Approximation to Scattered Data by Radial Basis Function

By

Jungho Yoon

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

(MATHEMATICS)

at the

UNIVERSITY OF WISCONSIN – MADISON

1998

Abstract

A new multivariate approximation scheme to scattered data on arbitrary bounded domains in $\mathbb{R}^d$ is developed. The approximant is selected from a space spanned (essentially) by corresponding translates of the 'shifted' thin-plate spline ('essentially', since the space is augmented by certain functions in order to eliminate boundary effects). The approximation scheme is derived from the optimal approximation scheme of de Boor and Ron on uniform grids, using the conversion method to the scattered centers developed by Dyn and Ron, and going to the limit with that scheme as the mesh size of the uniform grid tends to zero.

The scheme is constructed in two steps. In the first one, the information on the scattered centers is 'interpolated' using an algorithm of David Levin. The output of the first step is used as input in the above scheme. The scheme is shown to provide spectral approximation order, i.e., approximation order that depends only on the smoothness of the approximand. It applies to noisy data as well as to noiseless data, but its main advantage seems to be in the former case. We suggest an algorithm for the new approximation scheme with a detailed description (in a MATLAB-like program). Some numerical examples are presented, as well as comparisons with Wahba's thin-plate smoothing spline approximation.

Acknowledgements

I wish to express my sincere gratitude to my Advisor Amos Ron. He has given me much valuable advice and insight in my research. Thank him for his support, constructive criticism, patience, and encouragement. A special thanks goes to Prof. Carl de Boor. Whenever I asked him, he was always willing to help me. Thank him for his good advice, support, and interest in my work. In particular, I appreciate the other committee members, professors G. Wahba, D. Shea and A. Seeger. I also would like to express my appreciation to the professors Suk-Geun Hwang, Hong-Oh Kim, Sang-Dong Kim, Sang-Hun Lee, and Young-Soo Park for their help and encouragement during the course of my research.

I want to thank my mother. Her unconditional love and prayers have been essential to this task. I am indebted to my parents-in-law. I always remember her love, prayer and warm heart. I am so grateful to my brothers (in-law) and sisters who continued support and care with love during this long and arduous process. I express also my deep appreciation to numerous people who gave me emotional support. They are Pastors Chi-Dae Kim, Pil-Soo Lee, and my friends Jong-Sik Lee, Jae-Gon Lee, and Sun-Hee Yoo.

I wish to acknowledge my wife Yoon Kyung whose support, assistance, encouragement, and understanding have been my energy, especially, whenever I was discouraged. Thank you, Yoon Kyung, with all my heart. I love you and will never forget these years that have drawn us closer together.

Contents

Abstract **i**

Acknowledgements **ii**

1 Introduction **1**

1.1 Statement of the Problem 1

1.2 An Outline of Our Approach 8

2 Approximation Scheme into the Span of Scattered Shifts **17**

2.1 The Pseudo-Shift $\phi(\cdot, t)$ 17

2.2 Approximation Scheme 22

2.3 Approximation Power of the Scheme R_Ξ 40

3 Approximation by Scattered Data on a Bounded Domain **46**

3.1 Approximation in the Interior of Bounded Domains 46

3.2 Error Estimates near the Boundary 55

3.3 Smoothing Noisy Data 59

4 Computational Aspects **62**

4.1 On the Coefficients of Pseudo-Shift $\phi(\cdot, t)$ 62

4.1.1 Preliminary 62

4.1.2 Factorization 64

4.1.3 Algorithm Details : The Coefficients $(A(t, \xi))_{\xi \in \Xi}$ for $\phi_c(\cdot, t)$. 71

4.2 A General Algorithm for the Scheme $\tilde{R}_\Xi$ 74

4.2.1 Formulation for the Construction of $\tilde{R}_{\Xi} f$ 74

4.2.2 A General Algorithm for $\tilde{R}_{\Xi} f$ 79

4.3 Numerical Results 81

Bibliography **100**

Chapter 1

Introduction

1.1 Statement of the Problem

Given $\Xi \subset \Omega \subset \mathbb{R}^d$ and values $f|_\Xi$ (possibly contaminated) of some unknown function f, our objective is to construct a function $s : \Omega \to \mathbb{R}$ such that, in some sense, s approximates f:

$$s \approx f, \quad \text{on} \quad \Omega.$$

This problem is usually referred to as **scattered data approximation** and has many important applications. There are cases where the domain Ω is a rectangle and the points Ξ are uniformly gridded. There are, however, many other practical instances where Ω is of irregular shape and/or where the points Ξ are irregularly distributed on Ω. A large number of ideas have been proposed for the solution of this problem. In order for an approximation scheme to scattered data to be useful in practice, certain requirements should be satisfied:

(A) **Numerical Stability** : An approximation scheme should be local on the sense that the contribution to the approximant's value at a point x by the data value at $\xi \in \Xi$ decreases (fast!) as the distance between x and ξ increases. At the same time, many of the approximation methods use, and for good reasons, basis functions that are neither compactly supported, nor decay at ∞. Hence, in order to circumvent this initial instability, a 'localization process' is necessary. The localness of the scheme also ensures that 'boundary effects' do not spill over into the interior of the domain.

(B) **Boundary Effects** : Approximation near the boundary is a difficult problem. Because the data is usually given only inside the domain, the boundary effect is very serious, and eventually, one must lose some order of accuracy. Even worse, some of the pertinent radial basis functions have no local nature, hence special care is necessary for the approximation near the boundary.

(C) **Approximation Power** : It is basic to require that the approximant s approximates f better as the point set Ξ becomes dense in Ω. In most cases the approximation power is quantified by the asymptotic rate at which the error decays. The problem is how we can get the largest possible asymptotic rate when the function f is smooth.

(D) **Noise** : Noisy data arises in many scientific applications according to the model

$$y_\xi = f(\xi) + \epsilon_\xi, \quad \xi \in \Xi,$$

where, for example, ϵ_ξ's are independent noise with mean 0 and with (known or unknown) variance σ^2. In this case, the approximation scheme should have a smoothing effect.

One of the well known approaches for scattered data approximation is the use of piecewise-polynomials. In this case Ω has to be partitioned into suitable regions, different polynomials are employed on the different regions separately, and usually the pieces have to be joined in a smooth way. In the multivariate case, however, this problem is computationally expensive. For example, the evaluation of an approximant at a given point requires one to identify the polynomial piece relevant to the point.

Other techniques are based on forming suitable linear combinations of certain radially symmetric basis function. In particular, one may employ the translates along Ξ of one fixed such function ϕ. This approximation method has the general form

$$s(x) := \sum_{\xi\in\Xi} c_\xi\, \phi(x-\xi), \qquad x \in \Omega.$$

The set of scattered points Ξ in $\mathbb{R}^d$ by which a radial basis function ϕ is shifted is referred to as a set of "centers". The choices of ϕ common in the literature include:

$$\begin{aligned}
&\phi(x) = \exp(-c|x|^2), && c > 0, && \text{(Gaussian)}\\
&\phi(x) = |x|^\lambda \log|x|, && \lambda \in 2\mathbb{Z}_+,\ d \text{ even} && (\text{thin} - \text{plate spline})\\
&\phi(x) = (|x|^2 + c^2)^{\lambda/2}, && \lambda \in \mathbb{Z}_+,\ \lambda,\ d \text{ odd},\ c > 0, && \text{(multiquadric)}\\
&\phi(x) = (|x|^2 + c^2)^{\lambda/2}, && -d < \lambda < 0,\ \lambda \in \mathbb{Z},\ c > 0. && \text{(inverse multiquadric)}
\end{aligned}$$

In view of the discussion so far, one might wonder why compactly supported functions ϕ are not in this list (e.g., box splines). The answer is that, in general, since such basis function does not respect the geometry aspect of information, it can not yield a good approximation scheme. We refer the reader to [R2] for more details. Radial basis functions, in contrast, if handled correctly, are known to be suitable for any arrangement of the data locations. The initial approach to scattered data using radial basis functions has been focused on interpolation at the scattered points $\Xi \subset \mathbb{R}^d$. The specific interpolation scheme has the form

$$\begin{aligned}
&s(x) := \sum_{\xi\in\Xi} a_\xi \phi(x-\xi) + p_m(x),\\
&p_m \in \Pi_m, \quad \sum_{\xi\in\Xi} a_\xi p(\xi) = 0, \quad p \in \Pi_m.
\end{aligned} \tag{1.1}$$

The general conditions on ϕ that ensure the existence and a uniqueness of a solution of the above equations have been given by Micchelli [M]. The reader is also referred to the work of Madych and Nelson [MN1,2]: there, the approach of reproducing kernel Hilbert spaces is used. That approach is suitable for the approximation of functions that lie in the underlying Hilbert space (see also [MN3] and [WS]). More recently, M. J. Johnson [J2] established an asymptotic upper bound on the approximation order on the unit ball Ω of $\mathbb{R}^d$ for a basis function ϕ of the form $\phi = |\cdot|^{\lambda/2}$ for d, λ odd, and $\phi = |\cdot|^{\lambda/2}\log|\cdot|$ for d, λ even. Non-interpolation approximation methods were introduced in [BeP] and [BeD] for the case of a univariate multiquadric basis function.

Interpolation by translates of suitable radial basis functions is certainly an important approach towards solving the scattered data problem. However, it carries its own disadvantages. For example, for a large class of basis functions (including multiquadric and inverse multiquadric), the existing theory guarantees the interpolant to approximate well only for a very small class of approximands (see [MN2]). The approximands need to be extremely smooth for an effective error analysis. Another disadvantage of the interpolation method is that, with the increase in the number of centers, one needs to solve a large linear system which is very ill-conditioned. Most importantly, when the given data are contaminated, the interpolation method can not be used. All in all, there is an overwhelming need for approximation methods other than interpolation. A main concern of this study is to provide a new method for solving the scattered data problem, a method which is particularly effective for noisy data.

When considering approximation schemes for discrete data, one observes that there already exist many successful results that address that problem in the case

$\Xi = \mathbb{Z}^d$. In contrast less is known for the general case of Ξ in $\mathbb{R}^d$. Recently, Buhmann, Dyn and Levin [BuDL] were among the first to construct a non-interpolatory approximation scheme for infinitely many scattered centers and to analize its approximation power. Dyn and Ron [DR] then generalized the results of [BuDL]. In both papers, quasi-interpolation from radial basis function space with infinitely many centers Ξ were studied and both realized in the scattered case the same asymptotic approximation orders that were known on uniform grids. In particular, N. Dyn and A. Ron provide a general tool that allows us to convert *any* known approximation scheme on uniform grids to non-uniform grid, while preserving (to the extent that this is possible) the approximation orders known in the former case. Their results, however, requires certain properties of the basis function; see a further discussion in the sequel. The initial objective of this study was to check the possible practical value of the conversion methods of [DR]. The outcome, however, is far more reaching:

(1) The approximation scheme that is developed and analyzed here, while based on the general ideas of [DR], is intrinsically 'scattered', i.e., employs directly the scattered values of the approximand f, and the scattered shifts of the basis function ϕ.

(2) The potential numerical instability in the scheme is overcome by deriving a computationally stable 'local' algorithm for the computation and evaluation of the approximant.

(3) The scheme is adjusted to deal properly with bounded domains. This is done by adding a 'predictor step' to the algorithm.

(4) A MATLAB code was written, and initial numerical tests reveal that the

algorithm gives results comparable to, or better than, the state-of-art method for both clean data and noisy data.

(5) The algorithm is non-stationary while the method suggested in [DR] is stationary. (We explain these notions in the sequel.)

The following notations are used throughout this thesis. For a function $\phi : \mathbb{R}^d \to \mathbb{R}$ and $\Xi \subset \mathbb{R}^d$, we define

$$S_\Xi(\phi) := \text{closure } S_0(\phi)$$

under the topology of uniform convergence of compact sets, with

$$S_0(\phi) := \text{span}\{\phi(\cdot - \xi) : \xi \in \Xi\}.$$

Several function norms will be used. The L_∞-norm is used as the default norm, i.e.,

$$\|f\| := \|f\|_\infty := \|f\|_{L_\infty(\mathbb{R}^d)},$$

and the L_1-norm is denoted as

$$\|f\|_1 := \|f\|_{L_1(\mathbb{R}^d)}.$$

For $x = (x_1, \cdots, x_d)$ in $\mathbb{R}^d$, $|x|$ stands for its Euclidean norm:

$$|x| := \sqrt{x_1^2 + x_2^2 + \cdots + x_d^2},$$

and, for $x, y \in \mathbb{R}^d$, we use the abbreviations

$$x \cdot y := x_1 y_1 + \cdots + x_d y_d$$

and

$$[x \,.\,.\, y] := \{(1-t)x + ty \, : \, 0 \le t \le 1\};$$

also used is the notation

$$B_\eta := \{x \in \mathbb{R}^d : |x| < \eta\}.$$

Given $\alpha \in \mathbb{Z}_+^d := \{\beta \in \mathbb{Z}^d \,:\, \beta \geq 0\}$, we set

$$\alpha! := \prod_{j=1}^{d} \alpha_j, \quad D^\alpha := \frac{\partial^{|\alpha|_1}}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}, \quad |\alpha|_1 := \sum_{k=1}^{d} \alpha_k.$$

Finally, Π_k stands for the space of all polynomials of degree $\leq k$ in d variables and $C(\mathbb{R}^d)$ for the space of all continuous functions $f : \mathbb{R}^d \to \mathbb{C}$ (or $\mathbb{R}$) equipped with the topology of uniform convergence on compact sets.

For a given f continuous on $\mathbb{R}^d \setminus 0$, we say that f **has a singularity of exact order** k at the origin if there exist some constants $c_1,\ c_2 > 0$ such that $c_1 \leq |\cdot|^k |f| \leq c_2$ in some punctured neighborhood of the origin. The semi-discrete convolution is defined formally by

$$\phi *' : c \to \phi *' c := \sum_{\alpha \in \mathbb{Z}^d} c(\alpha)\phi(\cdot - \alpha).$$

The Fourier transform of an absolutely integrable function f is given by

$$\hat{f}(\theta) := \int_{\mathbb{R}^d} f(t)\, e_{-\theta}(t) dt, \qquad e_\theta : x \longmapsto e^{i\theta \cdot x}.$$

Also, for a function f absolutely integrable, we use the notation

$$f^\vee(\theta) := (2\pi)^{-d} \int_{\mathbb{R}^d} f(t)\, e_\theta(t) dt$$

for the inverse Fourier transform. We assume the reader to be familiar with the usual properties of Fourier transform. In particular, the Fourier transform can be uniquely extended to the space of tempered distributions on $\mathbb{R}^d$.

1.2 An Outline of Our Approach

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a smooth function, and let ϕ be a 'basis' function defined on $\mathbb{R}^d$. Suppose that we look for an approximant for f in the span of $\phi(\cdot - t)$, $t \in \mathbb{R}^d$. Then we may try to find the the exact solution f^* of the convolution equation

$$\phi * f^* = f.$$

However, this equation is not always solvable. In some cases, in order to make sure that there is a solution of this equation, the function f needs to be very smooth in a way that depends on the basis function ϕ. For example, if ϕ is the multiquadric, f should be a special type of a C^∞-function (e.g., a band-limited function). Thus, rather than solving the equation exactly, we approximate first the function f by

$$m^\vee * f$$

where m is a suitable cut-off function. Then, after substituting $m^\vee * f$ for f in above convolution equation, we look for a solution f^* of

$$\phi * f^* = m^\vee * f. \tag{1.2}$$

Under suitable assumptions on $\hat{\phi}$ and m, we find a solution in the expected form

$$f^* := \left(\frac{m\hat{f}}{\hat{\phi}}\right)^\vee.$$

Our real intent, however, is to approximate the function f from the space

$$S_\Xi(\phi) = \operatorname{span}\{\phi(\cdot - \xi) : \xi \in \Xi\}$$

since it will provide a 'local resolution of the approximand' based on the local density of the data. Thus, we are seeking an approximant of the form

$$s = \sum_{\xi \in \Xi} c_\xi(f)\phi(\cdot - \xi)$$

that approximates $m^\vee * f$ in some sense. Also, we need to keep in mind that only $f|_\Xi$ is actually available to us.

We say that the approximation maps $(L_h)_h$ with L_h mapping into the space $S_{h\Xi}(\phi)$ **provides approximation order** $k > 0$ if, for every admissible f,

$$\|f - L_h f\| = O(h^k)$$

as h tends to 0.

We assume throughout this study that the function ϕ, when considered as a tempered distribution, has a Fourier transform $\hat{\phi}$ that coincides on $\mathbb{R}^d \setminus 0$ with some continuous function while having a certain type of singularity (necessarily of finite order) at the origin; especially, we assume that $\hat{\phi} \neq 0$ on $\mathbb{R}^d \setminus 0$.

The general approach of this study is to obtain an approximation scheme from a space spanned by scattered shifts of ϕ by employing the following two step emthod. In the first step we assume that the approximand f is fully available (e.g., thus, we compute $\hat{f}$). During this step, we consider approximation on $\mathbb{R}^d$ by $S_\Xi(\phi)$ for an infinite set Ξ. The actual approximation scheme considered in this study is of the form

$$R_\Xi : f \mapsto \int_{\mathbb{R}^d} \phi(\cdot, t)\, f^*(t)\, dt \tag{1.3}$$

where f^* is the exact solution of (1.2) and $\phi(\cdot, t)$ is a kernel of the form

$$\phi(\cdot, t) := \sum_{\xi \in \Xi} A(t, \xi) \phi(\cdot - \xi), \tag{1.4}$$

where $\phi(\cdot, t)$ approximating $\phi(\cdot - t)$ in some sense. In this introductory stage, we assume that the sequence $(A(t, \xi))_{\xi \in \Xi}$ is finitely supported (for each $t \in \mathbb{R}^d$).

Further properties of the map $(t,\xi)\mapsto A(t,\xi)$ that are essential for the success of (1.3) will be analyzed later.

In the second, more practical, step, we derive a scheme for approximating functions on a bounded domain under the assumption that the (only known) data are the possibly contaminated values of f at Ξ, with Ξ a finite subset of the interior of the domain. The approximant is selected from a space spanned 'essentially' by the corresponding translates $\phi(\cdot-\xi)$, $\xi\in\Xi$, of the basis function ('essentially', since the space $S_\Xi(\phi)$ is augmented by other functions in order to eliminate boundary effects).

Let us discuss now in more detail the scheme alluded to in (1.3). The main question is how to find the coefficients $(A(t,\xi))_{\xi\in\Xi}$ for the kernel $\phi(\cdot,t)$, $t\in\mathbb{R}^d$, in (1.4). The construction of this scheme is based on the general tool developed in [DR] for converting an approximation scheme on a uniform grid to a non-uniform grid. The conversion method in [DR] starts with a known approximation scheme R of the form

$$R: f\to\sum_{\alpha\in\mathbb{Z}^d}\psi(\cdot-\alpha)\Lambda(f)(\alpha)$$

with Λ a bounded operator from $L_\infty(\mathbb{R}^d)\cap C(\mathbb{R}^d)$ into itself, and with the function ψ being a linear combination of shifts of the original function ϕ,

$$\psi := \phi *' \mu, \tag{1.5}$$

where $\mu:\mathbb{Z}^d\to\mathbb{R}$ decays fast around ∞ such that the sum in (1.5) converges. The function $\psi(\cdot-\alpha)$ is supposed to decay at some polynomial rate such that, at a minimum,

$$\sum_{\alpha\in\mathbb{Z}^d}|\psi(\cdot-\alpha)|\in L_\infty(\mathbb{R}^d).$$

Then one chooses for each shift $\psi(\cdot-\alpha)$ an approximation $\psi(\cdot,\alpha)$ from the space $S_\Xi(\phi)$, and by substituting $\psi(\cdot,\alpha)$ for $\psi(\cdot-\alpha)$, one obtains an approximation of the form

$$R_A : f \to \sum_{\alpha\in\mathbb{Z}^d} \psi(\cdot,\alpha)\Lambda(f)(\alpha)$$

with Λ as before. The function $\psi(\cdot,\alpha)$ thus lies in $S_\Xi(\phi)$. It is also assumed to satisfy

$$\sum_{\alpha\in\mathbb{Z}^d} |\psi(\cdot,\alpha)| \in L_\infty(\mathbb{R}^d).$$

It follows that the approximation scheme R_A is a bounded map from $L_\infty(\mathbb{R}^d)\cap C(\mathbb{R}^d)$ into $S_\Xi(\phi)$. The actual construction of $\psi(\cdot,\alpha)$ is done as follows. We first approximate each $\phi(\cdot-\alpha)$ by

$$\phi(\cdot,\alpha) := \sum_{\xi\in\Xi} A(\alpha,\xi)\phi(\cdot-\xi)$$

as in (1.4), and then define $\psi(\cdot,\alpha)$ by

$$\psi(\cdot,\alpha) := \sum_{\beta\in\mathbb{Z}^d} \mu(\beta-\alpha)\phi(\cdot,\beta), \tag{1.6}$$

which is a localization of the function $\phi(\cdot,\alpha)$. Under some suitable conditions of ϕ, it is shown in [DR] that the scheme R_A provides the same approximation order as R does, provided the scheme is stationary.

As a matter of fact, this method provides a general tool for deriving a scheme for a scattered set Ξ from a scheme on a uniform mesh, instead of approximating the function f directly from the space $S_\Xi(\phi)$. Since the present state-of-art in the area of approximation on uniform grids is quite satisfactory, it gives hope for finding new approaches into the unyielding scattered case. However, when we convert the gridded scheme to the non-uniform case, we are faced with the issue of choosing the density of the uniform grid ($h\mathbb{Z}^d$) corresponding the scattered center set Ξ: A method for selecting the density h associated with a given Ξ is not given in [DR]. At the outset of this work we found that the conversion method of [DR] converges to a limit as $h \to 0$; i.e., if $(R_h : C(\mathbb{R}^d) \cap L_\infty(\mathbb{R}^d) \to S_{h\mathbb{Z}^d}(\phi))_{h>0}$ are scaled versions of some uniform grid approximation R, and if $(R_{A,h})_h$ are counterparts of $(R_h)_h$ obtained by the [DR]-method, then $R_{A,h} \to R_\Xi$ as h tends to 0, with R_Ξ a new approximation scheme. Consequently, we obtain an approximation scheme in (1.3) that is independent of any uniform grid issue. Furthermore, we will see that this scheme provides spectral approximation order, (i.e., approximation order that depends only on the smoothness of the function f we approximate).

In order to discuss approximation orders, we measure the density of Ξ by

$$\bar{h} := \sup_{x\in\mathbb{R}^d} \inf_{\xi\in\Xi} |x-\xi|. \tag{1.7}$$

We also choose the basis function to be the 'shifted thin-plate spline'

$$\phi_c(x) := \begin{cases} (|x|^2+c^2)^{\lambda/2}, & \lambda \in \mathbb{Z}_+^d,\ \lambda,\ d \text{ odd}, \\ (|x|^2+c^2)^{\lambda/2}\log(|x|^2+c^2)^{1/2}, & \lambda \in \mathbb{Z}_+^d,\ \lambda,\ d \text{ even}, \end{cases} \tag{1.8}$$

and make the following further assumptions:

(a) With f^* the exact solution of $\phi_c * f^* = m^\vee * f$, the approximation scheme is of the form

$$R_\Xi : f \mapsto \int_{\mathbb{R}^d} \phi_c(\cdot,t)\, f^*(t)\, dt$$

with $\phi_c(\cdot,t) := \sum\limits_{\xi\in\Xi} A(t,\xi)\phi_c(\cdot-\xi)$. The coefficients $(A(t,\xi))_{\xi\in\Xi}$ for $\phi_c(\cdot,t)$ are chosen to satisfy $\sum_{\xi\in\Xi} A(t,\xi)p(\xi) = p(t)$, $p \in \Pi_n$, for some sufficiently large n.

(b) The mollifier $m^\vee$ depends on the density of Ξ, i.e., $m = m(\bar{h})$ with $\bar{h}$ as given in (1.7). In particular, for every admissible function f of smoothness class k, we assume that $f - m * f = o(\bar{h}^{rk})$ for some $0 < r < 1$.

The following serves as a prototype for one of the main results in this paper.

Prototype 1. *Let ϕ_c in (1.8) be our basis function and let c satisfy the relation $c = \rho\bar{h}^r$ for $0 < r < 1$ and $\rho > 0$ with $\bar{h}$ in (1.7). Under the conditions (a-b) described above, for every admissible function f of smoothness class k, we have*

$$\|f - R_\Xi f\| = o(\bar{h}^{rk}). \tag{1.9}$$

The exact smoothness conditions on f will be explained later on.

Numerical Example. We assume that $\Xi \subset [-3, 3]^2$. Choosing

$$f(x, y) = \exp(-(x^2 + y^2)) - \left[\frac{\sin(x)\sin(y)}{xy}\right]^5,$$

we approximate f by $R_\Xi f$ and measure the error on $[-1, 1]^2$ (in order to avoid 'boundary effects'). The set of scattered centers Ξ is generated via a random number generator in MATLAB. A comparison between the new scheme R_Ξ and thin-plate spline (TPS) interpolation is given in Figures 1.1-1.3 where the contour lines of the original function, of the TPS interpolant, and of the output of the scheme R_Ξ are given. The errors in the max norm are 0.1682 for TPS interpolation and 0.0397 by R_Ξ.

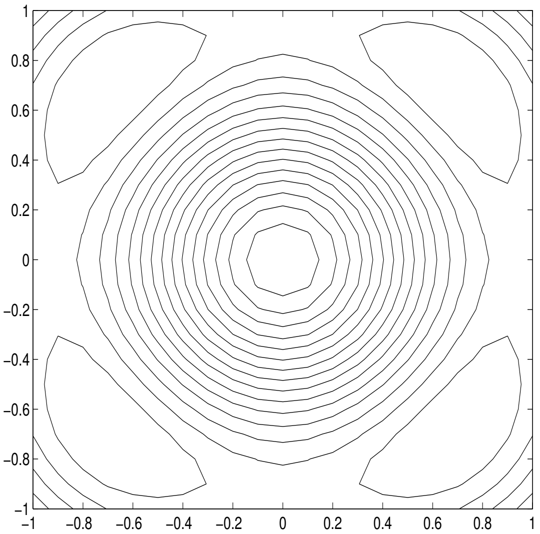


Figure 1.1: The given function

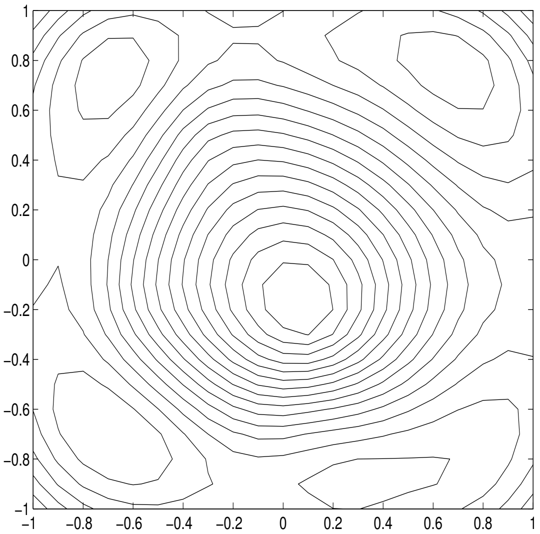


Figure 1.2: TPS interpolation

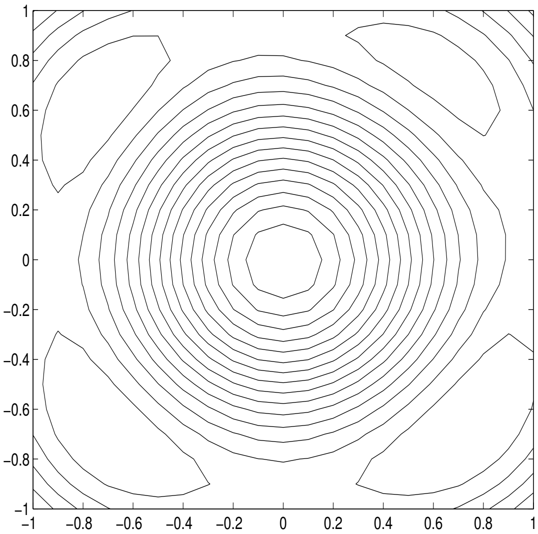


Figure 1.3: The new scheme

We now turn to the practical case of approximating a function f using only the scattered data (ξ, y_ξ) ($y_\xi = f(\xi)$ or $y_\xi = f(\xi) + \epsilon_\xi$), $\xi \in \Xi \subset \Omega$, Ω bounded. In this case, a special technique is necessary in order to eliminate the boundary effects. As one may observe, the scheme R_Ξ in (1.3) it is not local, because the kernel $\phi_c(\cdot,t)$ in (1.3) grows polynomially as $|x|$ goes to ∞. Hence, it is natural to localize the scheme R_Ξ first. To this end, letting

$$\tilde{\Omega}_\delta := \{y = x + z : x \in \Omega,\ |z| \le \delta\} = \Omega + B_\delta,$$

we augment the space $S_\Xi(\phi_c)$ by adding to Ξ points from $\Omega_\delta \setminus \Omega$, and extrapolate $f|_{\Omega_{2\delta}}$ using $f|_\Xi$:

$$f_C(t) := \sum_{\xi\in\Xi} C(t,\xi) f(\xi), \quad t \in \Omega_{2\delta}.$$

The augmented set of centers will be denoted by $\tilde{\Xi}$, and our approximation is actually from the space $S_{\tilde{\Xi}}(\phi_c)$. Precisely, our scattered approximation scheme for bounded domains is of the form

$$\tilde{R}_\Xi : f \mapsto \int_{\tilde{\Omega}_\delta} \psi(\cdot,t)\, \tilde{f}_C(t)\, dt \tag{1.10}$$

where $\psi(\cdot,t)$ is given by (1.6) and

$$\tilde{f}_C := \int_{\Omega_{2\delta}} m_{\phi_c}(\cdot - \theta) f_C(\theta) d\theta$$

with a mollifier m_{ϕ_c} depending on ϕ_c and m in (1.2).

With the scheme $\tilde{R}_\Xi$ in(1.10), we describe the development with some more details as follows:

(a) The function ψ satisfies the condition $\sup_x (1+|x|)^{-d-q}|\psi(x)| < \infty$ for a positive integer q.

(b) The mollifier in $m^\vee * f$ depends on the density of $\tilde{\Xi}$, i.e., $m = m(\bar{h})$ with $\bar{h}$ the density of $\tilde{\Xi}$ as in (1.7). In particular, for every admissible function f of smoothness class k, we assume that $f - m * f = o(h^{rk})$ for some $0 < r \le 1$.

(c) The coefficients $(A(t,\xi))_{\xi\in\Xi}$ for the pseudo-shift $\phi_c(\cdot,t)$ are chosen to satisfy $\sum_{\xi\in\Xi} A(t,\xi)p(\xi) = p(t)$ with $p \in \Pi_n$ for some sufficiently large n.

(d) The coefficients $(C(t,\xi))_{\xi\in\Xi}$ for $f_C = \sum_{\xi\in\Xi} C(t,\xi)f(\xi)$ are chosen to satisfy $\sum_{\xi\in\Xi} C(t,\xi)p(\xi) = p(t)$ with $p \in \Pi_{k-1}$, with k as in (b).

Prototype 2. *Let ϕ_c in (1.8) be the basis function in (1.8) and let the parameter c in (1.8) satisfy the relation $c = \rho\bar{h}^r$, $0 < r < 1$, for some $\rho > 0$. Assume that the parameter $\delta = \delta(\bar{h})$ in the definition of $\tilde{\Omega}_\delta$ decreases to zero as $\bar{h}$ tends to 0, but slower than $\bar{h}^r$, i.e., $\bar{h}^r/\delta \to 0$ as $h \to 0$. Under the conditions (a-d) above, for any admissible function f of smoothness class k, we have*

$$\|f - \tilde{R}_\Xi f\|_{L^\infty(\Omega)} \le O(\,(\bar{h}^r/\delta)^q) + o(\delta^k).$$

Since the approximation scheme is local, even though the set Ξ might be large, the scheme is suitable for implementation by a parallel algorithm.

Next, we provide an algorithm for the approximation scheme $\tilde{R}_\Xi$. The crucial part of any such algorithm is a method for constructing a suitable coefficient sequence $(A(t,\xi))_{\xi\in\Xi}$ for the pseudo-shift $\phi_c(\cdot,t)$ in (1.4). Specifically, 'Gauss elimination by degree' which is introduced by de Boor and Ron [BR2], is applied to a linear system generated by some basis of a polynomial space. We give a detailed description (in a MATLAB-like program) of the calculation of the coefficients $(A(t,\xi))_{\xi\in\Xi}$. Also, a general algorithm for the scheme $\tilde{R}_\Xi$ is given. Finally, some numerical examples on a bounded domain in $\mathbb{R}^2$ are provided. Since the scheme $\tilde{R}_\Xi$ is applied to noiseless data as well as noisy data, we explore both of the cases.

Chapter 2

Approximation Scheme into the Span of Scattered Shifts

2.1 The Pseudo-Shift $\phi(\cdot, t)$

For a given discrete infinite set Ξ in $\mathbb{R}^d$, our construction of an approximation scheme from $S_\Xi(\phi)$ is based on the approximation of each shift $\phi(\cdot - t)$, $t \in \mathbb{R}^d$, by a linear combination

$$\phi(\cdot, t) := \sum_{\xi \in \Xi} A(t, \xi)\phi(\cdot - \xi). \tag{2.1}$$

We referred to $\phi(\cdot, t)$ as a "pseudo-shift" of ϕ. For every $t \in \mathbb{R}^d$, the coefficients $(A(t, \xi))_{\xi \in \Xi}$ must satisfy the following condition, [DR];

Central Condition: For some $m_A > d$, the function $\phi(\cdot, t)$ of (2.1) satisfies

$$|\phi(x - t) - \phi(x, t)| \leq c(1 + |x - t|)^{-m_A}, \quad x \in \mathbb{R}^d, \tag{2.2}$$

with c independent of x and t.

Note that this is an active condition on A: in all the examples of radial basis function ϕ that were considered in [DR] and are considered here, the function ϕ itself grows at ∞. The central condition pertains to the most fundamental property of the basis functions ϕ that we study: while ϕ itself grows at ∞, a suitable linear combination of translates of ϕ may decay at ∞. Specifically, by applying a suitable difference operator to ϕ, a suitable bell-shaped function is obtained:

$$\psi := \sum_{\alpha \in \mathbb{Z}^d} \mu(\alpha)\phi(\cdot - \alpha). \tag{2.3}$$

The coefficients $\mu : \mathbb{Z}^d \to \mathbb{R}$ are called a **localization sequence**. In our study, μ is assumed to have finite support (generally a milder condition is imposed on μ) and the localized function ψ is assumed to satisfy the condition

$$\sup_{x \in \mathbb{R}^d} (1 + |x|)^{m_\psi} \psi(x) < \infty \tag{2.4}$$

for some $m_\psi > d$. Our localized pseudo-shifts, $\psi(\cdot, t)$, are then defined simply by

$$\psi(\cdot, t) := \sum_{\alpha \in \mathbb{Z}^d} \mu(\alpha)\phi(\cdot, t + \alpha). \tag{2.5}$$

Clearly, $\psi(\cdot, t)$, $t \in \mathbb{R}^d$, is a function in $S_\Xi(\phi)$, and it follows directly from (2.5) and Central Condition (2.2) that the difference $\psi(\cdot - t) - \psi(\cdot, t)$ satisfies the decay condition

$$|\psi(x - t) - \psi(x, t)| \leq c'(1 + |x - t|)^{-m_A}, \tag{2.6}$$

where m_A is as in (2.2), and c' is independent of t and x. We refer to [DR] for more details.

Recall that, for any fixed $t \in \mathbb{R}^d$, we assume

$$A_t : \xi \to A(t, \xi), \quad \xi \in \Xi, \tag{2.7}$$

to be finitely supported. Let

$$\Xi_t, \qquad t \in \mathbb{R}^d,$$

be the support of A_t. We choose the centers $\Xi_t \subset \Xi$ to be some 'close neighbors' of t and assume that $\#\Xi_t < C$ for some constant $C > 0$. We require that Ξ_t be

total for Π_n, i.e.,

$$(p|_{\Xi_t} = 0, \qquad p \in \Pi_n) \qquad \text{implies} \qquad p \equiv 0.$$

This, of course, implies also that $\#\Xi_t$ should be no smaller than

$$\dim \Pi_n(\mathbb{R}^d) = \binom{n+d}{n}.$$

We will first discuss sufficient conditions on the coefficient $(A(t,\xi))_{\xi\in\Xi}$ which imply the Central Condition (2.2), and we will also study other properties of the coefficient sequence $(A(t,\xi))_{\xi\in\Xi}$. We start with a lemma that is a simplified version of a result from [DR].

Lemma 2.1. *Let Ξ_t, $t \in \mathbb{R}^d$, be total for Π_n. Then $\sum_{\xi\in\Xi} A(t,\xi)\, p(\xi) = p(t)$ for every $p \in \Pi_n$ if and only if*

$$\sum_{\xi\in\Xi} A(t,\xi)(t-\xi)^\alpha = \delta_{\alpha 0}. \tag{2.8}$$

Proof. Assume that $\sum_{\xi\in\Xi} A(t,\xi)p(\xi) = p(t)$. If we choose $p(x) = (t-x)^\alpha$, then

$$\sum_{\xi\in\Xi} A(t,\xi)(t-\xi)^\alpha = \delta_\alpha.$$

Conversely, for any polynomial $p \in \Pi_n$, let

$$p(\xi - t) = \sum_{|\alpha|_1 \le n} \frac{D^\alpha p(0)}{\alpha!}(\xi - t)^\alpha.$$

Then, by assumption,

$$\sum_{\xi\in\Xi} A(t,\xi)p(\xi - t) = \sum_{\xi\in\Xi} A(t,\xi) \sum_{|\alpha|_1 \le n} \frac{D^\alpha p(0)}{\alpha!}(\xi - t)^\alpha = p(0).$$

This implies that

$$\begin{aligned}\sum_{\xi\in\Xi} A(t,\xi)p(\xi) &= \sum_{|\alpha|_1\le n} \frac{t^\alpha}{\alpha!}\sum_{\xi\in\Xi} A(t,\xi)D^\alpha p(\xi-t)\\ &= \sum_{|\alpha|_1\le n} \frac{t^\alpha}{\alpha!} D^\alpha p(0) = p(t),\end{aligned}$$

which completes our proof. ■

Combining Lemmas 1 and 2 we obtain:

Theorem 2.2 *Let $(A(t,\xi))_{\xi\in\Xi}$, $t\in\mathbb{R}^d$, be the coefficients for $\phi(\cdot,t)$. We assume that*

(a) The Fourier transform $\hat\phi$ belongs to $C^\infty(\mathbb{R}^d\setminus 0)$, and that each $D^\nu\hat\phi$, $\nu\in\mathbb{Z}^d_+$, is summable around ∞, and that each $D^\nu\hat\phi$ (calculated on $\mathbb{R}^d\setminus 0$) has a singularity of order $|\nu|_1+k$ at the origin.

(b) The set $\{A(t,\xi)(t-\xi)^j : t\in\mathbb{R}^d\}$ of functions on Ξ lies in $\ell_1(\Xi)$ and is bounded there for all $j<s$,

(c) For all $p\in\Pi_n$ and for some $n\in(k,s)$, the coefficients $(A(t,\xi))_{\xi\in\Xi}$ satisfies

$$\sum_{\xi\in\Xi} A(t,\xi)p(\xi) = p(t).$$

If Ξ_t is total for Π_n, then we have the relation

$$|\phi(x-t)-\phi(x,t)| \le \mathrm{const}\,(1 + |x-t|)^{-m_A} \tag{2.9}$$

with $m_A = n-k+d$ and const *independent of x and t.*

Proof. Under the conditions (a-b), the result can be derived directly from the result of [DR] (cf. the Theorem 2.7.1 in this article) if the linear system (2.8) holds. Hence, the result is immediate from the Lemma 2.1. ■

It is remarkable that the coefficient sequence $(A(\cdot,\xi))_{\xi\in\Xi}$ can be chosen independently of the basis function ϕ itself and only needs to respect the order of the singularity of $\hat{\phi}$ at the origin. If some different basis function has the same order of singularity, we can apply the same coefficient sequence $(A(t,\xi))_{\xi\in\Xi}$ to construct $\phi(\cdot,t)$. Basic functions that satisfy the assumptions in Theorem 2.2 include truncated power functions ($d=1$), multiquadric, inverse multiquadric, ('shifted') thin-plate spline, and so on. In particular, we will concentrate in the next section on the 'shifted' thin-plate spline function.

Corollary 2.3 *Under the conditions of Theorem 2.2, if the basis function ϕ is a piecewise polynomial on $\mathbb{R}$, i.e., $\phi(x) = |x|^n$ or x_+^n, $x \in \mathbb{R}$, with n a positive integer, then, for every $t \in \mathbb{R}$, the difference $\phi(\cdot - t) - \phi(\cdot, t)$ is compactly supported and* $\operatorname{supp}(\phi(\cdot - t) - \phi(\cdot,t))$ *is the convex hull of the set Ξ_t.*

In view of the above discussion, we introduce the notion of 'admissible coefficients' $(A(\cdot,\xi))_{\xi\in\Xi}$.

Definition. The coefficients $(a(\cdot,\xi))_{\xi\in\Xi}$ are termed **admissible** for Π_n if they satisfy the following three conditions:

(a) There exists $c_1 > 0$ such that, for any $t \in \mathbb{R}^d$, $a(t,\xi) = 0$ whenever $|t-\xi| > c_1\bar{h}$, with $\bar{h}$ the density of Ξ as in (1.7).

(b) The set $\{(a(t,\xi))_{\xi\in\Xi} : t \in \mathbb{R}^d\}$ is bounded in $\ell_1(\Xi)$.

(c) For every $t \in \mathbb{R}^d$, $\sum_{\xi\in\Xi} a(t,\xi)\delta_\xi = \delta_t$ on Π_n, i.e.,

$$\sum_{\xi\in\Xi} a(t,\xi)\, p(\xi) = p(t), \quad \forall p \in \Pi_n. \tag{2.10}$$

Remark : Assuming that the coefficients $(A(t,\xi))_{\xi\in\Xi}$, $t \in \mathbb{R}^d$, are admissible for Π_n, we note that the linear system in (2.9) is invariant under the dilation and translation on $\mathbb{R}^d$ and Ξ. Hence, WLOG, we assume the following conditions in this study:

$$(A(ct,\xi))_{\xi\in\Xi} = (A(t,\xi))_{\xi\in\Xi/c}, \quad c > 0. \tag{2.11}$$

2.2 Approximation Scheme

Among the basis functions that satisfy Theorem 2.2, we have chosen to focus on functions that are obtained from the fundamental solution of the iterated Laplacian by the shifting $|x| \mapsto (|x|^2 + c^2)^{1/2}$, $c > 0$. Thus we consider the following radial basis function

$$\phi_c(x) = \begin{cases} (|x|^2 + c^2)^{\lambda/2}, & \lambda \in \mathbb{Z}_+,\ \lambda,\ d \text{ odd}, \\ (|x|^2 + c^2)^{\lambda/2} \log(|x|^2 + c^2)^{1/2}, & \lambda \in \mathbb{Z}_+,\ \lambda,\ d \text{ even}. \end{cases} \tag{2.12}$$

In the univariate case, $\phi_1(x) = (|x|^2 + 1)^{1/2}$ is called Hardy's multiquadric.

The generalized Fourier transform of ϕ_c is the functions [GS]:

$$\hat{\phi}_c(\theta) = \begin{cases} \tilde{c}(\lambda)\tilde{K}_{(d+\lambda)/2}(c|\theta|)|\theta|^{-\lambda-d}, & \lambda,\ d \text{ odd}, \\ \frac{d}{d\beta}\tilde{c}(\lambda)\tilde{K}_{(d+\lambda)/2}(c|\theta|)|\theta|^{-\lambda-d}, & \lambda,\ d \text{ even}, \end{cases} \tag{2.13}$$

where $\tilde{c}(\beta)$ is of the form

$$\tilde{c}(\beta) = 2^{\frac{\beta}{2}+1}(2\pi)^{d/2}/\Gamma(-\frac{\beta}{2}), \tag{2.14}$$

and $\tilde{K}_\nu(|t|) := |t|^\nu K_\nu(|t|)$ with $K_\nu(|t|)$ the modified Bessel function of order ν. The following properties of $\tilde{K}_\nu$ are related to our analysis [AS]:

$$\tilde{K}_\nu(|t|) > 0, \quad \tilde{K}_\nu(|t|) \approx 2^{\nu-1}\Gamma(\nu) \quad (t \to 0), \tag{2.15}$$

$$\tilde{K}_\nu(|t|) \approx \sqrt{\frac{\pi}{2}} t^{\nu-\frac{1}{2}} e^{-|t|} \quad (t \to \infty),$$
$$\tilde{K}_\nu(|t|) \in C^{2\nu-1}(\mathbb{R}^d) \cap C^\infty(\mathbb{R}^d \setminus 0), \quad \nu \in \mathbb{Z}_+^d.$$

For a smooth function f, we consider the approximation scheme defined by

$$R_\Xi f(x) := \int_{\mathbb{R}^d} \phi_c(x,t) \left(\frac{\hat{f}}{\hat{\phi}_c}\right)^\vee (t)dt \tag{2.16}$$

where $\phi_c(\cdot,t)$ is a kernel of the form

$$\phi_c(\cdot,t) = \sum_{\xi\in\Xi} A(t,\xi)\phi(\cdot-\xi).$$

This scheme is intrinsic in the sense that it employs directly the scattered shifts of the basis function.

We hope that the approximation is getting better as the center set Ξ becomes 'dense'. In order to study this, we measure the density $\bar{h}$ of Ξ as in (1.7):

$$\bar{h} = \sup_{x\in\mathbb{R}^d} \inf_{\xi\in\Xi} |x-\xi|. \tag{2.17}$$

Let $(A(t,\xi))_{\xi\in\Xi}$ be admissible for Π_n. Then, for every function f such that $\hat{f}\hat{\phi}_c^{-1} \in L_1(\mathbb{R}^d)$, we will prove that the scheme R_Ξ provides the approximation error

$$\|f - R_\Xi f\| = O(\bar{h}^{n+1}).$$

However, in order for the integral in (2.16) to make sense, the function f need to be extremely smooth, and we do not want to impose such smoothness condition on our approximand. So, in this section, we will discuss how to apply the above approximation scheme to functions in a larger space. Also, we will discuss in detail the process by which the scheme R_Ξ is derived.

When we are looking for an approximant from the space $S_\Xi(\phi_c)$ in terms of the conversion method discussed earlier, it is essential to choose a good approximation scheme on the uniform grid. In the paper [BR1], C. de Boor and A. Ron introduce an optimal approximation scheme from the spaces spanned by shifts of a basis function. One can observe in this paper that spectral approximation order can be obtained if the basis function is smooth and satisfies certain other conditions.

Let f be a smooth function in some smoothness space (the specific function space will be defined later). The approximation scheme in [BR1] is of the form

$$R_h : f \to \sum_{\alpha\in\mathbb{Z}^d} \psi_{c/h}(\cdot/h-\alpha) f_h^*(\alpha)$$

with f_h^* the bounded analytic function

$$f_h^*(x) := (2\pi)^{-d} \int_{\mathbb{R}^d} \frac{e_{hx}(\theta)}{\hat{\psi}_{c/h}(h\theta)} \sigma(h\theta)\hat{f}(\theta)d\theta, \tag{2.18}$$

where $\sigma : \mathbb{R}^d \to [0\ldots1]$ is a nonnegative C^∞-cutoff function whose support σ lies in some ball B_η, $\eta > 0$; furthermore $\sigma = 1$ on $B_{\eta/2}$ and that $\|\sigma\|_\infty = 1$. Here and hereafter we assume that $\hat{\psi}_c \neq 0$ on B_η such that $\sigma/\hat{\psi}_c$ is well-defined.

Then, for any $h > 0$, the scattered center variant $R_{A,h}$ of R_h is defined by replacing $\psi_{c/h}(\cdot-\alpha)$ by $\psi_{c/h}(\cdot,\alpha)$ in accordance with the conversion method in [DR]. Hence, we arrive at an approximant from $S_\Xi(\phi_c)$ of the form

$$R_{A,h}f := \sum_{\alpha\in\mathbb{Z}^d} \psi_{c/h}(\cdot/h,\alpha) f_h^*(\alpha) \tag{2.19}$$

where $\psi_{c/h}$ is defined as in (2.5) with

$$\phi_{c/h}(\cdot/h,\alpha) = \sum_{\xi\in\Xi} A(\cdot,\xi/h)\, \phi_{c/h}(\cdot/h-\xi/h), \tag{2.20}$$

which ensures that $R_{A,h}f$ is an element of $S_\Xi(\phi_c)$.

Remark: The function $\psi_{c/h}(\cdot/h,\alpha)$ in (2.19) is obtained by an application of the same matrix $(A(\cdot,\xi))_{\xi\in\Xi}$ to the scattered shifts of $\phi_{c/h}(\cdot/h-\alpha)$. However, a careful examination of the function $R_{A,h}f$ reveals that $R_{A,h}f$ is not in the space $S_\Xi(\phi_c)$, but in the space $S_{h\Xi}(\phi_c)$. Since our goal is to approximate from the space $S_\Xi(\phi_c)$ the fact that the dilated center set $h\Xi$ is employed as h changes should not occur. For this reason, at each h-level, we employ the center set $h^{-1}\Xi$ in the construction of $R_{A,h}f$. Then, for each $\alpha \in \mathbb{Z}^d$, the pseudo-shift $\phi_{c/h}(\cdot/h, \alpha)$ should be of the form in (2.20).

With these remarks, let us turn to the discussion of choosing a uniform grid density corresponding to the given scattered center set Ξ. Actually, given a set of scattered centers Ξ, we need to choose a density h of the uniform grid corresponding to Ξ. So, in the following results, we observe the relation $R_h - R_{A,h}$ with a fixed set Ξ and a ladder of uniform grids $(h\mathbb{Z}^d)_h$.

At this stage, we are first interested in the approximation of functions in the space

$$\mathcal{F}^k_{\phi_c} := \{f : \mathbb{R}^d \to \mathbb{R} \ : \ \sum_{|\nu|_1=k} (\,)^\nu \hat{f}\hat{\phi}_c^{-1} \in L_1(\mathbb{R}^d)\}$$

and we use the notation

$$\|f\|'_{k,1} := \sum_{|\nu|_1=k} \|(\,)^\nu \hat{f}\hat{\phi}_c^{-1}\|_1.$$

Lemma 2.4. *Let ϕ_c and $\phi_c(\cdot, t)$, $t \in \mathbb{R}^d$, be as above. Assume that the Central Condition (2.2) holds. Then there exist a constant $C > 0$ independent of x and h such that*

$$h^d \sum_{\alpha\in\mathbb{Z}^d} |\phi_c(x - h\alpha) - \phi_c(x, h\alpha)| < C$$

for any $0 < h \leq 1$ and $x \in \mathbb{R}^d$.

Proof. From the Central Condition (2.2), we have

$$\sum_{\alpha\in\mathbb{Z}^d} h^d|\phi_c(x-h\alpha)-\phi_c(x,h\alpha)| \le \text{const} \sum_{\alpha\in\mathbb{Z}^d} h^d(1+|x-h\alpha|)^{-m_A}$$

with $m_A > d$ as in (2.2) and const independent of x and $h\alpha$. Letting $\mathcal{B}_{h,x}$ be the set $\{\alpha\in\mathbb{Z}^d : |x-h\alpha|<h\}$, we have the relation

$$\begin{aligned}\sum_{\alpha\in\mathbb{Z}^d}(1+|x-h\alpha|)^{-m_A} &= \sum_{\alpha\in\mathcal{B}_{h,x}}(1+|x-h\alpha|)^{-m_A} + \sum_{\alpha\in\mathbb{Z}^d\backslash\mathcal{B}_{h,x}}(1+|x-h\alpha|)^{-m_A}\\ &\le 2^d(1+\sum_{\alpha\in\mathbb{Z}^d_+\backslash 0}(1+|h\alpha|)^{-m_A}) = 2^d\sum_{\alpha\in\mathbb{Z}^d_+}(1+|h\alpha|)^{-m_A}\end{aligned}$$

Since

$$h^d\sum_{\alpha\in\mathbb{Z}^d_+}(1+|h\alpha|)^{-m_A} \to \int_{\mathbb{R}^d_+}(1+|t|)^{-m_A}\,dt$$

as h tends to 0, we get the lemma's claim. ■

The following result is from [DR].

Lemma 2.5. *Assume that $\hat\phi$ is continuous on $\mathbb{R}^d\backslash 0$ and has a singularity of order $> k$ at the origin for some positive integer k. Let $(\mu(\alpha))_{\alpha\in\mathbb{Z}^d}$ be the localization sequence in* (2.3) *and assume that the localization ψ satisfies the condition (2.4). Assume also that the linear functional*

$$\mu : p \longmapsto \sum_{\alpha\in\mathbb{Z}^d}\mu(-\alpha)p(\alpha) \tag{2.21}$$

is well-defined on Π_k (i.e., the above sum converges absolutely for every $p\in\Pi_k$). Then μ annihilates Π_k.

The following proposition provides a clue for the relation between Ξ and $h\mathbb{Z}^d$ in terms of the conversion method.

Proposition 2.6. *Let ϕ_c, ψ_c, $\phi_c(\cdot,t)$, $\psi_c(\cdot,t)$, $t\in\mathbb{R}^d$, R_h and $R_{A,h}$ be as above. Let the coefficients $(A(t,\xi))_{\xi\in\Xi}$ for $\phi_c(\cdot,t)$ be admissible for Π_n. Then, for every function $f\in\mathcal{F}^k_{\phi_c}$ with $k=\lambda+d+1$, we have*

$$\|(R_h-R_{A,h})f\| = \sum_{\alpha\in\mathbb{Z}^d}(\phi_c(x-h\alpha)-\phi_c(x,h\alpha))\sum_{|\nu|_1=\lambda+d}\mu_\nu D^\nu f_h^*(\alpha)+O(h)$$

where f_h^* is as in (2.18) and

$$\mu_\nu := \sum_{\beta\in\mathbb{Z}^d}\mu(\beta)(-\beta)^\nu/\nu!, \quad \nu\in\mathbb{Z}^d_+. \tag{2.22}$$

Moreover,

$$\|(R_h-R_{A,h})f\|\le c,$$

with c a constant independent of h and Ξ.

Proof. First, from the definitions of R_h and $R_{A,h}$, we can write $(R_h-R_{A,h})f(x/h)$ as follows:

$$\begin{aligned}(R_h-R_{A,h})f(x/h) &= \sum_{\beta\in\mathbb{Z}^d}(\psi_{c/h}(x/h-\beta)-\psi_{c/h}(x/h,\beta))f_h^*(\beta)\\ &= \sum_{\beta\in\mathbb{Z}^d}\sum_{\alpha\in\mathbb{Z}^d}(\phi_{c/h}(x/h-\alpha)-\phi_{c/h}(x/h,\alpha))\mu(\alpha-\beta)f_h^*(\beta)\end{aligned}$$

with $\phi_{c/h}(\cdot/h,\alpha)$ as in (2.20). Due to the Central Condition (2.2), the above double sum converge absolutely, and summation by parts implies that

$$\begin{aligned}(R_h-R_{A,\bar{h}})f(x/h) &= \sum_{\alpha\in\mathbb{Z}^d}(\phi_{c/h}(x/h-\alpha)-\phi_{c/h}(x/h,\alpha))\sum_{\beta\in\mathbb{Z}^d}\mu(\alpha-\beta)f_h^*(\beta)\\ &= \sum_{\alpha\in\mathbb{Z}^d}(\phi_{c/h}(x/h-\alpha)-\phi_{c/h}(x/h,\alpha))(\mu *' f_h^*)(\alpha). \end{aligned}\tag{2.23}$$

Here, we claim that

$$\phi_{c/h}(\cdot/h-\alpha)-\phi_{c/h}(\cdot/h,\alpha) = h^{-\lambda}(\phi_c(\cdot-h\alpha)-\phi_c(\cdot,h\alpha)). \tag{2.24}$$

For the case that d is odd, this equation follows immediately from the explicit formula in (2.12); if d is even, the relation between $\phi_{c/h}$ and ϕ_c is as follows:

$$\phi_{c/h} = h^{-\lambda}\phi_c(h\cdot) - (|\cdot|^2 + (c/h)^2)^{\lambda/2}\log h.$$

Accordingly,

$$\begin{aligned}\phi_{c/h}(x/h-\alpha) - \phi_{c/h}(x/h,\alpha) = h^{-\lambda}\Big[\sum_{\xi\in\Xi} A(\alpha,\xi/h)(\phi_c(x-h\alpha) - \phi_c(x-\xi)) \\ - \log h \sum_{\xi\in\Xi} A(\alpha,\xi/h)((|x-h\alpha|^2+c^2)^{\lambda/2} - (|x-\xi|^2+c^2)^{\lambda/2})\Big]\end{aligned} \tag{2.25}$$

Since $(A(\cdot,\xi/h))_{\xi\in\Xi} = (A(h\cdot,\xi))_{\xi\in\Xi}$ by (2.11) and $(|\cdot|^2+c^2)^{\lambda/2}$ is a polynomial of degree λ, it follows from the condition on the matrix $(A(\cdot,\xi))_{\xi\in\Xi}$ in (2.10) that the second sum in the above equation is identically zero. This establishes the identity (2.24).

Thus, by (2.23) and (2.24), we arrive at the relation

$$(R_h - R_{A,\bar{h}})f(x/h) = h^{-\lambda}\sum_{\alpha\in\mathbb{Z}^d}(\phi_c(x-h\alpha) - \phi_c(x,h\alpha))(\mu *' f_h^*)(\alpha). \tag{2.26}$$

Recalling the definition of the linear functional μ in (2.21), we have the identity

$$(\mu *' f_h^*)(\alpha) = \mu(f_h^*(\alpha+\cdot)) = \sum_{\beta\in\mathbb{Z}^d}\mu(-\beta)f_h^*(\alpha+\beta).$$

Since f_h^* is real analytic, its Taylor polynomial of degree $\lambda+d-1$ around $\vartheta = y$ is well-defined:

$$T_y f_h^*(s) := \sum_{|\nu|_1<\lambda+d} D^\nu f_h^*(y)(s-y)^\nu/\nu!, \qquad y,\, s\in\mathbb{R}^d.$$

Then it follows from Lemma 2.5 that

$$\mu(T_\alpha(f_h^*(\alpha+\cdot))) = \sum_{\beta\in\mathbb{Z}^d}\mu(-\beta)\,T_\alpha f_h^*(\alpha+\beta) = 0.$$

Hence, we have

$$(\mu *' f_h^*)(\alpha) = \sum_{|\nu|_1=\lambda+d} \mu_\nu D^\nu f_h^*(\alpha) + \sum_{|\nu|_1=\lambda+d+1} \sum_{\beta\in\mathbb{Z}^d} \mu(-\beta)\beta^\nu D^\nu f_h^*(\beta_\alpha)/\nu! \quad (2.27)$$

where $\beta_\alpha \in [\alpha\,..\,\beta]$, and μ_ν is in (2.22). From the explicit formula of f_h^* in (2.18), it is clear that the function $D^\nu f_h^*$ is well-defined, and especially, it has the form

$$D^\nu f_h^*(t) = \frac{h^{|\nu|_1}}{(2\pi)^d} \int_{\mathbb{R}^d} \frac{e_{ht}(\theta)}{\hat{\psi}_{c/h}(h\theta)} (i\theta)^\nu \sigma(h\theta) \hat{f}(\theta)\, d\theta. \quad (2.28)$$

This leads to the inequality

$$\|D^\nu f_h^*\| \leq \text{ const } h^{|\nu|_1} \, \|f\|'_{|\nu|_1,1} \quad (2.29)$$

with const independent of h. From (2.27) and (2.29), we obtain

$$|(\mu *' f_h^*)(\alpha)| \leq \mathrm{c}_1 h^{\lambda+d} \, \|f\|'_{\lambda+d,1} + \mathrm{c}_2 h^{\lambda+d+1} \, \|f\|'_{\lambda+d+1,1}. \quad (2.30)$$

Therefore, combining (2.30) with (2.26), we arrive at the bound

$$\|(R_h - R_{A,h})f\| \leq K \, (\mathrm{c}_1 \, \|f\|'_{\lambda+d,1} + \mathrm{c}_2 \, h \|f\|'_{\lambda+d+1,1}) \quad (2.31)$$

where

$$K := \sup_{x,h} \, h^d \sum_{\beta\in\mathbb{Z}^d} |\phi_c(x-h\beta) - \phi_c(x,h\beta)| < \infty$$

by Lemma 2.4. ■

From the view point of finding optimal h associated with a given set of scattered centers Ξ in terms of the error $(R_h - R_{A,\bar{h}})$, the bound of the form $c_1 + c_2 h$ does not provide us with any preferable value of h: the parameter h does not play any major role in our error estimation. Hence, we let the uniform mesh size go to 0 for the purpose of minimizing the upper bound of errors. Next, we describe the convergence property of $R_h - R_{A,h} f$ as h tends to 0.

Lemma 2.7. *Let ϕ_c, μ and ψ_c be as above. Then, as h tends to 0, we have the following convergence property:*

$$\hat{\psi}_{c/h}(h\theta) \to (-i)^{\lambda+d}\hat{\phi}_c(\theta) \sum_{\alpha\in\mathbb{Z}^d} \mu(-\alpha)(\alpha\cdot\theta)^{\lambda+d}/(\lambda+d)!, \qquad \theta\in\mathbb{R}^d.$$

Proof. From the definition of $\psi_{c/h}$, we can compute its Fourier transform term by term. Defining a trigonometric polynomial function

$$\tau := \sum_{\alpha\in\mathbb{Z}^d} \mu(\alpha)e_{-\alpha},$$

we have

$$\hat{\psi}_{c/h} = \tau\hat{\phi}_{c/h}$$

pointwise on $\mathbb{R}^d\setminus 0$. We then simply obtain the following relation from the expression (2.13)

$$\hat{\psi}_{c/h}(h\cdot) = \tilde{c}_\lambda \tilde{K}_{(\lambda+d)/2}(c\cdot)\tau(h\cdot)/|h\cdot|^{\lambda+d}$$

with $\tilde{c}(\lambda)$ from (2.14). Invoking the localization condition of $\psi_{c/h}$ in (2.4), we know that $\hat{\psi}_{c/h}(h\cdot)$ is continuous everywhere, especially, at the origin. Hence, $\tau(h\cdot)$ has a zero of order $\lambda+d$ at the origin. Consequently, when h converges to 0, we have

$$\tau(h\theta)/|h\theta|^{\lambda+d} \to (-i)^{\lambda+d}\sum_{\alpha\in\mathbb{Z}^d}\mu(-\alpha)(\alpha\cdot\theta)^{\lambda+d}/(\lambda+d)!/|\theta|^{\lambda+d}$$

with $\theta\in\mathbb{R}^d\setminus 0$, and this completes our proof. ■

Theorem 2.8. *Assume that the Central Condition (2.2) holds, and let R_h and $R_{A,h}$ be as above. Then, for every function $f\in\mathcal{F}^0_{\phi_c}$, we have the convergence property*

$$(R_h - R_{A,\bar{h}})f(x/h) \longrightarrow f(x) - \int_{\mathbb{R}^d}\phi_c(x,t)\Big(\frac{\hat{f}}{\hat{\phi}_c}\Big)^{\vee}(t)dt$$

pointwise as h converges to 0.

Proof. We invoke Proposition 2.6. Using the explicit formula for $D^\nu f_h^*$ in (2.28), we estimate

$$
\begin{aligned}
&h^{-\lambda}\sum_{\alpha\in\mathbb{Z}^d}(\phi_c(x-h\alpha)-\phi_c(x,h\alpha))\sum_{|\nu|_1=\lambda+d}\mu_\nu D^\nu f_h^*(\alpha) \qquad (2.32)\\
&=\frac{i^{\lambda+d}}{(2\pi)^d}\int_{\mathbb{R}^d}\frac{\sigma(h\theta)\hat{f}(\theta)}{\hat{\psi}_{c/h}(h\theta)}\sum_{|\nu|_1=\lambda+d}\mu_\nu\theta^\nu\sum_{\alpha\in\mathbb{Z}^d}h^d(\phi_c(x-h\alpha)-\phi_c(x,h\alpha))e_{h\alpha}(\theta)d\theta.
\end{aligned}
$$

Note that

$$
\begin{aligned}
\sum_{|\nu|_1=\lambda+d}\mu_\nu\theta^\nu &= \sum_{|\nu|_1=\lambda+d}\sum_{\alpha\in\mathbb{Z}^d}\mu(-\alpha)\alpha^\nu\theta^\nu/\nu!\\
&= \sum_{\alpha\in\mathbb{Z}^d}\mu(-\alpha)\sum_{|\nu|_1=\lambda+d}\binom{\lambda+d}{\nu}(\alpha_1\theta_1)^{\nu_1}\cdots(\alpha_d\theta_d)^{\nu_d}/(\lambda+d)!\\
&= \sum_{\alpha\in\mathbb{Z}^d}\mu(-\alpha)(\alpha\cdot\theta)^{\lambda+d}/(\lambda+d)!.
\end{aligned}
$$

Here, we use the standard abbreviation

$$\binom{\lambda+d}{\nu}=\frac{(\lambda+d)!}{\nu_1!\cdots\nu_d!}.$$

Now the right hand side of the equation (2.32) can be rewritten as

$$
\begin{aligned}
&\frac{i^{\lambda+d}}{(2\pi)^d(\lambda+d)!}\int_{\mathbb{R}^d}\frac{\sigma(h\theta)\hat{f}(\theta)}{\hat{\psi}_{c/h}(h\theta)}\sum_{\alpha\in\mathbb{Z}^d}\mu(-\alpha)(\alpha\cdot\theta)^{\lambda+d} \qquad (2.33)\\
&\qquad\times\sum_{\beta\in\mathbb{Z}^d}h^d(\phi_c(x-h\beta)-\phi_c(x,h\beta))e_{h\beta}(\theta)d\theta.
\end{aligned}
$$

Though this formula may look very complicated, it tends to a simple expression as we take $h\to 0$. We first observe that, for any fixed $x\in\mathbb{R}^d$,

$$\sum_{\beta\in\mathbb{Z}^d}h^d(\phi_c(x-h\beta)-\phi_c(x,h\beta))e_{h\beta}(\theta)\to\int_{\mathbb{R}^d}(\phi_c(x-t)-\phi_c(x,t))e_t(\theta)\,dt.$$

A direct calculation using Lemmas 2.4 and 2.5 yields that the integrand in (2.33) is bounded uniformly by

$$\text{const}\left|\frac{\hat{f}}{\hat{\phi}_c}\right|\in L_1(\mathbb{R}^d).$$

Thus, by Lemma 2.7, we have the limit of the above equation as follows:

$$\lim_{h\to 0}(R_h - R_{A,\bar{h}})f(x/h) = \int_{\mathbb{R}^d} \left(\frac{\hat{f}}{\hat{\phi}_c}\right)(\theta) \int_{\mathbb{R}^d} (\phi_c(x-t) - \phi_c(x,t))\, e_t(\theta)\; dt d\theta$$

by Lebesgue Dominated Convergence Theorem. Since $f \in \mathcal{F}^0_{\phi_c}$ and $\phi_c(\cdot - t) - \phi_c(\cdot, t) \in L_1(\mathbb{R}^d)$, we deduce that

$$\begin{aligned}\lim_{h\to 0}(R_h - R_{A,\bar{h}})f(x/h) &= \int_{\mathbb{R}^d} (\phi_c(x-t) - \phi_c(x,t)) \left(\frac{\hat{f}}{\hat{\phi}_c}\right)^{\vee}(t)dt \\ &= f(x) - \int_{\mathbb{R}^d} \phi_c(x,t) \left(\frac{\hat{f}}{\hat{\phi}_c}\right)^{\vee}(t)dt\end{aligned}$$

by Fubini's theorem. ■

Remark. As a consequence of Theorem 2.8, we obtain an approximation scheme from $S_\Xi(\phi)$ which does not require an association with any uniform mesh. Clearly, we deduce from the above theorem that

$$\lim_{h\to 0} R_{A,h}f(x) = \int_{\mathbb{R}^d} \phi_c(x,t) \left(\frac{\hat{f}}{\hat{\phi}_c}\right)^{\vee}(t)dt = R_\Xi f(x). \tag{2.34}$$

Recalling the definition of the pseudo-shift $\phi_c(\cdot, t)$ in (2.1), $R_\Xi f$ has the explicit form

$$R_\Xi f(x) = \sum_{\xi\in\Xi} c_\xi(f)\phi_c(x-\xi)$$

where c_ξ is the linear functional

$$c_\xi : f \longmapsto \int_{\mathbb{R}^d} A(t,\xi)\Big(\frac{\hat{f}}{\hat{\phi}_c}\Big)^{\vee}(t)dt. \tag{2.35}$$

It is clear that $R_\Xi f$ is an element of $S_\Xi(\phi_c)$.

However, since the modified Bessel function $\tilde{K}_\nu(x)$ which is a part of $\hat{\phi}_c$ is decreasing exponentially fast as x tends to ∞, the space of functions $\mathcal{F}^0_{\phi_c}$ is very small and only extremely smooth function can belong to it: Even some infinitely differentiable functions are not in this space. Our interest in this study is in approximating functions from a larger smoothness space. Specifically, we are interested in approximating functions in the space

$$\widetilde{W}^k_\infty(\mathbb{R}^d)$$

of all functions f such that the Fourier transform $\hat{f}$ is a Radon measure, and the total mass $\|(1+|\cdot|^2)^{k/2}\hat{f}\|_1$ of $(1+|\cdot|^2)^{k/2}\hat{f}$ is finite. If $\hat{f} \in L_1(\mathbb{R}^d)$, then the norm $\|(1+|\cdot|^2)^{k/2}\hat{f}\|_1$ coincides with the $L_1(\mathbb{R}^d)$-norm of $(1+|\cdot|^2)^{k/2}\hat{f})$. The above induces a norm on $\tilde{W}^k_\infty(\mathbb{R}^d)$,

$$\|f\|_{k,\infty} := \|(1+|\cdot|^2)^{k/2}\hat{f}\|_1.$$

In order to apply the scheme R_Ξ to a function $f \in \widetilde{W}^k_\infty(\mathbb{R}^d)$, we first replace f by its smooth part

$$\sigma^\vee * f$$

where σ is the cutoff function in (2.18). Then $\sigma^\vee * f$ is a band-limited function, and we apply the scheme R_Ξ to $\sigma^\vee * f$ instead of f.

Since we hope that the approximation is getting better as the center set Ξ becomes denser, the cutoff function σ is dilated proportional to the density of Ξ. Thus, for a given set Ξ, we consider the approximation

$$f \approx \sigma_\omega^\vee * f = (\sigma_\omega \hat{f})^\vee$$

where $\omega := \omega(\bar{h})$ and

$$\sigma_\omega := \sigma(\omega\cdot).$$

It is important to note that we do not necessarily choose $\omega(\bar{h}) = \bar{h}$. We will discuss our choice in the next section,

In this case, there occurs some obstacles for this scheme to be used directly; First, since the basis function ϕ grows at some polynomial degree away from zero , it may cause to loose local property of the approximation scheme R_Ξ. Furthermore, we need to impose some extra conditions on f in order to make the above integration make sense. The function f is required to satisfy the condition $\hat{f} \in C^k(\mathbb{R}^d)$, $k > d + \lambda$. Thus, in order to circumvent those difficulties, a 'localization process' is necessary. The strategy to be used here is first to localize the kernel $\phi_c(\cdot - t)$ in the above convolution equation by applying a difference operator to ϕ_c, which construct a new bell-shaped kernel

$$\psi_c = \sum_{\alpha \in \mathbb{Z}^d} \mu(\alpha)\phi_c(\cdot - \alpha)$$

with favorable decay properties at ∞. Next, we then approximate the localized kernel $\psi_c(\cdot - t)$ from the space $S_\Xi(\phi_c)$.

The Fourier transform $\hat{\phi}_c$ of ϕ_c is very smooth off the origin. This means that in order to localize ϕ_c we only need to ensure that the Fourier transform $\hat{\psi}_c$ of the localized function ψ_c is smooth at the origin. Note that we also need to insist that $\hat{\psi}_c(0) \neq 0$. In other words, considering the localization condition on ψ_c, $\hat{\psi}_c$ is continuous everywhere, especially at the origin. Hence, the function

$$\tau = \sum_{\alpha \in \mathbb{Z}^d} \mu(\alpha)e_{-i\alpha}$$

with $\hat{\psi}_c = \tau\hat{\phi}_c$ has a high order zero at the origin. We note here that τ is a $2\pi\mathbb{Z}^d$-periodic function, and since the only singularity of $\hat{\phi}_c$ is at the origin and $\hat{\phi}_c \neq 0$

on $\mathbb{R}^d \setminus 0$, we can assume that τ does not vanish on some punctured neighborhood $\Omega \setminus 0$ of the origin. This ensures that $\hat{\psi}_c$ does not vanish on $\Omega \setminus 0$. Expressing the inverse Fourier transform of τ as

$$\tau^\vee = \sum_{\alpha \in \mathbb{Z}^d} \mu(\alpha)\, \delta_\alpha, \tag{2.36}$$

the above convolution equation (1.2) implies the relation

$$\sigma_\omega^\vee * f = \phi_c * \tau(\omega\cdot)^\vee * \Big(\sigma_\omega \frac{\hat{f}}{\tau(\omega\cdot)\hat{\phi}_c}\Big)^\vee = \left[\psi_{c/\omega} * \Big(\frac{\sigma_\omega \hat{f}}{\hat{\psi}_{c/\omega}}\Big)^\vee (\omega\cdot)\right](x/\omega) \tag{2.37}$$

where the equality holds by using the properties $(gh)^\vee = g^\vee * h^\vee$ and

$$\tau(\omega\cdot)\hat{\phi}_c = \omega^{-\lambda-d}\hat{\psi}_{c/\omega}(\omega\cdot). \tag{2.38}$$

Thus, substituting $\psi_{c/\omega}(\cdot/\omega, t)$ for $\psi_{c/\omega}(\cdot/\omega - t)$ in (2.37), we obtain our approximation scheme as following:

Definition 2.9 *With ϕ_c, ψ_c and $\psi_c(\cdot, t)$ as above, we define our approximation scheme R_Ξ by*

$$R_\Xi : f \mapsto \int_{\mathbb{R}^d} \psi_{c/\omega}(x/\omega, t)\Big(\frac{\sigma_\omega \hat{f}}{\hat{\psi}_{c/\omega}(\omega\cdot)}\Big)^\vee(\omega t)dt. \tag{2.39}$$

where $\psi_c(\cdot, t)$ is a localized kernel of the form

$$\psi_{c/\omega}(x/\omega, t) = \sum_{\alpha \in \mathbb{Z}^d} \mu(\alpha)\phi_{c/\omega}(x/\omega, t + \alpha).$$

In the error analysis, it is useful to divide $f - R_\Xi f$ into two parts,

$$f - R_\Xi f = ((\sigma_\omega \hat{f})^\vee - R_\Xi f) + (f - (\sigma_\omega \hat{f})^\vee). \tag{2.40}$$

In particular, for every function $f \in \widetilde{W}_\infty^k(\mathbb{R}^d)$, it is immediate that

$$\|f - (\sigma_\omega \hat{f})^\vee\| \le (2\pi)^{-d} \int_{\mathbb{R}^d} |(1 - \sigma_\omega(\theta))\hat{f}(\theta)| d\theta = o(\omega^k) \tag{2.41}$$

as ω tends to 0.

Usually, approximation power is quantified by approximation order. But, an error analysis can be carried out in terms of a given tolerance. By taking a sufficiently small ω in our scheme, we can make the tail error small enough to satisfy a (given) tolerance. Then, for a fixed ω, we choose the space $S_\Xi(\phi_c)$ such that the final error satisfies the required tolerance . The following results discuss this issue.

Lemma 2.10 *Let ϕ_c be as above. For $\nu \in \mathbb{Z}_+^d$ with $|\nu|_1 = n+1 > \lambda + d + 1$, $D^\nu \phi_c \in L_1(\mathbb{R}^d)$.*

Proof. It is sufficient to prove that the Fourier transform of $(\,)^\alpha D^\nu \phi_c$ with $|\alpha|_1 = d+1$ is bounded in $L_1(\mathbb{R}^d)$. We see that the Fourier transform of $(\,)^\alpha D^\nu \phi_c$ is

$$i^{|\alpha+\nu|_1} D^\alpha (\,)^\nu \hat{\phi}_c,$$

and hence by using Leibniz' rule, we only need to show that

$$D^\gamma (\,)^\nu D^{\alpha-\gamma} \hat{\phi}_c \in L_1(\mathbb{R}^d). \tag{2.42}$$

Since $D^{\alpha-\gamma}\hat{\phi}_c$ decays fast around ∞, the function in (2.42) is bounded in $L_1(N_\infty)$ for some neighborhood N_∞ of ∞.

Next, we note that the Fourier transform of $D^\nu \phi_c$ is $(i\cdot)^\nu \hat{\phi}_c$. Here, since the distribution $\hat{\phi}_c$ has the order of singularity $\lambda + d$ at the origin and θ^ν has a zero of order $n+1$ $(> \lambda + d + 1)$ at the origin, it follows that the Fourier transform of $D^\nu \phi_c$ extends to the entire $\mathbb{R}^d$. We also realize that the function in (2.42) has a singularity of order $\lambda + 2d - n$, and $\lambda + 2d - n < d$ by assumption. It implies that $D^\gamma (\,)^\nu D^{\alpha-\gamma}\hat{\phi}_c$ is integrable around origin. ■

Theorem 2.11 *Let $\phi_c(\cdot,t)$, $t \in \mathbb{R}^d$, and R_Ξ be as above. Let $(A(t,\xi))_{\xi\in\Xi}$ for $\phi_c(\cdot,t)$ be admissible for Π_n with $n > \lambda + d$. Then, for every $f \in \widetilde{W}^k_\infty(\mathbb{R}^d)$,*

$$\|f - R_\Xi f\| \le \text{const}\, \bar{h}^{n+1} + o(\omega^k)$$

with $\bar{h}$ the density of Ξ as in (1.7) and const *dependent on c and ω.*

Proof. We prefer to provide the proof of this theorem in the form of a separate lemma:

Lemma 2.12 *Let $\phi_c(\cdot,t)$, $t \in \mathbb{R}^d$, and R_Ξ be as above. Let $(A(t,\xi))_{\xi\in\Xi}$ for $\phi_c(\cdot,t)$ be admissible for Π_n with $n > \lambda + d$. Then, for every band-limited function f,*

$$\int_{\mathbb{R}^d} (\phi_c(\cdot - t) - \phi_c(\cdot,t)) \left(\frac{\hat{f}}{\hat{\phi}_c}\right)^{\vee} (t)dt \le \text{const}\, \bar{h}^{n+1}$$

with const *independent of Ξ, but dependent on f.*

Proof of Lemma. We first deduce the relation

$$\int_{\mathbb{R}^d} |(\phi_c(x - t) - \phi_c(x,t)) \left(\frac{\hat{f}}{\hat{\phi}_c}\right)^{\vee} (t)|dt \le \|\hat{f}\hat{\phi}_c^{-1}\|_1 \int_{\mathbb{R}^d} |\phi_c(x-t) - \phi_c(x,t)|dt$$

since $\|\hat{g}\| \le \|g\|_1$ for $g \in L_1(\mathbb{R}^d)$. Because, by assumption, $\sum_{\xi\in\Xi} A(t,\xi) = 1$, it is immediate that

$$\phi_c(x-t) - \phi_c(x,t) = \sum_{\xi\in\Xi} A(\xi,t)(\phi_c(x-t) - \phi_c(x-\xi)).$$

Let $T_{x-\xi}\phi_c$ be the Taylor polynomial for ϕ_c of degree n about $\vartheta = (x - \xi)$. Then $T_{x-\xi}\phi_c(x-t) - \phi_c(x-\xi)$ is also a polynomial of degree n in terms of variable t and

$$\phi_c(x-t) - \phi_c(x,t) = \sum_{\xi\in\Xi} A(\xi,t)[T_{x-\xi}\phi_c(x-t) - \phi_c(x-\xi) + R_n(t,\xi)]$$

with the remainder in integral form

$$R_n(t,\xi) := \int_0^1 \frac{(1-y)^n}{n!} \sum_{|\nu|_1=n+1} (t-\xi)^\nu D^\nu \phi_c(x-t+y(t-\xi))dy.$$

By the facts that $\sum_{\xi\in\Xi} A(\cdot,\xi)p(\xi) = p$ and that $A(t,\xi) = 0$ whenever $|t-\xi| \geq c_1\bar{h}$ with $\bar{h}$ the density of Ξ in (1.7) for some $c_1 > 0$, we have

$$\sum_{\xi\in\Xi} |A(t,\xi)||t-\xi|^\nu \leq \text{const}\, \bar{h}^{n+1}, \quad |\nu|_1 = n+1. \tag{2.43}$$

Thus, recalling the property $\#\Xi_t < C$ with Ξ_t the support of the map A_t in (2.7), we obtain the bound

$$\begin{aligned}
&\int_{\mathbb{R}^d} |\phi_c(x-t) - \phi_c(x,t)|dt \\
&= \int_{\mathbb{R}^d} |\int_0^1 \frac{(1-y)^n}{n!} \sum_{|\nu|_1=n+1} \sum_{\xi\in\Xi} A(t,\xi)(t-\xi)^{n+1} D^\nu \phi_c(x-t+y(x-\xi))dy|dt \\
&\leq \text{const}\, \bar{h}^{n+1} \sum_{\xi\in\Xi_t} \int_0^1 \frac{(1-y)^n}{n!} \int_{\mathbb{R}^d} | \sum_{|\nu|_1=n+1} D^\nu \phi_c(x-t+y(x-\xi))|\, dt\, dy \\
&\leq \text{const}_1\, \bar{h}^{n+1}
\end{aligned}$$

where the last inequality is implied by Lemma 2.10. We established the required result. ■

To complete the proof of the theorem, we need to show that

$$\sigma_\omega^\vee f - R_\Xi f = \int_{\mathbb{R}^d} (\phi_c(\cdot - t) - \phi_c(\cdot, t)) f^*(t)dt.$$

Invoking the relation(2.38), we have the identities

$$\begin{aligned}
(\sigma_\omega^\vee f - R_\Xi f)(x) &= \int_{\mathbb{R}^d} (\psi_{c/\omega}(x/\omega - t) - \psi_{c/\omega}(x/\omega, t)) \Big(\frac{\sigma_\omega \hat{f}}{\hat{\phi}_c}\Big)^\vee (t)dt \\
&= \omega^{-\lambda-d} \int_{\mathbb{R}^d} \tau_\omega^\vee * (\phi_c(x-\cdot) - \phi_c(x,\cdot))(t) \Big(\frac{\sigma_\omega \hat{f}}{\hat{\psi}_{c/\omega}(\omega\cdot)}\Big)^\vee (t)dt
\end{aligned}$$

$$= \omega^{-\lambda-d} \int_{\mathbb{R}^d} \tau_\omega^\vee * \Big(\frac{\sigma_\omega \hat{f}}{\hat{\psi}_{c/\omega}(\omega\cdot)} \Big)^\vee (t)(\phi_c(x-t) - \phi_c(x,t))dt$$
$$= \int_{\mathbb{R}^d} (\phi_c(x-\cdot) - \phi_c(x,\cdot))(t) f^*(t)dt$$

with $\tau^\vee$ is (2.36), which establishes our theorem. ■

Corollary 2.13 *Let $\phi_c(\cdot,t)$, $t \in \mathbb{R}^d$, and R_Ξ be as above. Let $(A(t,\xi))_{\xi\in\Xi}$ for $\phi_c(\cdot,t)$ be admissible for Π_n with $n > \lambda + d$. Then, for every function f such that $\hat{f}\hat{\phi}_c^{-1} \in L_1(\mathbb{R}^d)$, we have*

$$\|f - R_\Xi f\| \leq \text{const}\, \bar{h}^{n+1}$$

with const *dependent on c.*

Proof. Using the properties (2.13) and (2.15), this follows easily from the theorem above. ■

Remark 2.14 The localized pseudo-shift $\psi_{c/\omega}(\cdot/\omega,\alpha)$ is obtained by an application of the same matrix $(A(\cdot,\xi))_{\xi\in\Xi}$ to the scattered shifts of $\phi_{c/\omega}(\cdot/\omega - \alpha)$. However, a careful examination of the function $R_\Xi f$ reveals that $R_\Xi f$ is not in the space $S_\Xi(\phi_c)$, but in the space $S_{\omega\Xi}(\phi_c)$. Since our goal is to approximate from the space $S_\Xi(\phi_c)$ the fact that the dilated center set $h\Xi$ is employed as h changes should not occur. For this reason, at each h-level, we employ the center set $\omega^{-1}\Xi$ in the construction of $R_\Xi f$. Then, for each $t \in \mathbb{R}^d$, the pseudo-shift $\phi_{c/\omega}(\cdot/\omega,\alpha)$ should be of the form

$$\phi_{c/\omega}(x/\omega, t+\alpha) = \sum_{\xi\in\Xi} A(t+\alpha, \xi/\omega)\phi_{c/\omega}(x/\omega - \xi/\omega). \tag{2.44}$$

Then, the definition of $\psi_{c/\omega}(\cdot,t)$ leads to the explicit form

$$R_\Xi f(x) = \sum_{\xi\in\Xi} \phi_c(x-\xi) \sum_{\alpha\in\mathbb{Z}^d} \mu(\alpha) c_{\xi,\alpha}(f)$$

with

$$c_{\xi,\alpha}(f) := \int_{\mathbb{R}^d} A((\omega(t+\alpha),\xi)\Big(\frac{\sigma_\omega \hat{f}}{\hat{\psi}_{c/\omega}(\omega\cdot)}\Big)^{\vee}(\omega\, t)\, dt/\omega^\lambda.$$

It ensures that the approximant $R_\Xi f$ belongs to $S_\Xi(\phi_c)$.

2.3 Approximation Power of the Scheme R_Ξ

As the set Ξ becomes dense, we hope to approximate a function f better. First, we expect that the error $f-(\sigma_\omega \hat{f})^\vee$ is to be smaller as ω decreases to 0. However, since

$$f^* := \left(\frac{\sigma_\omega \hat{f}}{\hat{\phi}_c}\right)^\vee,$$

f^* can not be kept, in general, bounded as ω tends to zero since $\hat{\phi}_c$ decreases exponentially fast as observed in (2.13) and (2.15). Hence one of the important issues in our scheme is the choice of the proper parameter ω in accordance with the density of Ξ. In this section, we will study strategies for choosing the parameters c and ω according to the density of Ξ. Specifically, we will see how the parameters ω and c are interrelated.

Exploiting the relation $(gh)^\vee = g^\vee * h^\vee$, we have the identity

$$\begin{aligned}\left(\frac{\sigma_\omega \hat{f}}{\hat{\phi}_c}\right)^\vee(t) &= \frac{1}{\tilde{c}_\lambda}\int_{\mathbb{R}^d} f(t-\theta)\left(\frac{|\cdot|^{(\lambda+d)/2}\sigma_\omega}{\tilde{K}_{(\lambda+d)/2}(c\cdot)}\right)^\vee(\theta)d\theta \\ &= \frac{1}{\tilde{c}_\lambda}\int_{\mathbb{R}^d} f(t-c\theta)\left(\frac{\sigma_{\omega/c}|\cdot|^{(\lambda+d)/2}}{c^{\lambda+d}\tilde{K}_{(\lambda+d)/2}}\right)^\vee(\theta)d\theta \qquad (2.45)\end{aligned}$$

for a function $f \in \widetilde{W}^k_\infty(\mathbb{R}^d)$. Here, we note that the parameters c and ω can be managed simultaneously by controlling the ratio c/ω rather than controlling ω and c independently. For simplicity, we will use the abbreviation

$$\rho := \frac{c}{\omega}, \qquad \omega,\ c > 0.$$

Since supp$\sigma_{1/\rho} = B_{\rho\eta}$ with B_η =suppσ only the values of $\tilde{K}_{(\lambda+d)/2}$ on $B_{\rho\eta}$ are incorporated into the scheme regardless of the density of Ξ.

Furthermore, for $f \in \widetilde{W}^k_\infty(\mathbb{R}^d)$, $k \geq \lambda + d$, a variant of the expression in (2.45) is as follows:

$$\begin{aligned} f^* &= \frac{1}{\tilde{c}_\lambda}(|\cdot|^{\lambda+d}(1+|c\cdot|^2)^{k'/2}\hat{f})^\vee * \left(\frac{\sigma_\omega}{(1+|c\cdot|^2)^{k'/2}\tilde{K}_{(\lambda+d)/2}(c\cdot)}\right)^\vee(t) \qquad (2.46) \\ &= \frac{1}{\tilde{c}_\lambda}\int_{\mathbb{R}^d}(|\cdot|^{\lambda+d}(1+|c\cdot|^2)^{k'/2}\hat{f})^\vee(t-c\theta)\left(\frac{\sigma_{1/\rho}}{(1+|\cdot|^2)^{k'/2}\tilde{K}_{(\lambda+d)/2}}\right)^\vee(\theta)d\theta \end{aligned}$$

with $k' = k - \lambda - d$. Because $\|g\| \leq (2\pi)^{-d}\|\hat{g}\|_1$ for $\hat{g} \in L_1(\mathbb{R}^d)$, we estimate the bound

$$\begin{aligned} \|(|\cdot|^{\lambda+d}(1+|c\cdot|^2)^{(k-\lambda-d)/2}\hat{f})^\vee\| &\leq \|(|\cdot|^{\lambda+d}(1+|c\cdot|^2)^{(k-\lambda-d)/2}\hat{f})\|_1 \\ &\leq (1+c^{k-\lambda-d})\|f\|_{k,\infty}. \end{aligned}$$

From (2.46), we obtain that

$$\|f^*\| \leq \mathrm{const}_1\,(1+c^{k-\lambda-d})\;\|f\|_{k,\infty}$$

with const_1 dependent on ρ and k.

In a similar fashion, for a function $f \in \widetilde{W}^k_\infty(\mathbb{R}^d)$ with $k < \lambda + d$, we have the relation

$$c^{\lambda+d-k}f^* = \frac{1}{\tilde{c}_\lambda}\int_{\mathbb{R}^d}\left(\frac{|\cdot|^{\lambda+d-k}\sigma_{1/\rho}}{\tilde{K}_{(\lambda+d)/2}(c\cdot)}\right)(\theta)(1+|\cdot|^2)^{k/2}\hat{f})(\theta)e^{-it\cdot\theta}d\theta$$

Since $\lambda + d > k$, Lebesgue Dominated Convergence Theorem implies that $c^{\lambda+d-k}f^*$ tends to zero as $c \to 0$. Thus, we obtain that

$$\|f^*\| = o(c^{k-\lambda-d})$$

as $c \to 0$.

Therefore, the following lemma is established.

Lemma 2.15 *Let ϕ_c and f^* be defined as above. We assume that the parameters c and ω satisfy the relation $\rho := c/\omega$ for some fixed $\rho > 0$. Then, for every function $f \in \widetilde{W}^k_\infty(\mathbb{R}^d)$, we have*

$$\|f^*\| \leq \text{const} \begin{cases} (1 + c^{k-\lambda-d}) & \text{if } k \geq \lambda + d, \\ o(c^{k-\lambda-d}) & \text{if } k < \lambda + d, \end{cases}$$

with const *independent of c and ω, but dependent on ρ and k.*

Next, we discuss how the density of Ξ is related to the parameters c and ω.

Theorem 2.16 *Let $\phi_c(\cdot, t)$, $t \in \mathbb{R}^d$, and R_Ξ be defined as above. Let $\bar{h}$ be the density of Ξ in (1.7), and $(A(t,\xi))_{\xi\in\Xi}$ for $\phi_c(\cdot,t)$ be admissible for Π_n with $n > \lambda + d$. Assume further that the parameters c and ω satisfy the relation $\frac{c}{\omega} = \rho > 0$. If we choose $\omega = \bar{h}^r$ with $0 < r \leq 1$, then, for every $f \in \widetilde{W}^k_\infty(\mathbb{R}^d)$, we have*

$$\|f - R_\Xi f\| = o(\bar{h}^{rk}) + \begin{cases} O(\bar{h}^{(1-r)(n+1)+r(\lambda+d)}), & \text{if } k \geq \lambda + d, \\ o(\bar{h}^{(1-r)(n+1)+rk}), & \text{if } k < \lambda + d. \end{cases} \tag{2.47}$$

Remark 2.17 The reason of our choice $r \in (0,1)$ here is as following. When $k \geq \lambda + d$ and $\omega = h$, the approximation scheme becomes stationary, i.e., the approximation order is $\lambda + d$. In this theorem, however, the choice $\omega = h^r$ with $0 < r < 1$ induces nearly optimal approximation order $O(h^{rk})$ by taking sufficiently large n for a given r. Of course, if r is getting closer to 1, the degree of polynomial n should be larger enough and we need to solve larger linear system to get the approximation power $O(h^{rk})$. In case $k < \lambda + d$, there is no advantage in the choice of $r \in (0,1)$. Hence, taking $\omega = h$ brings approximation order $o(h^k)$.

Proof. For $f \in \widetilde{W}_\infty^k(\mathbb{R}^d)$, taking $\omega(\bar{h}) = \bar{h}^r$ with $0 < r \le 1$, it is clear from (2.41) that

$$\|f - (\sigma_\omega \hat{f})^\vee\| = o(\bar{h}^{rk}). \tag{2.48}$$

For the estimation of approximation error, we invoke Lemma 2.12 and then deduce by a change of variables the relation:

$$((\sigma_\omega \hat{f})^\vee - R_\Xi f)(x) = \int_{\mathbb{R}^d} (\phi_c(x - \bar{h}t) - \phi_c(x, \bar{h}t)) f^*(\bar{h}t)\bar{h}^d \, dt. \tag{2.49}$$

Using the property $\sum_{\xi\in\Xi} A(\cdot, \xi) = 1$ and the notation $\phi := \phi_1$, we claim that

$$\phi_c(x - \bar{h}t) - \phi_c(x, \bar{h}t) = c^\lambda \sum_{\xi\in\Xi} A(\bar{h}t, \xi)[\phi((x - \bar{h}t)/c) - \phi((x - \xi)/c)]. \tag{2.50}$$

For the case that d is odd, this equation follows immediately from the explicit formula in (2.12); if d is even, we first have the relation

$$\phi_c := c^\lambda \phi(\cdot/c) + (|\cdot|^2 + c^2)^{\lambda/2} \log c.$$

Correspondingly,

$$\begin{aligned}
& \sum_{\xi\in\Xi} A(\bar{h}t, \xi)(\phi_c(x - \bar{h}t) - \phi_c(x - h\xi)) \\
= \; & c^\lambda \sum_{\xi\in\Xi} A(\bar{h}t, \xi)[\phi((x - \bar{h}t)/c) - \phi((x - \xi)/c)] \\
& \qquad + \log c \sum_{\xi\in\Xi} A(\bar{h}t, \xi)[(|x - \bar{h}t|^2 + c^2)^{\lambda/2} - (|x - \xi|^2 + c^2)^{\lambda/2}].
\end{aligned}$$

As in the definition of ϕ_c, λ is even and $(|\cdot|^2 + c^2)^{\lambda/2}$ is polynomial. Hence, the second sum on the right hand side of the above equation is identically zero because, by assumption, $\sum_{\xi\in\Xi} A(t, \xi)p(\xi) = p(t)$, $t \in \mathbb{R}^d$, for $p \in \Pi_n$. This verifies our claim in (2.50).

Now, let $T_\theta f$ be the Taylor polynomial of f at θ of degree n, i.e.,

$$(T_\theta f)(s) = \sum_{|\nu|_1 \le n} D^\nu \phi(\theta)(s-\theta)^\nu/\nu! \qquad (s,\ \theta \in \mathbb{R}^d).$$

Then taking the Taylor expansion of $\phi((x-\xi)/c)$ about $\vartheta := (x-\bar{h}t)/c$ provides

$$\begin{aligned}\phi_c(x-\bar{h}t) - \phi_c(x,\bar{h}t) &= -c^\lambda \sum_{\xi\in\Xi} A(\bar{h}t,\xi) R_n(t,\xi) \\ &\quad + c^\lambda \sum_{\xi\in\Xi} A(\bar{h}t,\xi)[\phi((x-\bar{h}t)/c) - T_{(x-\bar{h}t)/c}\phi((x-\xi)/c)]\end{aligned} \tag{2.51}$$

where $R_n(t,\xi)$ is the remainder in the integral form

$$R_n(t,\xi) := \int_0^1 \frac{(1-y)^n}{n!} \sum_{|\nu|_1 = n+1} \left(\frac{\bar{h}t-\xi}{c}\right)^{n+1} D^\nu \phi\left(\frac{x-\bar{h}t+y(x-\xi)}{c}\right) dy.$$

Note that

$$\phi((x-\bar{h}t)/c) - T_{(x-\bar{h}t)/c}\phi((x-\xi)/c) = \sum_{0<|\nu|_1\le n} D^\nu\phi(\theta) c^{-|\nu|_1}(\bar{h}t-\xi)^\nu/\nu!.$$

Then, by the property that $\sum_{\xi\in\Xi} A(\cdot,\xi)p(\xi) = p$ for $p \in \Pi_n$ the second sum in (2.51) is identically zero. We arrive at the equality

$$\phi(x-\bar{h}t) - \phi(x,\bar{h}t) = c^\lambda \sum_{\xi\in\Xi} A(\bar{h}t,\xi) R_n(\phi((x-\xi)/c),(x-\bar{h}t)/c).$$

Now, recall $D^\nu\phi \in L^1(\mathbb{R}^d)$ with $|\nu|_1 = n+1$ and the relation in (2.43), i.e.,

$$\sum_{\xi\in\Xi} A(t,\xi)(t-\xi)^\nu \le \text{const}\, \bar{h}^{n+1}, \quad |\nu|_1 = n+1. \tag{2.52}$$

A similar calculation as in the proof of Lemma 2.12 yields that

$$\begin{aligned}&\int_{\mathbb{R}^d} |\phi(x-\bar{h}t) - \phi(x,\bar{h}t)|\bar{h}^d dt \\ &\le \text{const}\, \frac{\bar{h}^{n+d+1}}{c^{n+1-\lambda}} \sum_{\xi\in\Xi_{\bar{h}t}} \int_0^1 \frac{(1-y)^n}{n!} \int_{\mathbb{R}^d} | \sum_{|\nu|_1=n+1} D^\nu\phi\left(\frac{x-\bar{h}t+y(x-\xi)}{c}\right) |\, dt\, dy \\ &\le \text{const}'\, c^{\lambda+d-n-1}\bar{h}^{n+1}\end{aligned}$$

where Ξ_t is the support of the map $A(t,\cdot)$ and, by assumption, $\#\Xi_t < C$ for some $C > 0$. Thus, we derive from (2.49) the relation that

$$\|(\sigma_\omega \hat{f})^\vee - R_\Xi f\| \le \text{const}\, c^{\lambda+d-n-1}\bar{h}^{n+1}\|f^*\|.$$

Since $c = \rho\bar{h}^r$ by assumption, it is clear from Lemma 2.15 that

$$\|f^*\| \le \text{const}' \begin{cases} 1 & \text{if } k \ge \lambda + d, \\ o(\bar{h}^{r(k-\lambda-d)}) & \text{if } k < \lambda + d \end{cases}$$

for sufficiently small $\bar{h}$. Consequently, we get the theorem's claim. ■

Corollary 2.18 *Under the same conditions and notations as in the Theorem 2.16, let $\omega(\bar{h}) = \bar{h}$ and $c = \rho\bar{h}$ for some $\rho > 0$. Then, for every $f \in \widetilde{W}^k_\infty(\mathbb{R}^d)$,*

$$\|f - R_\Xi f\| = \begin{cases} O(\bar{h}^{\lambda+d}), & \text{if } k \ge \lambda + d, \\ o(\bar{h}^k), & \text{if } k < \lambda + d. \end{cases} \tag{2.53}$$

Chapter 3

Approximation by Scattered Data on a Bounded Domain

3.1 Approximation in the Interior of Bounded Domains

We discuss the more practical issue of approximating a function $f : \Omega \to \mathbb{R}$ using only its known values at finitely many scattered centers. Here, Ω is a bounded open set in $\mathbb{R}^d$. In this section, we look for the asymptotic rate of decay of the error on a closed subset of Ω, which is disjoint of the boundary of Ω, rather than on whole domain Ω. Hence, for $\delta > 0$, we define

$$\Omega_\delta := \{x \in \Omega \ : \ |x - y| \leq \delta \Rightarrow y \in \Omega\}.$$

We think of δ as either being fixed, or decreasing to 0 as the density of Ξ increases.

Some of our arguments deal with approximation to functions in the homogeneous L_∞-Sobolev space

$$w^k_\infty(\mathbb{R}^d), \qquad k \in \mathbb{Z}_+$$

of all functions whose derivatives of order k are bounded. We shall denote by $|\cdot|_{k,\infty}$ the homogeneous k-th order L_∞-Sobolev semi-norm on $w^k_\infty(\mathbb{R}^d)$, i.e.,

$$|f|_{k,\infty} := \sum_{|\alpha|_1 = k} \|D^\alpha f\|.$$

Our choice of radial basis function in this part is still the 'shifted thin-plate spline' ϕ_c, and its localized basis function ψ_c is assumed to satisfy the conditions

$$\sup_x (1+|x|)^{d+q}\psi_c(x) < \infty, \quad \hat{\psi}_c(0) \neq 0, \quad \text{and} \quad \hat{\psi}_c \in C^{d+q}(\mathbb{R}^d) \tag{3.1}$$

for some a positive integer q. In the case of approximation over $\mathbb{R}^d$, our approximant was of the form

$$R_\Xi f(x) = \int_{\mathbb{R}^d} \psi_{c/\omega}(x/\omega, t)\Lambda(f)(\omega t)\,dt \tag{3.2}$$

where Λ is an operator

$$\Lambda : f \longmapsto \int_{\mathbb{R}^d} \Big(\frac{\sigma_\omega}{\hat{\psi}_{c/\omega}(\omega\cdot)}\Big)^{\vee}(y-\theta)\, f(\theta)\,d\theta. \tag{3.3}$$

It is natural to truncate the above integrations properly to derive a suitable approximation on Ω. First, we need to pay attention to the construction of $\psi_c(\cdot,t)$, $t\in\Omega$, which approximates $\psi_c(\cdot-t)$, because our error estimation depends on the error $\psi_c(\cdot-t)-\psi_c(\cdot,t)$ (as a matter of fact, on $\phi_c(\cdot,t)-\phi_c(\cdot,t)$). Specifically, we note that the error is large when t is far away from all the centers while if there exist sufficiently many points from Ξ around t, we get a tight error bound. Since

$$\psi_c(\cdot/\omega, t) = \sum_{\alpha\in\mathbb{Z}^d} \mu(\alpha)\phi_c(\cdot/\omega, t+\alpha),$$

each $t+\alpha$ is supposed to be surrounded by centers from Ξ. Thus, we take $\psi_c(\cdot,t)$ into account only for $t\in\Omega_\delta$, and then, correspondingly, we truncate the function $\Lambda(f)$ as follows:

$$\Lambda(\chi_\Omega f)(t) = \int_\Omega \Big(\frac{\sigma_\omega}{\hat{\psi}_{c/\omega}(\omega\cdot)}\Big)^{\vee}(t-\theta)\, f(\theta)\,d\theta, \qquad t\in\Omega_\delta. \tag{3.4}$$

On the other hand, since the only known information of f is $f|_\Xi$, we approximate other function values on Ω:

$$f_C(\theta) := \sum_{\xi\in\Xi} C(\theta,\xi) f(\xi), \quad \theta \in \Omega. \tag{3.5}$$

Here, we assume that the coefficients $(C(\cdot,\xi))_{\xi\in\Xi}$ are admissible for Π_m for a positive integer m. From the definition of admissible coefficients, we note that the map $f \mapsto f_C$ is exact on Π_m. Thus, when we can show that f_C provides a "reasonable" approximation to f on Ω, we accept this approximant as an initial replacement for f. In order to discuss the approximation power, let us measure the density of Ξ as before by

$$\bar{h} := \sup_{x\in\Omega} \min_{\xi\in\Xi} |x-\xi|. \tag{3.6}$$

We describe the approximation properties of f_C to f on Ω in next lemma.

Lemma 3.1. *Let f_C be given by (3.5), and let the coefficients $(C(\cdot,\xi))_{\xi\in\Xi}$ for f_C be admissible for Π_m. Then, for $f \in w^k_\infty(\mathbb{R}^d) \cap C^k(\mathbb{R}^d)$ with $k > m$, we have*

$$\|f - f_C\|_{L_\infty(\Omega)} \le \text{const}\, \bar{h}^{m+1} |f|_{m+1,\infty}$$

with const *independent of* Ξ.

Proof. Since $\sum_{\xi\in\Xi} C(\cdot,\xi) = 1$ by definition, it is clear that

$$(f - f_C)(\theta) = \sum_{\xi\in\Xi} C(x,\xi)(f(\theta) - f(\xi)), \quad \theta\in\Omega.$$

Letting $T_\theta f$ be the Taylor polynomial of degree m of f about $\theta\in\Omega$, the function $\{f(\theta) - T_\theta f(y) : y \in \mathbb{R}^d\}$ is also a polynomial in Π_m that matches f and all its derivatives of degree up to m at $x = y$ and then, by the definition of admissible coefficients $(C(\cdot,\xi))_{\xi\in\Xi}$, we have

$$\sum_{\xi\in\Xi} C(\theta,\xi)(f(\theta) - T_\theta f(\xi)) = 0.$$

This induces the expression

$$|f(\theta) - f_C(\theta)| = |\sum_{\xi\in\Xi} C(\theta,\xi) \sum_{|\alpha|_1=n+1} (\theta-\xi)^\alpha D^\alpha f(\xi_\theta)|$$

with $\xi_\theta \in [\theta \,.\,.\, \xi]$. Further, due to the property $C(\theta,\xi) = 0$ for such ξ that $|\theta-\xi| > c_1\bar{h}$ for some $c_1 > 0$, we estimate that

$$\begin{aligned} |f(\theta) - f_C(\theta)| &= |\sum_{\xi\in\Xi} C(\theta,\xi) \sum_{|\alpha|_1=n+1} (\theta-\xi)^\alpha D^\alpha f(\xi_\theta)| \\ &\le \text{const}\, \bar{h}^{n+1} |f|_{n+1,\infty} \sum_{\xi\in\Xi} |C(\theta,\xi)|. \end{aligned}$$

Therefore, since

$$\|\sum_{\xi\in\Xi} |C(\cdot,\xi)|\|_{L_\infty(\Omega)} < \infty,$$

we obtain the lemma's claim. ■

With the definition of f_C in (3.5), we define an approximation scheme by

$$R_\Xi : f \mapsto \int_{\Omega_\delta/\omega} \psi_{c/\omega}(\cdot/\omega, t)\Lambda(\chi_\Omega f_C)(\omega t)dt$$

with the operator Λ in (3.3) and χ_Ω the characteristic function of Ω. Recall that the function $\psi_{c/\omega}(\cdot/\omega, t)$ is a linear combination of the pseudo-shifts

$$\phi_{c/\omega}(\cdot/\omega, t) = \sum_{\xi\in\Xi} A(t,\xi)\phi_{c/\omega}(\cdot/\omega - \xi/\omega).$$

Now, for $x \in \Omega_{2\delta}$, the error analysis is based on the estimate

$$|(f - R_\Xi f)(x)| \leq |(\sigma_\omega \hat{f})^\vee(x) - \int_{\Omega_\delta/\omega} \psi_{c/\omega}(x/\omega - t)\,\Lambda(f)(\omega t)dt| \tag{3.7}$$
$$+ |\int_{\Omega_\delta/\omega} (\psi_{c/\omega}(x/\omega - t) - \psi_{c/\omega}(x/\omega, t))\,\Lambda(f)(\omega t)dt|$$
$$+ |\int_{\Omega_\delta/\omega} \psi_{c/\omega}(x/\omega, t)(\Lambda(f) - \Lambda(\chi_\Omega f))(\omega t)dt|$$
$$+ |\int_{\Omega_\delta/\omega} \psi_{c/\omega}(x/\omega, t)(\Lambda(\chi_\Omega f) - \Lambda(\chi_\Omega f_C))(\omega t)dt|$$
$$+ |f(x) - (\sigma_\omega \hat{f})^\vee(x)|$$

with the operator Λ form (3.3).

We have observed earlier that the last term in (3.7) satisfies

$$\|f - (\sigma_\omega \hat{f})^\vee\| = o(\omega^k)$$

for every function $f \in \widetilde{W}^k_\infty(\mathbb{R}^d)$.

The next lemma treats the first term in (3.7).

Lemma 3.2. *Let ψ_c satisfy the conditions in (3.1) and Λ be as in (3.3). Assume further that the relation $c = \rho\,\omega$ holds for some $\rho > 0$. Then, for every $f \in L_\infty(\mathbb{R}^d)$,*

$$\|(\sigma_\omega \hat{f})^\vee - \int_{\Omega_\delta/\omega} \psi_{c/\omega}(\cdot/\omega - t)\Lambda(f)(\omega t)dt\|_{L_\infty(\Omega_{2\delta})} \leq \text{ const } (\omega/\delta)^q,$$

where q is as in (3.1).

Proof. Using the representation

$$f = \int_{\mathbb{R}^d} \phi_c(\cdot - t)\Big(\frac{\hat{f}}{\hat{\phi}_c}\Big)^\vee(t)dt,$$

we have the expression

$$(\sigma_\omega \hat{f})^\vee - \int_{\Omega_\delta/\omega} \psi_{c/\omega}(x/\omega - t)\Lambda(f)(\omega t)dt = \int_{\Omega'_\delta/\omega} \psi_{c/\omega}(x/\omega - t)\Lambda(f)(\omega t)dt \tag{3.8}$$

where Ω' indicates the complement set of Ω in $\mathbb{R}^d$. Since

$$
\begin{aligned}
|\Lambda(f)(y)| &= |\int_{\mathbb{R}^d} \left(\frac{\sigma_\omega}{\hat{\psi}_{c/\omega}(\omega\cdot)}\right)^\vee(\theta)\, f(y-\theta)\, d\theta| \\
&= |\int_{\mathbb{R}^d} \left(\frac{\sigma}{\hat{\psi}_\rho}\right)^\vee(\theta)\, f(y-\omega\theta)\, d\theta| \le \|f\|\, \|\left(\frac{\sigma}{\hat{\psi}_\rho}\right)^\vee\|_1,
\end{aligned}
$$

we can bound (3.8) by a constant multiple of

$$
\int_{\Omega'_\delta/\omega} |\psi_{c/\omega}(x/\omega - t)|\, dt.
$$

Furthermore, under the relation $\rho = c/\omega$, we have a bound of $\psi_{c/\omega}$ as in (3.1)

$$
\psi_{c/\omega} = \psi_\rho \le \text{ const}\, |\cdot|^{-d-q},
$$

with q in (3.1) and const independent of ω and c, but dependent on ρ. Since $|x-t| > \delta$ for $x \in \Omega_{2\delta}$ and $t \in \Omega'_\delta$, a direct calculation yields

$$
\begin{aligned}
\int_{\Omega'_\delta/\omega} |\psi_{c/\omega}(x/\omega - t)| dt &\le \text{const}\, \omega^q \int_{\Omega'_\delta} \frac{1}{(\omega + |x-t|)^{d+q}}\, dt \\
&\le \text{const}\, \omega^q \int_{B'_\delta} \frac{1}{(\omega + |t|)^{d+q}}\, dt \\
&= \text{const}\, \omega^q\, \sigma(S^{d-1}) \int_\delta^\infty \frac{r^{d-1}}{(\omega + r)^{d+q}}\, dr \le \text{const}_1 \left(\frac{\omega}{\delta}\right)^q
\end{aligned}
$$

where $S^{d-1} = \{x \in \mathbb{R}^d \,:\, |x| = 1\}$, and $\sigma(S^{d-1})$ is the measure associated with spherical coordinates. ■

Finding the approximation order of the second term in (3.7) was the focal point of our study in the case $\Omega = \mathbb{R}^d$ (see Section 2.3 and 2.4). We obtain the exact same approximation property on the bounded domain with the admissible coefficients for Π_n for the approximation

$$
\phi_c(\cdot, t) = \sum_{\xi \in \Xi} A(t,\xi)\phi_c(\cdot - \xi) \approx \phi(\cdot - t) \tag{3.9}
$$

where the coefficient sequence $(A(\cdot,\xi))_{\xi\in\Xi}$ is admissible for Π_n. In fact, for functions $f \in \widetilde{W}^k_\infty(\mathbb{R}^d)$, we have the following convergence property

$$\| \int_{\Omega_{\delta/\omega}} (\psi_{c/\omega}(x/\omega - t) - \psi_{c/\omega}(x/\omega, t))\, \Lambda(f)(\omega t) dt \|_{L_\infty(\Omega_{2\delta})}$$

$$= \begin{cases} O(\bar{h}^{(1-r)(n+1)+r(\lambda+d)}) & \text{if } k \geq \lambda + d, \\ o(\bar{h}^{(1-r)(n+1)+rk}) & \text{if } k < \lambda + d, \end{cases}$$

with $\bar{h}$ the density of Ξ in (3.6), in the case $\omega = \bar{h}^r$ $(0 < r \leq 1)$ and $c = \rho\omega$ for some $\rho > 0$. We note that the lemmas in this section are stated with respect to $w^k_\infty(\mathbb{R}^d) \cap C^k(\mathbb{R}^d)$. But, since $\widetilde{W}^k_\infty(\mathbb{R}^d)$ is continuously embedded into $w^k_\infty(\mathbb{R}^d) \cap C^k(\mathbb{R}^d)$, those results are valid for the smaller space $\widetilde{W}^k_\infty(\mathbb{R}^d)$.

A bound of the third term in (3.7) is obtained in next lemma.

Lemma 3.3. *Let ψ_c satisfy the conditions in (3.1), and let the operator Λ be defined as in (3.3). Assume that the relation $c = \rho\,\omega$ hold for some $\rho > 0$. Then, for every $f \in L_\infty(\mathbb{R}^d)$, we have*

$$\| \int_{\Omega_{\delta/\omega}} \psi_{c/\omega}(x/\omega - t)(\Lambda(f) - \Lambda(\chi_\Omega f))(\omega t)\, dt \|_{L_\infty(\Omega_{2\delta})} \;\leq\; const\,(\omega/\delta)^q,$$

where const *is independent of Ξ, and q is as in (3.1). Furthermore, if $\delta = \delta(\bar{h})$ and $\omega = \omega(\bar{h})$ such that $\omega/\delta \to 0$ as $\bar{h} \to 0$, then*

$$\| \int_{\Omega_{\delta/\omega}} \psi_{c/\omega}(x/\omega - t)(\Lambda(f) - \Lambda(\chi_\Omega f))(\omega t)\, dt \|_{L_\infty(\Omega_{2\delta})} \;\leq\; o(\omega/\delta)^q.$$

Proof. It is clear from the condition (3.1) that $\|\psi_{c/\omega}\|_1 = \|\psi_\rho\|_1 < \infty$, hence we estimate

$$\| \int_{\Omega_{\delta/\omega}} \psi_{c/\omega}(x/\omega - t)(\Lambda(f) - \Lambda(\chi_\Omega f))(\omega t)\, dt \|_{L_\infty(\Omega_{2\delta})}$$
$$\leq \|\psi_\rho\|_1 \|\Lambda(f) - \Lambda(\chi_\Omega f)\|_{L_\infty(\Omega_{2\delta})}.$$

For $y \in \Omega_\delta$, it follows from the explicit formula for $\Lambda(f)$ and $\Lambda(\chi_\Omega f)$ that

$$\begin{aligned}
(\Lambda(f) - \Lambda(\chi_\Omega f))(y) &= \int_{\Omega'} \Big(\frac{\sigma_\omega}{\hat{\psi}_{c/\omega}(\omega \cdot)} \Big)^{\vee} (y-\theta)\, f(\theta)\, d\theta \\
&= \int_{y-\Omega'} f(y-\theta) \Big(\frac{\sigma}{\hat{\psi}_\rho} \Big)^{\vee} (\theta/\omega)\, d\theta/\omega^d \\
&= \omega^q \int_{y-\Omega'} f(y-\theta)\, \frac{g(\theta/\omega)}{|\theta|^{\lambda+d}}\, d\theta,
\end{aligned}$$

with

$$g = |\cdot|^{q+d} \Big(\frac{\sigma}{\hat{\psi}_\rho} \Big)^{\vee}.$$

Hence, we have the bound of $\Lambda(f) - \Lambda(\chi_\Omega f)$ as following

$$\|\Lambda(f) - \Lambda(\chi_\Omega f)\| \le \omega^q\, \|f\| \int_{B'_\delta} \frac{|g(\theta/\omega)|}{|\theta|^{q+d}}\, d\theta = \Big(\frac{\omega}{\delta}\Big)^q \|f\| \int_{B'_1} \frac{|g(\delta\theta/\omega)|}{|\theta|^{q+d}}\, d\theta. \quad (3.10)$$

Since, as in (3.1),

$$\sigma \hat{\psi}_\rho^{-1} \in C^{d+q}(\mathbb{R}^d) \cap C^\infty(\mathbb{R}^d \setminus 0)$$

and $\theta \in B'_1$, it follows from the Riemann-Lebesgue lemma that $|g(\delta\theta/\omega)| \to 0$ as $\omega/\delta \to 0$, and it implies that the last integral in (3.10) tends to 0. ∎

A bound for the fourth term in (3.7) is deduced directly from Lemma 3.1 and relation (3.1).

Lemma 3.4. *Let ψ_c and the operator Λ be defined as above, and let the relation $c = \rho\,\omega$ hold for some $\rho > 0$. Assume that the coefficient sequence $(C(\cdot,\xi))_{\xi\in\Xi}$ for f_C is admissible for Π_m. Then, for every $f \in W^k_\infty(\mathbb{R}^d) \cap C^k(\mathbb{R}^d)$ with $k > m$, we have the relation*

$$\| \int_{\Omega_\delta/\omega} \psi_{c/\omega}(\cdot/\omega - t)(\Lambda(\chi_\Omega f) - \Lambda(\chi_\Omega f_C))(\omega t)\, dt \|_{L_\infty(\Omega_{2\delta})} \le O(\bar{h}^{m+1}).$$

Proof. For $y \in \Omega_\delta$, it follows from the explicit formula of $\Lambda(\chi_\Omega f)$ and $\Lambda(\chi_\Omega f_C)$ that

$$(\Lambda(\chi_\Omega f) - \Lambda(\chi_\Omega f_C))(y) \quad = \quad \int_\Omega \Big(\frac{\sigma_\omega}{\hat{\psi}_{c/\omega}(\omega\cdot)} \Big)^\vee (y-\theta)\,(f(\theta) - f_C(\theta))\,d\theta$$

with f_C in (3.5). Since

$$\int_\Omega \Big(\frac{\sigma_\omega}{\hat{\psi}_{c/\omega}(\omega\cdot)} \Big)^\vee (y-\theta)\,d\theta = \int_\Omega \Big(\frac{\sigma}{\hat{\psi}_\rho} \Big)^\vee (y/\omega - \theta)\,d\theta \le \| \Big(\frac{\sigma}{\hat{\psi}_\rho} \Big)^\vee \|_1 \le \infty,$$

by Lemma 3.1, we obtain the bound

$$\|\Lambda(\chi_\Omega f) - \Lambda(\chi_\Omega f_C)\|_{L_\infty(\Omega_\delta)} \le \text{const}\, \bar{h}^{m+1} |f|_{m+1,\infty}.$$

Invoking the condition (3.1), we see that $\|\psi_{c/\omega}\|_1 = \|\psi_\rho\|_1 < \infty$, and this leads to the estimation

$$\| \int_{\Omega_\delta/\omega} \psi_{c/\omega}(x/\omega - t)(\Lambda(\chi_\Omega f) - \Lambda(\chi_\Omega f_C))(\omega t)\,dt \|_{L_\infty(\Omega_{2\delta})} \le \text{const}\, \bar{h}^{m+1}.$$

■

Theorem 3.5. *Let ϕ_c, ψ_c, $\phi_c(\cdot,t)$, $t \in \Omega$, and R_Ξ be as above, and let ψ_c satisfy the conditions in (3.1). Let $f \in \widetilde{W}^k_\infty(\mathbb{R}^d)$ for a positive integer k. We assume further that the following hold:*

(a) *For $t \in \Omega$, the pseudo-shift $\phi_c(\cdot,t)$ is defined as in (3.9) and the coefficients $(A(t,\xi))_{\xi\in\Xi}$ for $\phi_c(\cdot,t)$ are admissible for Π_n with $n > \lambda + d$.*

(b) *The function f_C is defined as in (3.5), and the coefficients $(C(\cdot,\xi))_{\xi\in\Xi}$ for f_C are admissible for Π_{k-1} in $\Omega_{2\delta}$.*

(c) *The relation* $\rho = c/\omega > 0$ *holds with* $\omega(\bar{h}) = \bar{h}$.

Then, if $k < \lambda + d$ *and* $q > \lambda + d$ *with* q *in (3.1),*

$$\|f - R_\Xi f\|_{L_\infty(\Omega_{2\delta})} = o(\bar{h}^k) \tag{3.11}$$

where const *is independent of* Ξ, *but depends on* f *and* δ.

Theorem 3.6. *Let* ϕ_c, ψ_c, $\phi_c(\cdot, t)$, $t \in \Omega$, *and* R_Ξ *be as above, and let* ψ_c *satisfy the conditions in (3.1). Let* $f \in \widetilde{W}^k_\infty(\mathbb{R}^d)$ *for some* $k \in \mathbb{Z}_+$. *We assume the condition (a-b) in Theorem 3.5 above, and assume further that* $\rho = c/\omega > 0$ *with* $\omega(h) = h^r$, $0 < r \leq 1$, *Then, if* $k \geq \lambda + d$,

$$\|f - R_\Xi f\|_{L_\infty(\Omega_{2\delta})} \leq \mathrm{const}\, \bar{h}^{rq} + o(\bar{h}^{rk}) + O(\bar{h}^{(1-r)(n+1)+r(\lambda+d)}) \tag{3.12}$$

where q *is as in (3.1), and* const *is independent of* Ξ, *but depends on* δ.

Remark: Note that for the case studied in the above theorems, δ must be held fixed, so that the approximation orders in (3.11) and (3.12) are proved only for fixed subsets of Ω. Alternatively, Lemmas 3.2 and 3.3 show that one should restrict $\delta = \delta(\bar{h})$ by assuming it to decrease to 0 slower than $\omega(\bar{h})$ in order to get approximation order as high as possible. For example, in the case of $\omega(\bar{h}) = \bar{h}^r$ with $rq > \lambda + d$, $\delta(\bar{h})$ should be

$$\delta(\bar{h}) \geq c\,\bar{h}^{r-(\lambda+d)/q}$$

for the approximation order $\lambda + d$.

3.2 Error Estimates near the Boundary

Approximation near the boundary is a difficult problem. Because data information is unavailable outside of the domain, deterioration in fidelity of the approximation near the boundary is unavoidable. For example, an asymptotic upper bound on the approximation order of thin-plate interpolation on the unit ball in $\mathbb{R}^2$ is $O(\bar{h}^2)$ while $O(\bar{h}^4)$ is available inside the domain. The reader is referred to the paper [J2]. Actually, since the available information on f is only $f|_\Xi$ with $\Xi \subset \Omega$, some special care is necessary in order to eliminate the boundary effects. We approach this problem by adding additional pseudo-shifts $\psi_c(\cdot, t)$ around boundary. For the construction of these additional functions, we add new centers to Ξ. Accordingly, we need the function values at those new centers. First we define a superset

$$\tilde{\Omega}_\delta := \{y = x + z \,:\, x \in \Omega \text{ and } |z| \le \delta\} = \Omega + B_\delta \tag{3.13}$$

of Ω, and add some extra centers on $\tilde{\Omega}_{2\delta} \setminus \Omega$ to be used for the construction of $\psi_c(\cdot, t)$, $t \in \tilde{\Omega}_\delta$. We denote by

$$\tilde{\Xi}$$

the extended center set in $\tilde{\Omega}_{2\delta}$, and let

$$\bar{h} := \sup_{x \in \tilde{\Omega}_{2\delta}} \min_{\xi \in \tilde{\Xi}} |x - \xi|. \tag{3.14}$$

Then, for $t \in \tilde{\Omega}_{2\delta}$, we produce extrapolated values of f by applying the matrix $(C(t, \xi))_{\xi \in \Xi}$ to $f|_\Xi$ as follows:

$$f_C(t) = \sum_{\xi \in \Xi} C(t, \xi) f(\xi). \tag{3.15}$$

The coefficients $(C(\cdot, \xi))_{\xi \in \Xi}$ are assumed as before to be admissible for Π_m for some nonnegative integer m. Hence, the map $f \mapsto f_C$ has degree of polynomial reproduction $\ge m$.

In fact, extrapolation contains a strong element of uncertainty even under the best circumstances. We describe this approximation property of f_C to f on Ω in next lemma.

Lemma 3.7. *Let f_C be given by (3.15) and the coefficients $(C(\cdot,\xi))_{\xi\in\Xi}$ be admissible for Π_m. Let $\bar h$ be the density of $\tilde\Xi$ in (3.14). Then, for $f \in w^k_\infty(\mathbb{R}^d)\cap C^k(\mathbb{R}^d)$ with $k > m$, we have*

$$\|f - f_C\|_{L_\infty(\tilde\Omega_\delta)} \le [\mathrm{c}_1\delta + \mathrm{c}_2\bar h]^{m+1}|f|_{m+1,\infty}$$

with c_1 and c_2 independent of Ξ and δ.

Proof. By applying the same idea as in Lemma 3.1, we first get the expression

$$|f(\theta) - f_C(\theta)| \;=\; |\sum_{|\alpha|_1=m}\sum_{\xi\in\Xi} C(\theta,\xi)(\theta-\xi)^\alpha f^{(\alpha)}(\xi_\theta)|, \quad \theta\in\tilde\Omega_\delta,$$

with $\xi_\theta \in [\theta\,..\,\xi]$. Since $C(\theta,\xi) = 0$ whenever $|\theta-\xi| > c_1\bar h$ for some $c_1 > 0$, we deduce that

$$\sum_{\xi\in\Xi} C(\theta,\xi)|\theta-\xi|^\alpha \le [\mathrm{dist}\,(\theta,\Omega) + \mathrm{const}\;\bar h]^m \sum_{\xi\in\Xi}|C(\theta,\xi)|.$$

Therefore, since

$$\sup_{\theta\in\tilde\Omega_\delta}\sum_{\xi\in\Xi}|C(\theta,\xi)| < \infty,$$

we obtain the lemma's claim. ∎

Now, for $x\in\Omega$, we define our approximation scheme on Ω by

$$\tilde R_\Xi f(x) = \int_{\tilde\Omega_\delta/\omega} \psi_{c/\omega}(x/\omega, t)\,\Lambda(\chi_{\tilde\Omega_{2\delta}} f_C)(\omega t)\,dt$$

where

$$\Lambda(\chi_{\tilde\Omega_{2\delta}} f_C)(y) := \int_{\tilde\Omega_{2\delta}} \left(\frac{\sigma_\omega}{\hat\psi_{c/\omega}(\omega\cdot)}\right)^{\vee}(y-\theta)\,f_C(\theta)\,d\theta, \qquad y\in\tilde\Omega_\delta, \tag{3.16}$$

with $f_C(\theta)$ in (3.15).

Our approach for the error analysis is now done by estimating each of the terms in the error bound

$$\begin{aligned}|(f-\tilde{R}_\Xi f)(x)| \;\leq\;& |(\sigma_\omega \hat{f})^\vee - \int_{\tilde{\Omega}_\delta/\omega} \psi_{c/\omega}(x/\omega - t)\,\Lambda(f)(\omega t)dt| \qquad (3.17)\\ &+ |\int_{\tilde{\Omega}_\delta/\omega} (\psi_{c/\omega}(x/\omega - t) - \psi_{c/\omega}(x/\omega, t))\,\Lambda(f)(\omega t)dt|\\ &+ |\int_{\tilde{\Omega}_\delta/\omega} \psi_{c/\omega}(x/\omega, t)\,(\Lambda(f) - \Lambda(\chi_{\tilde{\Omega}_{2\delta}} f))(\omega t)dt|\\ &+ |\int_{\tilde{\Omega}_\delta/\omega} \psi_{c/\omega}(x/\omega, t)(\Lambda(\chi_{\tilde{\Omega}_{2\delta}} f) - \Lambda(\chi_{\tilde{\Omega}_{2\delta}} f_C))(\omega t)dt|\\ &+ |f - (\sigma_\omega \hat{f})^\vee(x)| \qquad (3.18)\end{aligned}$$

with the operator Λ in (3.3). The terms in this bound are exactly the same as in the previous discussion except the fourth term which is related to the extrapolation.

The next lemma treats the fourth term in (3.17). It is immediate from the Lemma 3.7 and relation (3.1).

Lemma 3.8. *Let ψ_c and Λ be defined as above, and let $\bar{h}$ be the density of $\tilde{\Xi}$ in (3.14). We assume that the relation $c = \rho\,\omega$ holds for some $\rho > 0$, and the coefficient sequence $(C(\cdot,\xi))_{\xi\in\Xi}$ for f_C is admissible for Π_{k-1} with a positive integer k. Then, for $f \in w^k_\infty(\mathbb{R}^d) \cap C^k(\mathbb{R}^d)$, we have*

$$\| \int_{\tilde{\Omega}_\delta/\omega} \psi_{c/\omega}(\cdot/\omega - t)(\Lambda(\chi_{\tilde{\Omega}_{2\delta}} f) - \Lambda(\chi_\Omega f_C))(\omega t)\,dt\|_{L_\infty(\Omega)} \;\leq\; O((\delta + \bar{h})^k).$$

Theorem 3.9. *Let ϕ_c, ψ_c and $\tilde{R}_\Xi$ be as above, and let ψ_c satisfy the conditions in (3.1). Let $\bar{h}$ be the density of $\tilde{\Xi}$ in (3.14), and assume that $\delta = \delta(\bar{h}) \to 0$ as $\bar{h}$ tends to 0. We assume further that the following hold:*

(a) *For $t \in \tilde{\Omega}_{2\delta}$, the pseudo-shift $\phi_c(\cdot, t)$ is as in (3.9), and the coefficient sequence $(A(t,\xi))_{\xi\in\Xi}$ for $\phi_c(\cdot,t)$ is admissible for Π_n with a integer $n > \lambda + d$.*

(b) *For* $t \in \tilde{\Omega}_{2\delta}$, *the function* f_C *is defined as in (3.5), and the coefficient sequence* $(C(t,\xi))_{\xi\in\Xi}$ *for* f_C *is admissible for* Π_{k-1}.

(c) *The relation* $\rho = c/\omega > 0$ *holds with* $\omega(\bar{h}) = \bar{h}^r$, $0 < r \le 1$, *and* $\omega/\delta \to 0$ *as* $\bar{h}$ *tends to* 0.

Then, for every $f \in \widetilde{W}_\infty^k(\mathbb{R}^d)$, *we obtain*

$$\|f - \tilde{R}_\Xi f\|_{L_\infty(\Omega)} \le O(\,(\bar{h}^r/\delta)^q) + o(\delta^k) + \begin{cases} O(\bar{h}^{(1-r)(n+1)+r(\lambda+d)}) & \text{if } k \ge \lambda + d, \\ O(\bar{h}^{(1-r)(n+1)+rk}) & \text{if } k < \lambda + d, \end{cases}$$

where q *is as in (3.1).*

3.3 Smoothing Noisy Data

We assume that the data $(y_\xi)_{\xi\in\Xi}$ arise according to the model

$$y_\xi = f(\xi) + \epsilon_\xi$$

where the ξ belong to $\Omega \subset \mathbb{R}^d$ and, for example, ϵ_ξ are independent normally distributed random variables with mean 0 and (known or unknown) variance σ^2. Here $f : \mathbb{R}^d \to \mathbb{R}$ is assumed to be a smooth function. It is desired to recover an estimate of f from given data $(\xi, y_\xi)_{\xi\in\Xi}$. So we apply the same scheme for the purpose of smoothing the noisy data. As in the general algorithm, our smoothing procedure with R_Ξ is implemented in three steps. First, we get a extrapolation f_C from the given data as in (3.15). Next, we smooth the data by the convolution

$$\Lambda(\chi_{\tilde{\Omega}_{2\delta}} f_C)(y) = \int_{\tilde{\Omega}_{2\delta}} \left(\frac{\sigma_\omega}{\hat{\psi}_{c/\omega}(\omega\cdot)}\right)^\vee (y-\theta)\, f_C(\theta)\, d\theta, \qquad y \in \tilde{\Omega}_\delta,$$

where $\tilde{\Omega}_\delta$ as in (3.13). Finally, we make another convolution

$$R_{\tilde{\Xi}}f(x) = \int_{\Omega_\delta/\omega} \psi_{c/\omega}(x/\omega, t)\,\Lambda(\chi_{\tilde{\Omega}_{2\delta}} f_C)(\omega t)dt.$$

As we observed already, we have smoothing parameters c and ω which are being adjusted according to the density of centers and noise $(\epsilon_\xi)_{\xi\in\Xi}$. As $c, \omega \to 0$, the function $\Lambda(\chi_{\tilde{\Omega}_{2\delta}} f_C)$ tends to the local interpolant f_C, which make the approximant lose some smoothness. Also, as c is getting bigger, the approximant becomes smoother hence may lose some 'details'. In fact, a good choice for the parameters c and ω can be interpreted as a balanced compromise between smoothness and fidelity of the approximant to the data.

The following example illustrates these steps.

Example. The given data is of the form

$$y_\xi = f(\xi) + \epsilon_\xi, \qquad \xi \in [-1, 1]^2$$

with ϵ_ξ independent random variables normally distributed with mean 0 and variance $\sigma = 0.05$. Here the underlying function is

$$f(x, y) = [B_2(1.5(x - .5)) - B_2(1.5(x + .5))] * \exp(-y^2)$$

with B_2 tensor-product quadratic spline. Figures 3.1-3.4 show the surfaces obtained in each of steps described above. Figure 3.1 displays the surface defined by the function f, and Figure 3.2 the surface obtained by interpolating the noisy data. Figure 3.3 presents the surface after removing the noise (Step 2). Finally we obtain the surface displayed in Figure 3.4. We will discuss this example in detail in the next section.

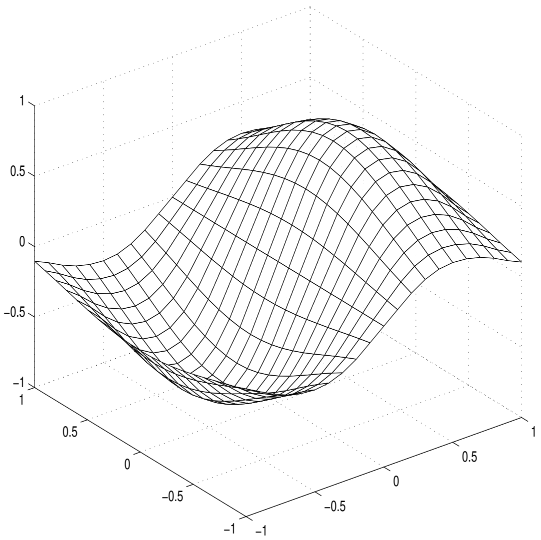


Figure 3.1: Original Function

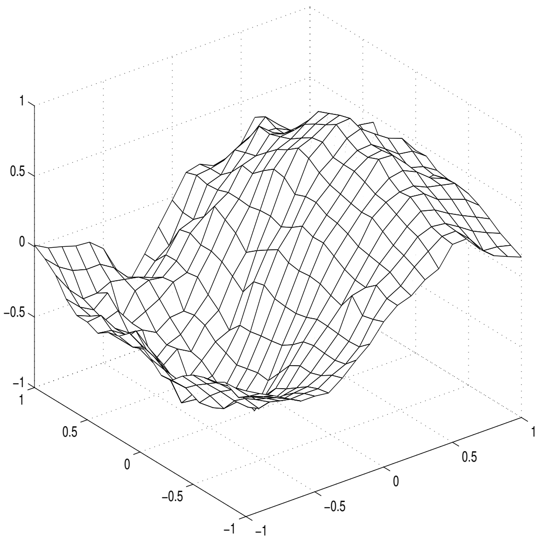


Figure 3.2: Step 1

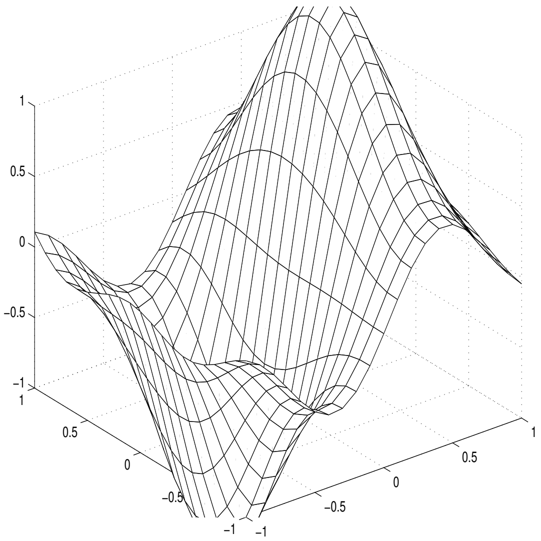


Figure 3.3: Step 2

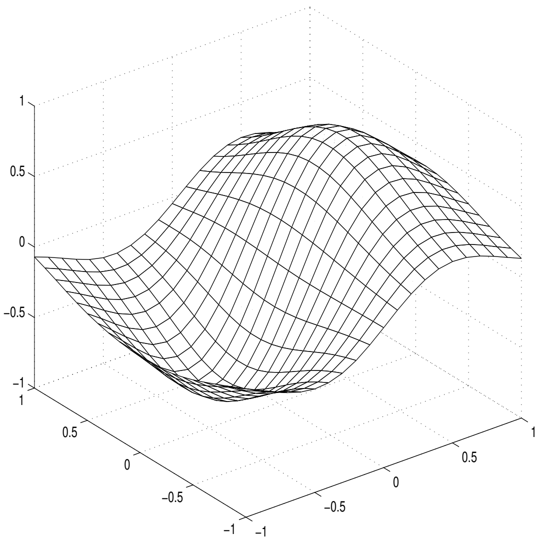


Figure 3.4: Step 3

Chapter 4

Computational Aspects

4.1 On the Coefficients of Pseudo-Shift $\phi(\cdot, t)$

4.1.1 Preliminary

In this section we describe an algorithm for the construction of the coefficient matrix for the pseudo-shift

$$\phi_c(\cdot, t) = \sum_{\xi \in \Xi} A(\cdot, \xi)\phi_c(\cdot - \xi).$$

This algorithm is also applied to find the coefficients $(C(t, \xi))_{\xi \in \Xi}$ for the function

$$f_C = \sum_{\xi \in \Xi} C(\cdot, \xi) f(\xi).$$

For $t \in \Omega$, the 'admissible' coefficients $(A(t, \xi))_{\xi \in \Xi}$ for $\phi_c(\cdot, t)$ are required to satisfy the linear system

$$\sum_{\xi \in \Xi} A(t, \xi) p(\xi) = p(t) \tag{4.1}$$

for $p \in \Pi_n$ with n greater than the order of singularity of $\hat{\phi}_c$ at the origin. Letting $q_1, \cdots, q_{n_d}$ a basis of a polynomial space $\Pi_n(\Omega)$, this condition (4.1) holds if and only if the coefficient matrix

$$\bar{a} := (A(t, \xi))_{\xi \in \Xi}$$

solves the linear system

$$E\bar{a} = \bar{b} \tag{4.2}$$

with

$$E = (q_j(\xi) \,:\, j = 1, \cdots, n_d,\ \xi \in \Xi_t)$$

and

$$\bar{b}^T := (q_j(t) \,:\, j = 1, \cdots, n_d).$$

Here, for $t \in \Omega$, we suggest to find the coefficient matrix $(A(t,\xi))_{\xi\in\Xi}$ for $\phi_c(\cdot, t)$ by a minimization problem

$$\begin{aligned} &\text{minimize} \quad \sum_{\xi\in\Xi_t} \eta(t,\xi) A^2(t,\xi) \\ &\text{subject to} \quad E\bar{a} = \bar{b} \end{aligned} \tag{4.3}$$

with a penalty function $\eta(t, \cdot)$. The choice of polynomial space Π_n requires that

$$\#\Xi_t > \dim\Pi_n(\mathbb{R}^d) = \binom{n+d}{d} =: n_d$$

with $\#\Xi_t > n_d$. In the univariate case, we can make this problem simple computationally by choosing some good basis functions of Π_n (e.g., Hermite or Lagrange polynomials), which makes the matrix E banded or triangular. However, in the multivariate case, since we do not know of a basis for Π_n that results in a matrix E with simple structure, we confront numerical difficulties caused by the conditioning of the matrix E. Eventually, the system (4.2) becomes ill-conditioned with the increase in the number of constraints. The well-known standard method to solve this problem is via Gauss elimination. However, Gauss elimination has to deal another numerical difficulty in providing a solution to the linear system in (4.3) when all the pivots available for the current step in the current column are all zero. For example, if all the centers in Ξ_t are on a straight line and t is not on this line, there may occur a controversy. In this case, we need interchange some of centers and recompute a part of the elimination. So the actual location and configuration of Ξ_t need to be taken into account.

C. de Boor and A. Ron proposed a particular elimination method, the so-called Gauss elimination by degree, which is more efficient for this problem (cf. [BR2]). It is designed for the construction of multivariate polynomial interpolation. Actually, it applies Gauss elimination degree-by-degree (not monomial-by-monomial) with partial pivoting to the Vandermonde matrix

$$V := (\xi^k)$$

by treating all the entries of a given degree as one entry. Hence, this method can be applied to the matrix E^T.

4.1.2 Factorization

For any $t \in \Omega$, let Ξ_t be the support of the sequence $(A(t,\xi))_{\xi\in\Xi}$, i.e.,

$$\phi_c(\cdot, t) = \sum_{\xi\in\Xi_t} A(t,\xi)\phi_c(\cdot - \xi).$$

We assume that $\#\Xi_t = m(> n_d)$ for all $t \in \Omega$ and let

$$\Xi_t = \{\xi_1, \xi_2, \cdots, \xi_m\}.$$

Choosing a set of functions $\{(t-\cdot)^\alpha\}_{|\alpha|\leq n}$ as a basis of Π_n, we denote the matrix E^T as follows,

$$E^T := ((t-\xi_p)^\alpha \;:\; p = 1, \cdots, m,\; 0 \leq |\alpha|_1 \leq n),$$

and correspondingly, the matrix $\bar{b}$ in (4.3) is changed to

$$\bar{b} := [1 \;\underbrace{0\cdots 0}_{(n_d-1)\ \text{terms}}]^T.$$

The strategy of Gauss elimination by degree begins by treating all the entries of a given degree as one entry: the (p,q)-entry of $\mathbf{E}^T$ is to be

$$\mathbf{E}^T(p,q) = (\xi_p^\alpha \ : \ |\alpha| = q).$$

Here and in the sequel, we shall use the notation $\mathbf{E}$ in lieu of E in order to emphasize the alternative point of view. Thus, the rows and columns of $\mathbf{E}^T$ are indexed by $\xi_p \in \Xi_t$ and $k = 0, \cdots, n$ respectively.

Now since the entries in $\mathbf{E}^T$ are considered not as scalars but as vectors, we make all the entries in the pivot column below the pivot row orthogonal to the pivot entry. In order to eliminate entries in column k of $\mathbf{E}^T$, a scalar product is defined as

$$(a,b)_k := \sum_{|\alpha|=k} a(\alpha)b(\alpha)\bar{\omega}(\alpha)$$

with $\bar{\omega}$ a weight function. Furthermore, since each entry in column k consists of

$$C(k+d-1;k) := \frac{(k+d-1)!}{k!\,(d-1)!}$$

monomials, after Gauss elimination by degree in column k, we want to have $C(k+d-1;k)$ nonzero orthogonal entries in the kth column below row

$$k^\circ := \sum_{j<k} C(j+d-1;j) + 1,$$

the first working position of elimination in the kth column. Hence, the ultimate goal is the factorization

$$E^T = L\,U$$

with L^T in row echelon form and U a block upper triangular non-singular matrix.

The algorithm is summarized as follows: Let $\mathbf{W}$ be the 'working array' which initially equals $\mathbf{E^T}$. At each column, say kth, we first put our working position at

k°-th row and go through below the row. At each step (let us assume that we start at k_jth row in column k), we look for the current row k_j or below the row such that, in order to alleviate the devastating interaction of rounding error, we find a largest non-trivial entry (relatively to the size of the corresponding entry or row of $\mathbf{E}^T$) and, if it is not on pivot position, interchange its row with row k_j of $\mathbf{W}$ to bring it into the pivot position $\mathbf{W}(\xi_{k_j}, k)$. Then we subtract the appropriate multiple of the pivot row $\mathbf{W}(\xi_{k_j}, :)$ from all subsequent rows in order to make $W(\xi_{k_i}, k)$ orthogonal to $\mathbf{W}(\xi_{k_j}, k)$ for all $k_i > k_j$. Then we proceed elimination by Gram-Schmidt process. Specifically, if we assume that orthogonal entries $w_1', \cdots, w_{(j-1)}'$ are already available in a column k, we can compute

$$w_j' := w_j - \sum_{i<j} w_i' \frac{< w_j, w_i' >}{< w_i', w_i' >} \tag{4.4}$$

for a next orthogonal entry, and thereby ensure that

$$< w_i', w_j' >= 0 \qquad (i < j)$$

while $w_j' \neq 0$. It may, of course, happen that all the pivots available for the current step in the current column k are zero before we obtain $C(k + d - 1; k)$ nonzero orthogonal entries. Then we have to replace some of the centers in Ξ_t by another centers in $\Xi \setminus \Xi_t$, and perform the calculation (4.4) again on the corresponding row until we obtain $C(k + d - 1; k)$ orthogonal entries in the column k.

On the other hand, Gauss elimination is usually performed to a square matrix, and it factors this matrix into a lower triangular matrix and a upper triangular matrix. However, since our matrix E^T is a $m \times n_d$ rectangle matrix with $m > n_d$, each step of elimination by degree is equivalent to factoring E^T into a $m \times m$

matrix and a $m \times n_d$ matrix. For example, the first step can be expressed as the following

$$\mathbf{E}^T = \tilde{L}_1 \mathbf{W}$$

with a matrix $\mathbf{W}$ of working array and a lower triangular matrix $\tilde{L}_1$ which is associated to making the entries in row 2 through row m orthogonal to the first row. Continuing this process, the final output of elimination by degree is a factorization of E^T in the form

$$\mathbf{E}^T = \tilde{L}\,\mathbf{W}. \tag{4.5}$$

with $\tilde{L}$ a $m \times m$ unit lower triangular matrix. Since E^T is a $m \times n_d$ matrix with $m > n_d$, the elimination is performed with n_d-th columns, and hence the matrix $\tilde{L}$ can be written in the form

$$\tilde{L} = [L \; L_0] \tag{4.6}$$

where the matrix L consists of the first n_d columns of $\tilde{L}$ which are associated with elimination procedures of matrix $\mathbf{E}^T$, but L_0 takes the last $m - n_d$ columns of $\tilde{L}$, and it is not associated with any elimination progress, which means $L_0(i,j) = 0$ for $i \neq j$ with $j > n_d$, and $L_0(i,j) = 1$ for $i = j$. Furthermore, the final output $\mathbf{W}$ is a row echelon matrix in the following sense. If we make ordering $k_1, \cdots, k_{n_d}$ of columns according to the degree of each entry and $\xi_1, \cdots, \xi_m$ of the rows, the last $m - n_d$ rows of matrix W are completely zero, and the leading entry (the first nonzero entry) in the nonzero row $W(\xi_j, :)$ is the entry $W(\xi_j, k_j)$ for all $j \leq n_d$. Hence, the matrix W is of the form

$$\begin{bmatrix} U \\ \bar{0} \end{bmatrix}$$

where U is a $n_d \times n_d$ block upper triangular square matrix, and $\bar{0}$ is a $(m-n_d)\times n_d$ zero matrix. In actual calculation, the two matrices L_0 and $\bar{0}$ can be ignored. Therefore, the factorization in (4.5) is replaced by

$$E^T = LU$$

We note that the matrix U do not have to be upper triangular since $k_1, \cdots, k_{n_d}$ need not be strictly increasing. But each entry in the diagonal entries are orthogonal to each other, hence U is invertible. With this factorization, we return to the original system (4.2)

$$\bar{b} = E\bar{a} = U^T\, L^T \bar{a}.$$

By substituting

$$(U^T)^{-1}\bar{b} =: \bar{b}', \tag{4.7}$$

the linear system (4.2) can be replaced by

$$L^T \bar{a} = \bar{b}'.$$

With the matrices L and $\bar{b}'$ at hand, we find the coefficient matrix $\bar{a} := (A(t,\xi))_{\xi\in\Xi_t}$ by minimizing the quadratic form

$$\sum_{\xi\in\Xi_t} \eta(t,\xi)A^2(t,\xi)$$

subject to the constraints

$$L^T \bar{a} = \bar{b}'.$$

Theorem 4.1. *For $t \in \Omega$, let η be a weight function and $D = 2\,Diag(\eta(t,\xi_i) : i = 1, \cdots, m)$. Let $\bar{a} = (A(t,\xi))_{\xi\in\Xi_t}$, and let L and $\bar{b}'$ be the matrices defined in (4.6) and (4.7). Then*

$$\bar{a} = D^{-1}L(L^T D^{-1} L)^{-1}\bar{b}'. \tag{4.8}$$

Proof. The method of Lagrange multipliers induces the following linear system.

$$D\bar{a} + L\bar{\Lambda} = 0 \ \text{ and } \ L^T\bar{a} = \bar{b}' \tag{4.9}$$

with $\bar{\Lambda}^T := [\lambda_1 \ \cdots \ \lambda_{n_d}]$ the matrix of Lagrange multipliers. Then it can be easily verified that the matrix for the linear system (4.9),

$$\begin{pmatrix} D & L \\ L^T & 0 \end{pmatrix},$$

is non-singular. Therefore, $\bar{a} = D^{-1}L(L^TD^{-1}L)^{-1}\bar{b}'$ and $\Lambda = (L^TD^{-1}L)^{-1}\bar{b}'$ solve (4.9). ∎

We shall be mostly interested in an optimality condition for the minimization problem. Here we suggest some examples.

Example 1. We adopt a penalty function of the form

$$\eta(t, \xi) = \eta_+(|t - \xi|)$$

where η_+ is an increasing function on $\mathbb{R}^d+$, and $\eta_+(0) = 0$. As a good choice of η_+, the following function is suggested in [L]:

$$\eta_+(|t - \xi|) = \exp(\frac{|t - \xi|^2}{h}) - 1$$

where h is the density of Ξ, i.e.,

$$h := \sup_{t\in\Omega} \min_{\xi\in\Xi} |t - \xi|. \tag{4.10}$$

Example 2. In this study, we want to find the coefficient matrix $(A(t,\xi))_{\xi\Xi_t}$ for $\phi_c(\cdot, t)$ which approximates $\phi_c(\cdot - t)$ in some sense. But the above example does not depend on the basis function ϕ_c. Hence, in order to minimize an upper bound of the error $\phi_c(\cdot - t) - \phi_c(\cdot, t)$, we consider the relation

$$\begin{aligned} |\phi_c(\cdot - t) - \phi_c(\cdot, t)|^2 &= |\sum_{\xi \in \Xi_t} A(t,\xi)[\phi_c(\cdot - t) - \phi_c(\cdot - \xi)]|^2 \\ &\leq \text{const} \sum_{\xi \in \Xi_t} A^2(t,\xi)\,[\phi_c(\cdot - t) - \phi_c(\cdot - \xi)]^2. \end{aligned}$$

Thus, our optimization problem becomes

$$\begin{aligned} &\text{minimize} \quad \sum_{\xi \in \Xi_t} A^2(t,\xi)\eta(t,\xi) \\ &\text{subject to} \quad L^T \bar{a} = \bar{b}' \end{aligned}$$

with L^T and $\bar{b}'$ in (4.6) and (4.7) respectively, and

$$\eta(t,\xi) = \|\phi_c(\cdot - t) - \phi_c(\cdot - \xi)\|^2_{L^\infty(\Omega)} \tag{4.11}$$

with Ω a bounded domain in $\mathbb{R}^d$.

Example 3. As an another example of a minimization condition for finding $\bar{a}$, we consider the following L_2-norm minimization:

$$\begin{aligned} &\text{minimize} \quad \|\phi_c(\cdot, t) - \phi_c(\cdot - t)\|_{L^2(\Omega)} \\ &\text{subject to} \quad L^T \bar{a} = \bar{b}'. \end{aligned} \tag{4.12}$$

The next theorem provides the solution of the problem in (4.12).

Theorem 4.2. *For $t \in \Omega$, let $\bar{a} = (A(t,\xi))_{\xi \in \Xi_t}$ and E be as above. Using the notation $\Xi_t = \{\xi_1, \xi_2, \cdots, \xi_m\}$, let $D = (2\int_\Omega \phi_c(x - \xi_i)\phi_c(x - \xi_j)dx \,:\, i,j = 1, \cdots, m)$. Then*

$$\bar{a} = D^{-1}\bar{d} - D^{-1}E(E^T D^{-1} E)^{-1}(E^T D^{-1}\bar{d} - \bar{b}). \tag{4.13}$$

where

$$\bar{d} = (2\int_\Omega \phi_c(x-\xi_j)\phi_c(x-t)dx : j = 1,\cdots,m).$$

Proof. The method of Lagrange multipliers induces the following linear system

$$D\bar{a} + L\bar{\Lambda} = \bar{d} \text{ and } L^T\bar{a} = \bar{b}' \tag{4.14}$$

with $\bar{\Lambda}^T := [\lambda_1 \ \cdots \ \lambda_n]$ the matrix of Lagrange multipliers. Then it can be easily verified that the matrix of the system (4.14),

$$\begin{pmatrix} D & L \\ L^T & 0 \end{pmatrix},$$

is non-singular. Thus, the matrices $\Lambda = (L^T D^{-1} L)^{-1}(L^T D^{-1}\bar{d} - \bar{b}')$ and $\bar{a}$ in (4.13) solve (4.14). ■

4.1.3 Algorithm Details : The Coefficients $(A(t,\xi))_{\xi\in\Xi}$ for $\phi_c(\cdot,t)$

Letting $d = 2$, we give here a MATLAB-like pseudo-program to construct coefficients $(A(t,\xi))_{\xi\in\Xi}$, $t \in \Omega$, for the pseudo-shift

$$\phi_c(\cdot,t) = \sum_{\xi\in\Xi} A(t,\xi)\phi_c(\cdot - \xi).$$

This algorithm selects a Ξ_t which is **total** for Π_n, and then finds an admissible coefficients $(A(\cdot,t))_{\xi\in\Xi_t}$ for Π_n. In this 'program,' we use the following convention.

The set of scattered centers Ξ is considered as $m \times 2$ matrix such that we use the notation

$$\Xi_t^1 := \Xi(:,1) \quad \text{and} \quad \Xi_t^2 := \Xi(:,2).$$

The matrices $\mathtt{E}^{\mathtt{T}}$ and $\mathtt{W}$ are denoted by $\mathbf{E}^{\mathbf{T}}$ and $\mathbf{W}$ respectively. In particular, since $C(k+d-1;k) = k+1$ with $d=2$, we note that $\mathtt{W(j,k)}$ is a vector with $(k+1)$ entries, indexed by $\{\alpha \in \mathbb{Z}^d : |\alpha| = k\}$, and $(k+1)$ orthogonal entries will be obtained in column k. All matrices mentioned in the 'program' other than $\mathtt{E}^{\mathtt{T}}$ and $\mathtt{W}$ are proper MATLAB matrices, i.e., have scalar entries. Correspondingly, for two vectors `a` and `b` (such as `W(i,k)`, `W(j,k)`) indexed by $\{\alpha \in \mathbb{Z}^d : |\alpha| = k\}$, `<a,b>` denotes the scalar product

$$< \mathtt{a}, \mathtt{b} >:= \sum_{|\alpha|=k} \mathtt{a}(\alpha) \cdot \mathtt{b}(\alpha) \frac{|\alpha|!}{\alpha}. \tag{4.15}$$

We borrow from MATLAB the notations : (i) `ones(m,n)` for the matrix of size $\mathtt{m} \times \mathtt{n}$ with all entries equal to 1; (ii) `eys(m,m)` for the identity matrix of order `m`; (iii) `a:b` for the vector with entries $\mathtt{a}, \mathtt{a}+1, \cdots, \mathtt{a}+\mathtt{m}$, with `m` the natural number for which $\mathtt{a}+\mathtt{m} \le \mathtt{b} < \mathtt{a}+\mathtt{m}+1$; (iv) `A*B` for the matrix product of the matrices `A` and `B`; (v) standard logical constructs like (`for` $j = 1 : n, \cdots,$ `end`), and (`if`$\cdots, \cdots,$ `end`); (vi) the construct (`while 1,` $\cdots$`, if` $\cdots$`, break, end,` $\cdots$`, end`), which is a loop exited only through the break; (vii) the construct `[m,i] = max(a)` to provide $\mathtt{m{:}=a(i){:}=\max_j a(j)}$; (viii) the command `function [a,b] = ft-name(x)` defines a new function called `ft-name`. The variables within the body of a function are all by default local; (ix) the relation operator `a==b` means that `a` is equal to `b` while `a=b` is used for the assignment statement. Furthermore, we use an occasional word to describe an action whose details seem clear.

% INPUT : $\Xi_t = [\Xi_t^1\ \Xi_t^2]$, m, n, tol, penalty function η

% OUTPUT : $\bar{\mathtt{a}} = \mathtt{D}^{-1} * \mathtt{L} * (\mathtt{L}^{\mathtt{T}} * \mathtt{D}^{-1} * \mathtt{L})^{-1}) * \mathtt{b}$ by Theorem 4.1

Select Ξ_t (from Ξ) which are closest to t, $\#\Xi_t = \mathtt{m} > \mathtt{n}$

```
nbr_og=0 ; k=0
Eᵀ(:,k) = ones(n,1) ; W(:,k) = Eᵀ(: .k)
L = eye(m,m)
for j=1:m
   while 1
      [m,i] = max_{i>j-1}<W(i,k),W(i,k)>/<Eᵀ(i,k), Eᵀ(i,k)>
      if m > tol
         nbr_og = nbr_og+1
         break
      end
      if nbr_og > k+1
         replace ξ_j by ξ, one of the closest centers to t from Ξ \ Ξ_t, i.e.,
         Eᵀ(ξ_j,:) = Eᵀ(ξ,:)
         [L_j,W_j]= RE-COMP(j,k,Eᵀ(j,:), W); break
         L(j,1:j-1)=L_j; W(j,k) = W_j;
      end
      k = k+1; nbr_og=0;
      Eᵀ(:,k) = [Eᵀ(1,k-1) * Ξ_t^1 ··· Eᵀ(k-1,k) * Ξ_t^2,  Eᵀ(k-1,k) * Ξ_t^2]
      W(:,k) = L^{-1} * Eᵀ(:,k)
   end
   if i > j, interchange i and j, end
   for i=j+1:n
      L(i,j) = <W(i,k), W(j,k)>/<W(j,k),W(j,k)>
```

```
    W(i,k) = W(i,k)-L(i,j)*W(j,k)
    end
  end
end
L = L(:,1:n) ; W = W(1:n,:)
```

$\mathtt{b} = (\mathtt{W}^{\mathtt{T}})^{-1}\ \mathtt{b}\ ;\ \mathtt{D} = \mathtt{Diag}(\eta(\mathtt{t},\xi))_{\xi\in\Xi_{\mathtt{t}}}$

$\bar{\mathtt{a}} = \mathtt{D}^{-1} * \mathtt{L} * (\mathtt{L}^{\mathtt{T}} * \mathtt{D}^{-1} * \mathtt{L})^{-1}) * \mathtt{b}$

function $[\mathtt{L_j}, \mathtt{W(j,q)}] = \mathtt{RE} - \mathtt{COMP}(\mathtt{j},\mathtt{k},\mathtt{E}^{\mathtt{T}}(\mathtt{j},:),\mathtt{W})$

```
q=0
for p=1:j-1
  if W(j,q)==0, q=q+1, end
```

$\quad \mathtt{L_j(p)} = <\mathtt{E}^{\mathtt{T}}(\mathtt{j},\mathtt{q}),\mathtt{W}(\mathtt{p},\mathtt{q})> / <\mathtt{W}(\mathtt{p},\mathtt{q}),\mathtt{W}(\mathtt{p},\mathtt{q})>$

$\quad \mathtt{W}(\mathtt{j},\mathtt{q}) = \mathtt{E}^{\mathtt{T}}(\mathtt{j},\mathtt{q}) - \mathtt{L_j}(\mathtt{p}) * \mathtt{W}(\mathtt{p},\mathtt{q})$

```
end
```

4.2 A General Algorithm for the Scheme $\tilde{R}_\Xi$

4.2.1 Formulation for the Construction of $\tilde{R}_\Xi f$

We now describe a general algorithm for the approximation scheme $\tilde{R}_\Xi$ on a bounded domain in $\mathbb{R}^2$ by using $\phi_c(x) = (|x|^2 + c^2)^{\lambda/2}\log(|x|^2 + c^2)^{1/2}$, the 'shifted' thin-plate spline. We assume that the given data (ξ, y_ξ) arise according the model

$$y_\xi = f(\xi) \quad \text{or} \quad y_\xi = f(\xi) + \epsilon_\xi$$

where ξ belong to $\Omega \subset \mathbb{R}^d$ and $(\epsilon_\xi)_{\xi\in\Xi}$ is a type of noise, for example, independent normally distributed random variables with mean 0 and variance σ^2.

Step 0 : [Initialization]

For the computational feasibility, we localize ϕ_c by an application of a difference operator, i.e., obtains a function of the form

$$\psi_c := \sum_{\alpha\in N} \mu(\alpha)\phi_c(\cdot - \alpha)$$

where N is finite subset of $\mathbb{Z}^d$. One example of N (to be used for our numerical examples in next section) for this linear combination is illustrated by stencils centered at the origin as shown in Figure 4.1.

$$\begin{array}{ccccccc}
 & & \times & \times & \times & & \\
 & \times & \times & \times & \times & \times & \\
\times & \times & \times & \times & \times & \times & \times \\
\times & \times & \times & & \times & \times & \times \\
\times & \times & \times & \times & \times & \times & \times \\
 & \times & \times & \times & \times & \times & \\
 & & \times & \times & \times & & \\
-3 & -2 & -1 & 0 & 1 & 2 & 3
\end{array}$$

Figure 4.1: Stencil for N (d=2, $\lambda = 2$)

Actually, the function ψ_c need to satisfy the condition in (3.1) with $q = 2\lambda + d - 1$. Letting

$$\tau(\theta) = \sum_{\alpha\in N} \mu(\alpha)e^{-\alpha\cdot\theta},$$

the sufficient condition for (3.1) is as follows:

$$D^{\alpha}(\tau - \hat{\phi}_c^{-1})(0) = 0, \quad |\alpha|_1 \le 2\lambda + 2d - 1 \tag{4.16}$$

(see the paper [DJLR]). It is not surprising that the localization involves the behavior near zero of $\hat{\phi}_c$. It follows that the localization sequence $(\mu(\alpha))_{\alpha\in N}$ satisfies

$$\sum_{\alpha\in N} \mu(\alpha)\alpha^{\beta} = D^{\beta}\hat{\phi}_c^{-1}(0)$$

with $|\beta|_1 \le 2\lambda + 2d - 1$. Let $N = \{\alpha_j : j = 1, \cdots, \#N\}$. Then the localization sequence

$$\bar{\mu} := (\mu(\alpha_j) \,:\, j = 1, \cdots, \#N)^T$$

is obtained by solving the linear system

$$M\,\bar{\mu} = \bar{c}$$

where

$$M = (\alpha_j^{\beta} : |\beta|_1 \le 2\lambda + 2d - 1, j = 1, \cdots, \#N)$$

and

$$\bar{c} = (D^{\beta}\hat{\phi}_c^{-1}(0) : |\beta|_1 \le 2\lambda + 2d - 1)^T.$$

Then ψ_c satisfies the decaying property

$$\sup_{x\in\mathbb{R}^2} (1 + |x|)^{2\lambda+2d-1}\psi_c(x) < \infty.$$

In particular, assuming N is as in Figure 4.1 with $\lambda = 2$ and $d = 2$, the graph of function ψ_c is symmetric about coordinate axes.

Step 1 : [Extrapolation]

Let $\tilde{\Omega}_{2\delta} \supset \Omega$ with $\delta > 0$ as in (3.13). For $x \in \tilde{\Omega}_{2\delta}$, we define an extrapolation f_C by applying the matrix $(C(\cdot, \xi))_{\xi\in\Xi}$ to the given data $(y_\xi)_{\xi\in\Xi}$, i.e.,

$$f_C(t) = \sum_{\xi\in\Xi} C(t,\xi) y_\xi \tag{4.17}$$

where y_ξ is a function value $f(\xi)$ or a noisy value $f(\xi) + \epsilon_\xi$. Like the algorithm for $(A(t,\xi))_{\xi\in\Xi}$, we suggest to find the coefficients $(C(t,\xi))_{\xi\in\Xi}$ for f_C by a minimizing problem

$$\begin{aligned} &\text{minimize} \quad \sum_{\xi\in\Xi_t} \eta(t,\xi) C^2(t,\xi) \\ &\text{subject to} \quad L^T \bar{c} = \bar{b}' \end{aligned} \tag{4.18}$$

with L and $\bar{b}'$ as in (4.6) and (4.7) respectively, and

$$\bar{c} := (C(t,\xi))_{\xi\in\Xi}.$$

Because the approximation should be local in the sense that its value at x depend on ξ which is close to x, we assign a high penalty to centers which are far from x. Having performed some numerical experimentations with several alternatives for the function η, we found that a good choice is

$$\eta(|t-\xi|) = \left[\exp(\frac{|t-\xi|^2}{h^2}) - 1\right] |t-\xi|^{2k} \tag{4.19}$$

with $k \in \mathbb{Z}_+$ and h the density of Ξ as in (4.10).

Step 2 : [Compute $\Lambda(\chi_{\tilde{\Omega}_{2\delta}} f_C)$]

We describe a subalgorithm to compute the function $(\sigma/\hat{\psi}_c)^\vee$. In fact, the function $\Lambda(\chi_{\tilde{\Omega}_{2\delta}} f_C)$ is a combination of $(\sigma/\hat{\psi}_c)^\vee$ and f_C.

The Fourier transform of ψ_c is of the form

$$\begin{aligned} \hat{\psi}_c(\theta) &= \hat{\phi}_c(\theta) \sum_{\alpha \in N} \mu(\alpha) e^{-i\alpha\cdot\theta} \\ &= \frac{d}{d\beta}\tilde{c}(\lambda) \frac{\tilde{K}_{(\lambda+d)/2}(\theta)}{|\theta|^{\lambda+d}} \sum_{\alpha\in N} \mu(\alpha)\cos(\alpha\cdot\theta) \end{aligned}$$

where $\tilde{K}_\nu$ is a modified Bessel function and

$$\tilde{c}(\beta) = 2^{\frac{\beta}{2}+1}(2\pi)/\Gamma(-\frac{\beta}{2}).$$

For example, in the case $\lambda = 2$ and $d = 2$, the constant $\frac{d}{d\beta}\tilde{c}(2)$ is computed as

$$\frac{d}{d\beta}\tilde{C}(2) = 4\pi$$

by using some tools like MATHEMATICA.

Next, we construct a C^∞-cutoff function σ as the tensor product of a one-variable C^∞-cutoff function σ^1 whose support lies in the ball B_M with $0 < M < 2\pi$, so that $\sigma = 1$ on $B_{M/2}$ and $\|\sigma\| = 1$. For $t \in \mathbb{R}$, let

$$g(t) := \begin{cases} C_0 \exp(-\frac{1}{1-|t|^2}), & t \in [-1\,..\,1] \\ 0, & t \in [-1\,..\,1]' \end{cases}$$

with

$$C_0 = \left[\int_{[-1\,..\,1]} \exp(-\frac{1}{1-|t|^2})dt\right]^{-1}.$$

We know that $g \in C^\infty(\mathbb{R}^d)$, $\|g\|_1 = 1$, and it has support in $[-1\,..\,1]$. Then

$$g_\epsilon = \epsilon^{-1}g(\cdot/\epsilon) \tag{4.20}$$

has support in $[-\epsilon\,..\,\epsilon]$ and $\|g_\epsilon\|_1 = 1$, so

$$\sigma^1 := \chi_{[-M+\epsilon\,..\,M-\epsilon]} * g_\epsilon$$

satisfies our requirements. Then the cutoff function σ is defined as

$$\sigma(x) = \sigma^1(x_1) \cdots \sigma^1(x_d), \qquad x = (x_1, \cdots, x_d).$$

In particular, if $d = 2$ and N is as in Figure 4.1, the function $(\sigma/\hat{\psi}_\rho)^\vee$ is simplified as

$$\left(\frac{\sigma}{\hat{\psi}_\rho}\right)^\vee (\theta) = \frac{1}{\pi^2} \int_{[0\,..\,M]^2} \frac{\sigma}{\hat{\psi}_\rho}(s) \cos(\theta \cdot s) ds.$$

4.2.2 A General Algorithm for $\tilde{R}_\Xi f$

Under the above setting and notation, we propose the following general algorithm for the scheme $\tilde{R}_\Xi f$ in the case $\lambda = 2$ and $d = 2$. Since it contains several numerical integration issues, a general policy of numerical integration should be set up in advance. In this algorithm, we use the notation

$$\tilde{\Omega}_\delta := \{y = x + z \,:\, x \in \Omega \text{ and } |z| \leq \delta\}.$$

Step 0 : [Initialization]

0.1 Choose the tuning parameters c and $\omega = h^r$, $0 < r \leq 1$, depending on the density of Ξ.

0.2 Compute the localization sequence $\bar{\mu} := (\mu(\alpha_j))_{\alpha_j \in N}$ by solving $\mu = M^{-1}\bar{c}$. If $d = 2$ and $\lambda = 2$, we may use the N in Figure 4.1.

0.3 Choose $\delta > 0$ and $\Omega_\delta \supset \Omega$.

0.4 Place additional centers in $\Omega_{2\delta} \setminus \Omega$ and construct an extended center set $\tilde{\Xi}$.

Step 1 : [Extrapolation]

1.1 For $t \in \tilde{\Omega}_\delta$, construct a linear system $\sum_{\xi\in\Xi} C(t,\xi)p(\xi) = p(t)$ (considering the given data).

1.2 Compute the coefficients $(C(\cdot,\xi))_{\xi\in\Xi}$ for f_C by using the algorithm in Section 4.1.

1.3 For $t \in \tilde{\Omega}_{2\delta}$, compute $f_C(t) = \sum_{\xi\in\Xi} C(t,\xi)f(\xi)$

Step 2 : [Compute $\Lambda(\chi_{\tilde{\Omega}_{2\delta}} f_C)$]

2.1 Let $\sigma_1 := \chi_{[-M+\epsilon..M-\epsilon]} * g_\epsilon$ with $0 < M < \pi$, $\epsilon > 0$, and g_ϵ be as in (4.20). Construct a cutoff function $\sigma(x) = \sigma^1(x_1)\sigma^1(x_2)$.

2.2 Compute $\hat{\psi}_c(\theta) = 4\pi|\theta|^{-4}\tilde{K}_2(\theta)\sum_{\alpha\in N}\mu(\alpha)\cos(\alpha\cdot\theta)$

2.3 Compute $(\sigma/\hat{\psi}_\rho)^\vee(t) = \frac{1}{\pi^2}\int_{[0..M]^2}\sigma(\theta)\hat{\psi}_\rho^{-1}(\theta)\cos(t\cdot\theta)d\theta$ (by numerical integration).

(d) Compute $\Lambda(\chi_{\tilde{\Omega}_{2\delta}} f_C)$ (by numerical interaction) from f_C and $(\sigma/\hat{\psi}_\rho)^\vee$.

Step 3 :

3.1 For $t \in \Omega_{2\delta}$, compute $(A(t,\xi))_{\xi\in\tilde{\Xi}/\omega}$ for $\phi_{c/\omega}(\cdot,t)$, using the algorithm in Section 4.1.

3.2 For $x \in \Omega$, compute $\psi_{c/\omega}(x/\omega, t)$.

3.3 Compute $\tilde{R}_\Xi f$ from $\psi_{c/\omega}(x/\omega,t)$ and $\Lambda(\chi_{\tilde{\Omega}_{2\delta}} f_C)$.

4.3 Numerical Results

In this section, we present some numerical results on the approximation of functions by using R_Ξ and $\tilde{R}_\Xi$ on a bounded domain in $\mathbb{R}^2$. Those new approximation schemes suggested in this study can be applied to noiseless data as well as noise data. From the theory and experience gained with this approximation, and in compare with other methods, we see that this scheme provides better approximants. In the following examples, all the scattered centers are generated by a random number generator in MATLAB, and we observe approximation of a function f on the domain $[-1,1]^2$. In Example 1-2, we assume that a set Ξ is given in a larger space containing $[-1,1]^2$ such that, for a given function f, we observe the approximation power of R_Ξ on $[-1,1]^2$ by using all the given scattered shifts of basis function. Currently, one of the most well-known approximation methods to scattered data is thin-plate spline (TPS) interpolation, hence some comparisons are given between these two schemes, and we see better approximations by using the scheme R_Ξ. In addition, one major advantage behind this comparison is that the scheme R_Ξ is local in the sense that its value at a point mainly depend on $\phi_c(\cdot - \xi)$ for which $\xi \in \Xi$ is close to the point. A well-known disadvantage of TPS interpolation is that, with the increase of the number of centers, it requires the computation with a huge matrix which is very ill-conditioned. However, the scheme R_Ξ requires to solve the linear system

$$\sum_{\xi \in \Xi_t} c(t,\xi) p(\xi) = p(t), \quad p \in \Pi_k, \tag{4.21}$$

which is independent of the size of Ξ. Furthermore, in case a huge set of scattered center is given, it is possible to do parallel computation by dividing the domain into several pieces.

Next, in Example 3-4, we consider approximation of a function f known only at finitely many centers with noise, i.e., the given data is of the form (ξ, y_ξ) with

$$y_\xi = f(\xi) + \epsilon_\xi, \quad \xi \in [-1,1]^2, \tag{4.22}$$

where ϵ_ξ's are independent noise with mean 0 and variance σ^2. Unlike the previous case, the noisy data (ξ, y_ξ) arise only inside of a domain. As discussed in Section 3.2, we use an extrapolation method to recover the function values on some domain including $[-1,1]^2$, and augment the space $S_\Xi(\phi_c)$ by adding some extra centers around $[-1,1]^2$. Then, letting $\tilde{\Xi}$ be the extended center set, we look for an approximant from the space $S_{\tilde{\Xi}}(\phi_c)$ with $\tilde{\Xi}$ by using the same scheme $\tilde{R}_\Xi$. Among the other approaches for smoothing noisy data, G. Wahba's thin-plate smoothing spline (TPSS) technique is widely used. So we provide some comparisons between $\tilde{R}_\Xi$ and TPSS, and we see better approximations by using the scheme $\tilde{R}_\Xi$.

Example 1 Let B_2 be a standard quadratic spline. We consider approximation of a C^1-function

$$f(x,y) = B_2(x)B_2(y)$$

on the square $[-1,1]^2$ from the space $S_\Xi(\phi_c)$, where Ξ consist of 200 scattered centers in $[-4.5,4.5]^2$ as in Figure 4.2. Comparisons between R_Ξ and TPS interpolation are given in Figures 4.4-4.5 through the surfaces and contour lines. The error distributions are also displayed in Figure 4.6-4.7. The sizes of the surfaces of error distributions look similar, but they have different scale relative each other. Even, in the contour lines of these errors, the distances between level lines are 0.1 by TPS interpolation, but 0.05 by R_Ξ. We obtain absolutely maximal errors 0.1423 by TPS interpolation and 0.0458 by R_Ξ. The coefficient $(A(t,\xi))_{\xi\in\Xi}$ for $\phi_c(\cdot,t)$ are computed by (4.8) with penalty function η in (4.11), which involves the solution of a linear system in (4.21) with $k = 7$. In particular, we use the tuning parameters $c = 1.5 * \omega$ with $\omega = .5^{\cdot 5}$.

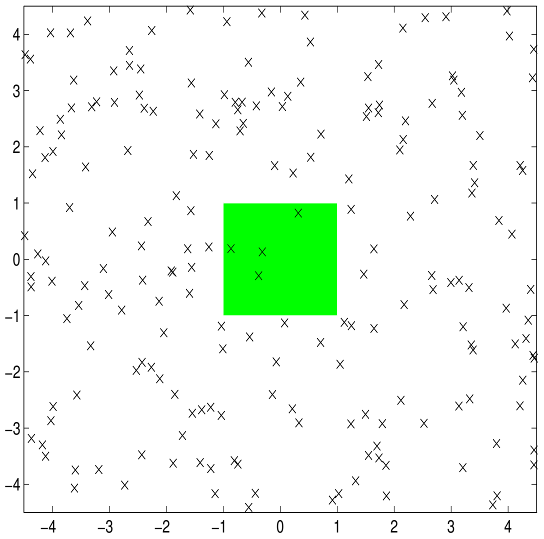


Figure 4.2: Center Set

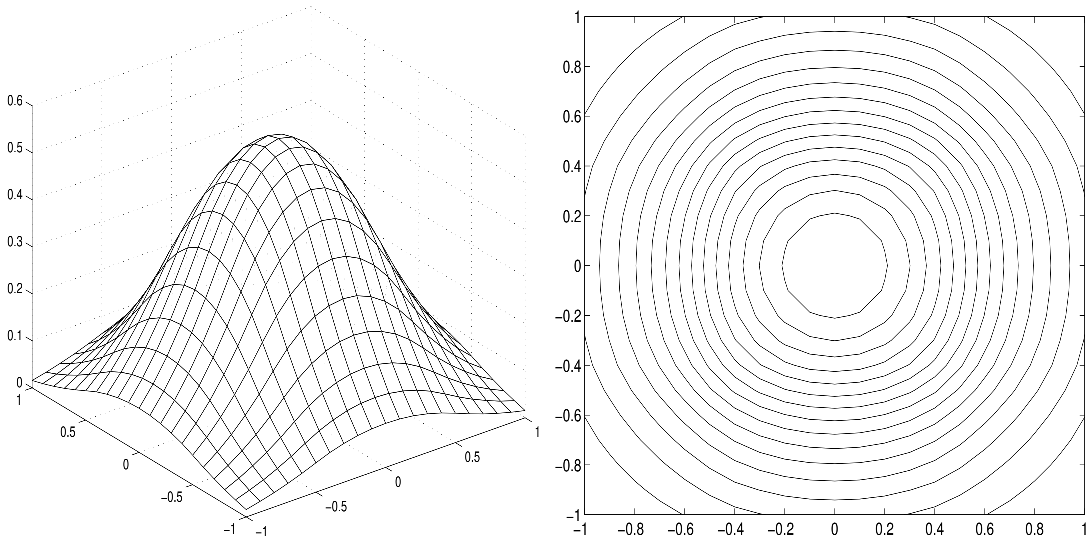


Figure 4.3: Original Function f

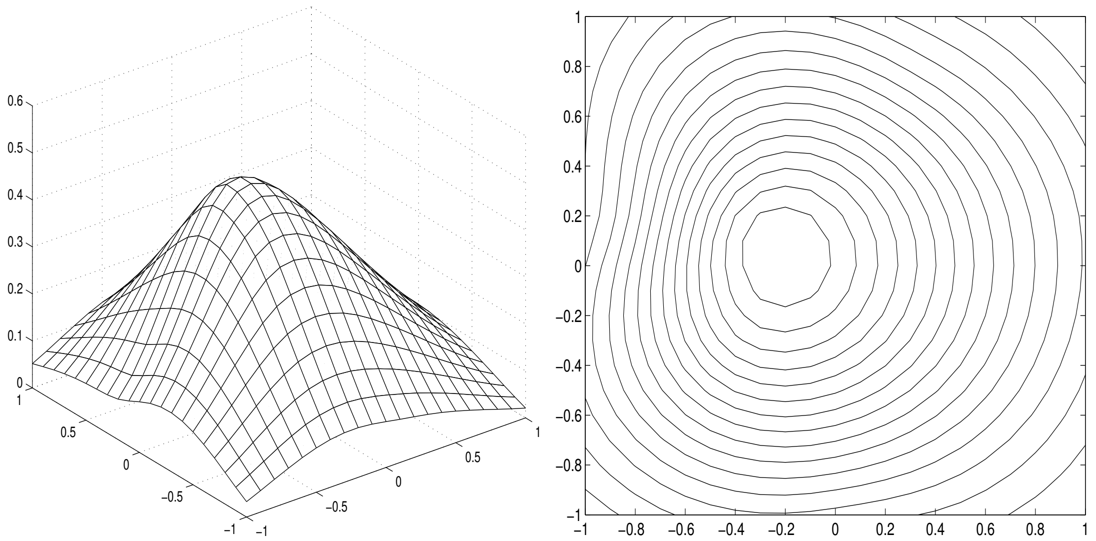


Figure 4.4: TPS interpolation

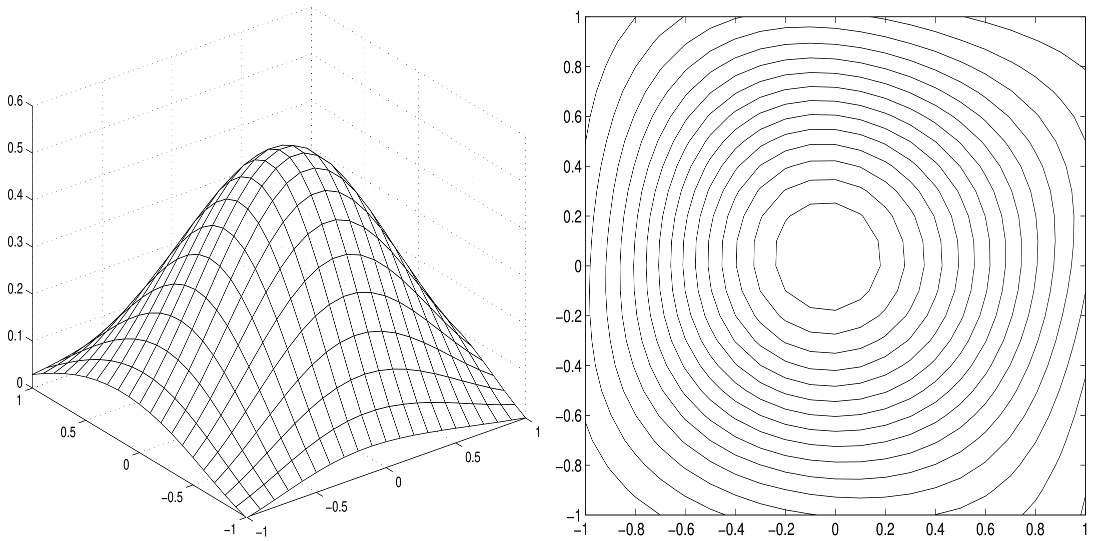


Figure 4.5: $R_\Xi f$

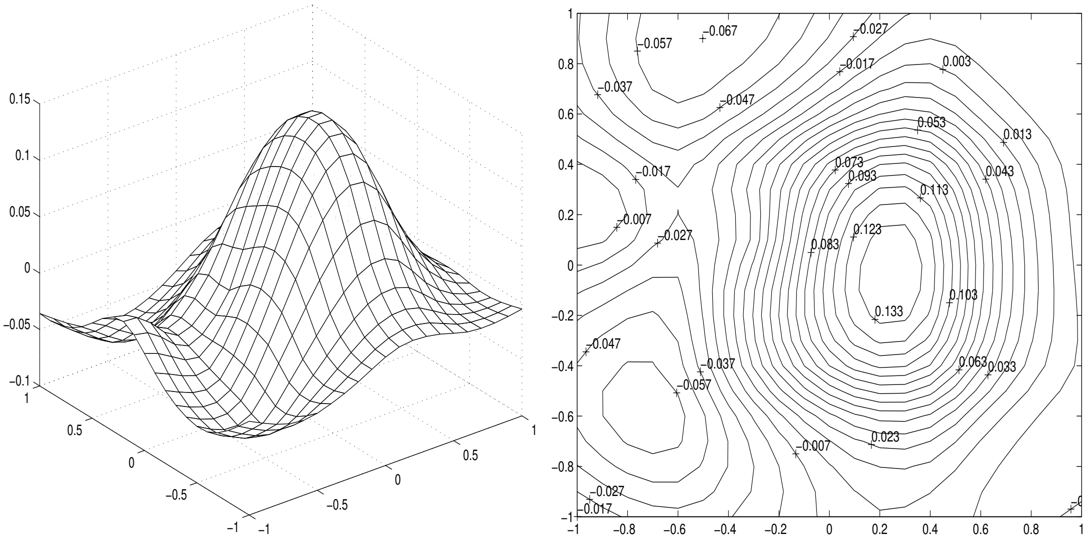


Figure 4.6: Errors by TPS interpolation, max : 0.1423

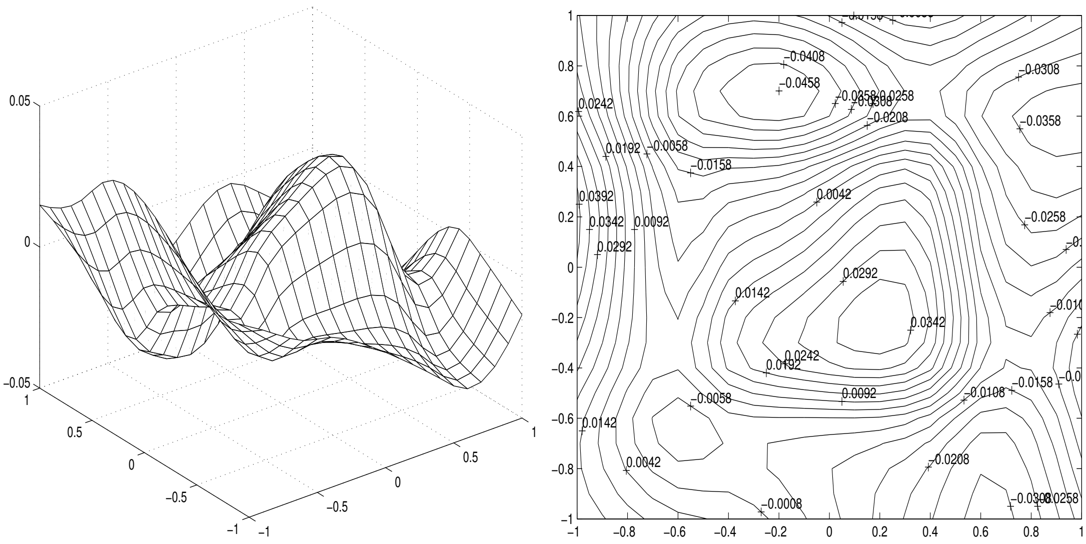


Figure 4.7: Errors by R_{Ξ}, max : 0.0458

Example 2 As a second example, we approximate a C^∞-function

$$f(x,y) = -\exp(-(x^2+y^2)) + \left[\frac{\sin(x)\sin(y)}{xy}\right]^5.$$

Another set of 200 scattered centers are generated in $[-3,3]$ (see Figure 4.8), and we also approximate f over the square $[-1,1]^2$. A comparison has been made between R_Ξ and TPS interpolation. Figures 4.10-4.11 show the contour lines of the original function and the approximants by TPS interpolation and R_Ξ. The error distributions are also displayed in Figure 4.12-4.13. As in the previous example, surfaces of the error distributions have different scales each other, and the distance between the contour lines are 0.1 by TPS interpolation, but 0.05 by R_Ξ. We obtain absolutely maximal errors 0.1682 by TPS interpolation and 0.0397 by R_Ξ. The coefficients $(A(t,\xi))_{\xi\in\Xi}$ for $\phi_c(\cdot,t)$ are computed by the same way as in Example 1. The values of the tuning parameters are $c = 2*\omega$ and $\omega = .5^{4/5}$. Actually, as we observed through the theory, the parameter c in the basis function ϕ_c can be chosen in the manner of $c = \rho\omega$ with $\rho > 0$ and $\omega = \bar{h}^r$, $0 < r \leq 1$. Also, it is recommended to choose $r < 1$ for the spectral approximation order, and ρ is chosen to depend on the smoothness of function f. More specifically, since ρ determines the function $\sigma_\rho \hat{f}/K_2^{-1}$ being involved in the construction of approximant and the decaying property of $\hat{f}$ is determined by the smoothness of f, a bigger number ρ (eventually, c) can be assigned in the implementation of $R_\Xi f$ as f becomes smoother. Fortunately, though we have an issue of choosing this tuning parameter, its effect to the approximation is not sensitive according to the smoothness of f. In this example, the approximant $R_\Xi f$ is computed in terms of $\rho = 2$, while $\rho = 1.5$ with a C^1-function f in previous example.

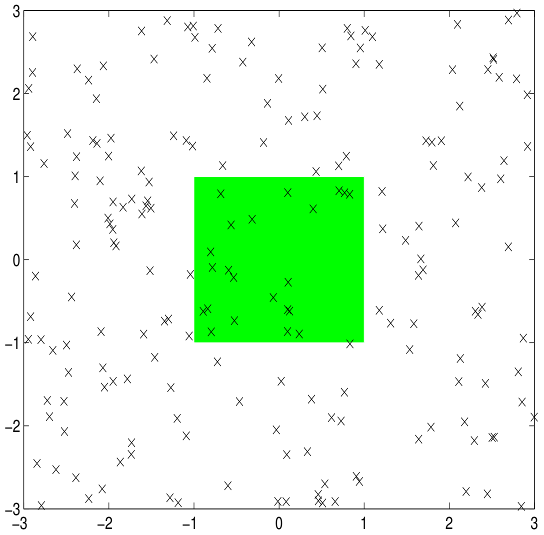


Figure 4.8: Center Set

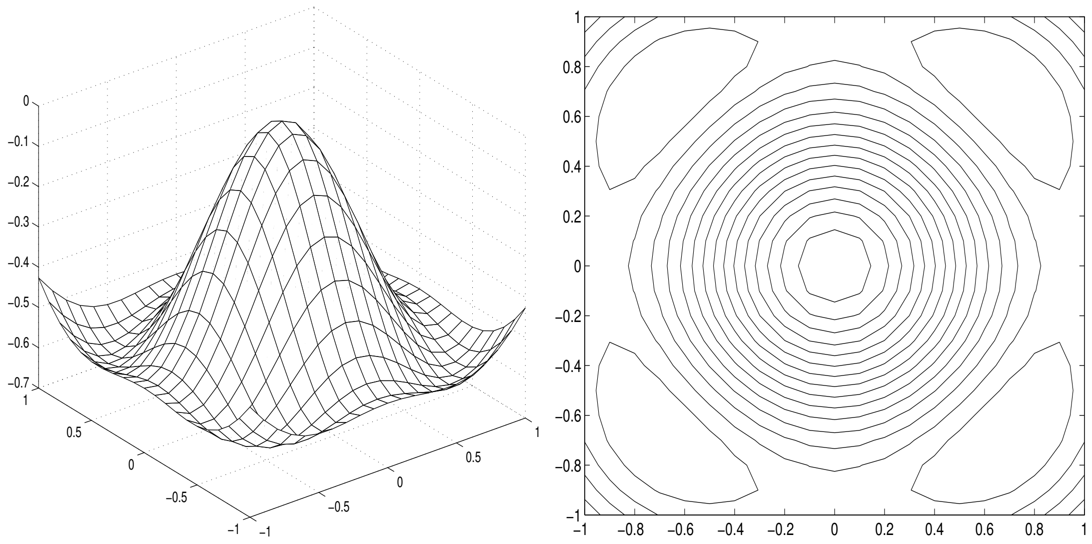


Figure 4.9: Original function f

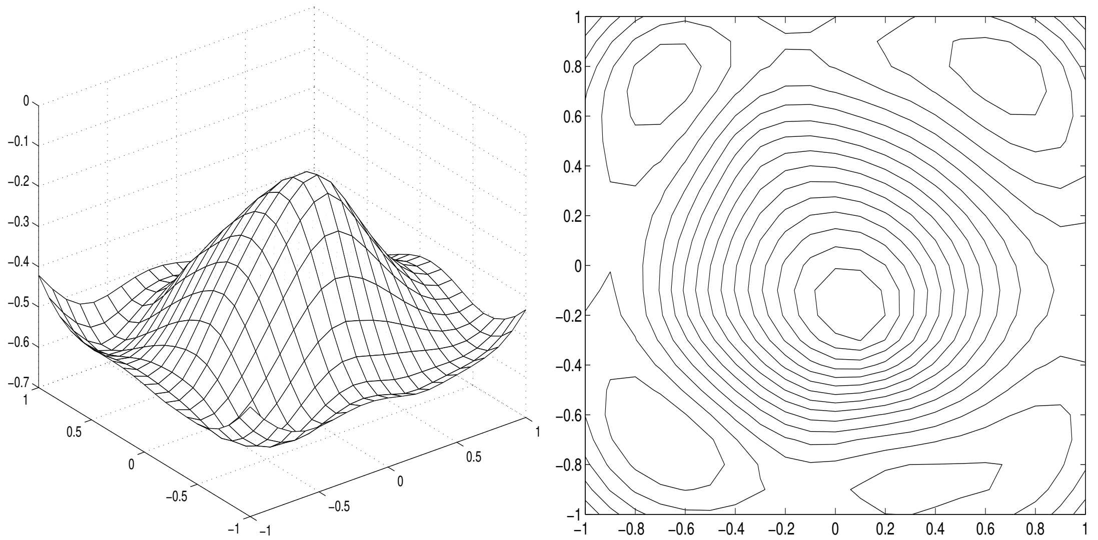


Figure 4.10: TPS interpolation

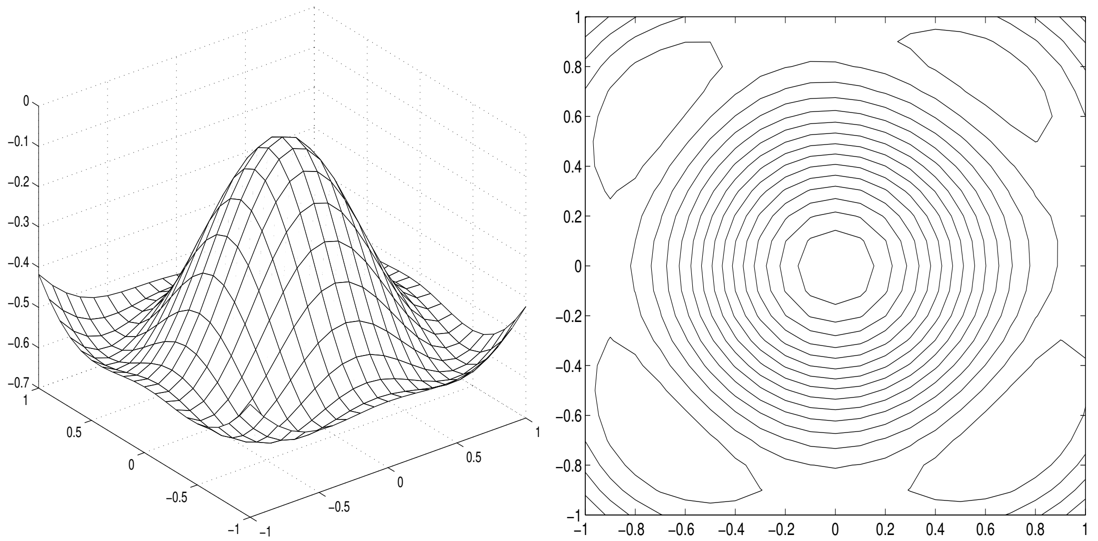


Figure 4.11: $R_{\Xi} f$

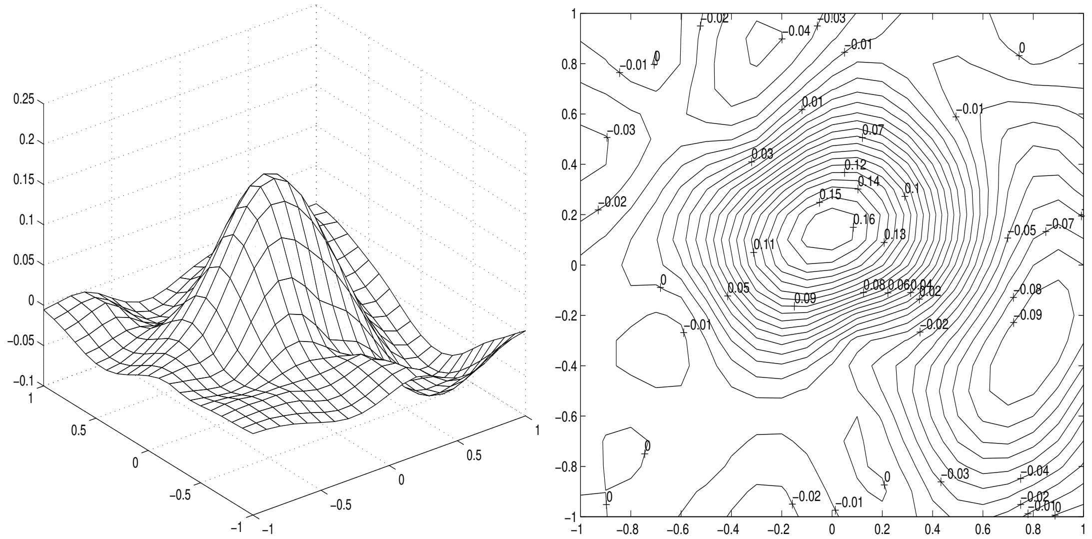


Figure 4.12: errors by TPS interpolation, Max. Error=0.1682

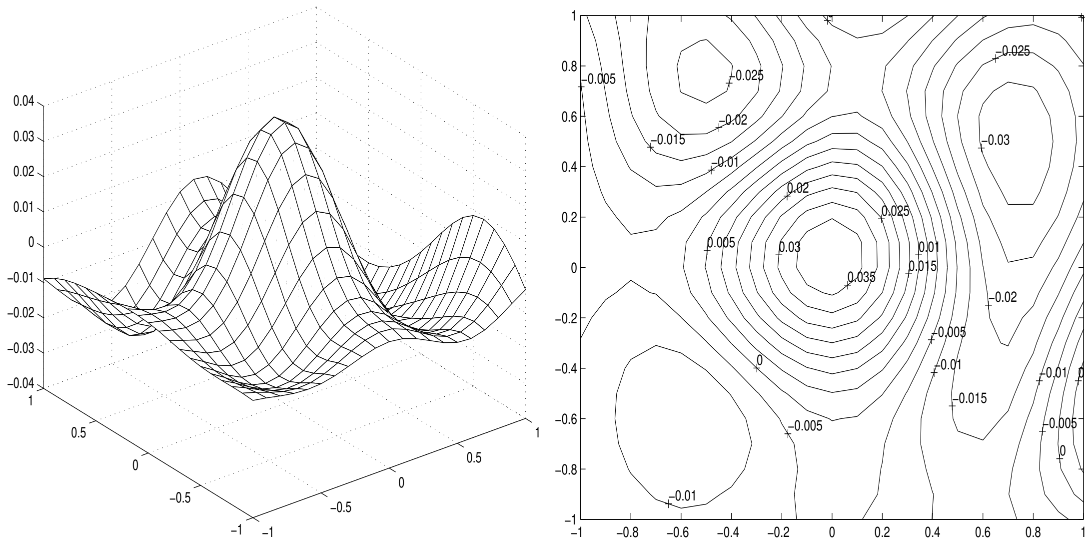


Figure 4.13: errors by R_Ξ, Max. Error=0.0397

Example 3 Given a noisy data $(\xi, y_\xi)_{\xi\in\Xi}$ with

$$y_\xi = f(\xi) + \epsilon_\xi, \quad \xi \in [-1,1]^2,$$

we construct an approximant to the function f. The underlying function is

$$f(x,y) = [B_2(1.5(x-.5)) - B_2(1.5(x+.5))] * \exp(-y^2)$$

with B_2 standard quadratic spline. Here, a set of 160 scattered centers Ξ is given in $[-1,1]^2$, shown in Figure 4.14 and ϵ_ξ's are independent normally distributed random variables with mean 0 and variance $\sigma = 0.05$. As before, the sets Ξ and $(\epsilon_\xi)_{\xi\in\Xi}$ come from a random number generator in MATLAB. The coefficient $(C(t,\xi))_{\xi\in\Xi}$ for the extrapolation f_C are computed by (4.8) with penalty function

$$\eta(t,\xi) = \left[\exp(\frac{|t-\xi|^2}{h^2}) - 1\right] |t-\xi|^2,$$

and $\bar{h}$ the density of Ξ as in (4.10). which involves the solution of a linear system in (4.21) with $k = 1$. The coefficient $(A(t,\xi))_{\xi\in\Xi}$ for $\phi_c(\cdot,t)$ are computed by the same way as in Example 1 with $k = 5$ in (4.21). In particular, because of the uncertainty of extrapolation outside of $[-1,1]^2$, we technically adopt a high tension $\omega = .1$ and assign $c = 7\omega$. Also, we use cutoff function σ such that supp$\sigma = [-M,M]^2$ with $M = .19\pi$. As a matter of fact, the tuning parameter c and ω control the trade-off between smoothness of approximant and fidelity of approximant to the data. But if the tension parameter ω is fixed for the technical issue, the parameter c is determined by ρ. In particular, the smoothing step

$$\Lambda(\chi_{\tilde{\Omega}_{2\delta}} f_C)(y) := \int_{\tilde{\Omega}_{2\delta}} \left(\frac{\sigma_\omega}{\hat{\psi}_{c/\omega}(\omega\cdot)}\right)^{\vee} (y-\theta)\, f_C(\theta)\, d\theta, \qquad y \in \tilde{\Omega}_\delta$$

with $\tilde{\Omega}_\delta$ as in (3.13), which is a key ingredient in the development of our approximation from noisy data, is influenced by the cutoff function σ with supp$\sigma = [-\eta, \eta]^2$, $\eta < 2\pi$. Then ρ and η become our tuning parameters on behalf of c and ω. However, the visual appearance of approximant is not sensitive to the choice of ρ and η. Comparisons between $\tilde{R}_\Xi$ and Wahba's thin-plate smoothing spline (TPSS) are made. Figures 4.15-4.17 give the surfaces and the contour lines of the original function f, the approximants by TPSS and $\tilde{R}_\Xi$ over the square $[-1, 1]^2$. We see that $\tilde{R}_\Xi$ provides better approximant, but also has smaller error.

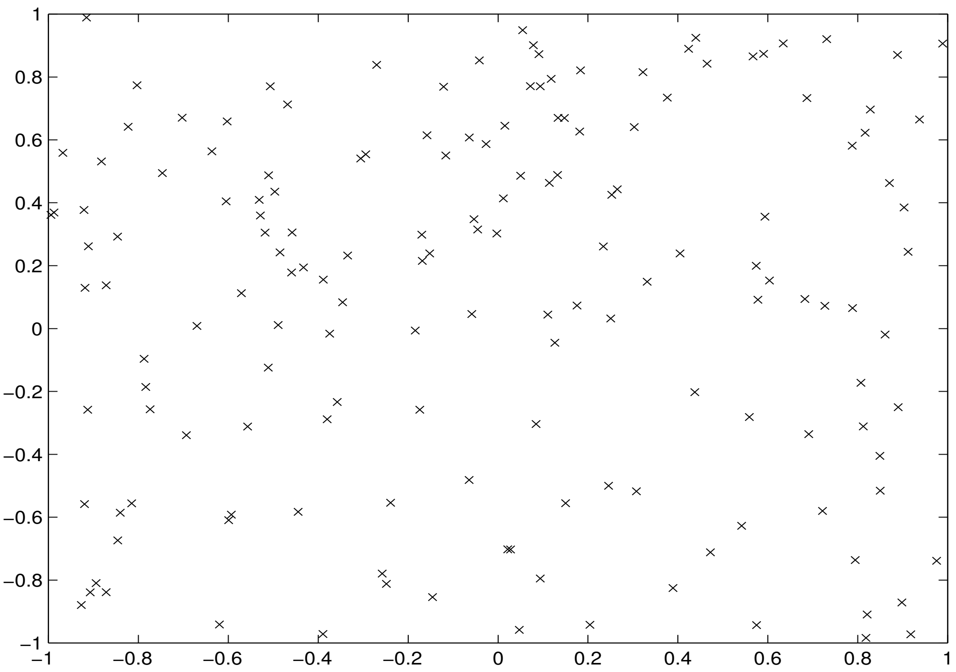


Figure 4.14: Scattered Points

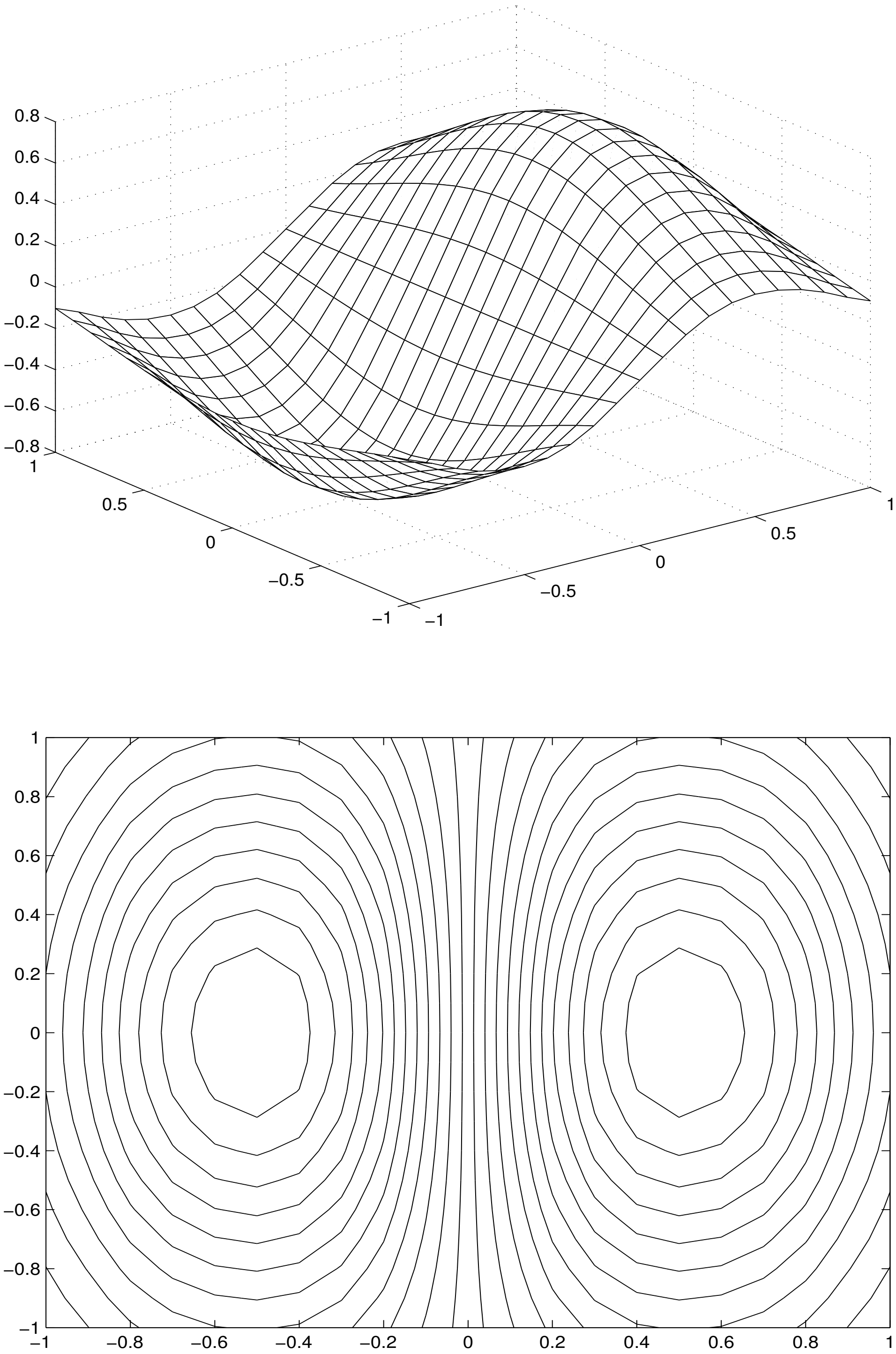


Figure 4.15: Original Function f

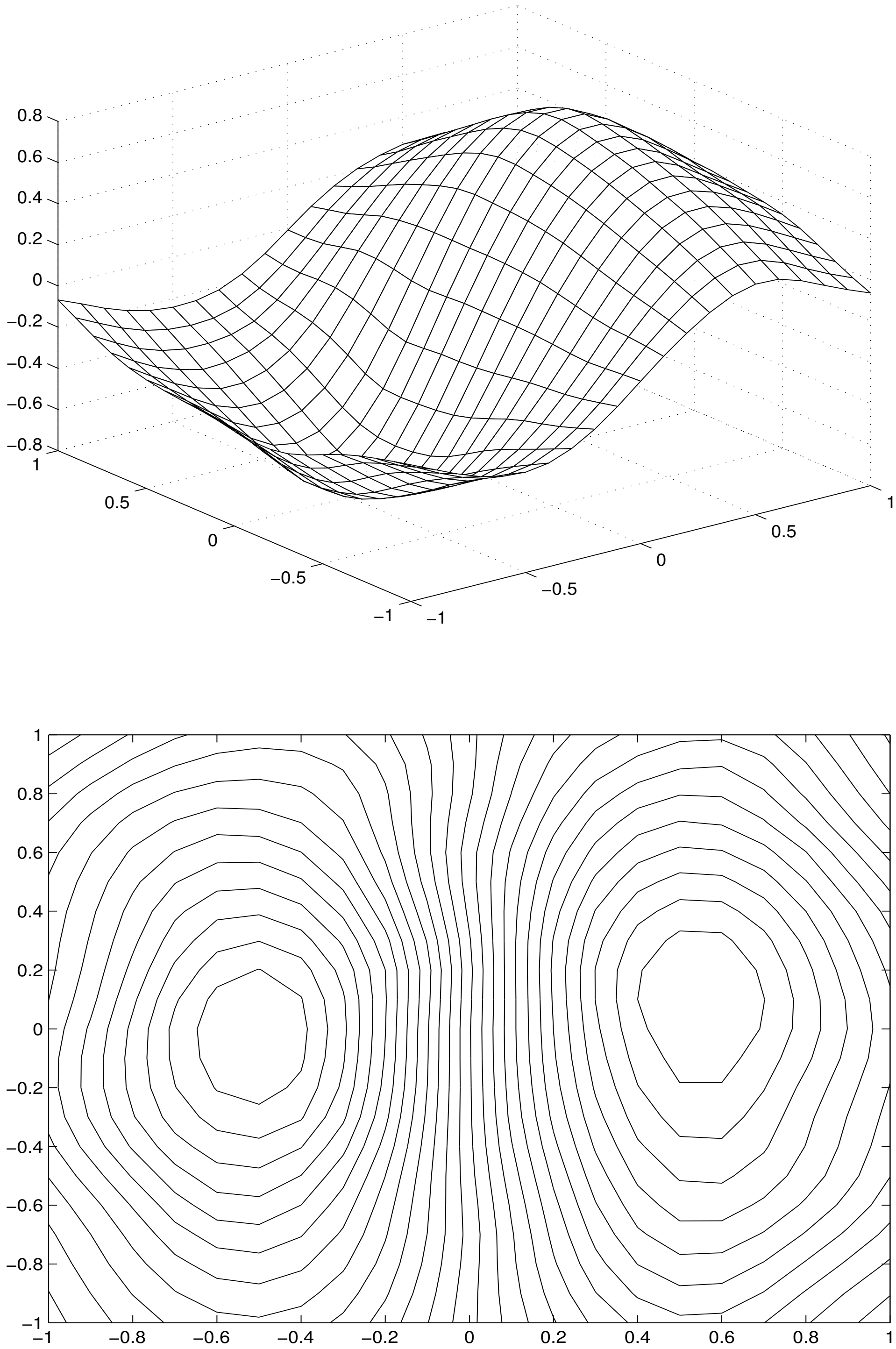


Figure 4.16: TPSS : Max. Error=0.1250

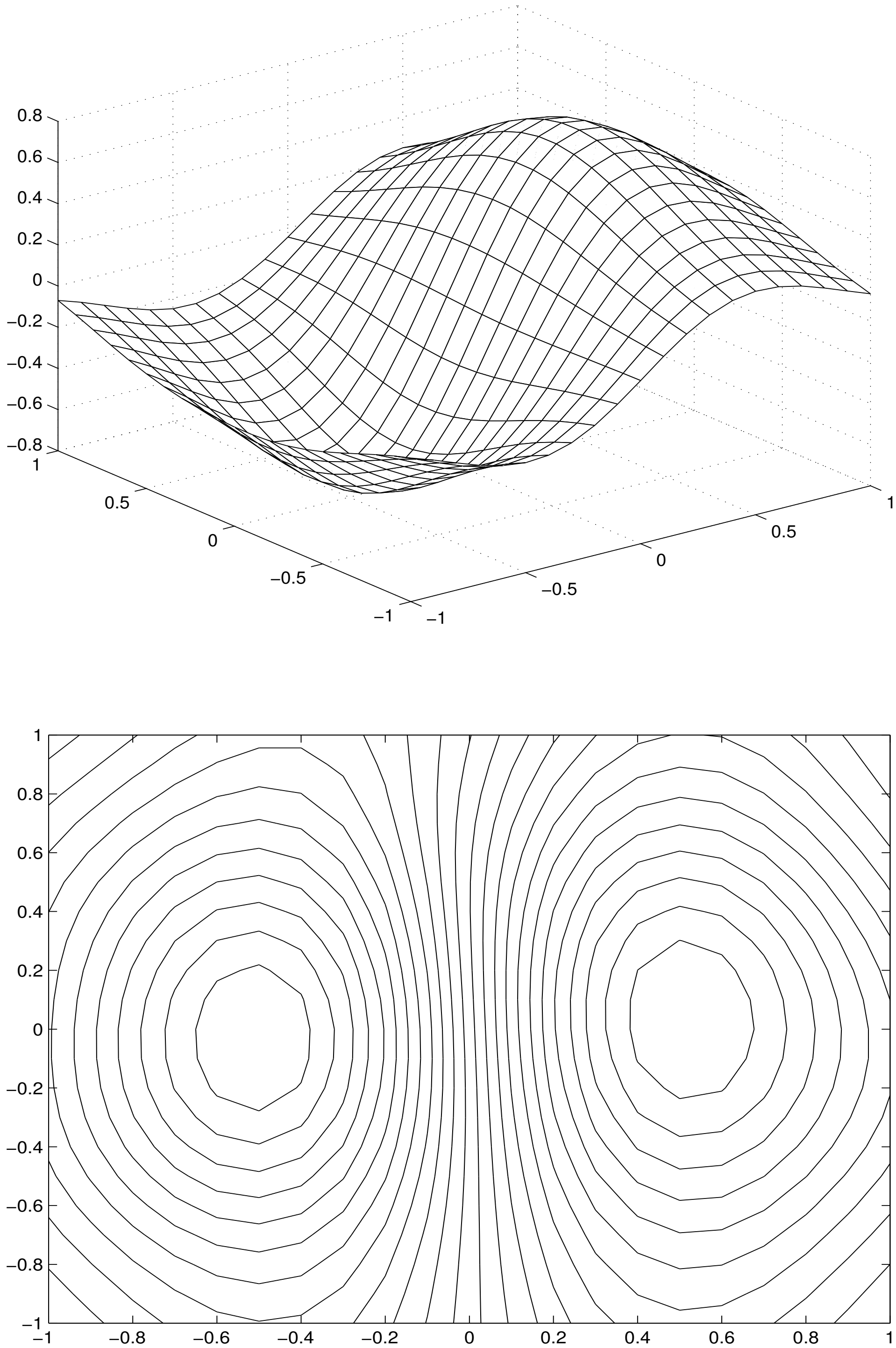


Figure 4.17: New scheme : Max. Error=0.0949

Example 4 Given a noisy data $(\xi, y_\xi)_{\xi\in\Xi}$ with

$$y_\xi = f(\xi) + \epsilon_\xi,$$

we construct an approximant to the function f. The underlying function is

$$f(x,y) = [1.4 * B_4(1.2(x-.8)) + B_3(1.2(x+.8))] * \exp(-y^2)$$

with B_k k-th order standard spline. Here, the same set Ξ and $(\epsilon_\xi)_{\xi\in\Xi}$ in Example 3 are used. The coefficients $(C(t,\xi))_{\xi\in\Xi}$ and $(A(t,\xi))_{\xi\in\Xi}$ for $f_C(t)$ and $\phi_c(\cdot,t)$ respectively are also computed by exactly same way as in Example 3. Our setting of tuning parameters are in the same spirit as in the previous example. We assign a high tension $\omega = .1$ and $c = 5\omega$. Also, we use cutoff function σ such that $\text{supp}\sigma = [-M, M]^2$ with $M = .21\pi$. Comparisons between $\tilde{R}_\Xi$ and Wahba's thin-plate smoothing spline (TPSS) are made. Figures 4.18-4.20 give the contour lines of original function f, the approximants by TPSS and R_Ξ over the square $[-1,1]^2$.

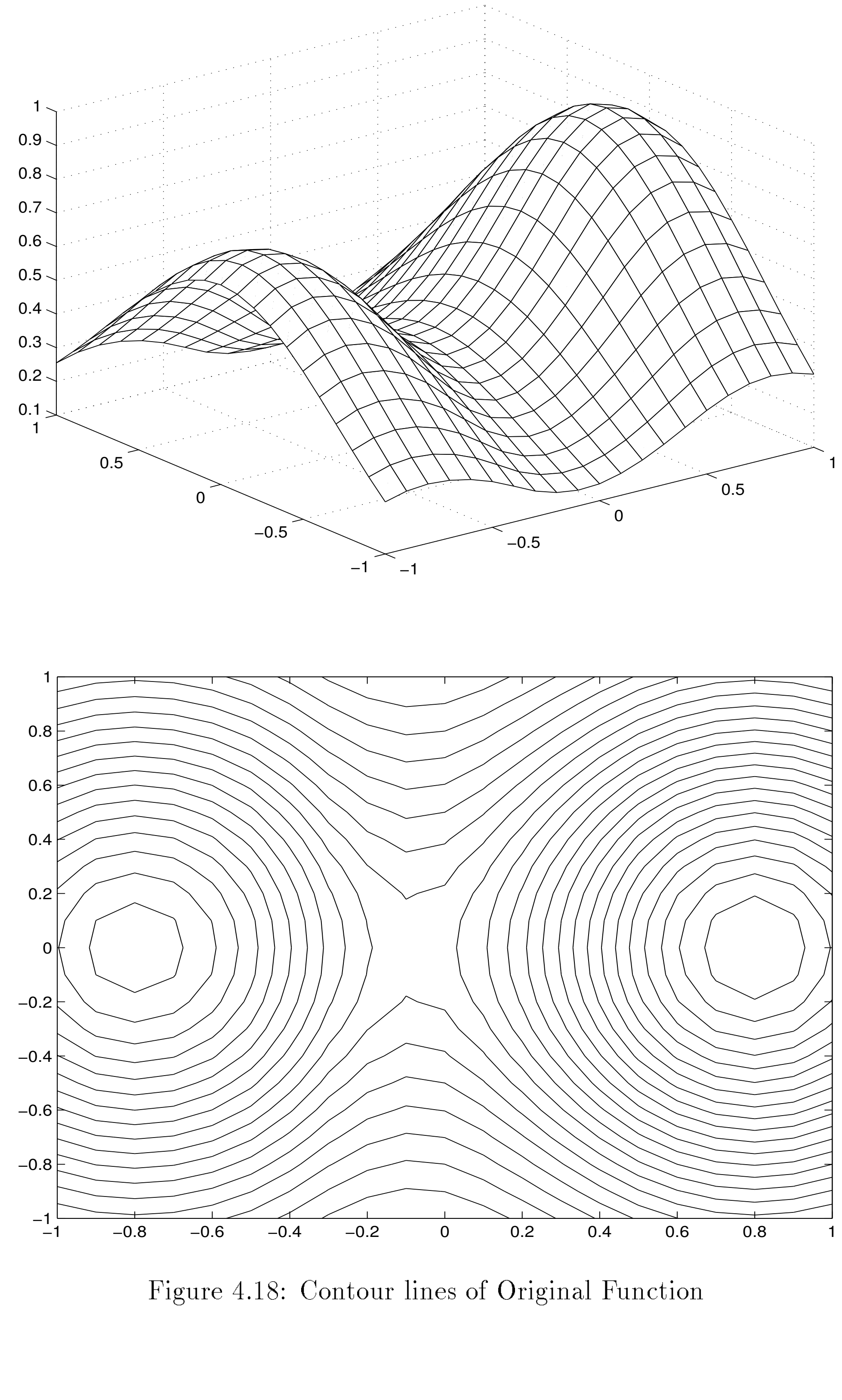
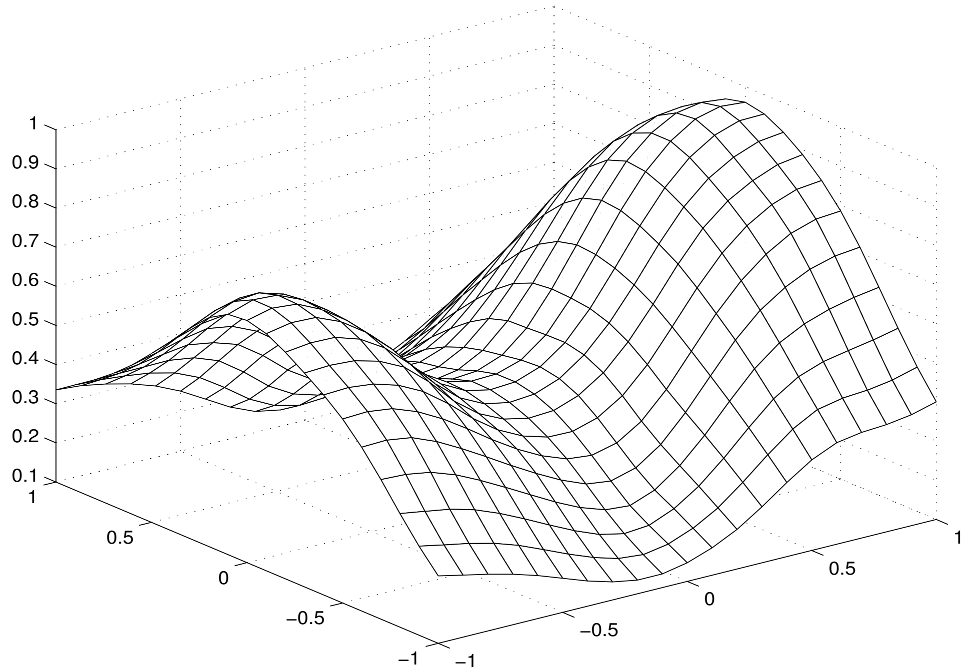


Figure 4.18: Contour lines of Original Function



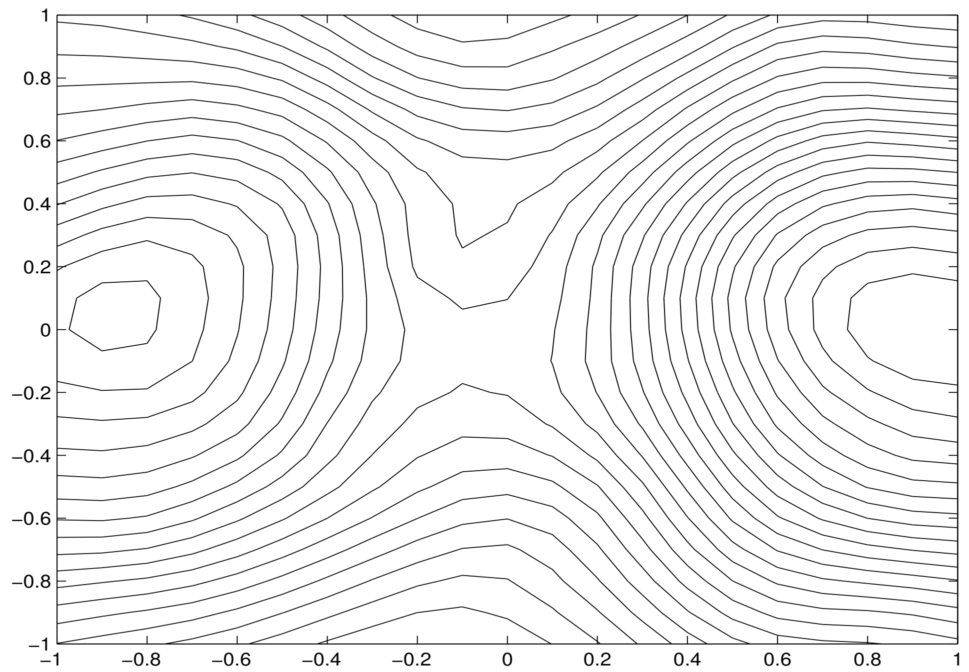
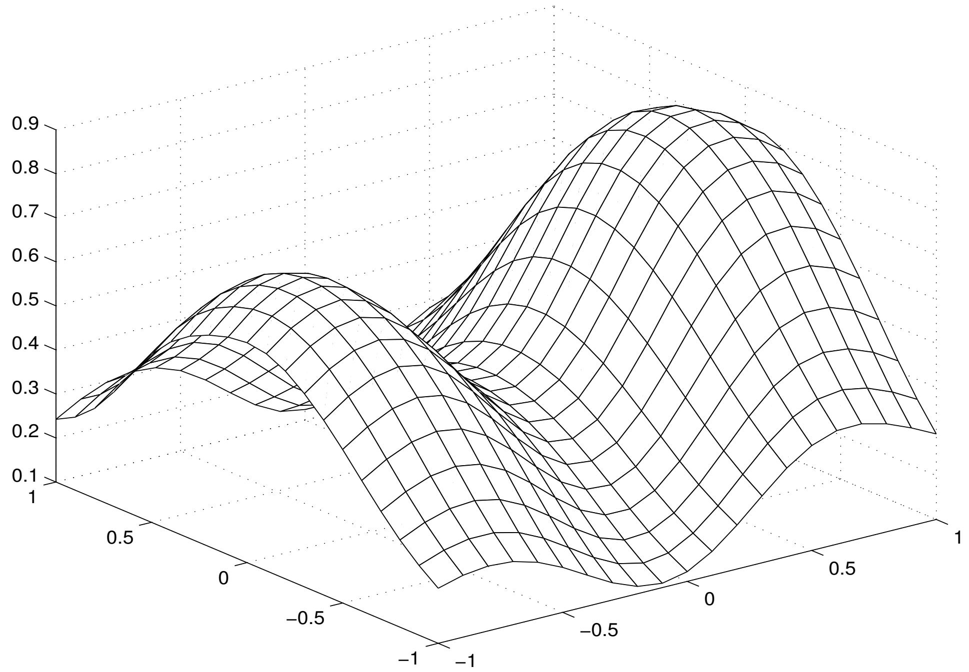


Figure 4.19: TPSS : Max. Error = 0.1136



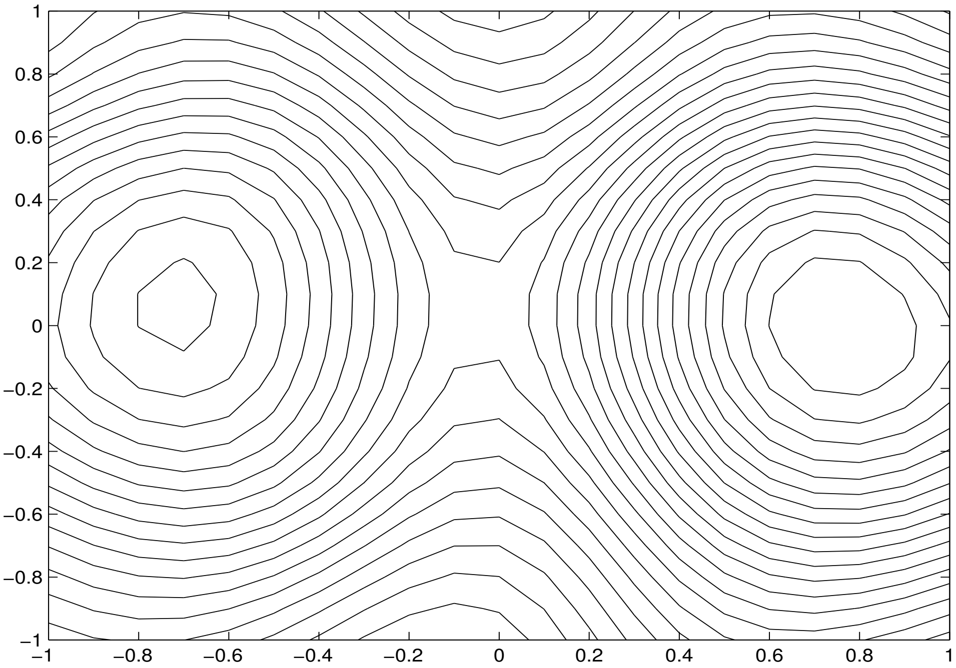


Figure 4.20: $R_{\Xi} f$: Max. Error=0.1010

Bibliography

[A] M.A. Al-Gwaiz, *Theory of Distributions*, Marcel Dekker, New York, 1992.

[AS] M. Abramowitz and I. Stegun, *A Handbook of Mathematical Functions*, Dover Publications, New York, 1970.

[BeD] R. K. Beatson and N. Dyn, *Multiquadric B-Splines*, J. Approx. Theory. **87** (1996), 1-24.

[BeL1] R. K. Beatson and W. A. Light, *Quasi-interpolation in the Absence of Polynomial Reproduction*, Numerical Methods of Approximation Theory (D. Braess and L. L. Shumaker eds.), Birkhäuser-Verlag, (1992), 21-39.

[BeL2] R. K. Beatson and W. A. Light, *Quasi-Interpolation by Thin-Plate Splines on a Square*, Constr. Approx. **9** (1993), 407-433.

[BeP] R. K. Beatson and M. J. D. Powell *Univariate Multiquadric Approximation: Quasi-Interpolation to Scattered Data*, Constr. Approx. **8** (1992), 275-288.

[B] C. de Boor, *Quasiinterpolants and approximation power of multivariable splines*, Computation of curves and surfaces (NATO Adv. Sci. Inst. Ser. C: Math. Phys. Sci., 307), (1990), 313–345.

[BDR] C. de Boor, R. DeVore, and A. Ron, *Approximation from shift-invariant subspaces of* $L_2(\mathbb{R}^d)$, Trans. Amer. Math. Soc. **341** (1994), 787–806.

[BR1] C. de Boor and A. Ron, *Fourier Analysis of the Approximation power of Principal Shift-Invariant Spaces*, Constr. Approx. **8** (1992), 427-462.

[BR2] C. de Boor and A. Ron, *Computational aspects of polynomial interpolation in several variables*, Math. Comp. **58** (1992), 705-727.

[Bu1] M. D. Buhmann, *Convergence of Univariate Quasi-Interpolation Using Multiquadrics*, IMA J. Num. Anal. **8** (1988), 365-383.

[Bu2] M. D. Buhmann, *Multivariate interpolation in odd-dimensional Euclidean spaces using multiquadrics*, Constr. Approx. **6** (1990), 21-34.

[Bu3] M. D. Buhmann, *On Quasi-Interpolation with radial basis function*, J. Approx. Th. **72** (1993), 103-130.

[Bu4] M. D. Buhmann, *New Developments in the Theory of Radial Basis Functions Interpolation*, Multivariate Approximation: From CAGD to Wavelets (K. Jetter, F.I. Utreras eds.), World Scientific, Singapore, (1993), 35-75.

[BuDL] M. D. Buhmann, N. Dyn, and D. Levin, *On quasi-interpolation with radial basis functions with scattered centres*, Constr. Approx. **11** (1995), 239-254.

[BuR] M. D. Buhmann and A. Ron, *Radial Basis Functions: L^p-approximation orders with Scattered centers*, Wavelets, Images and Surface Fitting, (P. J. Laurent, A Le Méhaué, and L. L. Schumaker eds.), (1995), 93-112.

[D] N. Dyn, *Interpolation and Approximation by Radial and Related Functions*, Approximation Theory VI, (C. K. Chui, L. L. Schumaker and J. Ward eds.), Academic press, (1989), 211-234.

[DJLR] N. Dyn, I.R.H. Jackson, D. Levin, and A. Ron, *On Multivariate Approximation by Integer Translates of a basis Function*, Israel Journal of Mathematics **78** (1992), 95-130.

[DL] N. Dyn and D. Levin, *Bell shaped basis functions for surface fitting*, Approximation Theory and Applications (Z. Ziegler Ed.), Academic Press, New York, (1981), 113-129.

[DR] N. Dyn and A. Ron, *Radial basis function approximation: from gridded centers to scattered centers*, Proc. London Math. Soc. **71** (1995), 76-108.

[GMM] A. Gray, G. B. Mathews, and T. M. Macrobert, *A Treatise on Bessel Function and Their Application to Physics*, MacMillan, London, 1992.

[GS] I.M. Gelfand and G.E. Shilov, *Generalized Functions*, Vol. 1, Academic Press, 1964.

[Ja] I. R. H. Jackson *An order of convergence for some radial basis functions*, IMA J. Numer. Anal. **9** (1989), 567-587.

[J1] M. Johnson, *An upper bound on the approximation power of principal shift-invariant subspaces*, Constr. Approx., to appear.

[J2] M. Johnson, *Bound on the Approximation Order of Surface Splines*, Constr. Approx., to appear.

[L] D. Levin, *Near-best scattered-data approximation in* $\mathbb{R}^d$, Math. Comp. (to appear).

[M] C. A. Micchelli, *Interpolation of Scattered Data: Distance Matrices and Conditionally Positive Functions*, Constr. Approx. **2** (1986), 11-22.

[MN1] W. R. Madych, S. A. Nelson, *Multivariate interpolation and conditionally positive function I*, Approximation Theory and its Application **4** (1988), no 4, 77-89.

[MN2] W. R. Madych, S. A. Nelson, *Error bounds for Multivariate interpolation*, Approximation Theory VI: Vol 2, (C. K. Chui, L. L. Schumaker and J. D. Ward eds.), Academic press, (1989), 413-416.

[MN3] W. R. Madych, S. A. Nelson, *Multivariate interpolation and conditionally positive function II*, Math. Comp. **54** (1990), 211-230.

[MN4] W. R. Madych, S. A. Nelson, *Bounds on Multivariate Polynomials and Exponential Error Estimates for Multiquadric Interpolation*, J. Approx. Theory. **70** (1992), 94-114.

[P1] M. J. D. Powell, *The Theory of Radial basis functions approximation in 1990*, Advances in Numerical Analysis Vol. II: Wavelets, Subdivision Algorithms and Radial Basis Functions (W.A. Light ed.), Oxford University Press, (1992), 105-210.

[P2] M. J. D. Powell, *The uniform convergence of thin plate spline interpolation in two dimensions*, Numer. Math. **9** (1994), 107-128.

[R1] A. Ron, *The L_2-Approximation Orders of Principal Shift-Invariant Spaces Generated by a Radial Basis Function*, Numerical Methods of Approximation Theory (D. Braess and L. L. Shumaker eds.), vol. **9** (1991), 245-268.

[R2] A. Ron, *Negative observations concerning approximations from spaces generated by scattered shifts of functions vanishing at* ∞, J. Approx. Theory **78** (1994), 364-372.

[Ru] W. Rudin, *Functional Analysis*, McGraw-Hill, New York, 1973.

[S] R. Schaback, *Approximation by Radial Basis functions with Finitely Many Centers*, Constr. Approx. **12** (1996), 331-340.

[SF] G. Strang and G. Fix, *A Fourier analysis of the finite element variational method*, Constructive Aspects of Functional Analysis (G.Geymonat, ed.), C.I.M.E., 1973, 793-840.

[SW] E.M. Stein and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton Univ. Press, Princeton, NJ, 1971.

[W] G. Wahba *Spline Models for Observational Data*, CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 59. SIAM, 1990.

[WS] Z. Wu and R. Schaback, *Local error estimates for radial basis function interpolation of scattered data*, IMA J. Numer. Anal. **13** (1993), 13-27.