Monografías de la Academia de Ciencias de Zaragoza. **20:** 111–135, (2002).

Testing Methods for $3D$ Scattered Data Interpolation

Mira Bozzini and Milvia Rossini

Department of Mathematics and Applications. Univ. of Milano Bicocca

via Bicocca degli Arcimboldi 8 I-20126 Milano Italy

email: bozzini@matapp.unimib.it, rossini@matapp.unimib.it

Abstract

This paper is concerned with the evaluation of methods for $3D$ scattered data interpolation. In particular, we discuss the computational performances in $3D$ of methods which give superior results for the two-dimensional case. The testing process was carried out by considering the accuracy, the graphical behaviour of the interpolant and the timing. In addition we have taken into account the computational efficiency and the sensitivity respect to the sample. Moreover, in order to evaluate the graphical behaviour, we present an evolutive visualization of the interpolant.

1 Introduction

The problem of constructing a smooth function $g(x, y, z)$, $g : Q \subset \mathbb{R}^3 \to \mathbb{R}$ which takes on certain prescribed values

$$g(x_i, y_i, z_i) = f_i, \quad i = 1, \ldots, N, \quad (x_i, y_i, z_i) \in Q, \tag{1}$$

arises in many applied fields. We mention some examples that can be useful for a computational analysis:

- behaviour of the temperature in a furnace,
- concentrations of a mineral in the soil,
- behaviour of precipitations in a geographical area,
- electroencephalogram (EEG).

In the literature there are methods which can be extended to any dimension d (for instance the radial basis methods) and the theoretic properties of some of these have been studied (see [17]), but little or nothing is known on the computational results for $d = 3$.

For the case $d = 2$, Franke has evaluated (see his well–known paper [8]) the numerical performances of a wide class of methods. For $d = 3$, there is the paper [13] which gives a first answer to the problem for data sets with a moderate dimension.

Our aim is to study the computational behaviour of methods to interpolate $3D$ scattered data. Taking into account [8], we have analyzed those methods which have shown, at least in two dimensions, the better performances.

In the testing process we have considered, as usual, the accuracy and, in addition, all those aspects that in the $3D$ case can give more problems than in two dimensions. Namely we have considered the stability and the computational costs. By the definition of two indices, we provide the computational efficiency of the different methods. In addition we think it was important to analyze the conditioning of the sample respect to the approximating function. Finally we provide a new visualization method, useful for the comparison of the graphical behaviour of the interpolants. For sake of brevity, we will report only a few examples. The full experimentation can be found in [5].

2 The testing process

2.1 The methods considered

In [8] it is shown that the methods giving the superior performances (see table 1 of [8]), belong to the classes of inverse distance weighted methods (the modified quadratic Shepard Method), triangle based methods (Nielson minimum norm network), radial basis methods (thin-plate splines, multiquadrics).

The triangle based methods need a triangulation of the convex hull of the point set and, to achieve a C^1 function, they need also a scheme for estimating some derivatives at the data points. The quality of the corresponding interpolating function (as measured in terms of visual appearance, smoothness, and accuracy) depends critically on the "accuracy" of the derivative estimates. Moreover, it is known that triangulations have to satisfy some optimal conditions that, at the moment, do not exist for three dimensions.

For these reasons, we have considered only inverse distance weighted methods (the modified quadratic Shepard method (MQSM)), and the standard radial basis functions

$$\phi(r) = r^3, \quad (2)$$
$$\phi(r) = r^2 log r, \quad (3)$$
$$\phi(r) = \sqrt{r^2 + c^2} \quad (4)$$
$$\phi(r) = (1-r)_+^4 (1+4r). \quad (5)$$

The reader can see the Appendix A for a short description of these methods.

2.2 Test functions

The methods have been tested on functions that exhibit a variety of behaviours. Namely, we have considered the $3D$ extension of the functions used in [8]

$$\begin{aligned}
\mathbf{F2}(x,y,z) &= (\tanh(9z-9x-9y)+1)/9 \\
\mathbf{F3}(x,y,z) &= \cos(6z)\left(1.25+\cos(5.4y)\right) / \left(6+6\left(3x-1\right)^2\right) \\
\mathbf{F4}(x,y,z) &= \exp\left\{-\frac{81}{16}\left[(x-0.5)^2+(y-0.5)^2+(z-0.5)^2\right]\right\}/3 \\
\mathbf{F5}(x,y,z) &= \exp\left\{-\frac{81}{4}\left[(x-0.5)^2+(y-0.5)^2+(z-0.5)^2\right]\right\}/3 \\
\mathbf{F6}(x,y,z) &= \sqrt[2]{64-81\left[(x-0.5)^2+(y-0.5)^2+(z-0.5)^2\right]}/9-0.5.
\end{aligned}$$

In addition we have considered some functions with particular behaviours. A "relative" of the two-dimensional Franke's function

$$\begin{aligned}
\mathbf{F1}(x,y,z) &= .75\exp\left[-\frac{(9x-2)^2+(9y-2)^2+(9z-2)^2}{4}\right] \\
&\quad +0.75\exp\left[-\frac{(9x+1)^2}{49}-\frac{(9y+1)^2}{10}-\frac{(9z+1)^2}{10}\right] \\
&\quad +0.5\exp\left[-\frac{(9x-7)^2+(9y-3)^2+(9z-5)^2}{4}\right] \\
&\quad -0.2\exp\left[-(9x-4)^2-(9y-7)^2-(9z-5)\right],
\end{aligned}$$

an extension of the sigmoidal function

$$\mathbf{F7}(x,y,z) = 1/\sqrt{1+2exp(-3(\sqrt{x^2+y^2+z^2}-6.7))},$$

and the peak function

$$\begin{aligned}
\mathbf{F8}(x,y,z) &= 50exp(-200((x-0.3)^2+(y-0.3)^2)) \\
&+ \quad exp(-50((x-0.5)^2+(y-0.5)^2)).
\end{aligned}$$

These functions have been considered in $Q = [0,1]^3$.

2.3 Data configurations

The test functions have been sampled at N scattered points of Q. Let

$$S_N = \{P_i(x_i, y_i, z_i) \in Q = [0,1]^3,\ i = 1, \ldots, N\}$$

be the point set.

The word "*scattered*" may have different meanings.

In [8], scattered means that the points of S_N are not assumed to satisfy any particular condition as spacing or density. (In a simulation process, the point are obtained from the generation of pseudo random numbers.)

In practice, many people use points that satisfy some condition. We can have

- *equidistributed scattered points*. For instance, they can be obtained by a pseudo random number generation so that one point falls in each subcube of side $1/\sqrt[3]{N}$.
- *Perturbed grid points*. If G_i are the points of a grid G, the points of S_N are

  $$P_i = G_i + e_i,$$

  where e_i are random variables with mean $E(e_i) = 0$ and variance (which measures the distortion from the grid points) $E(e_i^2) = \sigma^2$.

We call all these point sets *volumetric data*.

In many applications, the data points are scattered but with some structure. For instance, when we study mineral concentrations in the subsoil, we pick up the data by drillings. This mean that the points P_i are scattered along some straight lines of Q. In other situations, the points of S_N can be scattered on certain number of planes or, more generally, on some surfaces (for instance when we study the monthly or seasonal precipitations).

These kind of data (say *structured scattered data*) have been discussed in [6].

2.4 Comparison

It is now important to choose the characteristics on which the methods are to be evaluated and compared.

In our opinion, two fundamental aspects are the accuracy and the graphical behavior. In addition, we have also considered the computation times and the condition of the interpolation matrices.

Accuracy. We measure the accuracy with the root mean square error e_2 and the maximum error e_∞:

$$e_2 = \sqrt{\frac{\sum_{i=1}^{n_x}\sum_{j=1}^{n_y}\sum_{l=1}^{n_z}(g(x_i,y_j,z_l) - f(x_i,y_j,z_l))^2}{n_x n_y n_z}}, \tag{6}$$

$$e_\infty = \max_{(x_i,y_j,z_l)\in G} |g(x_i,y_j,z_l) - f(x_i,y_j,z_l)|, \tag{7}$$

where $g(x, y, z)$ is the interpolant function, $f(x, y, z)$ is the test function, and (x_i, y_j, z_l) are the points of a grid G of Q with size $n_x \times n_y \times n_z$.

These indices provide a global information on the resulting approximation.

Graphical visualization. The interpolant graphic gives an immediate indication on the approximation goodness. In fact, different methods could give the same accuracy but some may reproduce the function behaviour more faithfully than the others.

Since trivariate function representations describe hypersurfaces in $\mathbb{R}^4$, it is obvious that we cannot render them directly. It is possible to visualize their behaviour by displaying one or more projected surfaces in $\mathbb{R}^3$. For example, we can work with isoparametric surfaces which correspond to constant values of one variable x, y or z. Alternatively, we can create contour plots of hypersurfaces which correspond to constant function values. Surfaces of constant parameter values are considered in [12], [7], [1]. [15] and [14] suggest using a combination of three isoparametric surfaces $g(x_i, y, z)$, $g(x, y_j, z)$ and $g(x, y, z_k)$, associated with the value $g(x_i, y_j, z_k)$. This idea involves an axonometric view of the domain with the planes $x = x_i$, $y = y_j$ and $z = z_k$, along with another copy of the domain with the graphs of the three surfaces located on the faces of the cubical domain. The paper [14] contains another example of visualization named *slice viewer*. A survey of other techniques can be found in [11].

Although these techniques are very useful, they still provide a static representation which do not allow to have a dynamic vision of the problem we are studying.
Usually, in a real problem, the variation of one independent variable represents the evolution of some phenomenon. For instance, to study mineral concentrations in the subsoil means to study how the concentration changes in relation to depth.
Therefore, we believe that the visualization should be connected to the problem features and point out the evolution of the phenomenon with respect to the variable describing its variation.
For this reason, we have considered an evolutive representation. We pick up the variable describing the evolution (for instance z), we take some values of it $(z_1, \ldots, z_m)$, and we consider the evolution of the surfaces $g(x, y, z_k)$, $k = 1, \ldots, m$. (See figs. 1–10 in the Appendix).

3 Volumetric data

In this section we deal with the interpolant behavior when the sample dimension N increases.

First we will compare the method performances both with respect to the accuracy (§3.1) and to the graphical visualization (§3.2).

Then we will consider the condition numbers of the interpolation matrices and the computational efficiency in relation to the sample size N. Finally, in §3.4 we will discuss the sensitivity of the methods with respect to the point locations (for point sets S_N with the same size).

Here we have considered the case of scattered equidistributed data, because, for this case, the theoretical results for radial basis functions surely hold. Examples for the other point sets (scattered and perturbed grid points) can be found in [5].

3.1 Accuracy

The values of e_2 and e_∞ for N=343, 1000, 3375 are shown in tables 1, 2 and 3. These indices have been computed using the points of a grid of size $n_x = n_y = n_z = 21$.

The tables show that the methods give equivalent results. We can also say that, in general, the radial basis function (2) provides the better accuracy.

Moreover, the error decreases, as N increases, according to the theory [18].

Remark. For the multiquadric, we know that the optimal choice of c is still an open problem. There are some suggestions for the two–dimensional case, including methods based only on the points distribution. [20] suggests a value that takes account of how the points are dispersed in both the x and y directions

$$c = \sqrt{1/10 \max\{\max_{i,j} |x_i - x_j|, \max_{i,j} |y_i - y_j|\}}. \tag{8}$$

[10] uses

$$c = 0.851d, \tag{9}$$

where d is the average distance of the points to their near neighbors. Franke replaced d by $D/\sqrt{N}$, where D is the diameter of the minimal circle enclosing all data points and suggests to use

$$c = 1.25D/\sqrt{N}. \tag{10}$$

These techniques can be trivially extended to the $3D$ case and they provide more or less the same results (both for the accuracy and the graphical behaviour [5]). We have not considered techniques that take account of the function values f_i (see, for instance, [16]) because we believe they are too expensive for our case.

In tables 1, 2, 3, the c parameter has been computed using (9).

3.2 Graphical visualization

Let us now compare the graphical performances. For sake of brevity, we will show only the graphs related to the test functions $F1$ and $F2$ (which exhibit particular behaviors),

Method	e	$F1$	$F2$	$F3$	$F4$	$F5$	$F6$	$F7$	$F8$
r^3	e_2	4.6e-3	8.5e-3	4.2e-3	5.1e-4	6.5e-4	5.4e-3	2.1e-2	2.6
	e_∞	4.2e-2	7.8e-2	4.5e-2	5.6e-3	8.3e-3	1.3e-1	2.2e-1	42e
$r^2\log(r)$	e_2	5.9e-3	9.1e-3	6.9e-3	1.1e-3	1.3e-3	8.7e-3	2.2e-2	2.6
	e_∞	9.3e-2	7.5e-2	7.1e-2	1.4e-2	1.7e-2	1.8e-1	2.1e-1	42
$\sqrt{r^2+c^2}$	e_2	6.6e-3	9.2e-3	8.2e-3	8.3e-4	1.2e-3	1.2e-2	2.3e-2	2.7
	e_∞	1.5e-1	8.2e-2	8.2e-2	1.0e-2	1.5e-2	2.3e-1	2.2e-1	42
$(1-r)_+^4(1+4r)$	e_2	5.3e-3	8.5e-3	3.9e-3	7.3e-4	8.6e-4	1.2e-2	2.1e-2	2.8
	e_∞	1.2e-1	7.3e-2	4.6e-2	8.7e-3	8.4e-3	2.3e-2	2.2e-1	43
MQSM	e_2	9.4e-3	9.7e-3	5.9e-3	2.2e-3	2.1e-3	4.5e-3	2.3e-2	3.5
	e_∞	1.3e-1	9.1e-2	6.1e-2	1.4e-2	1.9e-02	1.2e-1	1.9e-1	46

Table 1: Errors for $\mathbf{N = 343}$.

and to $F5$ (for which we get the better accuracy). In fig. 1 their evolutions respect to the z–variable are shown.

We begin to consider the interpolants achieved with (2), which provides the better accuracy.

Figures 2, 3, 4 point out that, generally, samples of size $N = 343$ do not assure a good phenomenon reproduction. Only for $F5$, we get adeguate results.
In this case, for $N = 1000$, we have a remarkable improvement at the boundary (see fig. 2).
This does not happen for $F1$ and $F2$. For $F1$, the interpolant does not reproduce the correct behaviour near the boundary $z = 1$. The approximation in this zone becomes better when we consider $N = 3375$. But on the boundary $z = 1$ the function behavior has not yet been reproduced (fig. 3).

For $F2$ (fig. 4), we may observe an improvement from $N = 343$ to $N = 1000$, even if we have many oscillations that are eliminated, only in part, considering $N = 3375$.

Let us now consider the other methods. As already said, the interpolants achieved from the radial basis (3)–(5), provide, in mean, the same accuracy of (2), but their graphical behaviors can be worse. See, for instance, fig. 5 where the $F1$ interpolants are shown ($N = 1000$).

Finally, let us consider the modified quadratic Shepard method. The experimentation has shown that, near the boundary, it can perform worse than the other methods. But for samples with $N > 1000$, it provides approximations comparable to those given by (2) and with lower computational costs (see §3.3). In fig. 6 we show the graphs of the $F1$ MQSM interpolants ($N = 343, 1000, 3375$).

Method	e	$F1$	$F2$	$F3$	$F4$	$F5$	$F6$	$F7$	$F8$
r^3	e_2	1.5e-3	4.7e-3	1.3e-3	1.5e-4	1.1e-4	3.3e-3	7.8e-3	1.4
	e_∞	4.0e-2	7.3e-2	1.4e-2	1.8e-3	1.3e-3	9.6e-2	1.1e-1	34
$r^2\log(r)$	e_2	2.2e-3	5.4e-3	2.2e-3	4.4e-4	2.4e-4	5.3e-3	1.0e-2	1.5
	e_∞	5.3e-2	8.0e-2	2.0e-2	7.2e-3	3.5e-3	1.4e-1	1.3e-1	35
$\sqrt{r^2+c^2}$	e_2	2.9e-3	5.6e-3	3.0e-3	4.0e-4	2.7e-4	7.9e-3	1.1e-2	1.6
	e_∞	9.1e-2	8.3e-2	3.4e-2	6.7e-3	4.3e-3	1.8e-1	1.4e-1	35
$(1-r)_+^4(1+4r)$	e_2	1.8e-3	4.7e-3	9.7e-4	2.9e-4	2.5e-4	6.7e-3	8.4e-3	1.4
	e_∞	5.8e-2	7.0e-2	1.2e-2	4.5e-3	2.6e-3	1.7e-1	1.2e-1	33
MQSM	e_2	2.7e-3	4.8e-3	2.1e-3	4.6e-4	7.9e-4	2.7e-3	1.0e-2	1.1
	e_∞	3.9e-2	9.0e-2	2.7e-2	4.3e-3	8.7e-3	8.0e-2	1.0e-1	30

Table 2: Errors for $\mathbf{N = 1000}$.

We can evaluate the graphical performances also considering the absolute error graphs. We will see that this can be useful, but at the same, time misleading.
In fig. 7, the errors of the $F1$ interpolants are shown ($N = 343$, 1000, 3375, radial basis (2)).
In each graph, we have used the scale given by the maximum error obtained for $N = 343$. We can notice that the higher errors are at the extrema, while near the boundary $z = 1$, the approximation seems to be accurate. But we have remarked that, just in this zone, the interpolant does not reproduce the correct $F1$ behavior (see fig. 1 and 3).

3.3 Conditioning, computation times and efficiency

In this section we shall see how the condition number $K_2(A)$ of the interpolation matrix, the computation times and the efficiency change with regard to the sample dimension N.

• **Condition Number.** For the global methods (2)–(5), $K_2(A)$ increases as shown in table 4.
For the local method MQSM, we do not have this problem because we solve a linear system of small dimension ($n = 9$).

• **Computation times**. For the methods (2)–(5), the increase of sample size leads to a remarkable increase of computation times. This is due both to the solution of a linear system with full matrix of order N, and to interpolant evaluation. But efficient methods exist for the evaluation (see, for instance, [2]). Therefore, in table 5[1], we show only the system solution times (radial bases (2)). For MQSM, the times include also the interpolant

[1]The computations have been performed on a PC with a AMD 800 Mhz processor and 256 Mb Ram.

Method	e	$F1$	$F2$	$F3$	$F4$	$F5$	$F6$	$F7$	$F8$
r^3	e_2	3.7e-4	1.3e-3	5.2e-4	3.7e-5	2.2e-5	1.9e-3	2.4e-3	4.1e-1
	e_∞	8.0e-3	2.7e-2	9.5e-3	4.5e-4	2.6e-4	6.6e-2	3.9e-2	13
$r^2\log(r)$	e_2	8.3e-4	1.6e-3	1.1e-3	1.6e-4	4.6e-5	3.3e-3	3.7e-3	4.9e-1
	e_∞	3.5e-2	3.4e-2	1.7e-2	3.5e-3	6.6e-4	9.8e-2	6.0e-2	15
$\sqrt{r^2+c^2}$	e_2	1.8e-3	1.8e-3	1.7e-3	2.0e-4	7.6e-5	4.8e-3	4.6e-3	4.8e-1
	e_∞	8.0e-2	3.7e-2	2.6e-2	4.3e-3	1.1e-3	1.4e-1	7.8e-2	16
$(1-r)_+^4(1+4r)$	e_2	7.4e-4	1.2e-3	3.9e-4	9.5e-5	7.3e-5	3.3e-3	2.6e-3	4.1e-1
	e_∞	4.0e-2	2.7e-2	7.0e-3	2.1e-3	7.3e-4	1.1e-1	4.4e-2	13
MQSM	e_2	8.7e-4	1.6e-3	6.5e-4	1.4e-4	2.4e-4	1.5e-3	3.4e-3	5.6e-1
	e_∞	1.5e-2	2.4e-2	1.2e-2	1.3e-3	2.5e-3	4.9e-2	5.3e-1	15

Table 3: Errors for $\mathbf{N = 3375}$.

$\phi(r)$	$N = 343$	$N = 1000$	$N = 3375$
r^3	3.9e+06	2.8e+07	9.5e+08
$r^2\log(r)$	5.4e+04	3.4e+05	5.3e+06
$\sqrt{r^2+c^2}$	2.3e+05	9.0e+05	1.2e+07
$(1-r)_+^4(1+4r)$	7.7e+04	5.2e+05	1.7e+07

Table 4: Condition numbers for the radial basis interpolation matrices.

evaluation.

Le us notice that to solve the system with $N = 3375$, we need a computation time which is approximately equal to $1000t_{343}$. MQSM is more advantageous: in fact $t_{3375} \approx 5t_{343}$.

N	r^3	MQSM
343	0.55	1.04
1000	14.94	1.96
3375	547.28	5.5

Table 5: Computation times.

• **Computational efficiency.** The computational efficiency can be evaluated by two different indices.

1. Computational efficiency, defined as the inverse of the product between the compu-

tation time t_N and the relative error $e^r_\infty(N)$

$$Ef_N = \frac{1}{t_N e^r_\infty(N)}, \tag{11}$$

2. Ratio quality (maximum error $e_\infty(N)$) to cost (computation time t_N)

$$R_N = \frac{e_\infty(N)}{t_N}. \tag{12}$$

The values of (11) e (12) for the the radial basis (2) and MQSM are shown in table 6 (test function $F1$).

When we use a radial basis function method, the increase of N leads to high computation times not rewarded by an appreciable reduction of the error. This causes a considerable loss of efficiency.

According to definition (11), to use a sample of size $N = 1000$ is ten times more efficient than to use a sample with $N = 3375$. If we consider definition (12), using $N = 1000$ is approximately one hundred and eighty times more efficient than using $N = 3375$.

MQSM is more efficient. In fact we have $Ef_{3375} \approx Ef_{1000}$, while, according to (12), to use $N = 1000$ is five times more efficient than to use $N = 3375$.

Increasing the sample dimension is not always convenient. In fact, on one hand, the interpolant may not reproduce anyway the correct behavior in some zones and on the other we may have a considerable loss of efficiency.

Method	Efficiency	$N = 343$	$N = 1000$	$N = 3375$
r^3	Ef_N	25.97	0.99	0.14
	R_N	7.6e-2	2.6e-3	1.5e-5
MQSM	Ef_N	4.4	7.8	7.3
	R_n	1.2e-1	1.2e-2	2.7e-3

Table 6: Computational efficiency.

3.4 Sensitivity respect to the point set

We conclude our analysis by considering the method sensitivity with regard to the point sets S_N. That is we want to see how the point locations influence the interpolant problem solution.

Here, we report the results obtained for three different equidistributed point sets S^1_N , S^2_N e S^3_N, with $N = 343$, 1000, 3375 and show what we get for $F1$ and the radial basis (2). In [5] we have considered also scattered points and distorted grid points.

The error does not change very much: in fact it is a function of N. The graphs, indeed, point out that the point locations condition the local behavior of the interpolants, especially for $N = 343$ and $N = 1000$ (figs. 8, 9). For $N = 3375$ the graphical behavior is always the same (fig. 10). That is we have stability respect to the data set.

N	e	S_N^1	S_N^2	S_N^3
343	e_2	4.6e-3	4.0e-3	3.9e-3
	e_∞	7.0e-2	5.5e-2	6.1e-2
1000	e_2	9.5e-4	1.4e-3	1.4e-3
	e_∞	1.2e-2	3.3e-2	3.8e-2
3375	e_2	3.7e-4	2.5e-4	3.1e-4
	e_∞	8.0e-3	5.3e-3	5.1e-3

Table 7: Sensitivity respect to point sets. Errors for the radial basis function (2).

This analysis has also pointed out that there can be data configurations of moderate size which may provide an interpolant function with a good graphic behavior. See, for instance, fig. 8. The interpolant obtained from S_{343}^2 has a graph comparable to that of S_{1000}.

4 Conclusions

We have tested the computational performances of some methods for interpolating $3D$ scattered data and we have measured the performances, considering the accuracy (e_2, e_∞), the graphic behaviour, the computation times, and the computational efficiency.

In our opinion, the evaluation of the interpolant graphs is a fundamental aspect, even if quite subjective. Ratings by different persons will give somewhat different results.

From our study it is came out that

• In general, the global methods(2)–(5) provide better results than MQSM.

Among the considered radial basis functions, (2) gives the better results in terms of accuracy and graphic behavior (even if it has the worst condition number).

• The disadvantage of global methods is that an increase of N leads to high condition numbers and high computation times.

• Increasing the sample dimension is not always convenient. In fact, on one hand, the interpolant may not reproduce anyway the correct behavior in some zones and, on the other, we may have a considerable loss of efficiency.

• MQSM usually reproduces the qualitative features of test functions quite well. Near the boundary, it may perform worse than the methods (2)–(5). The interpolant is only C^1. But for samples of size $N > 1000$, it gives a graphic quality comparable to that provided by (2), with the advantage of being more and more efficient.

• For samples with moderate dimension, the interpolation problem solution is strongly influenced by the point locations. There can be data configurations of moderate size which may provide an interpolant with a good behavior.

This remark suggests a possible way to get a satisfactory solution with low costs. When we have large samples of size N, we can extract from it a subsample of dimension $\bar{N} << N$ which allows a correct reproduction of the unknown function.

On this subject, there are some techniques for one and two dimensions. (see [3], [4]). For three dimensions, the problem is still open.

References

[1] C.L Bajaj. Rational hypersurface display. *Adv. Comput. Math*, 68:117–127, 1990.

[2] R.K. Beatson and G.N. Newsam. Fast evaluation of radial basis functions: moment based methods. *SIAM J. Comput.*, 19:1428–1449, 1998.

[3] L. Bozzini, M. Lenarduzzi and R. Schaback. Adaptive interpolation by scaled multiquadrics. To appear in Adv. Comp. Math, 2001.

[4] M. Bozzini and L. Lenarduzzi. Selection of data and interpolation by multiquadrics translate with different elliptical shapes. IAMI 22, 2000.

[5] M. Bozzini and M. Rossini. A comparison of some methods for interpolation of trivariate data. Technical Report, 2001.

[6] M. Bozzini and M. Rossini. Approximating 3D "structured" scattered data. Preprint, 2002.

[7] T.A. Foley. Interpolation and approximation of 3-d and 4-d scattered data. *Comput. Math. Appl.*, 13:711–740, 1987.

[8] R. Franke. Scattered data interpolation. Test of some methods. *Mathematics of Computation*, 48:181–199, 1982.

[9] R. Franke and G. M. Nielson. Smooth interpolation of large sets of scattered data. *Internat. J. Numerical Method in Engineering*, 15:1691–1704, 1980.

[10] R.L. Hardy. Multiquadric equations of topography and other irregular surfaces. *J. Geophys. Res.*, 76:1905–1915, 1971.

[11] J. Hoschek and D. Lasser. *Computer aided geometric design*. A K Peters, Wellesley, Massachusetts, 1993.

[12] D. Lasser. Bernstein–Bézier representation of volumes. *Comput. Aided Geom. Design*, 2:145–150, 1985.

[13] G. M. Nielson. Scattered data modeling. *IEEE Computer Graphics & Applications*, 1:60–70, 1993.

[14] G.M. Nielson and al. Visualizing and modeling scattered multivariate data. *IEEE Computer Graphics & Applications*, 11:47–55, (1991).

[15] G.M. Nielson and B. Hamann. Techniques for the visualization of volumetric data. In A. Kaufman, editor, *Visualization'90*, pages 45–50. IEEE Computer Society Press, 1990.

[16] S. Rippa. An algorithm for selecting a good value for the parameter c in radial basis function interpolation. *Adv. Comput. Math*, 11(vol. 2–3):193–210, (1999).

[17] R. Schaback. Error estimates and condition numbers for radial basis function interpolation. *Advances in Computational Mathematics*, 3:251–264, 1995.

[18] R. Schaback. Improved error bounds for scattered data interpolation by radial basis functions. *Mathematics of Computation*, 68:201–216, 1999.

[19] R. Schaback and H. Wendland. Characterization and construction of radial basis functions. In N. Dyn, D. Leviatan, and D Levin, editors, *Eilat proceedings*. Cambridge University Press, 2000.

[20] S. E. Stead. Estimation of gradients from scattered data. *Rocky Mountain J. Math.*, 14:265–279, (1984).

A Description of the methods

A.1 Radial basis function interpolation

Let $\Omega \subset \mathbb{R}^d$ be a compact set, and let us denote the space of d-variate polynomials of order not exceeding m by $\mathbb{P}_m^d$. We consider multivariate interpolation by conditionally positive definite radial functions

$$\phi \; : \; \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$$

of order $m \geq 0$. This means that for all possible choices of sets

$$X = \{X_1, \ldots, X_N\} \subset \Omega$$

of N distinct points the quadratic form induced by the $N \times N$ matrix

$$A = (\phi(\|X_j - X_k\|_2))_{1 \le j,k \le N} \tag{13}$$

is positive definite on the subspace

$$V := \left\{ \alpha \in I\!R^N \ : \ \sum_{j=1}^{N} \alpha_j p(X_j) = 0 \text{ for all } p \in I\!P_m^d \right\}.$$

Note that $m = 0$ implies $V = I\!R^d$ because of $I\!P_m^d = \{0\}$, and then the matrix A in (13) is positive definite. The most prominent examples of conditional positive definite radial basis functions of order m on $I\!R^d$ are

$$\begin{array}{rcll}
\phi(r) & = & (-1)^{\lceil \beta/2 \rceil} r^\beta, \beta > 0, \beta \notin 2I\!N_0 & m \ge \lceil \beta/2 \rceil \\
\phi(r) & = & (-1)^{k+1} r^{2k} \log(r), \quad k \in I\!N & m \ge k+1 \\
\phi(r) & = & (c^2 + r^2)^{\beta/2}, \quad \beta < 0 & m \ge 0 \\
\phi(r) & = & (-1)^{\lceil \beta/2 \rceil} (c^2 + r^2)^{\beta/2}, \quad \beta > 0, \beta \notin 2I\!N_0 & m \ge \lceil \beta/2 \rceil \\
\phi(r) & = & e^{-\alpha r^2}, \quad \alpha > 0 & m \ge 0 \\
\phi(r) & = & (1-r)_+^4 (1+4r) & m \ge 0, d \le 3.
\end{array}$$

See e.g. [19] for a comprehensive derivation of the properties of these functions. Interpolation of real values $f_1, \ldots, f_N$ on a set $X = \{X_1, \ldots, X_N\}$ of N distinct scattered points of Ω by such a function $\phi(\cdot)$ is done by solving the $(N+q) \times (N+q)$ system

$$\begin{array}{rcl}
A\alpha & + & P\beta = f \\
P^T\alpha & + & 0 = 0
\end{array}$$

where $Q = \dim I\!P_m^d$ and

$$P = (p_i(X_j))_{1 \le j \le N, 1 \le i \le q}$$

for a basis $p_1, \ldots, p_q$ of $I\!P_m^d$. In fact, if the additional assumption

$$\text{rank } (P) = Q \le N$$

holds, then the system is uniquely solvable. The resulting interpolant has the form

$$s(x) = \sum_{j=1}^{N} \alpha_j \phi(\|X_j - x\|_2) + \sum_{i=1}^{q} \beta_i p_i(x)$$

with the additional condition $\alpha \in V$.

A.2 The modified quadratic Shepard Method

The ideas of the bivariate modified quadratic Shepard method extend directly to the trivariate case (see for instance [9] and [13]). This method has the general form

$$s(X) = \frac{\sum_{k=1}^{N} \frac{Q_k(X)}{\rho_k^2(X)}}{\sum_{k=1}^{N} \frac{1}{\rho_k^2(X)}}, \tag{14}$$

where

$$\frac{1}{\rho_k(X)} = \frac{(R_w - ||X - X_k||_2)_+}{R_w||X - X_k||_2} \tag{15}$$

for some constant R_w. $Q_k(X)$ are the quadratic polynomials, obtained by a weighted least squares fit and constrained to take on the value f_i at the trivariate points X_i. The weights in the least squares process are the same form as the weight functions (15), but with a different constant R_q.

We select two values N_q and N_w and define

$$R_w = \frac{D}{2}\sqrt{\frac{N_w}{N}}, \quad R_q = \frac{D}{2}\sqrt{\frac{N_q}{N}},$$

where $D = \max_{X_i, X_j \in X} ||X_i - X_j||_2$.

We consider the weight functions

$$\begin{aligned} \frac{1}{\rho_k(X)} &= \frac{(R_w - ||X - X_k||_2)_+}{R_w||X - X_k||_2} \\ \frac{1}{\nu_i(X)} &= \frac{(R_q - ||X - X_k||_2)_+}{R_q||X - X_k||_2}. \end{aligned}$$

We define the local basis

$$\begin{aligned} Q_k(x,y,z) &= f_k + a_{k2}(x - x_k) + a_{k3}(y - y_k) + a_{k4}(z - z_k) \\ &+ a_{k5}(x - x_k)^2 + a_{k6}(y - y_k)^2 + a_{k7}(z - z_k)^2 \\ &+ a_{k8}(x - x_k)(y - y_k) + a_{k9}(x - x_k)(z - z_k) \\ &+ a_{k10}(y - y_k)(z - z_k) \end{aligned}$$

solving the following least squares problem

$$\begin{aligned} \min_{a_{kj}\, j=2,\ldots,10} \sum_{i=1}^{N} \frac{1}{\nu_i^2(x_k, y_k, z_k)} \{ f_k &+ a_{k2}(x_i - x_k) + a_{k3}(y_i - y_k) \\ &+ a_{k4}(z_i - z_k) + a_{k5}(x_i - x_k)^2 \\ &+ a_{k6}(y_i - y_k)^2 + a_{k7}(z_i - z_k)^2 \\ &+ a_{k8}(x_i - x_k)(y_i - y_k) \\ &+ a_{k9}(x_i - x_k)(z_i - z_k) \\ &+ a_{k10}(y_i - y_k)(z_i - z_k) - f_i \}^2. \end{aligned}$$

The interpolant is locally determined, the influence of any point not extending further than a distance $R_w + R_q$ from each data point.

Assuming that the data are equidistributed, constant values for R_q and R_w (that is constant values for N_q and N_w)are appropriate. A good choice (suggested in the literature, see [13]) is $N_q = 54$, $N_w = 27$. If the data density is not reasonably uniform, we might want to let the radii R_q and R_w depend on i.

B Figures

For the graphs, we have chosen the evolutive representation described in §2. In the examples, the evolution variable is z. We have evaluated the interpolant $g(x, y, z)$ in the points of a grid of size $21 \times 21 \times 21$. In the pictures, we show the evolutions of some surfaces $g(x, y, z_k)$. Namely, we have considered $z_k = (k-1)/20$ with $k \in K$, $K = \{1, 3, 6, 9, 13, 16, 19, 21\}$ for $F1$, $F5$ and with $k \in K \setminus \{11\}$ for $F2$. In all the figures, except fig. 7 (in which the errors are shown), the surfaces are represented using a gray scale.

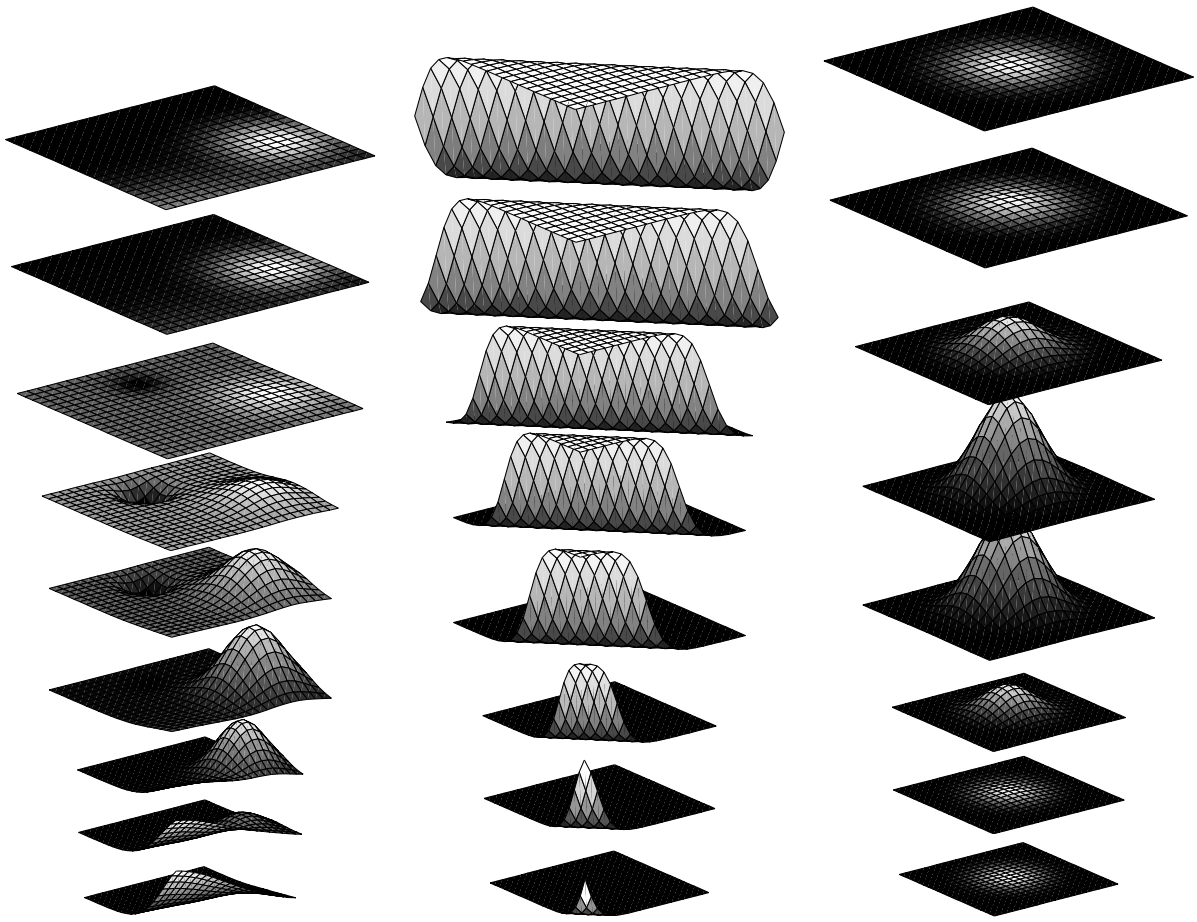

Figure 1: Exact functions. Left: $F1$, center: $F2$, right: $F5$

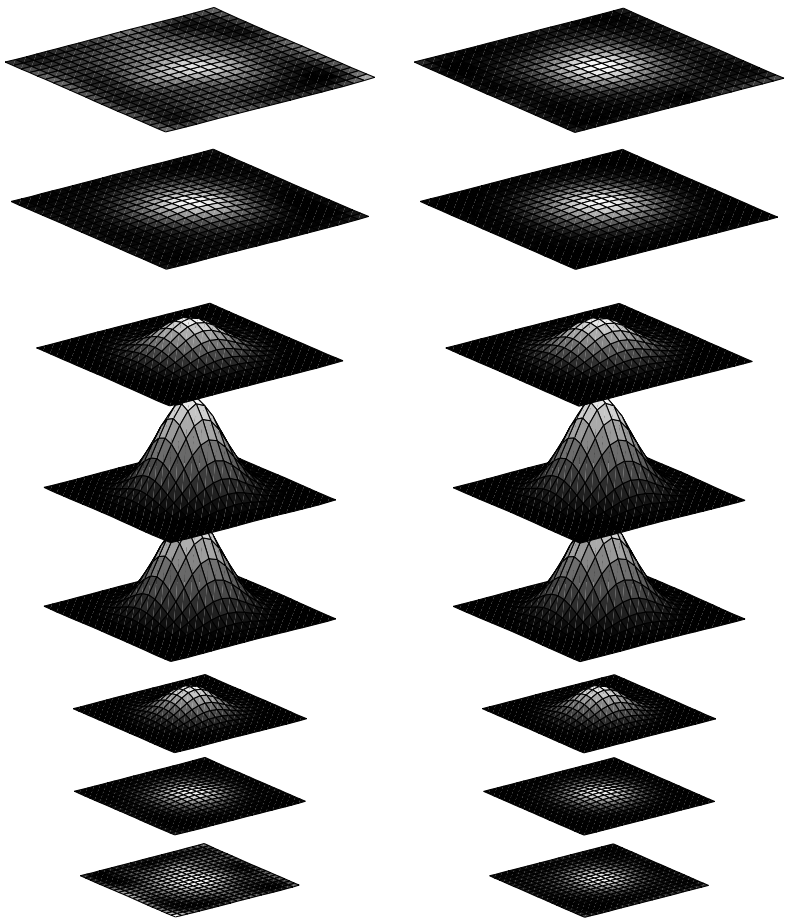

Figure 2: Test function $F5$. Interpolation with r^3. Left: $N = 343$, right: $N = 1000$

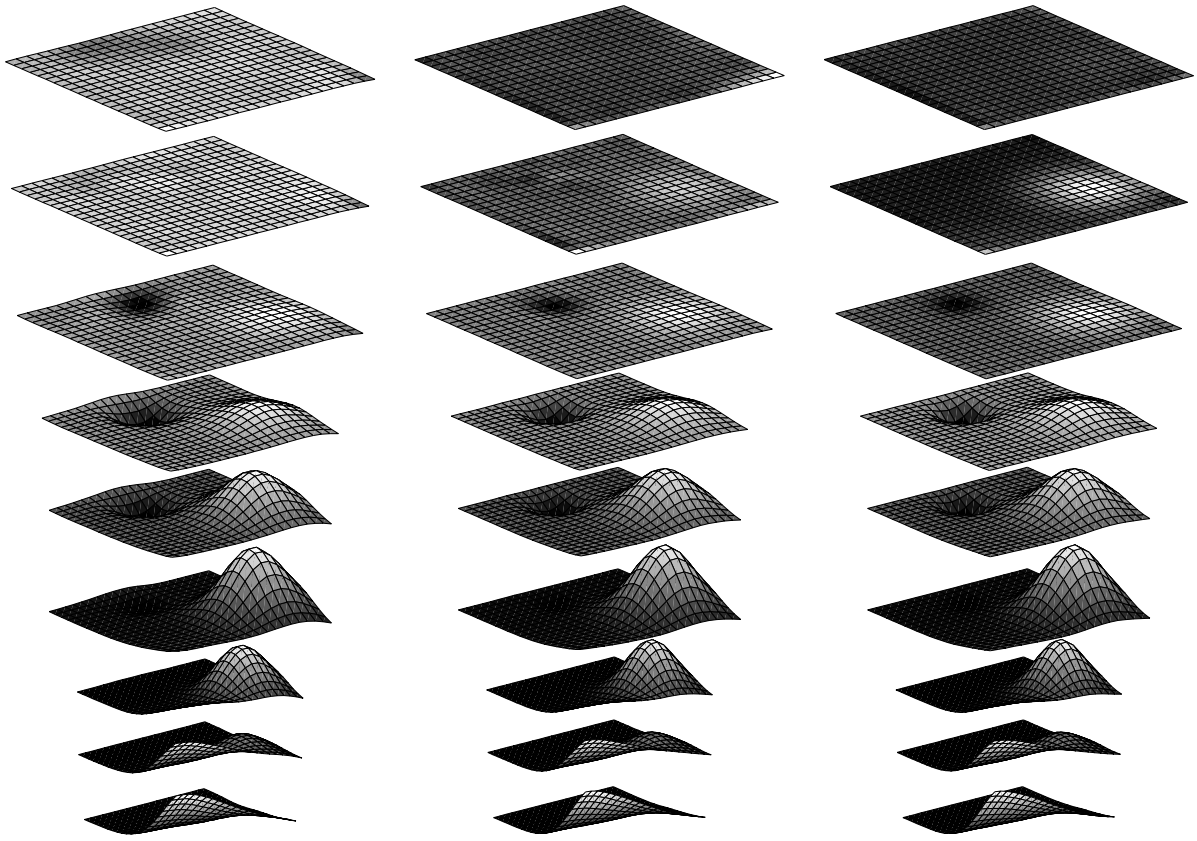

Figure 3: Test function $F1$. Interpolation with r^3. Left: $N = 343$, center: $N = 1000$, right: $N = 3375$

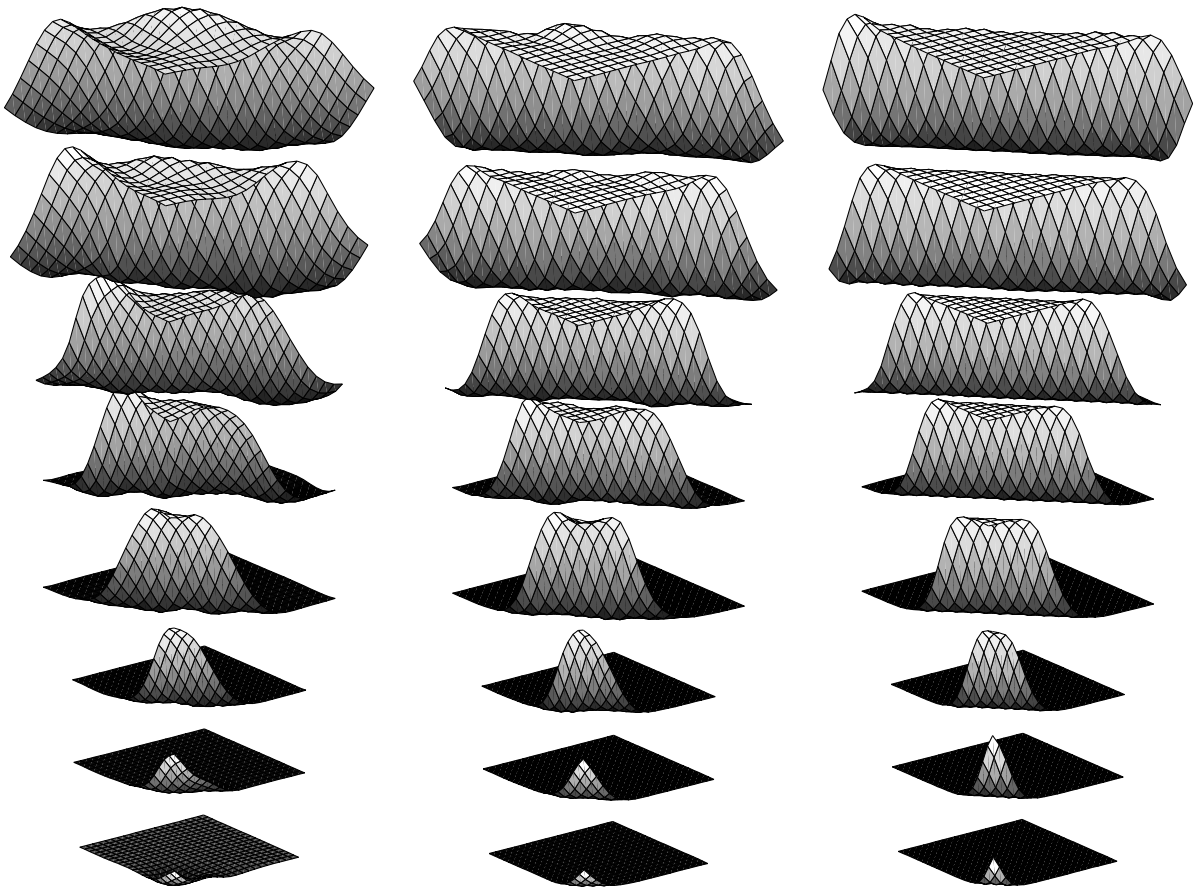

Figure 4: Test function $F2$. Interpolation with r^3. Left: $N = 343$, center: $N = 1000$, right: $N = 3375$

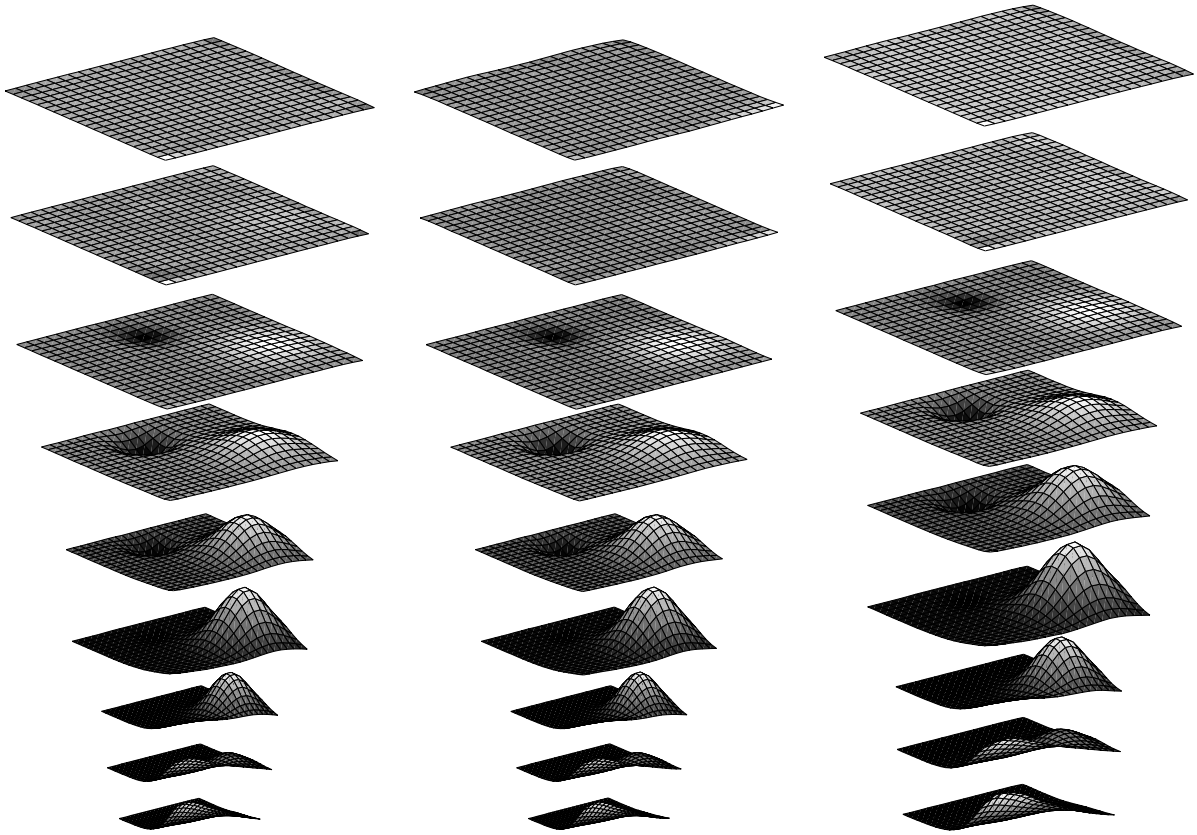

Figure 5: Test function $F2$. Interpolation with r^3. Left: $N = 343$, center: $N = 1000$, right: $N = 3375$

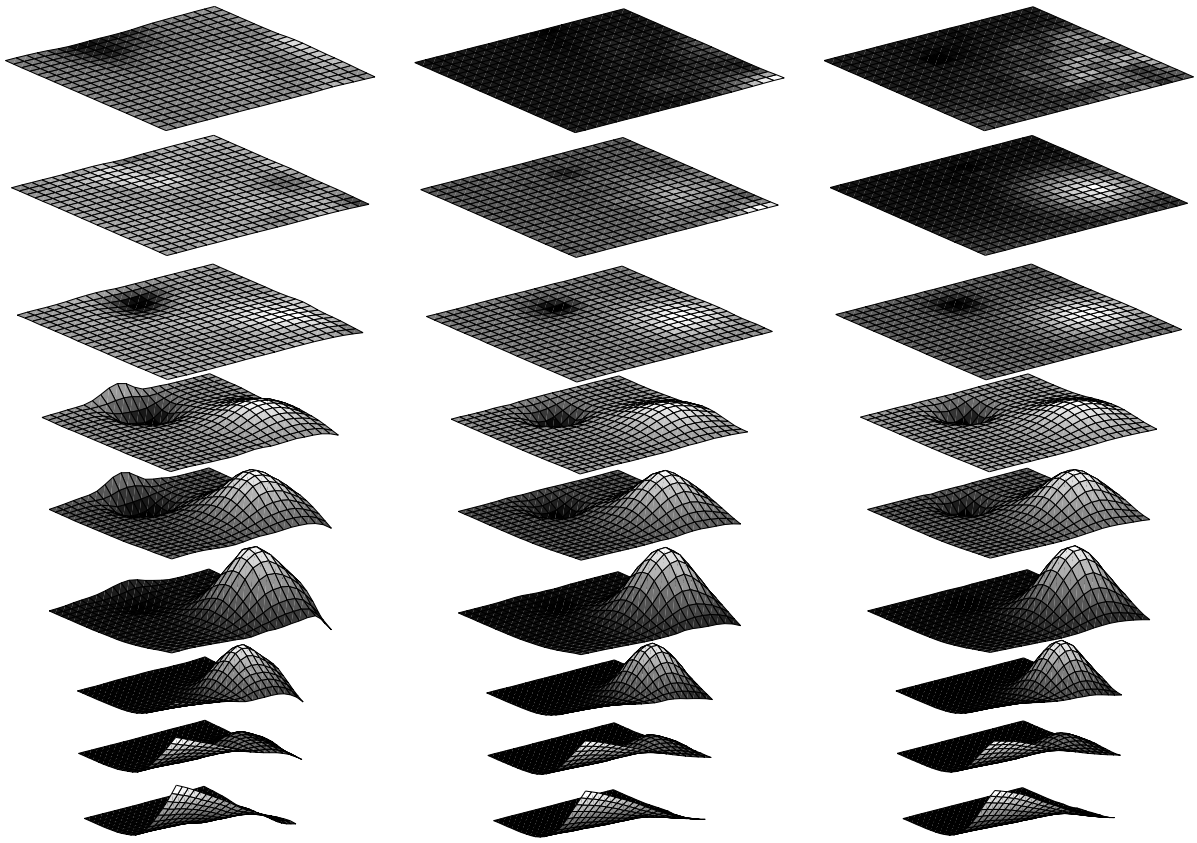

Figure 6: Test function $F1$. Interpolation with MQSM. Left: $N = 343$, center: $N = 1000$, right: $N = 3375$

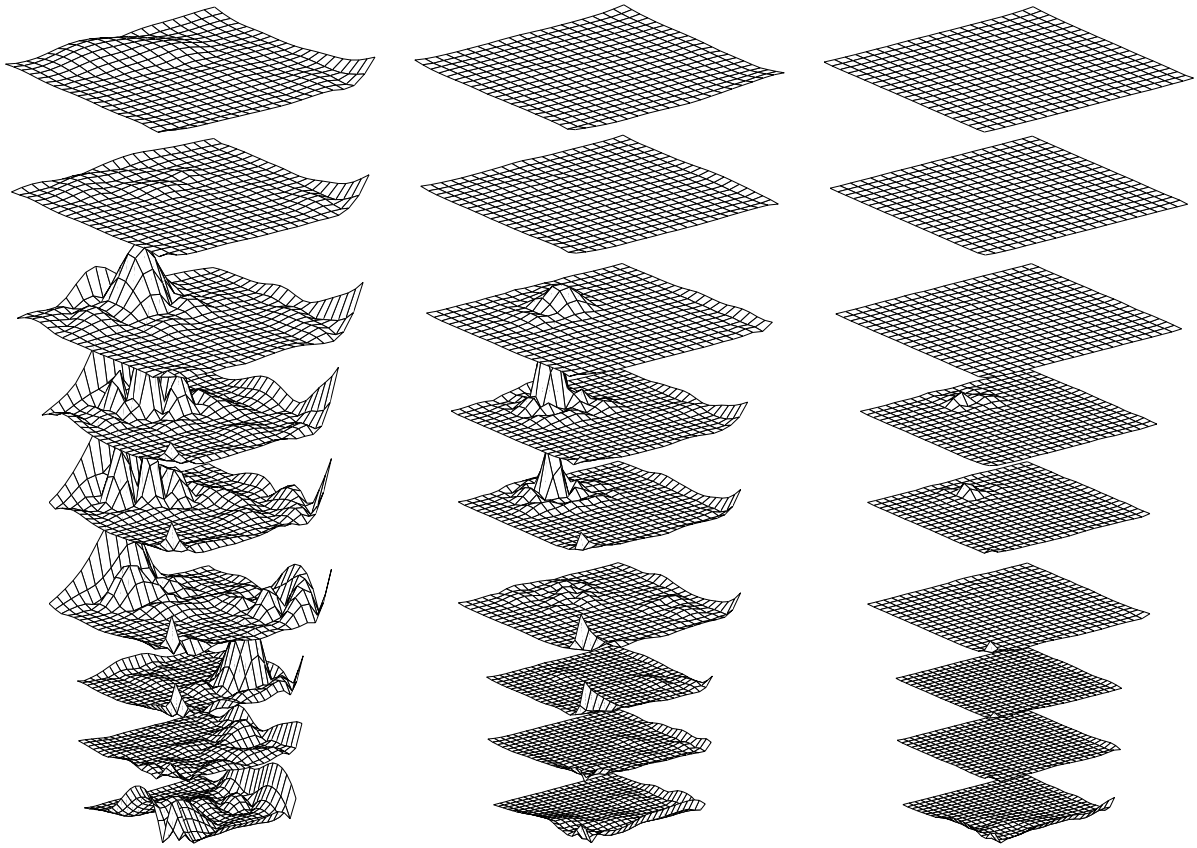

Figure 7: Test function $F1$. Absolute errors for the interpolants obtained with r^3. Left: $N = 343$, center: $N = 1000$, right: $N = 3375$

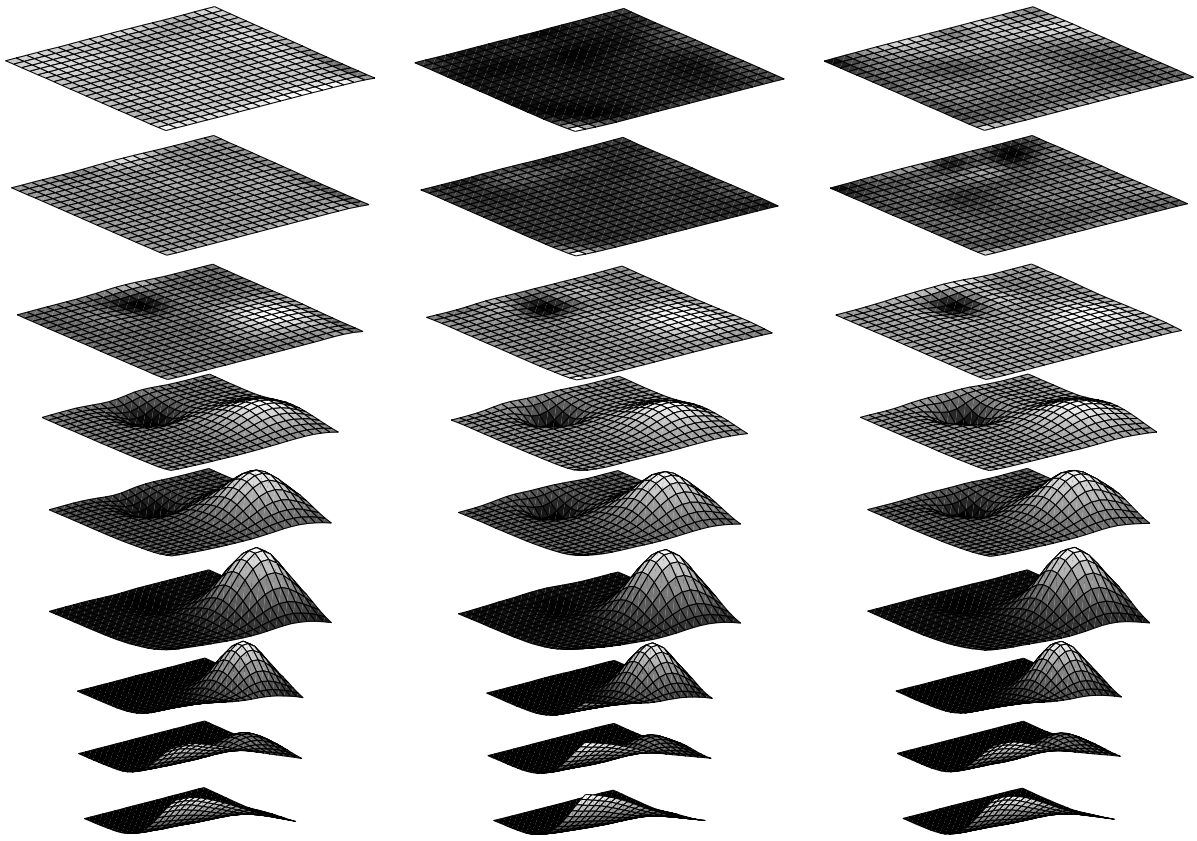

Figure 8: Test function $F1$. Interpolants on different sets. Left: S^1_{343}, center: S^2_{343}, right: S^3_{343}

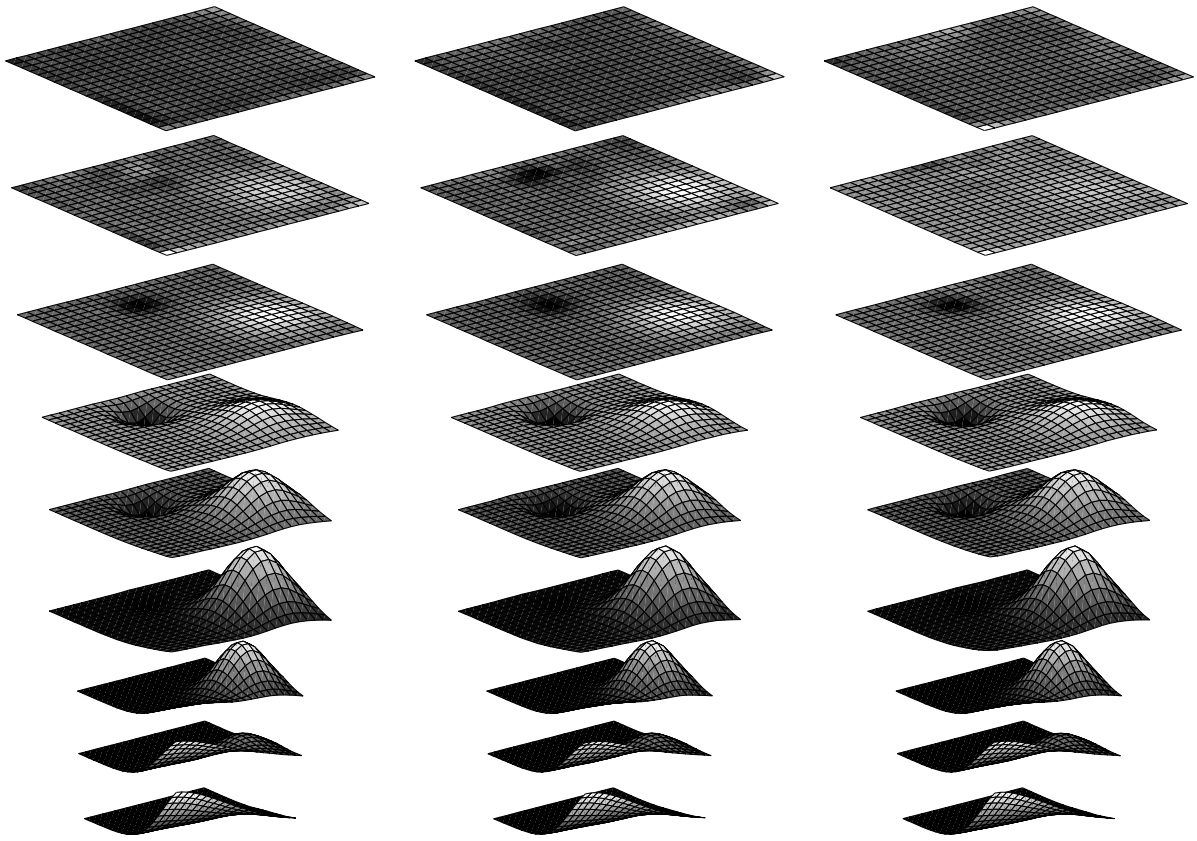

Figure 9: Test function $F1$. Interpolants on different data sets. Left: S^1_{1000}, center: S^2_{1000}, right: S^3_{1000}

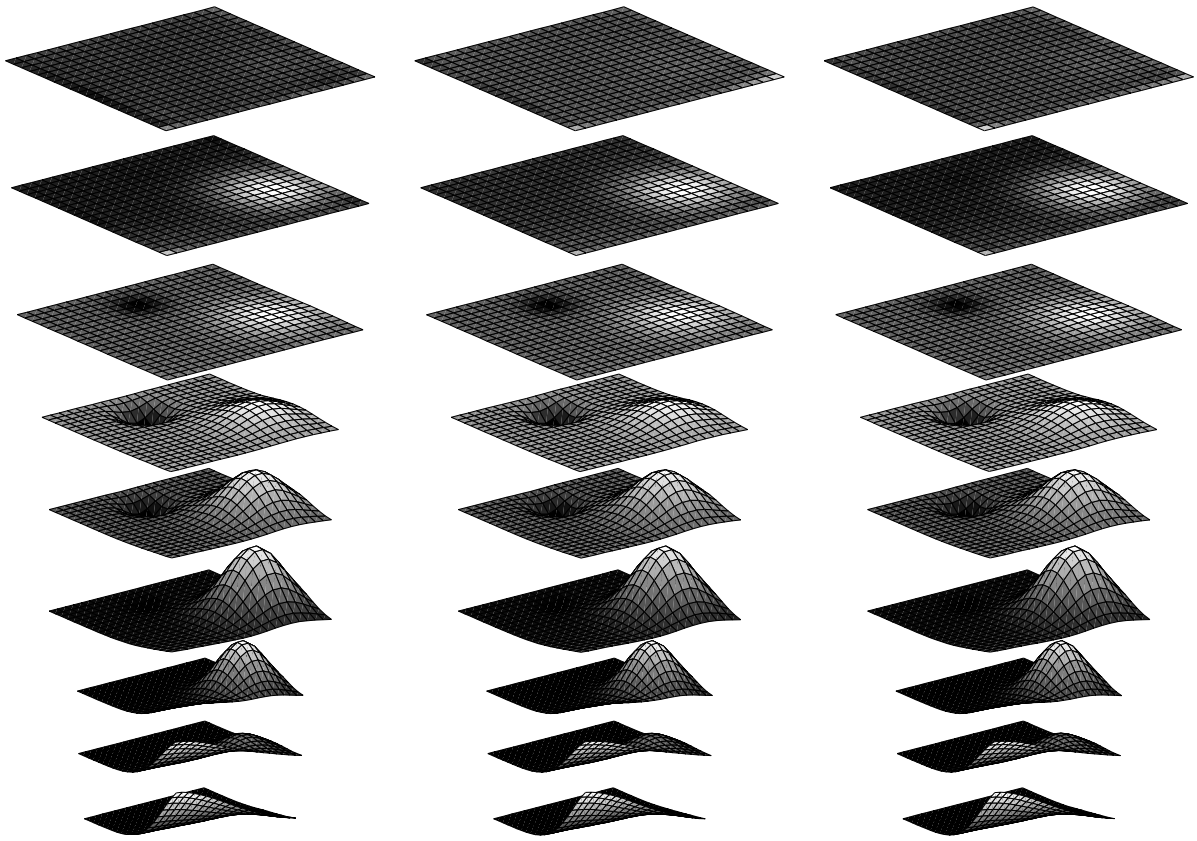

Figure 10: Test function $F1$. Interpolants on different data sets. Left: S^1_{3375}, center: S^2_{3375}, right: S^3_{3375}