthermore, we find estimates for the local interpolants by Theorem 3. Setting $h = C_{ae} h_{X,\Omega}$, $\tilde{\theta} = \theta$, $\tilde{\Omega} = \Omega_j$, and $\tilde{X} = X_j$, we see that $h_{X_j,\Omega_j} \leq h$ and that Ω_j satisfies an interior cone condition with angle θ and radius $r = \frac{C_r}{C_{ae}} h$. Since $C_r > C_0 C_{ae}$, Theorem 3 yields

$$|D^\alpha f(x) - D^\alpha s_{f,X_j}(x)| \leq C h_{X_j,\Omega_j}^{\frac{k+\nu}{2}-|\alpha|} |f|_{\mathcal{N}_\Phi(\Omega_j)} \leq C h_{X,\Omega}^{\frac{k+\nu}{2}-|\alpha|} |f|_{\mathcal{N}_\Phi(\Omega)}$$

for $x \in \Omega_j$ and $|\alpha| \leq k/2$ with C being independent of j. To apply (3) we need two more ingredients. Since every patch Ω_j satisfies an interior cone condition with radius $C_r h_{X,\Omega}$, we have $\delta_j = \operatorname{diam}(\Omega_j) \geq C_r h_{X,\Omega}$. Moreover, every $x \in \Omega$ is contained in at most K patches Ω_j. Hence, the error bound (3) leads to

$$\begin{aligned}
|D^\alpha f(x) - D^\alpha s_f(x)| &\leq \sum_{j \in I(x)} \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} C_{\alpha-\beta} \delta_j^{|\beta|-|\alpha|} \varepsilon_j(\beta) \\
&\leq K \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} C_{\alpha-\beta} C_r^{|\beta|-|\alpha|} h_{X,\Omega}^{|\beta|-|\alpha|} C h_{X,\Omega}^{\frac{k+\nu}{2}-|\beta|} |f|_{\mathcal{N}_\Phi(\Omega)} \\
&= C h_{X,\Omega}^{\frac{k+\nu}{2}-|\alpha|} |f|_{\mathcal{N}_\Phi(\Omega)}
\end{aligned}$$

for all $x \in \Omega$ and all $|\alpha| \leq k/2$. $\square$

§4. Data Structures for the Centers

To take advantage of the locality of our method, we have to show that this diminishes the computational complexity substantially. To this end we have to investigate data structures for both the covering Ω_j and the data sites X. Since this is a dual problem, we concentrate here on data structures for the data sites.

To be useful for our method, the data structure for X has to be able to answer range query problems efficiently. A range query problem is the following. Suppose we are given a set of points X in $\mathbb{R}^d$ and a range $R \subseteq \mathbb{R}^d$. Then, the task is to report all points $x_j \in X$ with $x_j \in R$. Fortunately, there exist several good data structures to handle this problem, in particular if R is an axis-parallel box. Examples are *kd*-trees, *bd*-trees, and range trees. We will give some information about the latter and about a very simple data structure, called fixed grid, which is very favourable in case of quasi-uniform data. Details on *kd*- and *bd*-trees can be found in [9] and [1], respectively.

For the fixed grid method, one first searches the bounding box BB of the given data $X = \{x_1, \ldots, x_N\}$. Then, a fine grid is defined on BB. This grid consists of $\lfloor N^{1/d} \rfloor^d$ axis-parallel boxes, also called cells. For each cell we keep a list that contains the indices of the points in that specific cell. If the cells have side length k, then the indices of the cell containing a specific point can efficiently be found by multiplying the components of the point by $1/k$ and taking integer parts, which is often faster than dividing by k.

After the data structure has been built, a range query can then simply be answered in two steps. In the first step we find all cells that have common points with the query region R. Then, we test the data points in these cells explicitly. To execute both steps, we will assume that it is possible in constant time to decide whether R and a cell are disjoint or not, and whether a specific point belongs to R or not. This is obviously true if R is a box or a ball or something similar.

As said before, the fixed grid method is favourable in case of quasi-uniform data. Quasi-uniformity means that the fill distance is asymptotically proportional to the separation distance $q_X = \min_{j \neq k} \|x_j - x_k\|_2$. If this is true, it is easy to see that both are also asymptotically behaving like $N^{-1/d}$. Moreover, it follows almost immediately that the number of points in each cell of the fixed grid is bounded by a fixed constant.

Proposition 6. *If the data sites X are quasi-uniform, the fixed grid data structure can be built in $\mathcal{O}(N)$ time and space. If R is a range of diameter proportional to $N^{-1/d}$, the range query problem for R can be answered in constant time.*

To apply this to our partition of unity method, we need an additional assumption on the locality of the covering. We will say that $\{\Omega_j\}$ is local

if there is a constant C_{ba} such that $\operatorname{diam}(\Omega_j) \leq C_{ba} h_{X,\Omega}$. This means that the cells are not too big. Hence, if $\{\Omega_j\}$ is local and regular, the size of each cell is proportional to $h_{X,\Omega}$. In case of a quasi-uniform data set this means that the number of cells M is proportional to the number of data sites N.

Theorem 7. *Let $X = \{x_1, \ldots, x_N\} \subseteq \Omega$ be quasi-uniform and $\{\Omega_j\}$ a regular and local covering for (Ω, X). Suppose $\{\Omega_j\}$ can be built in $\mathcal{O}(N)$ time and $\mathcal{O}(N)$ space and for every $x \in \Omega$ all K or fewer patches Ω_j with $x \in \Omega_j$ can be reported in constant time. Then the partition of unity method based on radial basis functions can be implemented in $\mathcal{O}(N)$ space with $\mathcal{O}(N)$ time needed for the prepocessing step. Furthermore, each evaluation of the global interpolant needs $\mathcal{O}(1)$ time.*

Proof: If we use the fixed grid to build our data structure for X, we know that this can be done in $\mathcal{O}(N)$ time and space. Since the number of centers in each patch is bounded by a constant, we need constant space and constant time for each patch to solve the local interpolation problem. Furthermore, the points in each patch can be reported in constant time. Since the number M of patches Ω_j is bounded by $\mathcal{O}(N)$, this adds up to $\mathcal{O}(N)$ space and time for solving all local problems in all patches. By assumption we can determine $I(x) = \{j : x \in \Omega_j\}$ in constant time, and the cardinality of $I(x)$ is also bounded by a constant. Thus, we have to add up a constant number of local interpolants to get the value of the global interpolant. This can be done in constant time. □

In case of data that is not quasi-uniform, the number of points in the cells may vary dramatically. Hence, the data structure might need more than linear space. In the worst case, all N points are concentrated in one of the cells. In this situation, the data structure needs $\mathcal{O}(N^2)$ space. Nonetheless, it still can be built in $\mathcal{O}(N)$ time. The time for a range query problem can also vary.

If R is always an axis-parallel box, another data structure is better suited for highly irregular data sites: the range tree. In one dimension a range tree is simply a balanced binary search tree, where all data sites are stored in the leafs and the inner nodes contain splitting information.

A d-dimensional range tree can be built in the following way. Firstly, a balanced binary search tree is built on the first coordinate of the points. This is the first-level or main tree. To each node ν of this tree we associate a canonical set of points $P(\nu)$ containing all the points stored in the leaves of the subtree rooted at ν. For each node ν we construct an associated data structure $\mathcal{T}_{\text{assoc}}(\nu)$. This second-level tree is a $(d-1)$-dimensional range tree for the points in $P(\nu)$ restricted to their last $d-1$ coordinates. This $(d-1)$-dimensional range tree is constructed recursively. The recursion stops when we have reached the last coordinate. Here, the points are stored in a one-dimensional range tree.

The query algorithm now works on each coordinate. Given a range $R = [\alpha_1, \beta_1] \times \ldots \times [\alpha_d, \beta_d]$, it selects $\mathcal{O}(\log N)$ canonical subsets whose union contains the points whose first coordinates lie in the range $[\alpha_1, \beta_1]$. On these canonical subsets it performs a range query on the second level tree. This gives subsets of points whose first and second coordinates lie in the range. Continuing recursively, this search procedure results in the points in the range.

A more thorough description of range trees and the following result can be found in [6].

Proposition 8. *Given N points in $\mathbb{R}^d$, a range tree can be built using $\mathcal{O}(N \log^{d-1} N)$ space and time. It is possible to report the points that lie in a rectangular query range in $\mathcal{O}(\log^d N + k)$ time, where k is the number of reported points.*

The query time can be reduced to $\mathcal{O}(\log^{d-1} N + k)$ if a technique called fractional cascading is employed.

§5. Stability

So far we know that our method inherits the convergence properties of the global radial basis function interpolation. We also know that in case of regular data it can be implemented efficiently leading to an interpolant that can be evaluated in constant time.

The next question that comes up naturally is stability. Unfortunately, the condition numbers of the interpolation matrices depend more on the separation distance than on the number of points. For most of the cells, the local separation distance q_{X_j} will be of the same size (or smallness) as the global separation distance q_X. Hence, our method seems to be as stable as the global one. Fortunately, this is not a problem in case of basis functions of moderate smoothness. For example, the compactly supported basis functions used as global functions of smoothness up to order 4 work nicely, and a direct solver for the local problems can be applied.

But for certain classes of basis functions including thin plate splines, we have a different possibility. Instead of using the basis function itself to form the interpolation matrix, a modified kernel can be used. If the basis function $\Phi(x, y)$ satisfies a homogeneous condition, one can use the kernel

$$\begin{aligned}\kappa(x, y) =& \Phi(x, y) - \sum_{j=1}^{Q} p_j(x)\Phi(\xi_j, y) - \sum_{k=1}^{Q} p_k(y)\Phi(x, \xi_k) \\ &+ \sum_{j,k=1}^{Q} p_j(x)p_k(y)\Phi(\xi_j, \xi_k).\end{aligned}$$

Here Q is the dimension of the linear space of d-variate polynomials of degree less than m, if Φ is conditionally positive definite of order m, the

points $\xi_1, \ldots, \xi_Q$ are different from those in X, and $\{p_j\}$ is a polynomial Lagrangian basis with respect to these points. For this kernel it can be shown that the condition number of the involved interpolation matrices is rather independent of the separation distance, but depends only on the number of points. Details can be found in [5].

§6. A Numerical Example

To demonstrate the numerical behaviour of our method, we give one "best case" example. The data sites are a regular $(2^j + 1)^2$ grid on $\Omega = [0, 1]^2$ for different values of j. The function values come from Franke's test function, which is a sum of four exponential terms. For the covering we have chosen overlapping axis-parallel boxes on a fine grid with a certain offset such that these boxes overlap each other and also the boundary of the domain. The partition of unity was formed by a Shepard function. The radial basis function was given by $\phi(r) = (1 - r)_+^6(35r^2 + 18r + 3)$. The theory for this basis function allows us to expect a convergence order of at least 2.5. The error was measured on a fine 2000×2000 grid. The results are

j	$N = (2^j + 1)^2$	L_∞	Order
3	81	4.51084×10^{-2}	
4	289	3.32996×10^{-3}	3.76
5	1089	2.81404×10^{-4}	3.56
6	4225	3.56702×10^{-5}	2.98
7	16641	4.49974×10^{-6}	2.99
8	66049	5.56833×10^{-7}	3.01
9	263169	6.83149×10^{-8}	3.03

References

1. Arya, S., D. Mount, N. Netanyahu, R. Silverman, and A. Wu, An optimal algorithm for approximate nearest neighbor searching, 5th Ann. ACM-SIAM Symposium on Discrete Algorithms, 1994, 573–582.
2. Beatson, R. K., and G. N. Newsam, Fast evaluation of radial basis functions: I, Comput. Math. Appl. **24** (1992), 7–19.
3. Beatson, R. K., G. Goodsell, and M. J. D. Powell, On multigrid techniques for thin plate spline interpolation in two dimensions, Lectures in Applied Mathematics **32** (1996), 77–97.
4. Beatson, R. K., and G. N. Newsam, Fast evaluation of radial basis functions: moment-based methods, SIAM J. Sci. Comp. **19** (1998), 1428–1449.

5. Beatson, R. K., S. Billings, and W. A. Light, Fast Solution of the radial basis function interpolation equations: domain decomposition methods, SIAM J. Sci. Comp. **22** (2000), 1717–1740.
6. Berg, M. de, M. van Kreveld, M. Overmars, and O. Schwarzkopf, *Computational Geometry*, Springer, New York, 1991.
7. Buhmann, M. D., Radial basis functions, in *Acta Numerica 2000*, Cambridge University Press, Cambridge, 2000, 1–38.
8. Faul, A. and M. J. D. Powell, Proof of convergence of an iterative technique for thin plate spline interpolation in two dimensions, Advances in Comp. Math. **11** (1999), 183–192.
9. Friedman, J., J. L. Bentley, and R. A. Finkel, An Algorithm for finding best matches in logarithmic expected time, ACM Transactions on Mathematical Software **3** (1977), 209–226.
10. Schaback, R., Native Hilbert spaces for radial basis functions I, in *New Developments in Approximation Theory*, Müller, M. W. et al. (eds.), Birkhäuser Verlag, Basel, 1999, 255–282.
11. Wu, Z., and R. Schaback, Local error estimates for radial basis function interpolation of scattered data, IMA J. of Num. Analysis **13** (1993), 13–27.

Holger Wendland
Institut für Numerische und Angewandte Mathematik
Universität Göttingen
Lotzestr. 16-18
37083 Göttingen
wendland@math.uni-goettingen.de