

Unsymmetric Meshless Methods for Operator Equations

Robert Schaback

Institut für Numerische und Angewandte Mathematik
Universität Göttingen
Lotzestraße 16-18
D-37083 Göttingen, Germany
Tel: ++49 +551-39-4501/02 Fax: ++49 +551-39-3944

Received: date / Revised version: date

Summary A general framework for proving error bounds and convergence of a large class of unsymmetric meshless numerical methods for solving well-posed linear operator equations is presented. The results provide optimal convergence rates, if the test and trial spaces satisfy a stability condition. Operators need not be elliptic, and the problems can be posed in weak or strong form without changing the theory. Non-stationary kernel-based trial and test spaces are shown to fit into the framework, disregarding the operator equation. As a special case, unsymmetric meshless kernel-based methods solving weakly posed problems with distributional data are treated in some detail. This provides a foundation of certain variations of the “Meshless Local Petrov-Galerkin” (MLPG) technique of S.N. Atluri and collaborators.

AMS Classification: 65N12, 65N35, 65N22, 65F22

1 Introduction

Since this paper has to turn very technical later, an outline of the basic arguments is necessary. We require six essential ingredients:

1. a well-posed linear *operator equation* to be solved,
2. *existence* of a solution,
3. *good approximability* of the exact solution by functions from a finite-dimensional *trial space*,

4. a well-posed *sampling* strategy for *testing* trial functions,
5. a *stable finite discretization* of the test sampling
6. a numerical method which *approximately minimizes* the residuals of the discretized test sampling over the admissible space of trial functions.

Then we can prove error bounds and convergence rates for the method in question, and the rates turn out to be best possible in certain cases. Applications cover unsymmetric methods in strong and weak form, and in particular this paper seems to be the first to provide a rigid mathematical foundation of certain variations of the “Meshless Local Petrov-Galerkin” (MLPG) technique of S.N. Atluri and collaborators [3] which already fills two books [2,1].

We shall explain the above ingredients now one by one, postponing examples to later sections because this requires plenty of details and obscures the basic line of argument. The paper [11] provided a similar framework, but restricted to problems in strong form, while [12] treated recovery of functions from weak data without considering operator equations.

1.1 Linear Operator Equations

We consider an equation

$$Lu = f \text{ for } L : \mathcal{U} \rightarrow \mathcal{F}, f \in \mathcal{F} \text{ given} \quad (1)$$

to be solved for $u \in \mathcal{U}$, where the *solution space* \mathcal{U} and the *data space* \mathcal{F} are normed linear spaces of functions, and L is a linear operator. The problem should be well-posed in the sense

$$L : \mathcal{U} \rightarrow \mathcal{F} \text{ is continuous and bijective.}$$

We assume existence of an exact solution $u^* \in \mathcal{U}$ which necessarily is unique under the above assumptions. But there are no other hypotheses on L , in particular there is no ellipticity, compactness, or self-adjointness assumed. Later, we have to pay a price for this by considering numerical methods that require some kind of optimization. In other words: if there is no hidden minimization in the analytic background, there should be one in the numerical technique.

Well-posedness of the operator equation (1) allows numerical methods to focus on residuals, because there always is a trivial error bound

$$\|u - u^*\|_{\mathcal{U}} \leq \|L^{-1}\| \cdot \|Lu - Lu^*\|_{\mathcal{F}} = \|L^{-1}\| \cdot \|Lu - f\|_{\mathcal{F}}$$

for a *trial* function u in terms of its *residual* $Lu - f$.

But we shall use the full spaces \mathcal{U} and \mathcal{F} only for theoretical purposes. They usually are too large to allow computations, since they often are L_2 or low-order Sobolev spaces. In addition, we shall use Sobolev spaces of negative order to prove certain convergence theorems for solutions of weakly posed problems. Our actual computations work on subspaces U and $F = L(U)$ of \mathcal{U} and \mathcal{F} , respectively, which inherit the corresponding norms. The problem $Lu = f$ will be posed such that a function $u^* \in U$ solves it for $f \in F$.

1.2 Test, Trial, and Symmetry

The rest of the paper will make a clear distinction between the *test* and *trial* side of the problem (1). *Trial* functions $u \in U \subseteq \mathcal{U}$ are candidates for an approximate solution, leading each to a *residual* $Lu - f$ which can be numerically *tested* for being small or zero. If this testing is done by function evaluation of the residual, it can be called *strong* testing, in contrast to *weak* testing which makes inner products of the residual with *test functions* small. Both variations are completely independent of how trial functions are supplied. The attributes *strong* and *weak* are used here exclusively to distinguish between different testing strategies. The notions of weak and strong *solutions* of partial differential equations are closely related, but different.

Symmetric methods like the standard finite element technique have a close link between the test and trial side, while *unsymmetric* methods uncouple these. We shall focus on the unsymmetric case here, and our abstract framework will not distinguish between strong and weak testing. However, we later focus on a class of weak unsymmetric techniques as our major example.

1.3 Trial Approximations

No matter how testing is done, the quality of a numerical method for solving a well-posed linear operator equation $Lu = f$ will always depend on how well the trial functions u are able to make the residual norm $\|Lu - f\|_{\mathcal{F}}$ or the error norm $\|u - u^*\|_{\mathcal{U}}$ small. Thus the convergence rate of an algorithm will mainly be determined by an *approximation property* of the *trial* side, and be independent of the test side. If certain features of the data f of the exact solution u^* are not modeled by data Lu of trial functions, there is no hope to get a useful method. For instance, it is questionable to refine discretizations or meshes if the addition of some special functions into the trial

space could do the job. Trial spaces should always allow adaptive enrichment by exotic trial functions, and this is another argument to uncouple the trial from the test side.

Therefore we model the trial side by an *approximation property*

$$\|u - \Pi_r u\|_{\mathcal{U}} \leq \epsilon(r) \|u\|_U \text{ for all } u \in U \subset \mathcal{U} \quad (2)$$

with projectors

$$\Pi_r : U \rightarrow U_r \subset U \subset \mathcal{U}$$

mapping U onto special finite-dimensional *trial subspaces* U_r of U . The bound (2) is assumed to hold on the full *regularity* subspace U of \mathcal{U} which we equip with a strong norm $\|\cdot\|_U$ for this purpose. Instead of the standard notation h for a discretization parameter, we use r for the trial and s for the test side, but we follow standard techniques by viewing the trial side as a scale of spaces U_r with approximation errors $\epsilon(r) \rightarrow 0$ for $r \rightarrow 0$ when approximating fixed functions from the fixed regularity subspace U .

Note that (1) and (2) already imply that there is a good, but numerically unknown candidate for a useful approximate solution, namely the approximation $u_r^* := \Pi_r u^*$ of the exact solution. It has the error bound

$$\|u^* - u_r^*\|_{\mathcal{U}} \leq \epsilon(r) \|u^*\|_U \quad (3)$$

since we assume that the exact solution u^* is sufficiently regular in the sense that $u^* \in U$ holds instead of only $u^* \in \mathcal{U}$.

This error bound is both a guideline and a goal for what follows. No numerical method should be as stupid as to discard $u_r^* := \Pi_r u^*$ when going for small residuals, and this is why *residual minimization* techniques must be successful for any well-posed linear operator equation with useful trial spaces. The main practical problem is that plain minimization of $\|Lu - f\|_{\mathcal{F}}$ has to take place in the norm of \mathcal{F} , which usually is not numerically accessible. The data space \mathcal{F} is determined by the well-posedness of the operator equation, and it will in many cases be a Cartesian product of Sobolev trace spaces, whose norms are hard to handle numerically. *Testing*, as discussed right now, can be seen as a workaround, enabling to assure small residuals without working directly with the norm of \mathcal{F} .

1.4 Testing

Having weak methods in mind, we now focus on the fourth ingredient of our list at the beginning of Section 1. We assume that *testing* is carried out via a linear, continuous, and bijective map

$$A : F \rightarrow T$$

mapping a practically accessible subspace F of the full data space \mathcal{F} onto a normed *test space* T . The idea is that a problem $Lu = f$, if to be solved in weak form, is not solved pointwise but rather as $\Lambda Lu = \Lambda f$ where the map Λ generates weak data $\Lambda f \in T$ of $f \in F$ in some *test space* T .

In general, the map Λ will evaluate an infinite number of linear *functionals*. They can take the form of point evaluations for strong testing, and they can be integrals against test functions for weak testing. Since we want to make sure that residuals depend continuously on the test data and vice versa, we assume continuity and bijectivity of the test map Λ , no matter how it is defined.

Even in case of *strong* testing like in collocation methods, the spaces T and F will not coincide in general, because they will often carry different norms, in particular if the norm in F is numerically unavailable, e.g. for Sobolev spaces of high regularity. The test map Λ will then often be an embedding into a test space like $L_\infty(\Omega)$.

With the test map Λ at hand, residual minimization can now be carried out in T instead of F , leaving all aforementioned arguments valid. However, from an abstract point of view, the introduction of the test map Λ is superfluous, because the operator equation (1) can be replaced by $\Lambda Lu = f$ with continuous dependence in a numerically accessible data space. But this eliminates the basic difference between F and T :

- The topology of the data space F , as inherited from \mathcal{F} , is solely determined by analytic properties of the operator equation, and is independent of numerical techniques, while
- the test space T depends on how the operator equation is tackled numerically, e.g. by strong or weak methods.

There for we do not eliminate Λ .

1.5 Test Discretizations

Up to here, we assumed an infinite number of test data, enough to identify the exact solution u^* via the full set of its data ΛLu^* . This is numerically infeasible and requires discretization. We do this by introducing “forgetful” projectors

$$\Pi_s : T \rightarrow T_s$$

mapping the full test space T onto finite-dimensional spaces T_s by taking only a finite subset of test data. It is reasonable to assume that these projectors are uniformly bounded, and to have in mind

that they form a scale for $s \rightarrow 0$ with a discretization parameter s acting like the standard h . In such a case, $u = 0$ should hold if all $\Pi_s \Lambda Lu$ are zero for all s , but we shall not assume this.

Instead, we assume a *stability condition*

$$\|\Lambda Lu_r\|_T \leq 2\beta(s)\|\Pi_s \Lambda Lu_r\|_{T_s} \text{ for all } u_r \in U_r \quad (4)$$

which needs some explanation. It links the trial side to the test side in a specific way, i.e. it bounds the norm $u_r \mapsto \|\Lambda Lu_r\|_T$ on U_r from above by a *discrete* norm, which is possible on finite-dimensional spaces. It implies that a trial function is zero if it has zero discrete data, and thus the inequality can often be satisfied by making the test discretization fine enough with respect to the trial discretization. The same reason lets the factor in (4) be only dependent on s . In standard applications, the connection to r comes as an additional requirement, making (4) valid only for a range of s that crucially depends on r .

We call a combination of trial and test discretizations *uniformly stable*, if $\beta(s)$ in (4) can be replaced by a constant. It will turn out below that growth of $\beta(s)$ for $s \rightarrow 0$ spoils optimality of error bounds. Thus it is a major problem for all applications to model the discretization of the data space in such a way that the discrete norms tend towards the non-discrete norms without loss, at least on the data provided by the trial space. This is not as easy as it sounds, e.g. for Sobolev spaces involving high-order derivatives. We shall address this question later.

1.6 Numerical Methods

The right-hand side of the stability condition (4) suggests an unsymmetric system of linear equations

$$\Pi_s \Lambda Lu_r = \Pi_s \Lambda f \quad (5)$$

to be solved for a trial function $u_r \in U_r$. Even if written in square form by choosing the same degrees of freedom on the trial and the test side, the system may not be solvable. For instance, this occurs [6] for Kansa's unsymmetric collocation technique [7] even if trial spaces are used that lead to nonsingular matrices for interpolation at scattered data. Atluri's meshless local Petrov-Galerkin method (MLPG) generates similar matrices, and since the Kansa technique is a special case of MLPG restricted to strong testing, there is no hope to prove exact solvability for the MLPG matrix, either.

Thus we only go for *approximate* solutions of the system, and we know that $u_r^* = \Pi_r u^*$ solves it to quite some accuracy

$$\|\Pi_s \Lambda(Lu_r^* - f)\|_{T_s} = \|\Pi_s \Lambda L(u_r^* - u^*)\|_{T_s} \leq \|\Pi_s \Lambda L\| \epsilon(r) \|u^*\|_U$$

dictated by the approximation power of the trial space U_r within the regularity subspace U . We thus only require that the numerical method is clever enough to produce some trial function $u_{r,s}^* \in U_r$ with

$$\|\Pi_s \Lambda(Lu_{r,s}^* - f)\|_{T_s} \leq C \|\Pi_s \Lambda L\| \epsilon(r) \|u^*\|_U \quad (6)$$

with some constant $C > 1$. Optimization of the discrete residuals in T_s will do, but any other technique is allowed which does not discard good approximate solutions. This means that unsymmetric methods like Kansa's or Atluri's techniques must take some care in solving the system (5) approximately. Usually, any numerical solution with small residuals will do in practice.

We now have finished the list of ingredients we started with, and the following section will prove an error bound leading later to convergence rates. The rest of the paper will show how this framework can be applied to unsymmetric methods solving a well-posed distributional operator equation using weak testing. In particular, a variation of the MLPG method of S.N. Atluri and his collaborators will get a solid mathematical foundation, explaining its success in applications, in particular for cases with non-smooth data.

2 General Results

Theorem 1 *Under the assumptions of Section 1 including the regularity condition $u^* \in U$ for the true solution u^* , there is an error bound*

$$\|u^* - u_{r,s}^*\|_{\mathcal{U}} \leq \left(C + 2\beta(s) \|(\Lambda L)^{-1}\| \right) \|\Pi_s \Lambda L\| \epsilon(r) \|u^*\|_U \quad (7)$$

for any approximate numerical solution $u_{r,s}^*$ of the system (5) with tolerance (6).

Proof: We combine everything into

$$\begin{aligned} \|u^* - u_{r,s}^*\|_{\mathcal{U}} &\leq \|u^* - \Pi_r u^*\|_{\mathcal{U}} + \|\Pi_r u^* - u_{r,s}^*\|_{\mathcal{U}} \\ &\leq \epsilon(r) \|u^*\|_U + \|\Pi_r u^* - u_{r,s}^*\|_{\mathcal{U}} \\ \|\Pi_r u^* - u_{r,s}^*\|_{\mathcal{U}} &\leq \|(\Lambda L)^{-1}\| \|\Lambda L(\Pi_r u^* - u_{r,s}^*)\|_T \\ &\leq 2\beta(s) \|(\Lambda L)^{-1}\| \|\Pi_s \Lambda L(\Pi_r u^* - u_{r,s}^*)\|_{T_s} \\ &\leq 2\beta(s) \|(\Lambda L)^{-1}\| \left(\|\Pi_s \Lambda L(\Pi_r u^* - u^*)\|_{T_s} \right. \\ &\quad \left. + \|\Pi_s \Lambda L(u^* - u_{r,s}^*)\|_{T_s} \right) \\ &\leq 2\beta(s) \|(\Lambda L)^{-1}\| \|\Pi_s \Lambda L\| \epsilon(r) \|u^*\|_U \\ &\quad + C \|\Pi_s \Lambda L\| \epsilon(r) \|u^*\|_U \end{aligned}$$

and in total we get (7). \square

But the stability condition (4) needs some additional theory. We have a special technique to prove stability, making a detour via subspaces \tilde{U} of \mathcal{U} and \tilde{T} of T , respectively. The two spaces should again be admissible for well-posedness of the operator equation (1) and the test sampling Λ in the sense

$$\Lambda L : \tilde{U} \rightarrow \tilde{T} \text{ is continuous and bijective.}$$

Furthermore, on T and \tilde{T} we use an inequality

$$\|v\|_T \leq C (\alpha(s)\|v\|_{\tilde{T}} + \beta(s)\|II_s v\|_{T_s}) \text{ for all } v \in T \quad (8)$$

bounding a weaker norm in terms of a stronger norm and some function values. It will have the effect that full test data $v := \Lambda Lu \in T$ are small in a weak norm, provided that they are bounded in a strong norm and small for finite many cases. Inequalities like (8) are known as Poincaré-Friedrichs inequalities in other circumstances, but we shall call them *sampling inequalities* because they describe the behavior of a projector II_s taking a finite sample of data from elements v of T . The standard behavior of the constants is

$$\alpha(s) \rightarrow 0 \text{ and } \beta(s) \rightarrow \infty \quad (9)$$

if the test discretization gets finer for $s \rightarrow 0$.

Next, we need an inverse inequality of the form

$$\|u_r\|_{\tilde{U}} \leq \gamma(r)\|u_r\|_{\mathcal{U}} \text{ for all } u_r \in U_r \quad (10)$$

on the trial space which bounds a strong norm by a weak one, thus leading to constants $\gamma(r)$ which tend to infinity for $r \rightarrow 0$. Such inequalities always exist by norm equivalence on finite-dimensional spaces. Note that this is independent of the test side.

Finally, and in view of (9), we require the test discretization to be fine enough to satisfy

$$C\alpha(s)\gamma(r)\|\Lambda L\|_{\tilde{U} \rightarrow \tilde{T}}\|(\Lambda L)^{-1}\|_{T \rightarrow U} \leq \frac{1}{2}. \quad (11)$$

Theorem 2 *Under the assumptions made above, the stability condition (4) is satisfied.*

Proof: Just reorder the result of

$$\begin{aligned} \|\Lambda L II_r u\|_T &\leq C (\alpha(s)\|\Lambda L II_r u\|_{\tilde{T}} + \beta(s)\|II_s \Lambda L II_r u\|_{T_s}) \\ &\leq C (\alpha(s)\|\Lambda L\|_{\tilde{U} \rightarrow \tilde{T}}\|II_r u\|_{\tilde{U}} + \beta(s)\|II_s \Lambda L II_r u\|_{T_s}) \\ &\leq C (\alpha(s)\|\Lambda L\|_{\tilde{U} \rightarrow \tilde{T}}\gamma(r)\|II_r u\|_{\mathcal{U}} + \beta(s)\|II_s \Lambda L II_r u\|_{T_s}) \\ &\leq C (\alpha(s)\|\Lambda L\|_{\tilde{U} \rightarrow \tilde{T}}\gamma(r)\|(\Lambda L)^{-1}\|_{T \rightarrow U}\|\Lambda L II_r u\|_T \\ &\quad + \beta(s)\|II_s \Lambda L II_r u\|_{T_s}). \square \end{aligned}$$

Later sections will focus on specific operator equations and numerical methods. But in order to show the wide applicability of the above framework to general operator equations, we first work out some general tools for the trial and test side before we select operators.

3 Kernels

Because they will occur later on both the test and the trial side, we collect some results on meshless kernel-based methods here. For background details we refer the reader to two recent books of M.D.B. Buhmann [5] and H. Wendland [14] and a survey article [13] on applications.

We define a (translation-invariant and positive definite) *kernel* $K : \mathbf{W}^d \rightarrow \mathbf{W}$ to be a function with a well-defined Fourier transform \hat{K} on \mathbf{W}^d satisfying

$$c_K(1 + \|\omega\|_2^2)^{-\kappa} \leq \hat{K}(\omega) \leq C_K(1 + \|\omega\|_2^2)^{-\kappa} \text{ for all } \omega \in \mathbf{W}^d. \quad (12)$$

Note that κ controls the smoothness of the kernel. Even for compactly supported kernels like the widely used ones of Wendland [16], the smoothness parameter κ usually is at least a half-integer, and since $\kappa > \frac{d}{2}$ ensures continuity, we shall always assume

$$\mathbf{N} \ni 2\kappa > d$$

for all kernels we consider here.

On the trial side, one can use translates $K(y - \cdot)$ for fixed *centers* $y \in \mathbf{W}^d$ to generate trial functions as translates of K . On the test side, weak data of a function u on some domain Ω can be obtained by a convolution-type integral

$$\lambda_y(u) := \int_{\Omega} u(t)K(y - t)dt. \quad (13)$$

But since we keep the trial and test side independent, we use a trial kernel R with smoothness parameter ρ and a test kernel S with smoothness parameter σ instead of K and κ . In both cases, we shall not scale or dilate the kernel. Instead, we vary the centers y occurring above in order to generate many test functionals or trial functions. In Approximation Theory, this is called a *non-stationary* approach, while a *stationary* approach links translations to dilations like in finite elements.

3.1 Kernel-based Trial Spaces

Let us first look at the trial side and work towards the approximation property (2) in Sobolev spaces $\mathcal{U} := W_2^\mu(\Omega)$ and $U := W_2^m(\Omega)$ for $m \geq \mu$. Finite-dimensional trial spaces U_r can be generated by

$$U_r := \text{span} \{R(\cdot - y) : y \in Y_r\}$$

for finite sets Y_r of translations of the trial kernel R . For work on a bounded domain $\Omega \subset \mathbf{W}^d$ we assume $Y_r \subset \Omega$ with *fill distance*

$$r := h(Y_r, \Omega) := \sup_{x \in \Omega} \min_{y \in Y_r} \|x - y\|_2$$

and *separation distance*

$$q := q(Y_r) := \frac{1}{2} \min_{x, y \in Y_r, x \neq y} \|x - y\|_2.$$

If this is done for a full scale of spaces, we assume a uniformly bounded *mesh ratio*

$$0 < c \leq \frac{h(Y_r, \Omega)}{q(Y_r)} \leq C.$$

This is no problem because users can often choose the set Y_r ad libitum.

We now cite two theorems concerning the approximation behavior of trial spaces spanned by non-stationary kernel translates.

Theorem 3 [10] *Under the above assumptions on the trial space, the projector*

$$\Pi_r : U := W_2^m(\Omega) \subset \mathcal{U} = W_2^\mu(\Omega) \rightarrow U_r \subset U$$

defined by plain interpolation in Y_r has an error bound (2) with

$$\epsilon(r) \leq Cr^{m-\mu} \tag{14}$$

under the conditions

$$0 \leq \mu \leq m \leq \rho, \quad [m - 1] > \frac{d}{2}. \tag{15}$$

for the regularity subspace $U := W_2^m(\Omega)$. \square

The somewhat unnatural right-hand condition in (15) can possibly be replaced by $m > \frac{d}{2}$ by future work refining the techniques of [10]. Theorem 3 defined the projector via strong data, but there also is a result concerning weak data.

Theorem 4 [12] *The projector Π_r defined by the best $L_2(\Omega)$ approximation has an error bound (2) with (14) for $m = 0$ under the conditions*

$$\mathbf{N}_0 \ni \mu + 2\rho < \mu + 2\rho + \frac{d}{2} < 2\rho - 1. \quad (16)$$

for the regularity subspace $U := W_2^0(\Omega) = L_2(\Omega) \subset \mathcal{U} := W_2^\mu(\Omega)$.
□

Note that both results provide optimal rates. Due to the low regularity in the second case, the convergence takes place in negative norms, using negative values of μ . A somewhat more natural condition close to (16) would be

$$-2\rho \leq \mu < -\frac{d}{2}$$

but since Theorem 4 currently relies on Theorem 3, this has to wait until (15) is extended to $m > \frac{d}{2}$.

It is conjectured that (2) holds with the optimal rate (14) between the spaces $U := W_2^m(\Omega)$ and $\mathcal{U} := W_2^\mu(\Omega)$ in many other cases also, but to determine the full range of admissible $\mu \leq m$ for a given kernel smoothness ρ is an open research problem.

For later use, we can get rid of the assumption $\mu + 2\rho \in \mathbf{N}_0$ in (16) in favor of

$$\|u - \Pi_r u\|_{W_2^\mu(\Omega)} \leq Cr^{-\lceil \mu \rceil} \|u\|_{L_2(\Omega)}$$

under the restrictions

$$-2\rho \leq \lceil \mu \rceil < -1 - \frac{d}{2} \quad (17)$$

by applying Theorem 4 for $\lceil \mu \rceil$ instead of μ and using 2ρ being an integer.

Before we go over to the test side, we should deal with the inverse inequality (10) needed for stability analysis. We use

$$\tilde{U} := W_2^n(\Omega) \subset U = W_2^m(\Omega)$$

with some $n > m$, and there clearly is some norm equivalence constant $\gamma(r) = \gamma(r, n, m, R, \Omega)$ with

$$\|u_r\|_{W_2^n(\Omega)} \leq \gamma(r) \|u_r\|_{W_2^m(\Omega)} \text{ for all } u_r \in U_r$$

which can be expected to be of the form

$$\gamma(r) \simeq r^{m-n} \quad (18)$$

for trial center distributions with bounded mesh ratio, but this is an open research problem. The assertion is true in the case $\Omega = \mathbf{W}^d$ [10],

but for bounded domains [11] there currently is only a suboptimal bound of the form $\gamma(r) \leq Cr^{-\rho}$ for the range

$$\frac{d}{2} < m \leq n \leq \rho.$$

To derive local inverse theorems for weak norms is a major challenge. But we shall try to get away with just using the existence of the inverse inequality. This restricts convergence results to qualitative assertions like “*If the test discretization is fine enough, then..*”, and future inverse theorems will replace this by quantitative results.

3.2 Testing

Testing depends on how the test data of residuals are sampled and how the error is measured. This concerns the choice of $\Lambda : F \rightarrow T$, how T is normed, discretized via $\Pi_s : T \rightarrow T_s$, and how the norm on T_s is defined. In view of our introduction, we need well-posedness of the sampling map Λ and a practically useful sampling inequality of the form (8). Later, if paired with a suitable trial space, we have to prove stability of the discretization. Since a systematic theory of testing seems to be missing, we outline it here and start with some easy examples.

3.3 Strong Testing

Let us first consider *strong* testing, i.e. we work with function values directly and avoid numerical integration. This requires some regularity in the choices of spaces T and F , but the map Λ will be a trivial embedding. The discretization Π_s will just take a finite sample $\Pi_s v := v|_{Y_s}$ of a function $v \in T$ on a finite set Y_s . This leads to the standard strong test scenario

$$F := W_2^\mu(\Omega) =: T \xrightarrow{\Pi_s} T_s := \mathbf{W}^{|Y_s|}$$

for $\mu > \frac{d}{2}$, and there are no problems with well-posedness if we avoid to define T differently. But we still are free to choose the ℓ_∞ or the ℓ_2 norm on $T_s = \mathbf{W}^{|Y_s|}$. Since we shall minimize discrete residuals in the end, this choice has consequences for our numerical procedures: we have to choose between linear optimization or least-squares. Solving linear systems of equations exactly will never work safely for unsymmetric problems.

The basic tool for proving sampling inequalities is a very useful result of Wendland and Rieger [15].

Theorem 5 *Let $\Omega \subset \mathbf{W}^d$ be a bounded piecewise smooth domain with an interior cone condition, and pick two parameters $n \in \mathbf{W}$ and $\mu \in \mathbf{N}_0$ with*

$$0 \leq \mu \leq \mu < \mu + \frac{d}{2} < \lfloor n - 1 \rfloor.$$

Then there are positive constants C and s_0 such that for every finite subset Y_s of Ω with fill distance $s \leq s_0$ and every $u \in W_2^n(\Omega)$ the inequality

$$\|f\|_{W_2^\mu(\Omega)} \leq C \left(s^{n-\mu} \|f\|_{W_2^n(\Omega)} + s^{-\mu} \|II_s f\|_{\infty, Y_s} \right)$$

holds.

A similar inequality comes from a parallel paper [9] of Madych and takes the form

$$\|f\|_{L_2(\Omega)} \leq C \left(s^n \|f\|_{W_2^n(\Omega)} + s^{d/2} \|II_s f\|_{\ell_2(Y_s)} \right) \quad (19)$$

holding for $n > d/2$. This deals with the discrete least-squares case and should be applied for T_s being normed by $s^{d/2} \|\cdot\|_{\ell_2(Y_s)}$ which cares for the discrete norms tending to the L_2 norm for data coming from smooth functions. In both cases of discrete norms, we have uniform boundedness of the projectors II_s because of $n > \frac{d}{2}$.

Like in the previous section, future research should provide results for larger choices of the parameters n , μ and for different choices of discrete data and their norms.

For applications, we need more general norms on the left-hand sides. If we apply Theorem 5 to $\lceil \mu \rceil$, we get a weaker inequality

$$\|f\|_{W_2^\mu(\Omega)} \leq C \left(h^{n-\lceil \mu \rceil} \|f\|_{W_2^n(\Omega)} + h^{-\lceil \mu \rceil} \|f\|_{\infty, Y} \right) \quad (20)$$

for the range

$$0 \leq \mu \leq \lceil \mu \rceil < \lceil \mu \rceil + \frac{d}{2} < \lfloor n - 1 \rfloor$$

which allows non-integer μ .

3.4 Full Weak Testing

We now turn to weak testing. Here, already the definition of the sampling map Λ defining the full weak data is debatable. In theory, one can define Λ via a dense total set of functionals, but the practical meaning of “weak data” consists of taking inner products

$$(\Lambda(f))(v) := \int_{\Omega} f(t)v(t)dt$$

of a *data function* $f \in F$ against many *test functions* v which usually are compactly supported. A canonical scenario thus takes the dual $T := V^*$ of a space V of test functions and defines $\Lambda : F \rightarrow T = V^*$ as above. Well-posedness is no problem if $V = F^*$ is chosen, e.g. for Sobolev spaces.

But discretization and its analysis towards a sampling inequality (8) needs more information, if it should lead to useful results. Thus we leave the approach via dual spaces to future work and focus on a more specific testing strategy, i.e. convolution-type integration (13) against a test kernel S with a small support radius δ_S . We rewrite this as a genuine convolution

$$\int_{\Omega} f(t)S(\cdot - t)dt = \int_{\mathbf{W}^d} (\mathcal{Z}_{\Omega}f)(t)S(\cdot - t)dt = (\mathcal{Z}_{\Omega}f) * S$$

of the test kernel S with the zero extension $\mathcal{Z}_{\Omega}f$ of the data function f outside its domain Ω of definition. The convolution is supported on the extended domain

$$\Omega_S := \{y \in \mathbf{W}^d : \text{dist}(y, \Omega) \leq \delta_S\}.$$

Note that this kind of testing takes a fixed scale of the test kernel (this is called *non-stationary* in other contexts, e.g. in quasi-interpolation), but it has to sample at test centers y outside the domain as long as the support of the translated kernel $S(\cdot - y)$ still hits the interior of the domain. There are plenty of other testing strategies, e.g. *stationary* ones which keep the kernel scale variable with the discretization parameter s , but they are often hard to discretize systematically. We leave this to future work and consider discretization of the above testing method instead.

Theorem 6 [12] *Weak non-stationary testing is well-posed for*

$$\begin{aligned} F &:= W_2^m(\Omega), \\ T &:= W_2^{m+2\sigma}(\Omega_S), \\ \Lambda f &:= (\mathcal{Z}_{\Omega}f) * S \text{ for all } f \in F \end{aligned}$$

provided that m is restricted to imply

$$\mathcal{Z}_{\Omega}f \in W_2^m(\mathbf{W}^d) \text{ for all } f \in W_2^m(\Omega)$$

i.e. at least for $m \leq 0$.

Indeed, by some easy Fourier transform arguments [12] the convolution map $f \rightarrow (\mathcal{Z}_{\Omega}f) * S$ provides a norm equivalence

$$c\|f\|_{W_2^m(\Omega)} \leq \|(\mathcal{Z}_{\Omega}f) * S\|_{W_2^{m+2\sigma}(\Omega_S)} \leq C\|f\|_{W_2^m(\Omega)} \quad (21)$$

under the above assumptions. \square

Inequality (21) also holds for more general m and more special f whenever the conditions $\mathcal{Z}_\Omega f \in W_2^m(\mathbf{W}^d)$ and $f \in W_2^m(\Omega)$ are satisfied. This allows larger m but at the expense of restricting the data functions to those who vanish smoothly at the boundary. Both ways of interpreting (21) thus take the limiting effect of the boundary into account. It is highly interesting to study testing strategies which fight the boundary effect, e.g. extending f first by some kernel-based method to a smooth function $\mathcal{E}_\Omega f$ outside Ω and then to sample weak data as $(\mathcal{E}_\Omega f) * S$.

The reader should be aware that (12) rules out testing against characteristic functions. More generally, a norm equivalence like (21), as needed for well-posedness of testing, cannot hold for non-stationary testing, if the Fourier transform of the kernel has zeros, because then the convolution map can vanish on nonzero functions whose spectrum is contained in the zeros of the spectrum of the kernel. As a univariate example, weak data obtained as integrals over intervals of length δ will be identically zero for all functions $\sin(2\pi(t - t_0)/\delta)$.

3.5 Discrete Weak Testing

Discretization will simply take a finite subset Y_s of the extended domain Ω_S having fill distance s there, and it will consider finitely many weak data by restricting the convolution map $f \rightarrow (\mathcal{Z}_\Omega f) * S$ to Y_s . If we define $T = W_2^{m+2\sigma}(\Omega_S)$ as above, this defines projectors

$$\Pi_s : T \rightarrow \mathbf{W}^{|Y_s|} =: T_s \text{ under } \|\cdot\|_\infty$$

for which we need a sampling inequality (8) to prove stability later. To this end, we can use results from [12] to transfer (20) to weak data.

Theorem 7 *Under the above notation, and if the parameters satisfy*

$$-2\sigma \leq m \leq [m] < [m] + \frac{d}{2} < [n - 1],$$

weak discrete non-stationary kernel testing satisfies a sampling inequality of the form

$$\|f\|_{W_2^m(\Omega)} \leq C \left(s^{n-[m]} \|f\|_{W_2^n(\Omega)} + s^{-[m]-2\sigma} \|(\mathcal{Z}_\Omega f) * S\|_{\infty, Y_s} \right) \quad (22)$$

for all $f \in W_2^n(\Omega)$, $\mathcal{Z}_\Omega f \in W_2^n(\mathbf{W}^d)$.

In fact, if S is a test kernel as above, we can use the norm equivalence (21) to prove the assertion. \square

Note the serious penalty factor $s^{-[m]-2\sigma}$ for the integration error of the discrete weak data, increasing with the smoothness of the test kernel. This is in accordance with the strong effect of integration errors on high-order finite element methods.

A weak version of the sampling inequality (19) is

$$\|f\|_{W_2^{-2\sigma}(\Omega)} \leq C \left(s^n \|f\|_{W_2^{n-2\sigma}(\Omega)} + \|(\mathcal{Z}_\Omega f) * S\|_{T_s} \right) \quad (23)$$

for all $f \in W_2^{n-2\sigma}(\Omega)$, $\mathcal{Z}_\Omega f \in W_2^{n-2\sigma}(\mathbf{W}^d)$ under the hypothesis $n > \frac{d}{2}$, where we used the properly scaled ℓ_2 norm on T_s . Note how this matches with (22) in case $m = -2\sigma \in \mathbf{Z}$ except for a slight difference in the admissible parameters. To eliminate the boundary effect, we should apply this only when $n - 2\sigma \leq 0$, which altogether implies

$$\frac{d}{2} < n \leq 2\sigma \quad (24)$$

for the applicability range of (23). If we start with a fixed kernel S , such an n always exists due to our standard condition $d < 2\sigma \in \mathbf{N}$ on the smoothness of kernels.

Uniform boundedness of the weak data projectors Π_s follows, if the functionals

$$\lambda_y(f) : f \mapsto \int_\Omega f(t) S(\cdot - t) dt = (\mathcal{Z}_\Omega f) * S$$

are uniformly bounded on $F := W_2^m(\Omega)$ for arbitrary $y \in \Omega_S$, and this clearly holds due to (21) and Sobolev embedding, if we assume

$$m + 2\sigma > \frac{d}{2} \quad (25)$$

under the above notation. Again, we can use $d < 2\sigma \in \mathbf{N}$ to find that the convolution functionals are uniformly bounded whenever $m \geq -\frac{d}{2}$, in particular for $m = 0$, no matter which test kernel we take.

3.6 Stability

If we want to satisfy the condition (11) for kernel-based non-stationary trial spaces and test strategies, we have $\alpha(s) \rightarrow 0$ for $s \rightarrow 0$ e.g. like in $\alpha(s) \simeq s^{n-[m]}$, and can cope with any $\gamma(r)$ from the trial side if we take the discretization parameter s of the test side small enough.

Theorem 8 *Standard kernel-based non-stationary discretizations along the above lines are always stable if the test discretization is fine enough.*
□

In the ideal case where (18) and (20) take their optimal forms with behavior $\gamma(r) \simeq r^{m-n}$ and $\alpha(s) \simeq s^{n-m}$, we can get away by choosing the test discretization parameter s proportional to the trial discretization parameter r . This is all to be hoped for, because there must always be at least as many degrees of freedom on the test side as on the trial side. Since the current state-of-the-art for the inverse inequality only has $\gamma(r) \simeq r^{-\rho}$ for strong settings and with no useful generalizations known, we cannot keep s and r proportional.

But we can even get uniform stability if we focus on cases of sampling inequalities (8) with constant $\beta(s)$. We have those in several cases. For strong testing we can take Theorem 5 for $\mu = 0$ and (19), but note that currently the test and data spaces T and F are only L_2 in these cases, while there is additional regularity behind the scenes. For weak sampling, we can take either (23) or (22) for $n \leq 0$.

Theorem 9 *If kernels and discretization parameters are chosen properly, there are uniformly stable non-stationary kernel-based discretizations of the data space $L_2(\Omega)$.* □

Note that this section on kernel-based trial and test strategies did not depend on the operator equations to be solved.

4 Example

Now we shall set this machinery to work for a special class of problems, but the reader will see that the scope of our framework extends far beyond this case. Since the paper [11] contains a simplified theory dealing with methods based on strong testing, we can confine ourselves here to weak testing, including a variation of Atluri's MPLG method.

Even if a boundary-value problem is fixed, there are plenty of ways to choose the kernels and the relevant spaces for the mathematical analysis. We have several possibilities to proceed:

- we can try to choose all parameters in an optimal way to get good convergence rates,
- we can consider parameters as already chosen by the user and figure out whether and how fast the discrete residual minimization method will converge for the chosen setting,
- we can try to specify the range of admissible choices to guarantee convergence at all,

- we can describe all of this for the current state-of-the-art of auxiliary tools,
- we can show the power of the framework by describing what happens if all tools were available in theoretically optimal form in the future.

We shall focus on the third and fourth case, but at certain places we shall comment on the other issues, too. The reader will finally be able to use this framework in other situations, hopefully.

4.1 Poisson Problem

We consider a Poisson problem

$$\begin{aligned} -\Delta u &= f^\Omega && \text{in } \Omega \\ u &= f^D && \text{in } \Gamma^D \subseteq \Gamma := \partial\Omega \\ \frac{\partial u}{\partial n} &= f^N && \text{in } \Gamma^N \subset \Gamma \end{aligned} \quad (26)$$

on a bounded domain $\Omega \subset \mathbf{W}^d$ with piecewise smooth Lipschitz boundary $\Gamma := \partial\Omega$ such that the standard trace theorems hold. We define

$$\begin{aligned} \mathcal{U} &:= W_2^\mu(\Omega) \\ L &:= (L^\Omega, L^D, L^N) \\ Lu &:= (-\Delta u, u|_{\Gamma^D}, \frac{\partial u}{\partial n}|_{\Gamma^N}) =: (L^\Omega u, L^D u, L^N u) \\ \mathcal{F} &:= W_2^{\mu-2}(\Omega) \times W_2^{\mu-1/2}(\Gamma^D) \times W_2^{\mu-3/2}(\Gamma^N) \\ &=: \mathcal{F}^\Omega \times \mathcal{F}^D \times \mathcal{F}^N. \end{aligned}$$

This setting leads to well-posedness [4,8] of the problem (1) in the sense of bijectivity of $L : \mathcal{U} \rightarrow \mathcal{F}$ for all μ , and it is independent of numerical methods. But note that we deliberately allow generalized functions in case of small or negative μ . We shall nowhere use the specific form of the Laplace operator, such that Δ can be replaced by any linear second-order differential operator in what follows, provided that there is well-posedness in the above form.

4.2 Trial Side

We take a trial strategy with a kernel R having a smoothness parameter ρ with $d < 2\rho \in \mathbf{N}$. Then we proceed exactly as in Section 3.1 in the context of Theorem 4. The restriction (17) is assumed to be satisfied for a sufficiently large ρ . Our regularity subspace will be $U := L_2(\Omega)$ because using $W_2^m(\Omega)$ with $m > 0$ does not improve

the outcome of Theorem 4. If future results provide approximation theorems with negative norms on the left and positive norms on the right-hand side, we can make better use of $U = W_2^m(\Omega)$ for positive m , using the same framework.

4.3 Test Side

We have three different equations with three different data to test, and thus we shall employ three different test kernels. These cases will be distinguished using Ω , D , and N as sub- or superscripts. At this point we do not follow the MLPG strategy, which combines all weak equations into one “local weak form”. We prefer to test the differential equation and each boundary equation separately, but using the same trial space.

To arrive at a final bound with an explicit convergence order, we have to make sure to work with a uniformly stable test discretization. If not, the current lack of inverse theorems for kernel-based trial approximation in weak norms lets the test parameter s be uncontrollably linked to the trial parameter r , excluding quantitative, but still enabling qualitative convergence results. This problem will hopefully be obsolete in the future. For the time being, we have to be satisfied with sampling inequalities (23) or (22), but note that the latter requires the (possibly superfluous) assumption $\mu = \lceil \mu \rceil = -2\sigma \in \mathbf{Z}$ to render uniform stability. To keep things simple, we shall base our weak testing on the inequality (23), though (22) would work as well for integer $\mu = -2\sigma$. In both cases, we get uniform stability from Theorem 9. However, this comes at a price: the test spaces and their norms are fixed now via the test kernels.

4.4 Domain Sampling

Weak non-stationary sampling in the domain is done with a test kernel S^Ω with smoothness parameter $\sigma_\Omega > \frac{d}{2}$ and support radius δ_Ω . Following Section 3.4 we have to go into a larger domain and define

$$\begin{aligned} \Omega_\Omega &:= \{y \in \mathbf{W}^d : \text{dist}(y, \Omega) \leq \delta_\Omega\} \supset \Omega \\ T^\Omega &:= L_2(\Omega_\Omega) \\ \Lambda^\Omega &: W_2^{-2\sigma_\Omega}(\Omega) \supseteq F^\Omega \rightarrow T^\Omega := \Lambda^\Omega(F^\Omega) \subseteq L_2(\Omega_\Omega) \\ \Lambda^\Omega(f) &:= (\mathcal{Z}_\Omega f) * S^\Omega \end{aligned}$$

and this weak sampling is well-posed in the sense

$$\begin{aligned} c\|f\|_{W_2^{-2\sigma_\Omega}(\Omega)} &\leq \|(\mathcal{Z}_\Omega f) * S^\Omega\|_{L_2(\Omega_\Omega)} \leq C\|f\|_{W_2^{-2\sigma_\Omega}(\Omega)} \\ c\|f\|_{F^\Omega} &\leq \|A^\Omega(f)\|_{T^\Omega} \leq C\|f\|_{F^\Omega} \end{aligned}$$

with no formal restriction on σ_Ω at this point. Note that we do not apply integration by parts, as is usually done for finite element analysis. We do not restrict our theory to even-order self-adjoint differential operators.

If we apply (23) with $f = -\Delta u$ and a transient parameter n_Ω there, obeying the inclusion

$$f^\Omega \in F^\Omega := W_2^{-2}(\Omega) \subseteq W_2^{-2\sigma_\Omega}(\Omega) \subseteq \mathcal{F}^\Omega = W_2^{\mu-2}(\Omega)$$

we should assume

$$\mu - 2 + 2\sigma_\Omega \leq 0 < \frac{d}{2} < -2 + 2\sigma_\Omega \quad (27)$$

to make also (25) valid for $m = -2$ there.

4.5 Boundary Sampling

Weak sampling on the boundary needs some additional work. We parameterize the boundary parts via domains in \mathbf{W}^{d-1} . For Dirichlet data, we should do weak testing in $L_2(\Gamma^D)$ via a bijective smooth parameterization $\varphi_D : \Omega^D \rightarrow \Gamma^D$ on a domain $\Omega^D \subset \mathbf{W}^{d-1}$ which also simplifies numerical integration. The test data for f are

$$\int_{\Gamma^D} f(t)v(t)dt = \int_{\Omega^D} f(\varphi_D(s))v(\varphi_D(s))|\nabla\varphi_D(s)|ds$$

for certain test functions v which are parameterized by $t = \varphi_D(s)$. Then we specialize to test functions of the form $v_z(t) := v_z(\varphi_D(s)) := S_s^D(z-s)$ with a kernel S^D on \mathbf{W}^{d-1} with smoothness $\sigma_D > \frac{d-1}{2}$ and scale δ_D , but where z now is a test point in the parameterization domain Ω^D . This yields

$$\int_{\Gamma^D} f(t)v_z(t)dt = \int_{\Omega^D} f(\varphi_D(s))S^D(z-s)|\nabla\varphi_D(s)|ds$$

and we see that this generates weak data $(\mathcal{Z}_{\Omega^D} u_D) * S^D$ on the extended domain

$$\Omega_D := \{y \in \mathbf{W}^{d-1} : \text{dist}(y, \Omega^D) \leq \delta_D\}$$

for the new function $f_D(s) := f(\varphi_D(s))|\nabla\varphi_D(s)|$, bringing us back to the standard situation. The boundary parameterization φ_D must be

assumed to be extendable to Ω_D without losing smoothness. To avoid complications with corners or periodicity conditions, we simply split the boundary conditions into several smooth non-periodic parts, if necessary. Then the parameterization should be extendable for each smooth piece of the boundary. We define

$$\begin{aligned} T^D &:= L_2(\Omega_D) \\ \Lambda^D &: W_2^{-2\sigma_D}(\Gamma^D) \supseteq F^D \rightarrow T^D := \Lambda^D(F^D) \subseteq L_2(\Omega_D) \\ \Lambda^D(f) &:= (\mathcal{Z}_{\Omega^D} f_D) * S^D. \end{aligned}$$

The well-posedness of weak testing is then expressed as

$$c\|u_D\|_{F^D} \leq \|(\mathcal{Z}_{\Omega^D} u_D) * S^D\|_{T^D} = \|\Lambda^D(u)\|_{T^D} \leq C\|u_D\|_{F^D}$$

with no additional restriction on σ_D at this point.

We now want to apply (23) again, and we can also satisfy (25) if we use the inclusion

$$F^D := W_2^{-1/2}(\Gamma^D) \subseteq W_2^{-2\sigma_D}(\Gamma^D) \subseteq \mathcal{F}^D := W_2^{\mu-1/2}(\Gamma^D)$$

and assume

$$\mu - \frac{1}{2} + 2\sigma_D \leq 0 < \frac{d-1}{2} < -\frac{1}{2} + 2\sigma_D, \quad (28)$$

where the right-hand part is always satisfied due to $d \geq 2$ and $2\sigma_D > d-1$. The recovery of Neumann data proceeds along the same lines with notation

$$\begin{aligned} \Omega_N &:= \{y \in \mathbf{W}^{d-1} : \text{dist}(y, \Omega^N) \leq \delta_N\} \\ T^N &:= L_2(\Omega_N) \\ \Lambda^N &: W_2^{-2\sigma_N}(\Gamma^N) \supseteq F^N \rightarrow T^N := \Lambda^N(F^N) \subseteq L_2(\Omega_N) \\ \Lambda^N(f) &:= (\mathcal{Z}_{\Omega^N} f_N) * S^N \end{aligned}$$

using a test kernel S^N with scale δ_N and smoothness σ_N . The well-posedness is expressed as

$$c\|u\|_{F^N} \leq \|\Lambda^N(u)\|_{T^N} \leq C\|u\|_{F^N}$$

and we need inclusions

$$F^N := W_2^{-3/2}(\Gamma^N) \subseteq W_2^{-2\sigma_N}(\Gamma^N) \subseteq \mathcal{F}^N := W_2^{\mu-3/2}(\Gamma^N)$$

and the condition

$$\mu - \frac{3}{2} + 2\sigma_N \leq 0 < \frac{d-1}{2} < -\frac{3}{2} + 2\sigma_N. \quad (29)$$

We can summarize these three weak non-stationary sampling strategies into

$$\begin{aligned} T &:= T^\Omega \times T^D \times T^N \\ \Lambda &:= \Lambda^\Omega \times \Lambda^D \times \Lambda^N \\ \Lambda &: F \rightarrow T, \text{ bijective.} \end{aligned}$$

Note that the differential operator and the boundary operators do not occur at all. We just used weak sampling on certain Sobolev spaces. This is very much in the spirit of the approximations via restrictions and prolongations used in [4].

4.6 Stability

This usually is the hardest task in the analysis of unsymmetric methods, but we have cared for uniform stability by choosing the right parameters in the previous sections. Looking back at sections 1.5 and 3.6, we always have a uniform stability condition (4) if the test discretization is fine enough. This statement can be made more precise once there are quantitative inverse theorems for weak norms.

4.7 Convergence

We now assemble this to prove the final error bound (7) for a certain range of parameters. Choosing the regularity space $U = L_2(\Omega)$ and a smooth trial kernel R with smoothness parameter ρ satisfying $d < 2\rho \in \mathbf{N}$ restricts the admissible range for μ to (17). For any such μ , we can hope for an error bound of the form $r^{-\mu}$ for the error measured in the norm of $\mathcal{U} = W_2^\mu(\Omega)$, provided that the test discretization is fine enough and if there are no additional conditions on μ that come up when fixing the rest.

The right-hand sides of (27), (28), and (29) are automatically satisfied for $d \geq 4$, but for small dimensions we still have to care for

$$\begin{aligned} 2\sigma_\Omega &> 2 + \frac{d}{2}, \\ 2\sigma_D &> \frac{d}{2}, \\ 2\sigma_N &> 1 + \frac{d}{2}, \end{aligned}$$

the condition on $2\sigma_D$ being always satisfied due to our general condition $d - 1 < 2\sigma_D \in \mathbf{N}$. These restrictions are not serious and will easily be satisfiable. However, the conditions

$$\begin{aligned} -2\rho - 1 &< \mu \leq 2 - 2\sigma_\Omega \\ -2\rho - 1 &< \mu \leq \frac{1}{2} - 2\sigma_D \\ -2\rho - 1 &< \mu \leq \frac{3}{2} - 2\sigma_N \end{aligned} \tag{30}$$

on μ do not allow arbitrarily smooth test kernels or arbitrarily rough trial kernels due to

$$\begin{aligned} \max(d, 2 + \frac{d}{2}) &< 2\sigma_\Omega < 2\rho + 3 \\ d - 1 &< 2\sigma_D < 2\rho + \frac{3}{2} \\ \max(d - 1, 1 + \frac{d}{2}) &< 2\sigma_N < 2\rho + \frac{5}{2}. \end{aligned} \quad (31)$$

This means that the test kernels are not allowed to be much smoother than the trial kernel. This makes sense, because if the test kernels smoothen the data very much, a rough trial kernel cannot repair this.

Under the above conditions, there is always a μ satisfying (30), leading to a valid error bound. Note that it is a valid strategy to take all kernels to be the same and not too rough.

Let us look at two extreme cases. First, we take the test kernels rather smooth. So let us focus at the case of identical kernels, i.e. $\rho = \sigma_\Omega = \sigma_D = \sigma_N$. This leaves not much leeway

$$-1 < \mu - 2\rho \leq \frac{1}{2}$$

for μ , but $\mu = -2\rho$ will still work, bringing us close to (23) and (16). Convergence will be restricted to norms with smallest possible Sobolev index due to the excessive data smoothing induced by the test kernels. The final result in the sense of (7) is of the optimal form

$$\|u^* - u_{r,s}^*\|_{W_2^{-2\rho}(\Omega)} \leq Cr^{2\rho} \|u^*\|_{L_2(\Omega)}$$

in a strongly negative norm.

The other extreme case is to take the test kernels rough. This allows a wider range of μ , defined by (30), and leading to the final result

$$\|u^* - u_{r,s}^*\|_{W_2^\mu(\Omega)} \leq Cr^{-\mu} \|u^*\|_{L_2(\Omega)}.$$

depending on the choice of a negative μ . If the user wants a maximal possible μ , i.e. a strongest possible error norm, the test kernels have to be taken as rough as possible, while the smoothness of the trial kernel can be chosen large without damage.

5 Summary

Our example showed that the abstract framework can be applied to a case which had no solid convergence theory so far, i.e. to a meshless local weak unsymmetric method similar to the MLPG technique. For reasons to be explained elsewhere, we did not use the ‘‘local

weak form” of the MLPG literature, but rather treated each part of the boundary value problem by a different test strategy, keeping the parts linked via a common trial space. Numerically, our method sets up an unsymmetric linear system consisting of three groups of test equations, one each in Ω , Γ^D , and Γ^N . The system is solved by least-squares minimization of residuals. Testing is done weakly by convolution with three different kernels S^Ω , S^D , and S^N , respectively. The given Poisson problem (26) is solved in a distributional sense, assuming the solution to be only in $L_2(\Omega)$ with distributional data

$$f^\Omega \in W_2^{-2}(\Omega), f^D \in W_2^{-1/2}(\Gamma^D), f^N \in W_2^{-3/2}(\Gamma^D).$$

Consequently, error bounds and convergence results can only be expected in negative Sobolev norms, i.e. in some space $W_2^\mu(\Omega)$ with negative μ . If the trial space uses a discretization parameter r , one can expect optimal order error bounds with the behavior $r^{-\mu}$, and this is actually achieved.

But this result comes at a price that has to be specified. First, it holds only if the test discretization is “fine enough” in a sense that cannot be quantified until some inverse theorems on kernel-based trial spaces are available in weak local norms. Second, the choice of test and trial kernels controls the range of admissible μ via (30) and (31). It is allowed to take the trial kernels arbitrarily smooth, but the smoothness of the allowed test kernels is roughly bounded above by the smoothness of the trial kernel. Taking smooth kernels will result in strongly negative μ , and conversely it is only possible to achieve moderately negative μ by choosing rough test kernels.

The abstract framework will allow generalizations to other forms of testing and to other types of operator equations. But there are some gaps in the necessary tools, leading to various precisely formulated new research problems.

References

1. S. N. Atluri. *The meshless method (MLPG) for domain and BIE discretizations*. Tech Science Press, Encino, CA, 2005.
2. S. N. Atluri and S. Shen. *The Meshless Local Petrov-Galerkin (MLPG) Method*. Tech Science Press, Encino, CA, 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. J.P. Aubin. *Approximation of elliptic boundary-value problems*. Pure and Applied Amathematics, vol. XXVI. Wiley-Interscience, 1972.

5. M. D. Buhmann. *Radial Basis Functions*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, 2004.
6. Y. C. Hon and R. Schaback. On unsymmetric collocation by radial basis functions. *Appl. Math. Comput.*, 119(2-3):177–186, 2001.
7. E. J. Kansa. Application of Hardy’s multiquadric interpolation to hydrodynamics. In *Proc. 1986 Simul. Conf., Vol. 4*, pages 111–117, 1986.
8. J.L. Lions and Magenes E. *Problèmes aux limites non homogènes at applications vol. 1*. Travaux et recherches mathématiques. Dunod, 1968.
9. W. R. Madych. An estimate for multivariate interpolation. II. *J. Approx. Theory*, 142(2):116–128, 2006.
10. Francis J. Narcowich, Joseph D. Ward, and Holger Wendland. Sobolev error estimates and a Bernstein inequality for scattered data interpolation via radial basis functions. *Constr. Approx.*, 24(2):175–186, 2006.
11. R. Schaback. Convergence of unsymmetric kernel-based meshless collocation methods. *SIAM J. Numer. Anal.*, 45(1):333–351 (electronic), 2007.
12. R. Schaback. Recovery of functions from weak data using unsymmetric meshless kernel-based methods. *Applied Numerical Mathematics*, 58:726–741, 2007.
13. R. Schaback and H. Wendland. Kernel techniques: from machine learning to meshless methods. *Acta Numerica*, 15:543–639, 2006.
14. H. Wendland. *Scattered Data Approximation*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, Cambridge, 2004.
15. H. Wendland and C. Rieger. Approximate interpolation with applications to selecting smoothing parameters. *Numer. Math.*, 101:643–662, 2005.
16. Holger Wendland. Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Adv. Comput. Math.*, 4(4):389–396, 1995.