PADOVA UNIVERSITY PRESS

Dolomites Research Notes on Approximation

Proceedings of *Multivariate Approximation: Theory, Algorithms & Applications 2017*, Volume 11 · 2018 · Pages 3–14

Interpolating functions with gradient discontinuities via Variably Scaled Kernels

Milvia Rossini[a]

Abstract

In kernel–based methods, how to handle the scaling or the choice of the *shape parameter* is a well–documented but still an open problem. The shape or scale parameter can be tuned by the user according to the applications, and it plays a crucial role both for the accuracy of the method and its stability. In [7], the Variably Scaled Kernels (VSKs) were introduced. The idea is to vary the scale in a continuous way by letting the scale parameter be an *additional coordinate*. In this way a *scale function* $c(x)$ is introduced and this allows varying scales without leaving the firm grounds of kernel–based interpolation. In [7] several numerical experiments are devoted to show that the method performs better than the standard fixed–scale basis and the proper use of VSKs, i.e. a proper choice of the scale function c, can lead to a more stable and better shape–preserving approximation. When the scale function c is chosen to depend on critical shape properties of the data, the interpolant reproduces the underlying phenomenon in a much more faithful way. This is the starting point of this paper. The goal here is to recover from scattered data, in one and two dimensions, functions with discontinuities. The paper is devoted to gradient discontinuities that in general are more difficult to handle. The idea is to use translates of basis functions that change their smoothness locally according to the position of the discontinuities and we show that this can be achieved by suitable choices of the scale function.

1 Introduction

Variably Scaled Kernels (VSKs) were introduced in [7] with the aim to give a new technique to handle the problem of the choice of the scale or shape parameter in kernel–based interpolation problems. There, the authors consider native spaces whose kernels allow for a change the kernel scale of a d–variate interpolation problem locally, depending on the requirements of the application. They define a *scale function* c on the domain $\Omega \subset \mathbb{R}^d$ to transform an interpolation problem from data locations x_j in $\mathbb{R}^d$ to data locations $(x_j, c(x_j))$ and to use a fixed–scale kernel on $\mathbb{R}^{d+1}$ for interpolation there. The $(d+1)$–variate solution is then evaluated at $(x, c(x))$ for $x \in \mathbb{R}^d$ to give a d–variate interpolant with a varying scale.

A large number of examples in [7], shows how this can be done in practice to get results that are better than the fixed–scale technique, with respect to both condition and error. In fact they work quite well in cases that are spoiled by a considerable instability of the standard method and they can significantly improve the recovery quality by preserving shape properties and particular features of the underlying function. Moreover, the background theory coincides with fixed–scale interpolation on the sub–manifold of $\mathbb{R}^{d+1}$ given by the points $(x, c(x))$ of the graph of the scale function c.

VSKs were recently used also in neural networks problems [32], in solving magnetohydrodynamic equations [14], and for approximating the solution of elliptic partial derivative problems [13]. We will show here that they are a useful tool also when recovering unknown non-regular functions from set of scattered data.

Surfaces with discontinuities appear in many scientific applications. By function with discontinuities, we mean that the function or its partial derivatives are discontinuous across planar curves. In the case of function discontinuities, these curves are known as *faults* or *edges*, while when dealing with gradient discontinuities they are called *gradient faults* and sometimes *oblique faults*. In the last decades, the problem has become an attractive field for research, (see e.g. [3, 9, 5, 23, 24, 28, 29, 30]) especially within applications arising from different branches of science going from image processing to geophysics.

The aim of this paper is to show that a proper choice of the scaling function c allows for a faithful interpolation of scattered data sampled from unknown two-variate functions which exhibit some discontinuous behaviour. We focus on functions with gradient faults that we believe is the more difficult situation.

The outline of the paper is as follows. In Section 2 we give some preliminary notions. In Section 2.1 we briefly review the main theoretical aspects of kernel-based interpolation and in Section 2.2, we discuss the role of the shape parameter when kernel are scaled. In Section 2.3 we describe the VSKs and in Section 3 we show how the scale function can be chosen when interpolating function with gradient faults. In Section 4 we show several numerical results. Finally, the last section deals with conclusions.

[a]Dipartimento di Matematica e Applicazioni, Università degli Studi Milano Bicocca, Milano, Italy

2 Background

2.1 Preliminaries

The reader can refer to [37] and [17] for more details on the topic. A symmetric *kernel*

$$K : \Omega \times \Omega \to \mathbb{R}$$

defined on a domain $\Omega \subset \mathbb{R}^d$ is very useful for a variety of purposes going from interpolation or approximation to PDE solving, if certain *nodes* or *centers*

$$X =: \{x_1, \dots, x_N\} \subset \mathbb{R}^d$$

are used to define *kernel translates*

$$K(\cdot, x_j)$$

as *trial functions*.

If the kernel is positive definite, i.e. the *kernel matrices* with elements $K(x_i, x_j)$, $1 \le i, j \le N$ are positive definite for all choices of nodes, there is a *native* Hilbert space $\mathcal{H}$ in the background in which the kernel is reproducing, i.e.

$$g(x) = (g, K(\cdot, x))_{\mathcal{H}} \quad \forall x \in \Omega, \quad g \in \mathcal{H}.$$

The use of reproducing kernels in Hilbert spaces leads to various optimality properties and applications. Interpolation of values

$$f := \{f_1, \dots, f_N\} \quad \text{on} \quad X := \{x_1, \dots, x_N\}$$

proceeds via solving a linear system

$$A_X a = f$$

with interpolation matrix

$$A_X := (K(x_j, x_k))_{1 \le j,k \le N},$$

which is positive definite. The coefficient vector $a \in \mathbb{R}^N$ then allows the interpolant function to be written as

$$s_{X,f}(x) := \sum_{j=1}^{N} a_j K(x, x_j). \tag{1}$$

We remember that, by adding a lower degree polynomial to (1), the existence of a solution is assured also for conditionally positive definite kernel. In the sequel we will consider only strictly positive definite functions.

If the kernel K is *translation–invariant* on $\mathbb{R}^d$, it is of the form

$$K(x, y) = \Phi(x - y) \quad \text{for all} \quad x, y \in \mathbb{R}^d.$$

If the kernel is *radial*, i.e. of the form

$$K(x, y) = \phi(\|x - y\|_2)$$

for a scalar function

$$\phi : [0, \infty) \to \mathbb{R},$$

the function Φ is called *Radial Basis Function (RBF)*.

It is well-known that the main strength of RBFs is their ability to approximate scattered data without using triangularizations, re-meshing and other geometric tools.

When dealing with scattered data and kernel-based interpolation, two indicators of data regularity have to be considered: the *separation* and *fill distances*.

Definition 2.1. The separation distance is

$$q(X) = \frac{1}{2} \min_{i \ne k} \|x_i - x_k\|_2 . \tag{2}$$

The quantity $q(X)$ represents the radius of the largest ball that can be centred at every point in X such that no two balls overlap.

Definition 2.2. The fill distance is

$$h(X, \Omega) = \sup_{x \in \Omega} \left(\min_{x_k \in X} \|x - x_k\|_2 \right). \tag{3}$$

The quantity (3) is a measure of data distribution and indicates how well the data fill out the domain Ω. It denotes the radius of the largest empty ball that can be placed among the data locations. They are respectively related to the interpolation matrix conditioning and to the accuracy of the solution.

It is worthwhile to remember the *uncertainty or trade-off principle:* when we use the standard approach to kernel interpolation problem, that is the solution of a linear system, there is always a conflict between the theoretical accuracy that one can obtain and the numerical stability (see e.g. [35]). For well distributed data (i.e. $q(X) \approx h(X, \Omega)$) a small error bound will necessarily imply a large condition number.

2.2 Scaled Kernels

It is well-known that kernels on $\mathbb{R}^d$ can be *scaled* by a positive factor δ. This turns to have a new kernel

$$K(x,y;\delta) := K(x/\delta, y/\delta) \ \forall x,y \in \mathbb{R}^d. \tag{4}$$

In case of a radial kernel supported on the unit ball (e.g. the Wendland kernels), the support of the scaled kernel has support radius δ. In general, large δ increase the condition of kernel matrices, while small δ let the translates turn into sharp peaks which approximate functions badly, if separated too far from each other.

The shape or scale parameter can be tuned by the user (according to the applications) and it plays an important role both for the accuracy of the method and its stability. How to choose the scale parameter δ is a well–documented but still an unsolved problem that started with Hardy many years ago with multiquadrics. When working with a fixed scale one has several possibilities. The most popular are to pick the parameter by some ad hoc criterion as done by Hardy and later by Franke (see [19] and the reference therein),or to choose the parameter by some optimal criterium based for instance on a variant of the cross validation approach (leave-one-out) [33] and on its extension applied in the setting of iterated approximate moving least squares [19].

A special case of scaling is the *flat limit* $\delta \to \infty$ (see e.g. [15, 21, 26, 27, 36]). In this case the papers of Fornberg and co-workers suggest a way to stably compute very accurate generalized multiquadic and gaussian interpolant by using a complex Contour-Padé Algorithm, and its variants, which is capable to overcome the conditioning problem. Usually there is a value of the shape parameter which provides an optimal approximation error. Typically optimal values of the scale parameter give instability. For the Gaussian kernel there are well-established tools, such as RBF-QR methods, that allow to overcome the instability issues (see e.g. [18, 20]). In [12] is presented a *rescaled-method* based on a proper selection of the supports of compactly supported basis functions that allows to keep the ill-conditioning under control.

We will see in Section 2.3 that a crucial property of VSKs is to improve stability and this fact holds for any kernel. So they are very useful to overcome instability issues, as done in [13] where they are used in a hybrid technique to develop a stable computation of the solution of elliptic boundary value problems.

People considered also the possibility to have the scale of a kernel translate varying with the translation. This means working with functions

$$\phi(\|x-x_j\|_2/\delta_j),\ 1 \le j \le N$$

in the radial case, [8, 22, 25], or even with basic functions that vary with j, both in shape and scale [4].

In this case, it is easy to come up with examples that let interpolation fails for certain nonuniform choices of scale. But in [6] sufficient conditions for the unique solvability of such interpolation processes are given.

2.3 Variably Scaled Kernels

So far, it is clear the importance of the scale parameters or scale parameter vectors and in order to make more effective their potential we can generalize the vector case by introducing a continuous scale function. In [7], the case of varying scales at the centers is generalized by letting the scale parameter be an additional coordinate. This allows varying scales in a continuous way without leaving the well-estabilished theory of kernel–based interpolation. It turns out that this approach can be fully understood as the standard fixed–scale method applied to a certain sub-manifold of $\mathbb{R}^{d+1}$.

Definition 2.3. Let K be a kernel on $\mathbb{R}^{d+1}$. If a *scale function*

$$c\ :\ \mathbb{R}^d \to (0,\infty)$$

is given, we define a *Variably Scaled Kernel (VSK)* on $\mathbb{R}^d$ by

$$K_c(x,y) := K((x,c(x)),(y,c(y)))\ \forall x,y \in \mathbb{R}^d. \tag{5}$$

If K is positive definite on $\mathbb{R}^{d+1}$, so is K_c on $\mathbb{R}^d$. Then if the chosen kernel K is positive definite, interpolation of values

$$f : \{f_1,\dots,f_N\} \quad \text{on} \quad X = \{x_1,\dots,x_N\}$$

proceeds as usual by solving a linear system

$$A_{c,X}\, a = f$$

with positive definite kernel matrix

$$A_{c,X} := (K_c(x_j,x_k))_{1\le j,k\le N}.$$

A VSK interpolant is then written as

$$s_{c,X,f}(x) := \sum_{j=1}^{N} a_j K_c(x,x_j) = \sum_{j=1}^{N} a_j K((x,c(x)),(x_j,c(x_j))). \tag{6}$$

If the kernel is radial, and of the form $K(x,y) = \phi(\|x-y\|_2^2)$, the interpolant is

$$s_{c,X,f}(x) := \sum_{j=1}^{N} a_j \phi(\|x-x_j\|_2^2 + (c(x)-c(x_j))^2).$$

Note that the interpolant is identical to the standard kernel interpolant if the scale function is constant.

The scale function $c(x)$ introduces a map

$$C : x \mapsto (x, c(x)) \tag{7}$$

that maps the space $\mathbb{R}^d$ into a d-dimensional submanifold $C(\mathbb{R}^d)$ of $\mathbb{R}^{d+1}$ and a set

$$X = \{x_1, \ldots, x_N\} \subset \Omega \subset \mathbb{R}^d$$

of nodes goes into

$$C(X) \subset C(\Omega) \subset C(\mathbb{R}^d) \subset \mathbb{R}^{d+1}.$$

As a consequence, interpolation by the fixed-scaled kernel K takes place on $\mathbb{R}^{d+1}$ at the point set

$$C(X) = \{(x_1, c(x_1)), (x_2, c(x_2)), \ldots, (x_N, c(x_N))\}.$$

This means that in $\mathbb{R}^{d+1}$ the kernel

$$K((x, c(x)), (y, c(y)))$$

is used and if we project points $(x, c(x)) \in \mathbb{R}^{d+1}$ back to $x \in \mathbb{R}^d$, the projection of the kernel turns into the VSK $K_c(x, y)$ on $\mathbb{R}^d$, and the analysis of error and stability of the variably–scale problem in $\mathbb{R}^d$ coincides with the analysis of a fixed–scale problem on a submanifold in $\mathbb{R}^{d+1}$. Moreover the fill distance (3) and the separation distance (2) will transform with C, and will roughly be multiplied by a factor related to the norm of the gradient of c or the Lipschitz constant L of c, depending on the regularity of c. Namely,

$$\begin{array}{rcl} \|C(x) - C(y)\|_2^2 & = & \|x - y\|_2^2 + (c(x) - c(y))^2 \\ & \leq & \|x - y\|_2^2 (1 + L)^2 \\ \|C(x) - C(y)\|_2^2 & \geq & \|x - y\|_2^2. \end{array}$$

This shows that distances will blow up with C, letting separation distances never decrease, thus enhancing stability. Fill distances may also blow up, increasing the usual error bounds. But this argument shows that one can successfully use the varying-scale technique on points x_j and x_k that have very small separation distance until one roughly gets that the transformed centers are approximately uniformly distributed, that is

$$q(C(X)) \approx h(X, \Omega) \approx h(C(X), C(\Omega)),$$

improving in this way the interpolation matrix condition number and preserving the accuracy. Note that we can also use $c(x)$ with discontinuities in its gradient.

3 Interpolating non-regular functions via VSKs

It is evident that kernel based methods are in principle not suitable for non regular surface reconstruction since the obtained interpolant or approximant are linear combinations of translates of smooth functions (at least continuous), so they cannot reproduce singularities along planar curves. Some strategies to use radial kernels even in this situation, have been developed. A very recent one [28], proposes an algorithm that first calculates sub-approximations on non-overlapping subdomains, then extends the sub-domains as much as possible and finally produces a global solution on the given domain by letting the sub-domains fill the whole domain. Consequently, there will be no Gibbs phenomenon along the boundaries of the sub-domains and the method detects faults and gradient faults as well with good accuracy. This algorithm works for fixed scattered input data of the function itself and it does not re-sample as happens, for instance in adaptive schemes, based on residual sub-sampling, that places more centers near singularities (see e.g. [16]). Another adaptive strategy [38] to improve accuracy in general and also near gradient discontinuities is based on the use of an error indicator which indicates the approximation quality at nodes inspected by the algorithm by comparing a global RBF interpolant and a local RBF interpolant. Other strategies combine a simplification (segmentation) of the dicontinuity curve and then recover the data for instance by D^m—spline (see e.g. [30] and [23]) or after detecting and approximating the discontinuity curve, involve a domain decomposition and a suitable blending of the obtained approximation (see e.g. [5]). In general some information on the position and behaviour of the fault and/or gradient fault is needed in order to avoid undue oscillations or over-smoothing (see e.g. [10]). The discontinuity curve can be detected and approximated beforehand by ad hoc methods. It is well-known that there is a huge literature on fault (edge) detection especially for gridded data (see e.g. [2, 34] and the references therein). There exists a number of papers developing algorithms for the curve detection and for its approximation from scattered data too (e.g. [1, 10, 11, 31]).

The VSKs contribute to fill the gap between the reconstruction of functions with discontinuities and kernels methods. The problem is the following: given a set of scattered data $X = \{x_i\}_{i=1}^N, Y = \{f(x_i)\}_{i=1}^N, X \subset \Omega \subset \mathbb{R}^2$, to recover unknown functions

$$f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$$

such that f or its gradient ∇f is discontinuous across a curve $\Gamma \in \Omega$. We assume that Γ has an explicit known equation, if this is not the case, one of the above mentioned techniques has to be applied in order to approximate it.

It is worth to remember that in seminal paper [24], the authors use the idea of the additional coordinate in order to interpolate faulted surfaces (i.e. surfaces with edges) and they give only heuristic justifications, but some promising numerical experiments and hints on the choice of what they call auxiliary function that in our case is the scaling function c.

Here we focus on gradient discontinuities which are more difficult to handle and less considered in the literature. The idea is very simple. It consists in using a scale function $c(x)$ that reproduces the shape of the discontinuity. The optimal situation would be to be able to relate its shape to the data, but this would be very hard to obtain and would require a considerable effort.

We start with the one dimensional case. Let the data site be in a bounded interval of $\mathbb{R}$, $X \subset [a,b]$ and let us assume that the derivative of the underlying function is discontinuous at $x^* \in [a,b]$. We fix the scale function to be

$$c(x) = \begin{cases} 1-3/2|x-x^*|/R+1/2|x-x^*|^3/R^3, & |x-x^*| < R, \\ 0, & \text{otherwise}, \end{cases} \tag{8}$$

where $2R$ is the support of $c(x)$ which goes to zero smoothly. Extensive experiments have shown that R had to be chosen less that $(b-a)/2$.

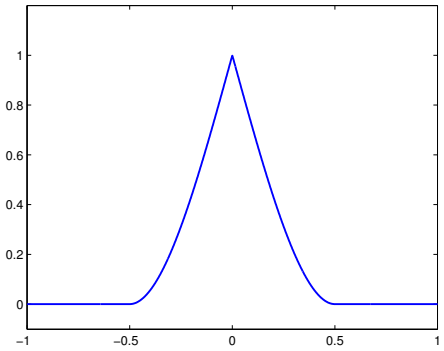

Figure 1: $c(x)$, $x^* = 0$

The shifts of the variably scaled kernel change their shape: they are no more radial functions and, as expected, exhibit, if the center is next to x^*, a discontinuity in the first derivative. In Fig. 2 the graphs of the C^2 Wendland VSK are shown for few selected centers.

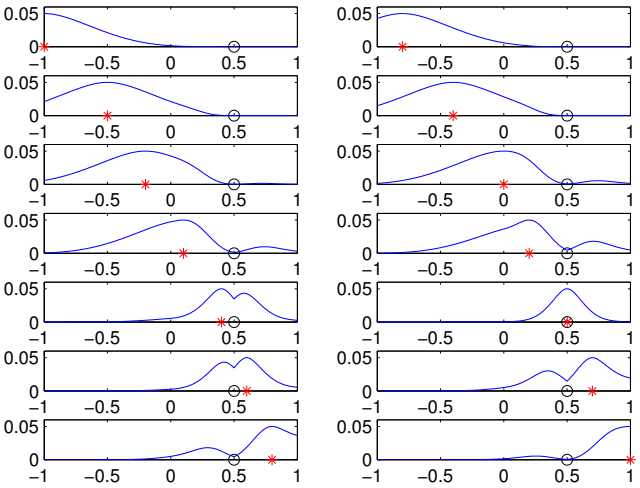

Figure 2: C^2 Wendland-VSK: $*$ is the shift and o the discontinuity point.

The basis function translates change their smoothness locally according to the position of the discontinuity. This mimics in some sense the effects of using univariate B-splines with multiple knots.

Let now turn to the bidimensional case and assume that the underlaying function $f(x,y)$ is defined on the unitary square $\Omega = [0,1]^2$ and that $y = \Gamma(x)$ is the explicit equation of the gradient fault. In this case we define the scale function as

$$c(x,y) = \begin{cases} 1-3/2|y-\Gamma(x)|/R+1/2|y-\Gamma(x)|^3/R^3, & |y-\Gamma(x)| < R, \\ 0, & \text{otherwise}. \end{cases} \tag{9}$$

Again c has support included in Ω and goes to zero smoothly at the boundaries. The numerical experiments suggest to choose $R < 0.5$.

4 Numerical Examples

The scale functions (8) and (9) have been combined with various fixed scaled radial kernels and tested on several functions, here we show the comparison between standard interpolation and VSK interpolation when C^2—Wendland kernels with support radius δ are used. In the univariate examples ($d = 1$) we have considered $N = 61$ scattered points in the interval $\Omega = [-1, 1]$, $\delta = 2$, and $R = 0.5$. while when $d = 2$, we have considered $N = 256$ scattered points in the unitary square, $\delta = 1$, and $R = 0.3$. Note that in univariate examples the discontinuity points do not belong to the data sites X as well in the two-variate cases where there is no need to have a more dense subset of data near the gradient fault.

The root mean square and maximum errors (e_2 and e_∞) in Tables 1-4 show that the VSK interpolation is more accurate. The graphs of the errors confirm that the error is reduced near the discontinuities, and the VSK recovery does not over-smooth near the gradient faults and does not present undue oscillations, providing a correct approximation, and an effective tool for the problem discussed in this paper.

4.1 $d = 1$

4.1.1 Example 1

The first test function

$$f_1(x) = \begin{cases} \exp(-0.5(5x+0.5)^2) & -1 \le x < 0 \\ \exp(-0.5(5x-0.5)^2) & 0 \le x \le 1, \end{cases}$$

(depicted in Fig. 3 together with the sampled data) has a first order derivative discontinuity at $x^* = 0$ and it is symmetric with respect tothe axis $x = 0$. The results are presented in Figs. 4, 5, and Table 1.

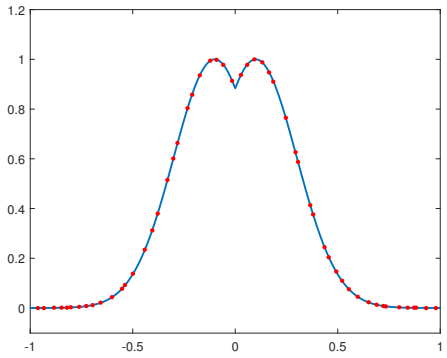

Figure 3: Test function and data points

Example 1	e_2	e_∞-error
Standard interpolant	2.110168e-03	2.692631e-02
VKS-interpolant	1.873396e-04	2.193475e-03

Table 1: Errors for Example 1

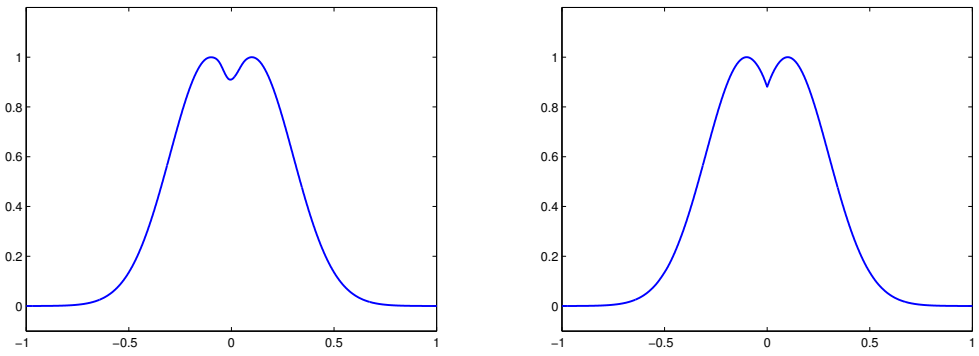

Figure 4: Left: standard interpolation. Right: VSK-interpolation

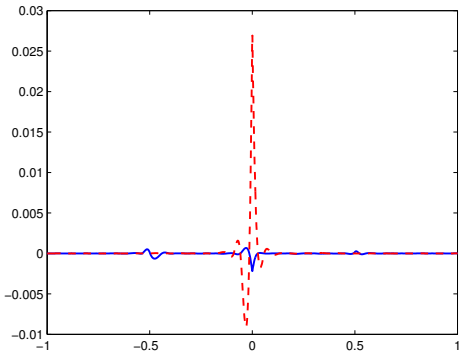

Figure 5: Blu line: VSK-error. Red line: Interpolant error

4.1.2 Example 2

The second test function

$$f_2(x) = \begin{cases} \sin(\pi x) & -1 \le x < 0.5 \\ -4x^2 + 2; & 0.5 \le x \le 1, \end{cases}$$

(depicted in Fig. 6 together with the sampled data) has a first order derivative discontinuity at $x^* = 0.5$ and differently from before it is not symmetric with respect to the axis $x = 0.5$. Even in this case the results are good as showed in Figs. 7, 9, Table 2, and in the enlarged particulars of Fig. 8.

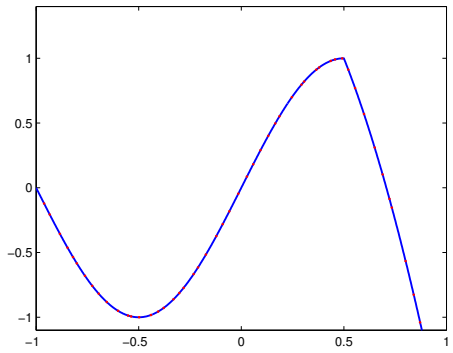

Figure 6

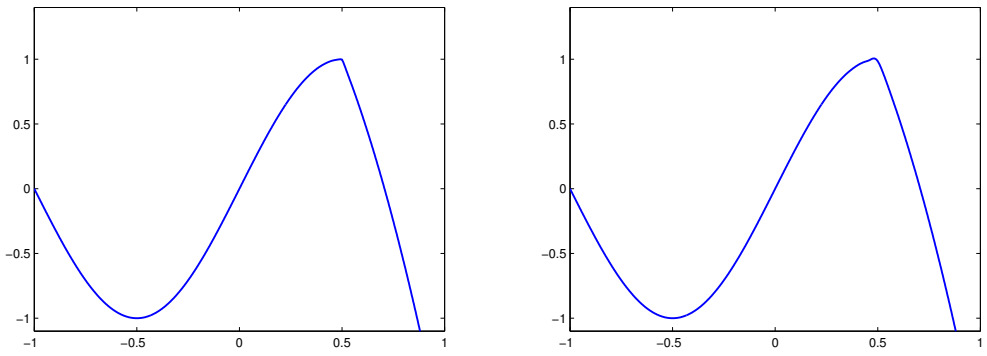

Figure 7: Right: VSK-interpolation. Right: Standard interpolation

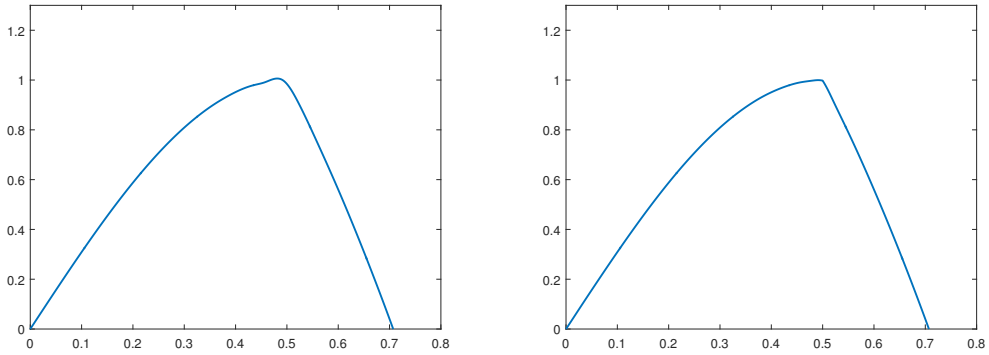

Figure 8: Zoom. Left: Standard Interpolant. Right: VSK-interpolant

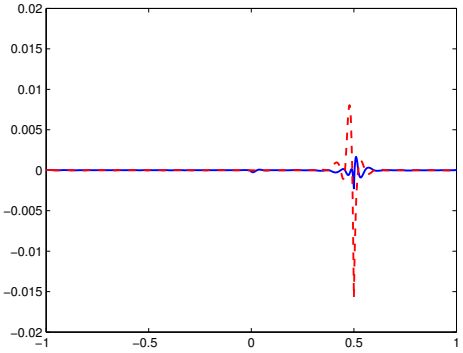

Figure 9: Blu line: VSK-error. Red line: Interpolant error

Example 2	e_2	e_∞-error
Standard interpolant	1.137331e-03	1.573100e-02
VKS-interpolant	1.728193e-04	2.250510e-03

Table 2: Errors for Example 2

4.2 $d = 2$

For each example we show the data sites together with the curve Γ in the domain $[0,1]^2$, and the scale function $c(x,y)$. It is important to note that no extra points are added near the discontinuity curve. The numerical examples confirm again what already observed in the one dimensional case: the VKS interpolant with scale function (9) is capable to reproduce faithfully the gradient discontinuities of the underlying function.

4.2.1 Example 3

The test function

$$f_3(x,y) = |y - 0.6 - 0.1\sin(2\pi x)|$$

has been sampled at the data sites of Fig. 10 where it is shown also the scaling function. As shown in Figs. 11 and 12, the VSK interpolant is effective and, as expected, does not over-smooth the approximant behaviour near the gradient fault.

Example 3	e_2	e_∞-error
Standard interpolant	2.184125e-03	2.648879e-02
VKS-interpolant	5.101804e-04	4.605954e-03

Table 3: Errors for Example 3

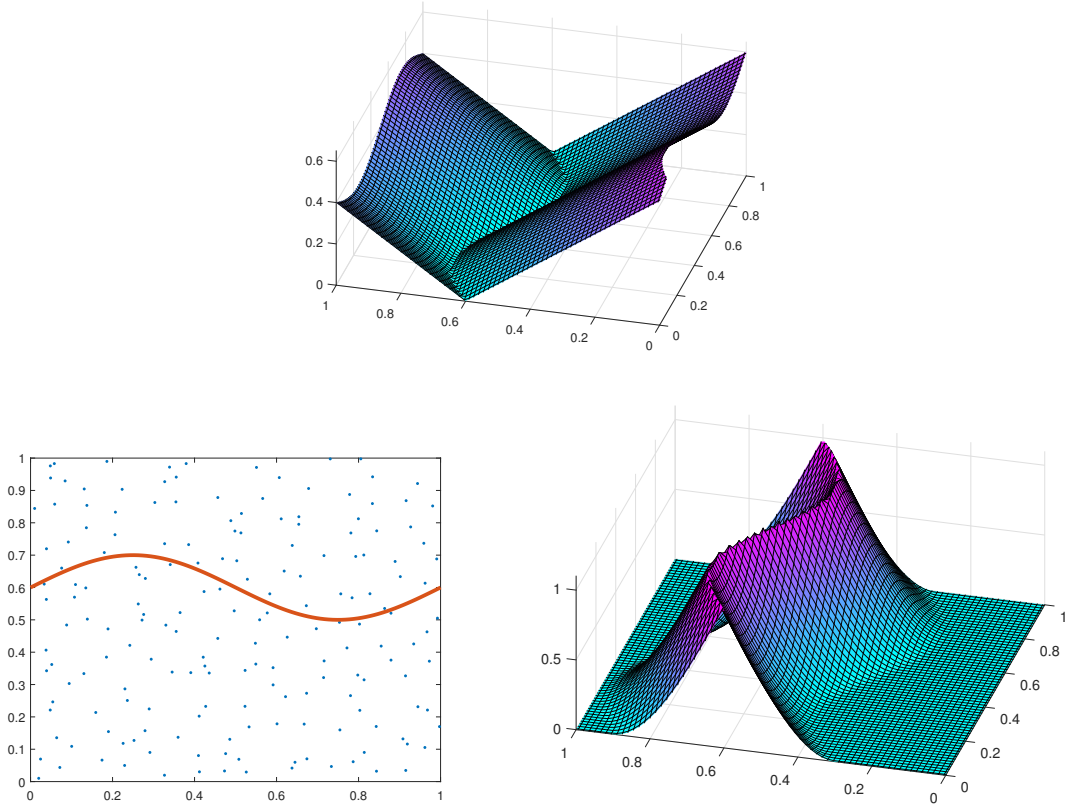

Figure 10: Test function f_3, data sites with Γ, and $c(x,y)$

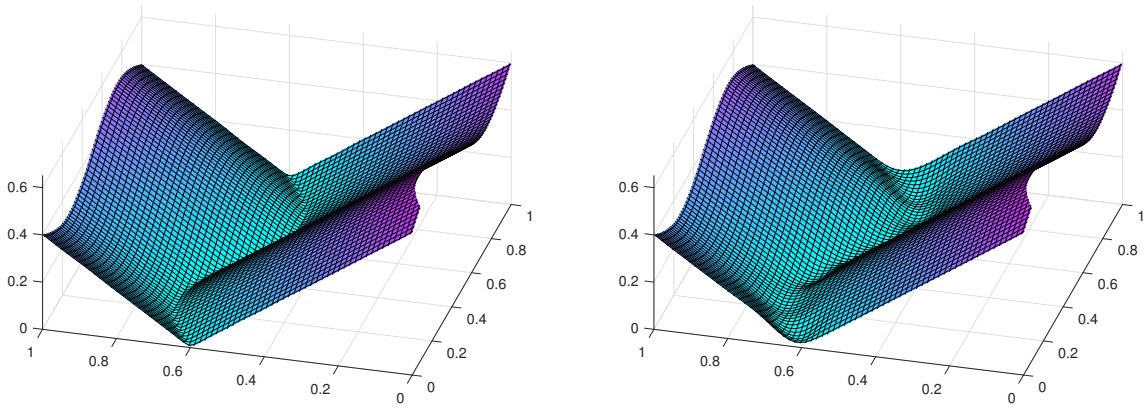

Figure 11: Right: VSK-interpolation. Right: Standard interpolation

4.2.2 Example 4

We conclude the section by showing that the VSK interpolant with c given by (9) provides excellent results also when the underlying function has very steep gradients near the gradient fault. To this end, we consider the test function

$$f_4(x,y) = |\tanh(9x-9y)|$$

sampled at the data sites depicted in Fig. 13. In this case the interpolation error is considerably reduced near Γ (see Fig. 15).

5 Conclusions

In this paper we have considered the problem of interpolating functions with gradient faults from scattered data. We have provided a simple strategy to choose a proper scaling function which depends on the discontinuity curve. The so obtained VSKs

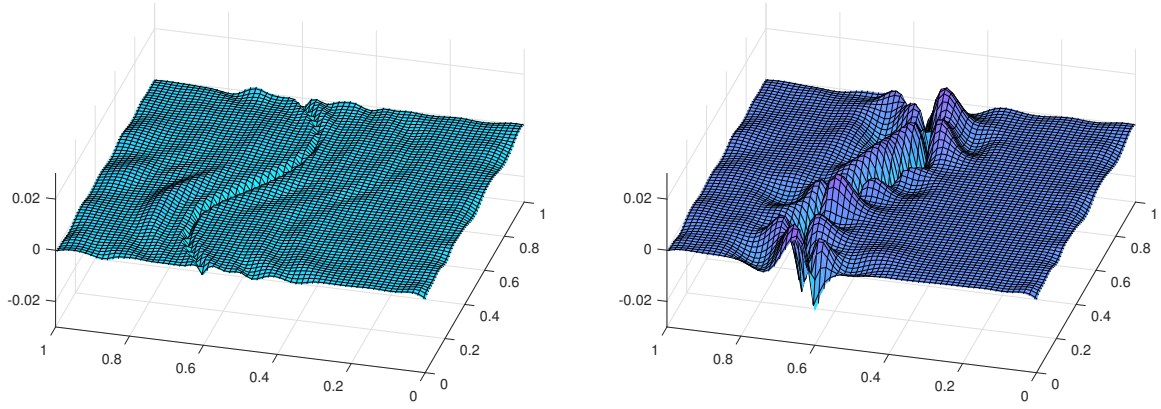

Figure 12: Right: VSK-interpolation error. Right: Standard interpolation error

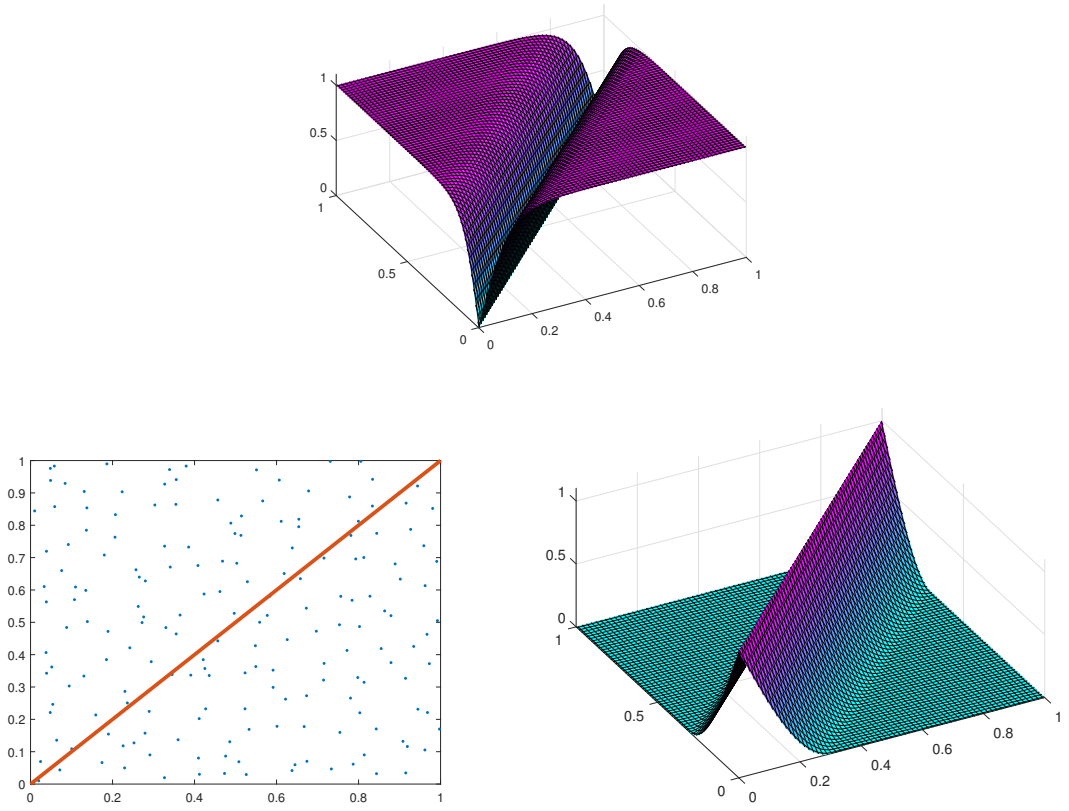

Figure 13: Test function f_3, data sites with Γ, and $c(x, y)$

Example 4	e_2	e_∞-error
Standard interpolant	3.531240e-02	2.811598e-01
VKS-interpolant	3.456239e-03	3.804824e-02

Table 4: Errors for Example 4

are capable to interpolate in an effective and faithful way the unknown function. This choice of the scale function c allows for a reduction of the interpolation error in the critical regions.

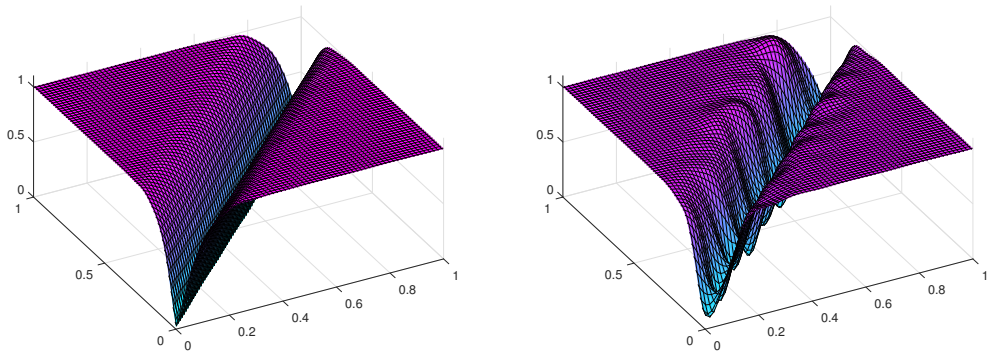

Figure 14: Right: VSK-interpolation. Right: Standard interpolation

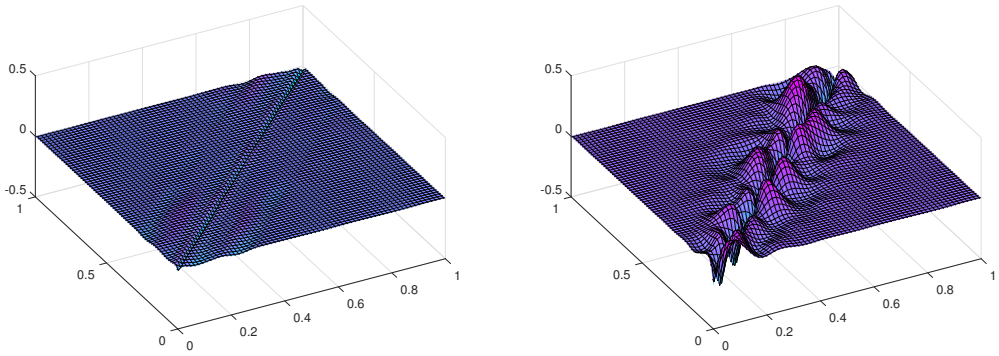

Figure 15: Right: VSK-interpolation error. Right: Standard interpolation error

References

[1] G. Allasia, R. Besenghi, and R. Cavoretto. Adaptive detection and approximation of unknown surface discontinuities from scattered data. *Simul. Model. Pract. Th.*, 17(6):1059–1070, 2009.

[2] F. Arandiga, A. Cohen, R. Donat, N. Dyn, and B. Matei. Approximation of piecewise smooth functions and images by edge-adapted (ENO-EA) nonlinear multiresolution techniques. *Appl. Comput. Harmon. Anal.*, 24(2):225–250, 2008.

[3] R. Besenghi, M. Costanzo, and A. De Rossi. A parallel algorithm for modelling faulted surfaces from scattered data. *Int. J. Comput. Numer. Anal. Appl.*, 3(4):419–438, 2003.

[4] M. Bozzini and L. Lenarduzzi. Bivariate knot selection procedure and interpolation perturbed in scale and shape. In *Curve and surface design (Saint-Malo, 2002)*, Mod. Methods Math., pages 1–10. Nashboro Press, Brentwood, TN, 2003.

[5] M. Bozzini, L. Lenarduzzi, and M. Rossini. Non-regular surface approximation. In *Mathematical methods for curves and surfaces*, volume 8177 of *Lecture Notes in Comput. Sci.*, pages 68–87. Springer, Heidelberg, 2014.

[6] M. Bozzini, L. Lenarduzzi, M. Rossini, and R. Schaback. Interpolation by basis functions of different scales and shapes. *Calcolo*, 41(2):77–87, 2004.

[7] M. Bozzini, L. Lenarduzzi, M. Rossini, and R. Schaback. Interpolation with variably scaled kernels. *IMA J. Numer. Anal.*, 35(1):199–219, 2015.

[8] M. Bozzini, L. Lenarduzzi, and R. Schaback. Adaptive interpolation by scaled multiquadrics. *Adv. Comput. Math.*, 16(4):375–387, 2002.

[9] M. Bozzini and M. Rossini. Approximating surfaces with discontinuities. *Math. Comput. Modelling*, 31(6-7):193–213, 2000.

[10] M. Bozzini and M. Rossini. The detection and recovery of discontinuity curves from scattered data. *J. Comput. Appl. Math.*, 240:148–162, 2013.

[11] A. Crampton and J. C. Mason. Detecting and approximating fault lines from randomly scattered data. *Numer. Algor.*, 39(1-3):115–130, 2005.

[12] S. De Marchi, A. Idda, and G. Santin. A rescaled method for RBF approximation. In *Approximation theory XV: San Antonio 2016*, volume 201 of *Springer Proc. Math. Stat.*, pages 39–59. Springer, Cham, 2017.

[13] S. De Marchi, A. Martínez, E. Perracchione, and M. Rossini. RBF-based partition of unity method for elliptic PDEs: Adaptivity and stability issues via VSKs. *Submitted*, 2017.

[14] M. Dehghan and V. Mohammadi. The method of variably scaled radial kernels for solving two-dimensional magnetohydrodynamic (MHD) equations using two discretizations: the Crank-Nicolson scheme and the method of lines (MOL). *Comput. Math. Appl.*, 70(10):2292–2315, 2015.

[15] T. A. Driscoll and B. Fornberg. Interpolation in the limit of increasingly flat radial basis functions. *Comput. Math. Appl.*, 43(3-5):413–422, 2002.

[16] T. A. Driscoll and A. R. H. Heryudono. Adaptive residual subsampling methods for radial basis function interpolation and collocation problems. *Comput. Math. Appl.*, 53(6):927–939, 2007.

[17] G. E. Fasshauer. *Meshfree approximation methods with MATLAB*, volume 6 of *Interdisciplinary Mathematical Sciences*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2007.

[18] G. E. Fasshauer and M. J. McCourt. Stable evaluation of Gaussian radial basis function interpolants. *SIAM J. Sci. Comput.*, 34(2):737–762, 2012.

[19] G. E. Fasshauer and J. G. Zhang. On choosing "optimal" shape parameters for RBF approximation. *Numer. Algor.*, 45(1-4):345–368, 2007.

[20] B. Fornberg, E. Larsson, and N. Flyer. Stable computations with Gaussian radial basis functions. *SIAM J. Sci. Comput.*, 33(2):869–892, 2011.

[21] B. Fornberg, G. Wright, and E. Larsson. Some observations regarding interpolants in the limit of flat radial basis functions. *Comput. Math. Appl.*, 47(1):37–55, 2004.

[22] B. Fornberg and J. Zuev. The Runge phenomenon and spatially variable shape parameters in RBF interpolation. *Comput. Math. Appl.*, 54(3):379–398, 2007.

[23] C. Gout, C. Le Guyader, L. Romani, and A.-G. Saint-Guirons. Approximation of surfaces with fault(s) and/or rapidly varying data, using a segmentation process, D^m-splines and the finite element method. *Numer. Algor.*, 48(1-3):67–92, 2008.

[24] D. Izquierdo, M. C. López de Silanes, M. C. Parra, and J.J. Torrens. CS-RBF interpolation of surfaces with vertical faults from scattered data. *Math. Comput. Simulation*, 102:11–23, 2014.

[25] E.J. Kansa and R.E. Carlson. Improved accuracy of multiquadric interpolation using variable shape parameters. *Comput. Math. Appl.*, 24(12):99–120, 1992.

[26] E. Larsson and B. Fornberg. Theoretical and computational aspects of multivariate interpolation with increasingly flat radial basis functions. *Comput. Math. Appl.*, 49(1):103–130, 2005.

[27] Y.J. Lee, G.J. Yoon, and J. Yoon. Convergence of increasingly flat radial basis interpolants to polynomial interpolants. *SIAM J. Math. Anal.*, 39(2):537–553, 2007.

[28] L. Lenarduzzi and R. Schaback. Kernel-based adaptive approximation of functions with discontinuities. *Appl. Math. Comput.*, 307:113–123, 2017.

[29] Y. Lipman and D. Levin. Approximating piecewise-smooth functions. *IMA J. Numer. Anal.*, 30(4):1159–1183, 2010.

[30] M. C. López de Silanes, M. C. Parra, M. Pasadas, and J. J. Torrens. Spline approximation of discontinuous multivariate functions from scattered data. *J. Comput. Appl. Math.*, 131(1-2):281–298, 2001.

[31] M. C. López de Silanes, M. C. Parra, and J.J. Torrens. Vertical and oblique fault detection in explicit surfaces. *J. Comput. Appl. Math.*, 140(1-2), 2002.

[32] M. Pazouki, S.S. Allaei, M. Hossein Pazouki, and D.P.F. Moller. Adaptive learning algorithm for RBF neural networks in kernel spaces. In *Proceedings of the International Joint Conference on Neural Networks*, pages 4811–4818. 2016.

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

[34] M. Rossini. Detecting discontinuities in two-dimensional signals sampled on a grid. *JNAIAM J. Numer. Anal. Ind. Appl. Math.*, 4(3-4):203–215, 2009.

[35] R. Schaback. Multivariate interpolation by polynomials and radial basis functions. *Constr. Approx.*, 21(3):293–317, 2005.

[36] R. Schaback. Limit problems for interpolation by analytic radial basis functions. *J. Comput. Appl. Math.*, 212(2):127–149, 2008.

[37] H. Wendland. *Scattered data approximation*, volume 17 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2005.

[38] Q. Zhang, Y. Zhao, and J. Levesley. Adaptive radial basis function interpolation using an error indicator. *Numer. Algor.*, 76(2):441–471, 2017.