Recovery of Functions From Weak Data Using Unsymmetric Meshless Kernel-Based Methods

Robert Schaback
University of Göttingen, Germany

March 21, 2006

Abstract

Recent engineering applications successfully introduced unsymmetric meshless local Petrov-Galerkin (MLPG) schemes. As a step towards their mathematical analysis, this paper investigates nonstationary unsymmetric Petrov-Galerkin-type meshless kernel-based methods for the recovery of L_2 functions from finitely many weak data. The results cover solvability conditions and error bounds in negative Sobolev norms with optimal rates. These rates are mainly determined by the approximation properties of the trial space, while choosing sufficiently many test functions ensures stability. Numerical examples are provided, supporting the theoretical results and leading to new questions for future research.

Keywords: Approximation, convolution, least squares, Petrov-Galerkin, overdetermined systems, error bounds, stability

Classification: 65D10, 65F20, 41A25, 41A30

1 Introduction

In the emerging field of *meshless methods* for solving partial differential equations, unsymmetric techniques for solving problems in *strong* form have quite some history beginning in 1986 with [10] within practical applications in engineering and science, and were mathematically underpinned in a recent paper [15]. They are special cases of *kernel-based* techniques which arise in many other areas as well [17]. Similar computational methods were introduced for problems in *weak* form [1, 2, 3] but they still deserve a thorough theoretical analysis.

As a first step, and building on [15], this paper looks at problems in *weak* form. However, it still avoids partial differential equations and concentrates instead on the direct recovery of L_2 functions from weak data.

To improve later applicability to the aforementioned techniques, it focuses on *unsymmetric meshless Petrov-Galerkin* methods and thus has to separate *test* and *trial* functions. But in order to get useful results, these functions are restricted to be *translations* and *dilations* of fixed *kernels*. Here, these kernels are positive definite and compactly supported, leading to sparse unsymmetric linear systems.

Weak data in the sense of this paper are generated by *convolution* of the solution and the data with scaled “test” kernels, independent of the trial space used. Consequently, the paper starts in Section 2 with an analysis of scaled kernels and the convolutions they introduce. Then Section 3 applies the results to prove *well-posedness* of recovery from weak data. *Least-squares approximation* is a symmetric special case, identifying test and trial functions, and Section 4 provides optimal error bounds in negative Sobolev norms for quite general choices of kernel-based trial and test functions. These results serve as a reference for what follows.

The paper turns to *unsymmetric* methods in Section 5, and in Section 6 it provides a stability condition which is necessary for any analysis of unsymmetric methods. Section 7 then combines the results on stability and well-posedness of weak recovery to derive error bounds and convergence rates for unsymmetric Petrov-Galerkin recovery techniques for L_2 functions from weak data, using meshless kernel-based trial and test functions.

The paper closes with two numerical experiments indicating typical features of recovery of functions from weak data:

- the trial spaces care for the approximation quality and deserve future research for their adaptive enrichment,
- the test spaces care for stability and require future research to guarantee a balance between noise reduction and preservation of local details.

2 Kernels and Convolutions

We consider a translation-invariant and Fourier-transformable *kernel* K on $I\!R^d$ with a generalized Fourier transform satisfying

$$c_K(1+\|\omega\|_2^2)^{-\kappa} \leq |\hat{K}(\omega)| \leq C_K(1+\|\omega\|_2^2)^{-\kappa} \text{ for all } \omega \in I\!R^d \qquad (1)$$

with some real number $\kappa > 0$ controlling the smoothness of the kernel. The kernel will be required to have support in the closed Euclidean unit ball $B(0,1) \subset I\!R^d$. Such kernels were provided bu Z.M. Wu [22] and H. Wendland [18], and the books [5, 19] together with the survey [17] contain a fairly complete account of the background information on kernels.

We scale the kernel by a fixed positive constant $k \leq 1$ and define *dilations* and *translations*

$$u_{y,k}(x) := k^{-d} K\left(\frac{x-y}{k}\right) =: K_k(x-y) \text{ for all } x, y \in I\!R^d. \qquad (2)$$

The function $u_{y,k}$ now is *centered* at y with support radius at most k, since the kernel has compact support in $B(0,1)$. The multiplier k^{-d} implies $\hat{K}_k(\omega) = \hat{K}(k\omega)$ for all $\omega \in I\!R^d$, and in particular the integral over K_k on $I\!R^d$ is independent of k. Functions of this form will occur later as *trial* and *test* functions, using different kernels. Since many application papers (e.g. [9, 21]) stress the importance of proper scaling, we want to track the influence of the dilation k carefully.

Since the kernel K is positive definite, it is the reproducing kernel of its *native* Hilbert space [13, 14] defined as the space of all generalized functions f on $I\!R^d$ with

$$\|f\|_K^2 := \int_{I\!R^d} \frac{|\hat{f}(\omega)|^2}{\hat{K}(\omega)} < \infty.$$

Under the above assumptions, this native space is norm-equivalent to a Sobolev space $W_2^\kappa(I\!R^d)$.

Lemma 1 *Let K be a positive definite kernel with $\hat{K} > 0$ on $I\!R^d$. For each function u in the global native space of K_k we have*

$$\frac{k^{2\kappa}}{C_K}\|u\|^2_{W_2^\kappa(I\!R^d)} \le \|u\|^2_{K_k} \le \frac{1}{c_K}\|u\|^2_{W_2^\kappa(I\!R^d)}.$$

Proof: We look at Fourier transforms and get

$$\begin{aligned}
\|u\|^2_{K_k} &= \int_{I\!R^d} \frac{|\hat{u}(\omega)|^2}{\hat{K}_k(\omega)} d\omega \\
&= \int_{I\!R^d} \frac{|\hat{u}(\omega)|^2}{\hat{K}(k\omega)} d\omega \\
&\le \frac{1}{c_K}\int_{I\!R^d} |\hat{u}(\omega)|^2 (1 + k^2\|\omega\|_2^2)^\kappa d\omega \\
&\le \frac{1}{c_K}\|u\|^2_{W_2^\kappa(I\!R^d)}, \\
\|u\|^2_{K_k} &\ge \frac{1}{C_K}\int_{I\!R^d} |\hat{u}(\omega)|^2 (1 + k^2\|\omega\|_2^2)^\kappa d\omega \\
&= \frac{k^{2\kappa}}{C_K}\int_{I\!R^d} |\hat{u}(\omega)|^2 \left(\frac{1}{k^2} + \|\omega\|_2^2\right)^\kappa d\omega \\
&\ge \frac{k^{2\kappa}}{C_K}\|u\|^2_{W_2^\kappa(I\!R^d)}. \qquad \Box
\end{aligned}$$

We shall use rather general forms of weak data later, but they will always be generated by convolution against a kernel. Thus we now take a look at the global behavior of the convolution map $f \mapsto f * K_k$.

Lemma 2 *For each generalized function $f \in W_2^\mu(I\!R^d)$ with $\mu \in I\!R$ we have $f * K_k \in W_2^{\mu+2\kappa}(I\!R^d)$ and*

$$c_K\|f\|_{W_2^\mu(I\!R^d)} \le \|f * K_k\|_{W_2^{\mu+2\kappa}(I\!R^d)} \le C_K k^{-2\kappa}\|f\|_{W_2^\mu(I\!R^d)}. \qquad (3)$$

Proof: Inspecting Fourier transforms for $g := f * K_k$ yields

$$\begin{array}{cl} & \|g\|^2_{W_2^{\mu+2\kappa}(I\!R^d)} \\ = & \int_{I\!R^d} |\hat{g}(\omega)|^2(1+\|\omega\|_2^2)^{\mu+2\kappa}d\omega \\ = & \int_{I\!R^d} \hat{f}^2(\omega)\hat{K}_k^2(\omega)(1+\|\omega\|_2^2)^{\mu+2\kappa}d\omega \\ = & \int_{I\!R^d} |\hat{f}(\omega)|^2\hat{K}^2(k\omega)(1+\|\omega\|_2^2)^{\mu+2\kappa}d\omega \\ \leq & C_K^2 \int_{I\!R^d} |\hat{f}(\omega)|^2(1+k^2\|\omega\|_2^2)^{-2\kappa}(1+\|\omega\|_2^2)^{\mu+2\kappa}d\omega \\ = & C_K^2 \int_{I\!R^d} |\hat{f}(\omega)|^2(1+\|\omega\|_2^2)^{\mu}\left(\frac{1+\|\omega\|_2^2}{1+k^2\|\omega\|_2^2}\right)^{2\kappa}d\omega \\ = & k^{-4\kappa}C_K^2 \int_{I\!R^d} |\hat{f}(\omega)|^2(1+\|\omega\|_2^2)^{\mu}\left(\frac{1+\|\omega\|_2^2}{\frac{1}{k^2}+\|\omega\|_2^2}\right)^{2\kappa}d\omega \\ \leq & k^{-4\kappa}C_K^2 \int_{I\!R^d} |\hat{f}(\omega)|^2(1+\|\omega\|_2^2)^{\mu}d\omega \\ = & k^{-4\kappa}C_K^2\|f\|^2_{W_2^{\mu}(I\!R^d)}. \end{array}$$

Similarly,

$$\begin{array}{rcl} \|g\|^2_{W_2^{\mu+2\kappa}(I\!R^d)} & = & \int_{I\!R^d} |\hat{g}(\omega)|^2(1+\|\omega\|_2^2)^{\mu+2\kappa}d\omega \\ & = & \int_{I\!R^d} |\hat{f}(\omega)|^2\hat{K}_k^2(\omega)(1+\|\omega\|_2^2)^{\mu+2\kappa}d\omega \\ & = & \int_{I\!R^d} |\hat{f}(\omega)|^2\hat{K}^2(k\omega)(1+\|\omega\|_2^2)^{\mu+2\kappa}d\omega \\ & \geq & c_K^2 \int_{I\!R^d} |\hat{f}(\omega)|^2(1+k^2\|\omega\|_2^2)^{-2\kappa}(1+\|\omega\|_2^2)^{\mu+2\kappa}d\omega \\ & = & c_K^2 \int_{I\!R^d} |\hat{f}(\omega)|^2(1+\|\omega\|_2^2)^{\mu}\left(\frac{1+\|\omega\|_2^2}{1+k^2\|\omega\|_2^2}\right)^{2\kappa}d\omega \\ & \geq & c_K^2 \int_{I\!R^d} |\hat{f}(\omega)|^2(1+\|\omega\|_2^2)^{\mu}\left(\frac{1+k^2\|\omega\|_2^2}{1+k^2\|\omega\|_2^2}\right)^{2\kappa}d\omega \\ & = & c_K^2\|f\|^2_{W_2^{\mu}(I\!R^d)}. \end{array}$$

□

The following is a localization of the previous lemma.

Lemma 3 *Let $f \in W_2^\mu(\Omega)$ be a generalized function with $\mu \in I\!R$ and an extension $\mathcal{Z}_\Omega f$ by zero outside Ω such that $\mathcal{Z}_\Omega f \in W_2^\mu(I\!R^d)$. The positive definite kernel K should have support in the unit ball and satisfy* (1)*. Then we have*

$$c_K\|f\|_{W_2^\mu(\Omega)} \leq \|(\mathcal{Z}_\Omega f) * K_k\|_{W_2^{\mu+2\kappa}(\Omega^k)} \leq C_K k^{-2\kappa}\|f\|_{W_2^\mu(\Omega)} \qquad (4)$$

where

$$\Omega^k := \{y \in I\!R^d \ : \ dist(y,\Omega) \leq k\}$$

is the support of the convolution.

Proof: This is the previous lemma applied to $\mathcal{Z}_\Omega f$. □

Lemma 3 covers two different situations:

- If f is globally smooth and compactly supported in Ω, both lemmas are the same, and they make sense for positive μ.
- But if f is not globally smooth, it will often have an extension $\mathcal{Z}_\Omega f \in L_2(I\!R^d)$ such that the above lemma works for $\mu = 0$ independent of the smoothness of f inside Ω.

We now want to check conditions on μ, k, and κ under which the test functions $u_{y,k}$ defined via (2) and (1) are in $W_2^\mu(I\!R^d)$.

Lemma 4 *If*

$$2\mu + d < 4\kappa \tag{5}$$

holds, each test function $u_{y,k}$ *for* $y \in I\!R^d$ *and* $k \in (0, 1]$ *lies in* $W_2^\mu(I\!R^d)$. *Its norm is bounded above by*

$$\|u_{y,k}\|^2_{W_2^\mu(I\!R^d)} \le C_K^2 \int_{I\!R^d} (1 + \|\omega\|_2^2)^\mu d\omega$$

if $\mu < -d/2$ *and*

$$\|u_{y,k}\|^2_{W_2^\mu(I\!R^d)} \le k^{-(d+2\mu+2\epsilon)} C_K^2 \int_{I\!R^d} (1 + \|\omega\|_2^2)^{-d/2-\epsilon} d\omega$$

if $\mu \ge -d/2$ *and if* $\epsilon > 0$ *is chosen to satisfy*

$$0 < \epsilon < 2\kappa - d/2 - \mu \tag{6}$$

which is possible due to (5).

We can combine these two inequalities in a somewhat sloppy notation by

$$\|u_{y,k}\|^2_{W_2^\mu(I\!R^d)} \le C(\epsilon) \cdot k^{-(d+2\mu+2\epsilon)_+}.$$

It may be surprising that the bound does not depend on the smoothness parameter κ of the kernel once the latter is smooth enough.

Proof: We evaluate Fourier transforms

$$\begin{aligned}
\|u_{y,k}\|^2_{W_2^\mu(I\!R^d)} &= \int_{I\!R^d} |\hat{w}_{y,k}|^2(\omega)(1 + \|\omega\|_2^2)^\mu d\omega \\
&= \int_{I\!R^d} |\hat{K}_k|^2(\omega)(1 + \|\omega\|_2^2)^\mu d\omega \\
&= \int_{I\!R^d} |\hat{K}|^2(k\omega)(1 + \|\omega\|_2^2)^\mu d\omega \\
&\le C_K^2 \int_{I\!R^d} \left(1 + k^2\|\omega\|_2^2\right)^{-2\kappa} (1 + \|\omega\|_2^2)^\mu d\omega
\end{aligned}$$

and the integral is well-defined for all positive k if we assume (5). To bound these integrals properly for all real μ, all $\kappa > 0$ and $0 < k \le 1$ with optimal powers of k, we first look at the case $\mu < -d/2$. Then

$$\begin{aligned}
& C_K^2 \int_{I\!R^d} \left(1 + k^2\|\omega\|_2^2\right)^{-2\kappa} (1 + \|\omega\|_2^2)^\mu d\omega \\
\le\ & C_K^2 \int_{I\!R^d} (1 + \|\omega\|_2^2)^\mu d\omega.
\end{aligned}$$

In case $\mu \geq -d/2$ we have to enforce convergence of the integral by

$$\begin{aligned}
& \int_{I\!R^d} \left(1+k^2\|\omega\|_2^2\right)^{-2\kappa} (1+\|\omega\|_2^2)^\mu d\omega \\
= & \int_{I\!R^d} \left(1+k^2\|\omega\|_2^2\right)^{-d/2-\mu-\epsilon} (1+\|\omega\|_2^2)^\mu \\
& \left(1+k^2\|\omega\|_2^2\right)^{-2\kappa+d/2+\mu+\epsilon} d\omega \\
\leq & \; k^{-d-2\mu-2\epsilon} \int_{I\!R^d} \left(1+\|\omega\|_2^2\right)^{-d/2-\mu-\epsilon} \\
& (1+\|\omega\|_2^2)^\mu \left(1+k^2\|\omega\|_2^2\right)^{-2\kappa+d/2+\mu+\epsilon} d\omega \\
\leq & \; k^{-d-2\mu-2\epsilon} \int_{I\!R^d} \left(1+\|\omega\|_2^2\right)^{-d/2-\epsilon} d\omega
\end{aligned}$$

provided that $-2\kappa + d/2 + \mu + \epsilon < 0$ and $\epsilon > 0$. This works with (5) if we pick any positive ϵ with (6). □

3 Nonstationary Testing

We now turn to the recovery of functions from weak data. In principle, weak data will be generated by convolution of a given generalized function f against a *test* kernel, while the approximating *trial* functions will be dilated translates of a *trial* kernel. Since we want to allow different kernels for the trial and test side later, we shall replace the notation $K,\ \kappa,\ k$ by $S,\ \sigma,\ s$ on the *te***S***t* side and $R,\ \rho,\ r$ on the *t***R***ial* side, applying the results of the previous section.

We first want to check if the recovery process is well-posed, i.e. that small weak data of a function imply a small norm of the function itself. We assume f to be given in a local Sobolev space $W_2^\mu(\Omega)$ with a well-behaving bounded domain $\Omega \subset I\!R^d$, and we use a test kernel S with smoothness σ and scale s. The trial side is not relevant in this section.

Classical weak data of f consist of

$$(f *_\Omega S_s)(y) := \int_\Omega f(x) S_s(y-x)dx =: (\mathcal{Z}_\Omega f * S_s)(y) \text{ for all } y \in I\!R^d$$

where $\mathcal{Z}_\Omega f$ is the zero extension of f to all of $I\!R^d$. Since the test kernel S has support in the unit ball, the data sites y can be confined to the larger domain

$$\Omega^s := \{y \in I\!R^d \ : \ dist(y,\Omega) \leq s\}$$

because the weak data will vanish outside Ω^s.

Note that we use a fixed scale s of the test kernel S, but allow all translates with respect to test centers $y \in \Omega^s$. This should be called *nonstationary testing*. It is fundamentally different from the “stationary” finite-element situation, where the discretization parameter h affets *both* the scaling and the translation simultaneously.

We assume $\mathcal{Z}_\Omega f$ to be in some global Sobolev space $W_2^m(I\!R^d)$ where we allow negative or zero values of m. Then Lemma 3 yields

$$c_S\|f\|_{W_2^m(\Omega)} \leq \|\mathcal{Z}_\Omega f * S_s\|_{W_2^{m+2\sigma}(\Omega^s)} \leq C_S s^{-2\sigma}\|f\|_{W_2^m(\Omega)},$$

proving

Theorem 1 *Recovery of functions from full nonstationary weak data is well-posed in the sense of the above inequality, if a function $f \in \mathcal{W}_2^m(\Omega)$ is recovered from all values of $\mathcal{Z}_\Omega f * S_s$ on Ω^s.* □

Of course, the above result is not useful in practice, because it requires infinitely many weak data. If we discretize on a set Y_s consisting of points in Ω^s with fill distance

$$h_s := \sup_{y\in\Omega^s}\min_{y\in Y_s}\|x-y\|_2$$

there, and for $\mu + 2\sigma$ being a nonnegative integer satisying

$$0 \leq \mu + 2\sigma < \mu + 2\sigma + d/2 < \lfloor m + 2\sigma - 1\rfloor, \tag{7}$$

we can apply a very useful result of [20] to get

$$\begin{array}{rcl} \|f *_\Omega S_s\|_{W_2^{\mu+2\sigma}(\Omega^s)} & \leq & C\Big(h_s^{m-\mu}\|f *_\Omega S_s\|_{W_2^{m+2\sigma}(\Omega^s)} \\ & & +h_s^{-\mu-2\sigma}\|f *_\Omega S_s\|_{\infty,Y_s}\Big)\,. \end{array} \tag{8}$$

By Lemma 2, this implies

$$\begin{array}{rcl} \|f\|_{W_2^{\mu}(\Omega)} & \leq & C\left(h_s^{m-\mu}s^{-2\sigma}\|f\|_{W_2^m(\Omega)}\right. \\ & & \left.+h_s^{-\mu-2\sigma}\|f *_\Omega S_s\|_{\infty,Y_s}\right)\,. \end{array} \tag{9}$$

The following sections will apply the above results to error functions $f-u$ where u is an unspecified trial function of at least the same smoothness as f. Then (9) implies

$$\begin{array}{rcl} \|f-u\|_{W_2^{\mu}(\Omega)} & \leq & C\left(h_s^{m-\mu}s^{-2\sigma}\|f-u\|_{W_2^m(\Omega)}\right. \\ & & \left.+h_s^{-\mu-2\sigma}\|(f-u) *_\Omega S_s\|_{\infty,Y_s}\right) \\ & \leq & C\left(h_s^{m-\mu}s^{-2\sigma}(\|f\|_{W_2^m(\Omega)}+\|u\|_{W_2^m(\Omega)})\right. \\ & & \left.+h_s^{-\mu-2\sigma}\|(f-u) *_\Omega S_s\|_{\infty,Y_s}\right) \end{array} \tag{10}$$

and leads to convergence theorems for all recovery processes which keep $\|u\|_{W_2^m(\Omega)}$ under control and recover the discrete data with satisfactory accuracy. The usual least-squares approximants have these properties. Other techniques will need additional arguments or some numerical regularization to keep a high derivative at bay. Note that most spline- and kernel-based interpolation methods have no problems with such an assumption.

Theorem 2 *Recovery of functions from discrete nonstationary weak data is well-posed in the sense of* (10)*, if a function $f \in \mathcal{W}_2^m(\Omega)$ is recovered from all values of $\mathcal{Z}_\Omega f * S_s$ on a discrete subset Y_s of Ω^s and if algorithms are used which keep the norms $\|u\|_{W_2^m(\Omega)}$ of the recovering trial functions uniformly bounded.* □

If f is not a globally smooth function with support in Ω, the extension $\mathcal{Z}_\Omega f$ will be only in $L_2(\mathbb{R}^d)$, and then we have to take $m = 0$ and negative μ with

$$-2\sigma \le \mu < \mu + d/2 < \lfloor 2\sigma - 1 \rfloor - 2\sigma < 0 \tag{11}$$

to get

$$\|f\|_{W_2^\mu(\Omega)} \le C \left(h_s^{-\mu} s^{-2\sigma} \|f\|_{L_2(\Omega)} + h_s^{-\mu - 2\sigma} \|f *_\Omega S_s\|_{\infty, Y_s} \right). \tag{12}$$

This is not too bad for this low regularity, because one cannot expect to reconstruct an L_2 function from discrete weak data with convergence in strong norms.

4 Least-squares Approximation

We now consider standard linear approximation in $L_2(\Omega)$ where test and trial functions coincide. For later use in a context where testing is done on a different set using a different kernel, we write everything in terms of a *trial* kernel R and associated parameters ρ, r. This approximates f by functions

$$u_{y,r}(x) := R_r(x - y)$$

which are superimposed to yield trial functions

$$u_r(x) := \sum_{y \in Y_r} \alpha_y u_{y,r}(x),$$

and which are used for testing also. Again, we use a *nonstationary* scenario where the scale r is fixed and convergence hopefully occurs by taking sufficiently many translates.

The optimal least-squares approximation u_r^* of the above form based on weak data of f satisfies

$$\begin{aligned} 0 &= (f - u_r^*, u_{y,r})_{L_2(\Omega)} \\ &= \int_\Omega R_r(x - y)(f - u_r^*)(x)dx \\ &= (\mathcal{Z}_\Omega(f - u_r^*) * R_r)(y) \text{ for all } y \in Y_r \end{aligned}$$

and trivially $\|u_r^*\|_{L_2(\Omega)} \le \|f\|_{L_2(\Omega)}$ bounds the approximant uniformly. Thus we can invoke (12) to get

Theorem 3 *Least-squares approximation by test and trial functions generated by a smooth compactly supported kernel R using weak data $(\mathcal{Z}_\Omega f * R_r)(y)$ sampled on a discrete set $Y_r \subset \Omega^r$ has an error bound*

$$\|f - u_r^*\|_{W_2^\mu(\Omega)} \le C h_r^{-\mu} r^{-2\rho} \|f\|_{L_2(\Omega)}$$

for the range of negative μ given by

$$-2\rho \le \mu < \mu + d/2 < \lfloor 2\rho - 1 \rfloor - 2\rho < 0 \tag{13}$$

if integration is carried out exactly. For numerical integration, an additional term has to be expected, consisting of the absolute integration error multiplied by a term of order $h_r^{-\mu-2\rho}$.

Note that the *nonstationary* (i.e. r fixed) case has an optimal convergence rate $-\mu$ for $h_r \to 0$ which does not depend on the smoothness or the scale of the kernel. As in other cases of kernel-based meshless approximations, the *stationary* case $h_r \simeq r$ cannot work for $r \to 0$. This is not surprising, because we assumed $f \in L_2(\Omega)$ only.

Of course, numerical integration is impossible for functions $f \in L_2(\Omega)$ for which we have no other information. But we feel in good company with the finite element literature here, since the integration of products of functions $f \in L_2(\Omega)$ with test functions is a standard ingredient which is only rarely questioned there. Special consideration of errors induced by numerical integration and spoiling the performance of weak recovery algorithms is provided by [4, 6, 7, 8], for instance, including finite element methods.

5 Unsymmetric Meshless Methods

We now generalize our reconstruction technique from the symmetric least-squares case to an unsymmetric Petrov-Galerkin-type strategy. We start again by taking trial functions which are linear combinations

$$u_r(x) := \sum_{y \in Y_r} \alpha_y R_r(x - y) \tag{14}$$

of translations and dilations of a *trial kernel* R taken at points y of a finite set $Y_r \subset I\!R^d$ of *trial centers*. This generates a trial space U_r we shall use later, but we shall also restrict the set Y_r to a domain Ω^r.

Testing is done by a kernel S with parameters σ, s and a set Y_s. We work or way towards unsymmetric methods, because we want to deal with unsymmetric meshless Petrov-Galerkin techniques like the MLPG [3] in future papers. This requires a thorough study of L_2 recovery by unsymmetric meshless methods first. We proceed like in the paper [15] dealing with unsymmetric strong collocation methods.

On both the trial and the test side, we shall stick to the *nonstationary* situation, keeping scales fixed and hoping for convergence when the fill distances h_r and h_s for the trial and test side tend to zero.

Our numerical procedures will recover the weak data via an approximate solution of the orthogonality equations

$$(\mathcal{Z}_\Omega(f - u^*_{r,s}) * S_s)(y) = 0, \; y \in Y_s \tag{15}$$

if $u^*_{r,s}$ is of the form (14), but now also depending on the test side. Note that this is not a standard least-squares problem on Ω for several reasons.

The system takes the form

$$
\begin{aligned}
(\mathcal{Z}_\Omega f * S_s)(y) &= ((\mathcal{Z}_\Omega u^*_{r,s}) * S_s)(y) \\
&= \int_\Omega u^*_{r,s}(x) \cdot S_s(y-x)dx \\
&= \sum_{z \in Y_r} \alpha_z \int_\Omega R_r(x-z)S_s(y-x)dx \\
&= \tilde{u}_{r,s}(y),\ y \in Y_s
\end{aligned} \tag{16}
$$

of a kernel-based meshless interpolation problem for a trial function

$$
\tilde{u}_{r,s}(y) := ((\mathcal{Z}_\Omega u^*_r) * S_s)(y) = \sum_{z \in Y_r} \alpha_z K_{r,s,\Omega}(z,y) \tag{17}
$$

based on the *localized convolution kernel*

$$
K_{r,s,\Omega}(z,y) := \int_\Omega R_r(x-z)S_s(y-x)dx
$$

which is neither symmetric nor positive definite nor explicitly accessible in general. We have $|Y_r|$ degrees of freedom and $|Y_s|$ equations. This models unsymmetric methods of Petrov-Galerkin type. For solving partial differential equations, these arise, for instance, in the meshless Petrov-Galerkin method due to Atluri [3] and collaborators. But at this point we confine ourselves to L_2 recovery and omit complications induced by differential equations.

6 Stability

Any argument for proving convergence of unsymmetric methods cannot bypass a property of the form

$$
\|(\mathcal{Z}_\Omega u_r) * S_s\|_{W_2^{\mu+2\sigma}(\Omega^s)} \le c(r,s,\mu)\|(\mathcal{Z}_\Omega u_r) * S_s\|_{\infty,Y_s} \text{ for all } u_r \in U_r. \tag{18}
$$

If such an inequality does not hold, there are nonzero trial functions solving the discrete homogeneous problem, spoiling any error bound. More precisely, this inequality bounds a "continuous" norm in terms of a discrete one on a finite-dimensional space, and it is a way of expressing stability of discretizations on both the test and trial side. The test side should contain enough test data to generate a "test norm" on the trial space, and then the above inequality follows from equivalence of norms on finite-dimensional spaces. Then the system (16) will have full rank $|Y_r| \le |Y_s|$, preventing numerical failure.

Theorem 4 *If kernels R and S with parameters ρ, r and σ, s are used for the trial and test side, respectively, the stability property* (18) *takes the form*

$$
\|(\mathcal{Z}_\Omega u_r) * S_s\|_{W_2^{\mu+2\sigma}(\Omega^s)} \le Ch_s^{-\mu-2\sigma}\|(\mathcal{Z}_\Omega u_r) * S_s\|_{\infty,Y_s} \tag{19}
$$

for all $u_r \in U_r$ and it holds under the assumptions (7),

$$
2m + d < 4\rho \tag{20}
$$

and

$$Ch_s^{m-\mu}s^{-2\sigma}\gamma(Y_r,m,\mu) < \frac{1}{2}, \tag{21}$$

where $\gamma(Y_r,m,\mu)$ is the constant in an inverse inequality

$$\|u_r\|_{W_2^m(\Omega)} \leq \gamma(Y_r,m,\mu)\|u_r\|_{W_2^\mu(\Omega)} \text{ for all } u_r \in U_r. \tag{22}$$

Proof: We start again with a bound like (8) from [20] to get

$$\begin{array}{rcl}\|(\mathcal{Z}_\Omega u_r) * S_s\|_{W_2^{\mu+2\sigma}(\Omega^s)} & \leq & C\left(h_s^{m-\mu}\|(\mathcal{Z}_\Omega u_r) * S_s\|_{W_2^{m+2\sigma}(\Omega^s)}\right.\\ & & \left.+h_s^{-\mu-2\sigma}\|(\mathcal{Z}_\Omega u_r) * S_s\|_{\infty,Y_s}\right)\\ & \leq & C\left(h_s^{m-\mu}s^{-2\sigma}\|u_r\|_{W_2^m(\Omega)}\right.\\ & & \left.+h_s^{-\mu-2\sigma}\|(\mathcal{Z}_\Omega u_r) * S_s\|_{\infty,Y_s}\right)\end{array} \tag{23}$$

for all trial functions $u_r \in U_r$. The range (7) of (8) applies again, but Lemma 4 allows all m with (20) to guarantee $U_r \subset W_2^m(\Omega)$. By norm equivalence on finite-dimensional spaces, the first term on the right-hand side will satisfy some inverse inequality of the form (22) which we cannot quantify more precisely, leaving evaluation of good constants $\gamma(Y_r,m,\mu)$ to future research.

We now continue from (22) by Lemma 2 to arrive at

$$\|u_r\|_{W_2^m(\Omega)} \leq C\gamma(Y_r,m,\mu)\|(\mathcal{Z}_\Omega u_r) * S_s\|_{W_2^{\mu+2\sigma}(\Omega^s)}$$

and require an additional assumption of the form (21) to turn (23) into (18) with $c(r,s,\mu) \leq Ch_s^{-\mu-2\sigma}$. □

As standard results [11, 16, 19] on kernels suggest, the expectable optimal form of $\gamma(Y_r,m,\mu)$ is

$$\gamma(Y_r,m,\mu) \simeq q_r^{\mu-m}$$

where q_r is the separation distance

$$q_r := \frac{1}{2}\min_{y\neq z,\ y,z\in Y_r}\|y-z\|_2$$

of Y_r, which is proportional to the fill distance h_r if the trial centers are not too wildly scattered. In such a case, the inequality (21) will be satisfied if the ratio h_s/h_r of fill distances stays bounded above by a sufficiently small constant determined by the other ingredients like smoothness and scaling of the kernels.

However, the upshot here is that the constant $\gamma(Y_r,m,\mu)$ does *not* depend on the *test* side. Of course, it will depend on the domain, the trial kernel parameters, and the distribution of the trial centers in Y_r.

The condition (21) can always be satisfied if the test discretization is "fine enough". It quantifies the statement at the beginning of this section, i.e. that the test side should contain enough test data to generate a "test norm" on the trial space.

Finally, note that the technical parameter m is transient in the sense that it does not directly appear in the final assertion (19).

7 Error Analysis

To analyze our class of unsymmetric meshless methods, we first fix the trial parameters and get a trial approximation u_r^* satisfying Theorem 3, serving as a reference. It is independent of the test side, because it is the (numerically unknown) least-squares approximation to the data by functions from the trial space.

This approximation generates weak test data satisfying

$$\begin{array}{rcl}\|\mathcal{Z}_\Omega(f-u_r^*)*S_s\|_{W_2^{\mu+2\sigma}(\Omega^s)} & \leq & C_S s^{-2\sigma}\|f-u_r^*\|_{W_2^\mu(\Omega)} \\ & \leq & Cs^{-2\sigma}h_r^{-\mu}r^{-2\rho}\|f\|_{L_2(\Omega)}\end{array} \qquad (24)$$

by Lemma 3 for μ and ρ restricted by (13). This implies that one can solve the system (15) by some trial function $u_{r,s}^*$ of the form (14) approximately to some accuracy

$$\|\mathcal{Z}_\Omega(f-u_{r,s}^*)*S_s\|_{\infty,Y_s} \leq \delta(r,s) \qquad (25)$$

with, for example,

$$\delta(r,s) = 2\cdot Cs^{-2\sigma}h_r^{-\mu}r^{-2\rho}\|f\|_{L_2(\Omega)}$$

taking the constant of (24). We now use (18) to proceed with the general error analysis as follows:

$$\begin{array}{rcl}\|f-u_{r,s}^*\|_{W_2^\mu(\Omega)} & \leq & C\|(\mathcal{Z}_\Omega(f-u_{r,s}^*))*S_s\|_{W_2^{\mu+2\sigma}(\Omega^s)} \\ & \leq & C\|(\mathcal{Z}_\Omega(f-u_r^*))*S_s\|_{W_2^{\mu+2\sigma}(\Omega^s)} \\ & & +C\|(\mathcal{Z}_\Omega(u_r^*-u_{r,s}^*))*S_s\|_{W_2^{\mu+2\sigma}(\Omega^s)} \\ & \leq & Cs^{-2\sigma}h_r^{-\mu}r^{-2\rho}\|f\|_{L_2(\Omega)} \\ & & +Cc(r,s,\mu)\|(\mathcal{Z}_\Omega(u_r^*-u_{r,s}^*))*S_s\|_{\infty,Y_s} \\ & \leq & Cs^{-2\sigma}h_r^{-\mu}r^{-2\rho}\|f\|_{L_2(\Omega)} \\ & & +Cc(r,s,\mu)\|(\mathcal{Z}_\Omega(u_r^*-f))*S_s\|_{\infty,Y_s} \\ & & +Cc(r,s,\mu)\|(\mathcal{Z}_\Omega(f-u_{r,s}^*))*S_s\|_{\infty,Y_s} \\ & \leq & Cs^{-2\sigma}h_r^{-\mu}r^{-2\rho}\|f\|_{L_2(\Omega)} \\ & & +Cc(r,s,\mu)s^{-2\sigma}h_r^{-\mu}r^{-2\rho}\|f\|_{L_2(\Omega)} \\ & \leq & Cc(r,s,\mu)s^{-2\sigma}h_r^{-\mu}r^{-2\rho}\|f\|_{L_2(\Omega)} \\ & \leq & Cr^{-2\rho}s^{-2\sigma}h_s^{-\mu-2\sigma}h_r^{-\mu}\|f\|_{L_2(\Omega)}\end{array}$$

if we assume $c(r,s,\mu) = Ch_s^{-\mu-2\sigma} \geq 1$ without loss of generality, keeping in mind that $\mu+2\sigma$ must be a nonnegative integer due to (7).

But we now have to make sure that the above bound implies convergence for a certain range of parameters. This will be done in a few successive steps, and from the user's point of view.

First, we fix the order ρ of the trial kernel R, and we want to make it large enough so that (20) provides some leeway for the transient parameter m, while (13) should provide leeway for μ to be picked later. Second, we fix m to satisfy (20). These two choices also define our leeway in (7), so that ρ and m should be not too small. Third, we fix the order σ of the test kernel. By (7) and (13) this gives a certain range of admissible values for μ. In particular, we are always allowed to choose $\mu = -2\sigma$ in case $\rho \geq \sigma$. We can take any μ in the admissible range now, but we should keep it negative and with a large absolute value.

So far, we have not chosen dilations r, s and discretizations Y_r and Y_s. The former do not seem to cause serious problems if chosen not too small, but note that large r, s require large point sets Y_r, Y_s discretizing large domains Ω^r, Ω^s, respectively.

If r and s are fixed, the user should first choose a fine and quasi-uniform set Y_r of trial centers, because its fill distance h_r in Ω^r will in the end drive the convergence rate. The final choice of test centers Y_s must then be made to satisfy (21) for a suitable fill distance h_s of Y_s in Ω^s. But in view of the term $h_s^{-\mu-2\sigma}$ in the final estimate, the choice of h_s should be not too small unless we pick $\mu = -2\sigma$ and $\rho \geq \sigma$.

For the latter case, the final error bound takes the form

$$\|f - u^*_{r,s}\|_{W_2^{-2\sigma}(\Omega)} \leq Cr^{-2\rho}s^{-2\sigma}h_r^{2\sigma}\|f\|_{L_2(\Omega)} \tag{26}$$

which is of optimal order concerning $h_r \to 0$.

For more general μ one has to wait for good quantitative results on the inverse inequality (22). If we work with the expectable condition $h_s \simeq Ch_r$ with a sufficiently small constant to make (21) valid for fixed choices of the other parameters, the final error bound turns out to be

$$\|f - u^*_{r,s}\|_{W_2^{\mu}(\Omega)} \leq Ch_r^{-\mu-2\sigma}h_r^{-\mu}\|f\|_{L_2(\Omega)}.$$

If we temporarily fix μ, the optimal rate $h_r^{-\mu}$ is always counteracted by $h_r^{-\mu-2\sigma}$ due to $\mu+2\sigma$ having to be a nonnegative integer. The best choice of σ for μ fixed will then be $2\sigma = -\mu$ getting us back to the previous situation.

Altogether, the interpretation of our results leads to the suggestion to pick $\mu = -2\sigma$ and $\rho \geq \sigma$ in all cases. We summarize this special case now.

Theorem 5 *If the kernel parameters are chosen to satisfy*

$$2\rho \geq 2\sigma \geq 2 + \frac{d}{2}, \tag{27}$$

there is an error bound of the form (26) *if the test discretization is fine enough to satisfy* (21). □

Proof: Under the above assumption, we have to show that all required inequalities can be satisfied for proper choices of $\mu = -2\sigma$ and m. The inequality (27) implies $4\sigma > 3 + d$ and thus

$$2\sigma - \tfrac{d}{2} - 1 \quad > \quad \tfrac{d}{2} + 2 - 2\sigma.$$

Then we can pick $m \in I\!R$ to satisfy

$$2\sigma - \frac{d}{2} > m > \frac{d}{2} + 2 - 2\sigma.$$

The left-hand side now proves (20) via

$$2m + d < 4\sigma \le 4\rho.$$

For $\mu = -2\sigma$ the right-hand side leads to

$$1 + \lfloor m - \mu \rfloor \ge m - \mu = m + 2\sigma \quad > \quad \tfrac{d}{2} + 2$$

and implies (7) via

$$\frac{d}{2} < \lfloor m - \mu \rfloor - 1 = \lfloor m + 2\sigma - 1 \rfloor.$$

Furthermore, (27) gives

$$-2\sigma + \frac{d}{2} \le -2 < \lfloor 2\sigma - 1 \rfloor - 2\sigma$$

and proves (11) and (13). □

The restriction $\rho \ge \sigma$ should be replaced in the future by something allowing discontinuous positive definite kernels. But there are hardly any examples and no theory for such, except for multiscale wavelet-related kernels by R. Opfer [12].

8 Numerical Examples

We now have to discuss techniques to solve the overdetermined and unsymmetric linear system (15) approximately. We know from (24) that there is a good approximate solution, but we have to make sure not to discard it. Thus any reasonable optimization routine will do the job, e.g. minimizing the residuals in the ℓ_∞ norm via linear optimization using a dual revised simplex method, but in practice a standard least-squares solver suffices.

Providing numerical examples supporting asymptotic results like (26) in a quantitative sense is a nontrivial task due to the negative norms involved. Furthermore, convergence in negative norms can take place in spite of visible Gibbs phenomena in the reproduction of discontinuous functions, because negative norms will iron out small local oscillations.

We provide two univariate examples on $\Omega = [-1, 1]$ with data in Table 1. In both cases, we tested with Wendland's C^2 radial basis function $\phi(r) = (1 - r)_+^4(1 + 4r)$ [18], but for trial functions we took also the Gaussian $\phi(r) = \exp(-2r^2)$. The latter has no compact support, but we used the standard scaling. The given function is $f(x) = x^2$ modified to be zero in $[-0.5, 0]$ in order to provide a discontinuity of f at -0.5 and of f'' at zero. At the boundary $x = 1$ we used extension by zero, while at

$x = -1$ we extended f itself, in order to study the chopping effect at the boundary. The reason is that in all possible cases one should try to extend the data beyond the domain instead of using a continuation by zero. The linear overdetermined systems were solved by standard least-squares. [1]

	Wendland	Gauss
r	0.20	0.10
s	0.01	0.05
h_r	0.05	0.10
h_s	0.01	0.05
$\|Y_s\|$	203	43
$\|Y_r\|$	49	23

Table 1: Example data

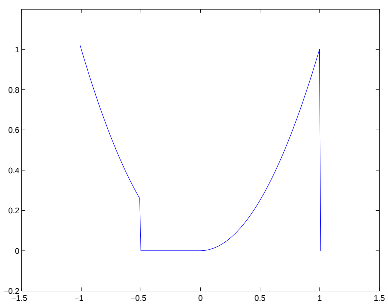

Figure 1: Given function

We see the expected Gibbs phenomena at the discontinuities in -0.5 and 1. These spread out if we solve the system by ℓ_∞ minimization of the residuals. This reveals that our error analysis via bounding $\|(\mathcal{Z}_\Omega f) * S_s\|_{\infty,Y_s}$ should be replaced by $\|(\mathcal{Z}_\Omega f) * S_s\|_{\ell_2,Y_s}$.

An advantage of unsymmetric methods is that they can add trial functions ad libitum without changing the test scenario. If step functions at 1 and -0.5 are added to the above example, the Gibbs phenomena disappear, and the approximations reproduce the given function with graphic accuracy, so that additional figures are not necessary.

The outcome of various other numerical experiments leads to the following conclusions:

[1] MATLAB© routines are available on http://www.num.math.uni-goettingen.de/schaback/research/group.html

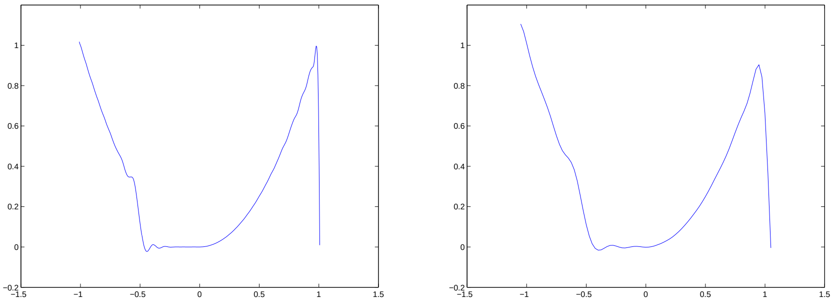

Figure 2: Reconstructions by different trial functions: Wendland (left) Gaussian (right)

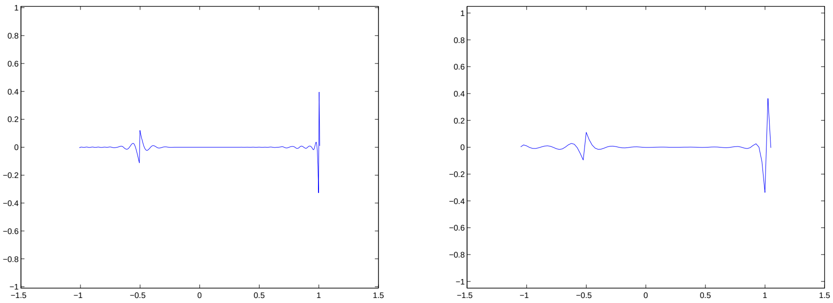

Figure 3: Errors: Wendland (left) Gaussian (right)

- Weak testing means that convolved data are reproduced by convolved trial functions. The smoothing effect of the convolution in weak nonstationary testing must be chosen very carefully, keeping a balance between smoothing the noise away and smoothing important details away. Key features which are smoothed away by testing with excessively smooth and wide kernels will not be recovered well by any trial space.
- If testing is done sensibly along the above lines, the reproduction quality depends mainly on the trial side, not on the test side. This practical observation is in accordance with our theory. Features of the data which cannot be modelled by the trial space will always be missed by the reproduction, no matter how testing is done. This means that peculiarities like known singularities should always be incorporated into the trial space.

Future research on unsymmetric meshless methods should exploit the freedom provided by separating the test from the trial side. In view of the two observations above, it has to provide good algorithms and theoretical foundations for

- enriching the trial space adaptively

- balancing testing adaptively between noise reduction and feature preservation.

References

[1] S. N. Atluri. *The meshless method (MLPG) for Domain and BIE discretization*. Tech Science Press, 2004.

[2] S. N. Atluri and S. Shen. *The Meshless Local Petrov-Galerkin (MLPG) Method*. 2002.

[3] S. N. Atluri and T. L. Zhu. A new meshless local Petrov-Galerkin (MLPG) approach in computational mechanics. *Computational Mechanics*, 22:117–127, 1998.

[4] I. Babuška, U. Banerjee, and J.E. Osborn. Meshless and generalized finite element methods: A survey of some major results. In *Meshfree Methods for Partial Differential Equations, Lecture Notes in Computational Science and Engineering 26*, pages 1–20. Springer, 2002.

[5] M.D. Buhmann. *Radial Basis Functions*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, 2004.

[6] J.S. Chen, C.T. Wu, S. Yoon, and Y. You. A stabilized conforming nodal integration for Galerkin mesh-free methods. *International Journal for Numerical Methods in Engineering*, 50:435–466, 2001.

[7] P. G. Ciarlet. Basic error estimates for elliptic problems. In *Handbook of numerical analysis, Vol. II*, Handb. Numer. Anal., II, pages 17–351. North-Holland, Amsterdam, 1991.

[8] J. Dolbow and T. Belytschko. Numerical integration of the Galerkin weak form in meshfree methods. *Comput. Mech.*, 23(3):219–230, 1999.

[9] R. E. Hagan and E. J. Kansa. Studies of the r parameter in the multiquadric function applied to ground water pumping. *J. Appl. Sci. Comput*, 1:266–281, 1994.

[10] E. J. Kansa. Application of Hardy's multiquadric interpolation to hydrodynamics. In *Proc. 1986 Simul. Conf., Vol. 4*, pages 111–117, 1986.

[11] F.J. Narcowich, J.D. Ward, and H. Wendland. Sobolev error estimates and a Bernstein inequality for scattered data interpolation via radial basis functions. 2006.

[12] R. Opfer. Multiscale kernels. Adv. Comput. Math. (to appear), 2006.

[13] R. Schaback. Native Hilbert spaces for radial basis functions I. In M.D. Buhmann, D. H. Mache, M. Felten, and M.W. Müller, editors, *New Developments in Approximation Theory*, number 132 in International Series of Numerical Mathematics, pages 255–282. Birkhäuser Verlag, 1999.

[14] R. Schaback. A unified theory of radial basis functions (Native Hilbert spaces for radial basis functions II). *J. Comp. Appl. Math.*, 121:165–177, 2000.

[15] R. Schaback. Convergence of unsymmetric kernel–based meshless collocation methods. Preprint Göttingen, 2005.

[16] R. Schaback and H. Wendland. Inverse and saturation theorems for radial basis function interpolation. *Math. Comp.*, 71(238):669–681 (electronic), 2002.

[17] R. Schaback and H. Wendland. Kernel techniques: From machine learning to meshless methods. 97 pages, to appear in Acta Numerica, 2006.

[18] H. Wendland. Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Advances in Computational Mathematics*, 4:389–396, 1995.

[19] H. Wendland. *Scattered Data Approximation.* Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, Cambridge, 2004.

[20] H. Wendland and C. Rieger. Approximate interpolation with applications to selecting smoothing parameters. *Numerische Mathematik*, 101:729–748, 2005.

[21] J. Wertz, E. J. Kansa, and L. Ling. The role of the multiquadric shape parameters in solving elliptic partial differential equations. *Comput. Math. Appl.*, pages 1133–1140, 2004.

[22] Z. Wu. Multivariate compactly supported positive definite radial functions. *Advances in Computational Mathematics*, 4:283–292, 1995.