

## SOLVING PDEs WITH INCOMPLETE INFORMATION\*

PETER BINEV<sup>†</sup>, ANDREA BONITO<sup>‡</sup>, ALBERT COHEN<sup>§</sup>, WOLFGANG DAHMEN<sup>†</sup>,  
RONALD DEVORE<sup>‡</sup>, AND GUERGANA PETROVA<sup>‡</sup>

**Abstract.** We consider the problem of numerically approximating the solutions to a partial differential equation (PDE) when there is insufficient information to determine a unique solution. Our main example is the Poisson boundary value problem, when the boundary data is unknown and instead one observes finitely many linear measurements of the solution. We view this setting as an optimal recovery problem and develop theory and numerical algorithms for its solution. The main vehicle employed is the derivation and approximation of the Riesz representers of these functionals with respect to relevant Hilbert spaces of harmonic functions.

**Key words.** optimal learning, partial differential equations, incomplete information, harmonic functions, measurements, finite elements

**MSC code.** 65N99

**DOI.** 10.1137/23M1546671

**1. Introduction.** The questions we investigate sit in the broad research area of using measurements to enhance the numerical recovery of the solution  $u$  to a PDE. The particular setting addressed in this paper is to numerically approximate the solution to an elliptic boundary value problem when there is insufficient information on the boundary value to determine a unique solution to the PDE. In place of complete boundary information, we have a finite number of data observations of the solution  $u$ . This data serves to narrow the set of possible solutions. We ask what is the optimal accuracy to which we can recover  $u$  and what is a near optimal numerical algorithm to approximate  $u$ . Problems of this particular type arise in several fields of science and engineering (see, e.g., [30, 4, 9] for examples in fluid dynamics), where a lack of full information on boundary conditions arises for various reasons. For example, the correct physics might not be fully understood [24, 26], or the boundary values are not accessible [13], or they must be appropriately modified in numerical schemes [10, 25]. Other examples of application domains for the results of the present paper can be found in the introduction of [11].

**1.1. A model for PDEs with incomplete data.** This paper is concerned with recovering a function  $u$  that is known to solve a specific PDE while lacking some

---

\*Received by the editors January 12, 2023; accepted for publication (in revised form) December 19, 2023; published electronically May 30, 2024. The U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes. Copyright is owned by SIAM to the extent not limited by these rights.

<https://doi.org/10.1137/23M1546671>

**Funding:** The work of the first author was supported by the ARO grant W911NF2010318. The work of the first and fourth authors was supported by the National Science Foundation (NSF) grant DMS-2038080. The work of the second author was supported by the NSF grant DMS-2110811. The work of the fourth author was supported by the NSF grant DMS-2012469 and the German Research Foundation (DFG) grant SFB 1481. The work of the fifth and sixth authors was supported by the NSF grant DMS-21340077 and the MURI ONR grant N00014-20-1-278.

<sup>†</sup>Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA (binev@math.sc.edu, dahmen@math.sc.edu).

<sup>‡</sup>Department of Mathematics, Texas A&M University, College Station, TX 77843, USA (bonito@tamu.edu, ronalddevore@tamu.edu, gpetrova@tamu.edu).

<sup>§</sup>Laboratoire Jacques-Louis Lions, Sorbonne Université, 4, Place Jussieu, 75005 Paris, France (albert.cohen@sorbonne-universite.fr).

of the data that would determine  $u$  uniquely. Specifically, we consider the model elliptic problem

$$(1.1) \quad -\Delta u = f \quad \text{in } \Omega, \quad u = g \quad \text{on } \Gamma := \partial\Omega,$$

where  $\Omega \subset \mathbb{R}^d$  is a bounded Lipschitz domain with  $d = 2$  or  $3$ . The Lax–Milgram theorem [31] implies the existence and uniqueness of a solution  $u$  from the Sobolev space  $H^1(\Omega)$  to (1.1), once  $f$  and  $g$  are prescribed in  $H^{-1}(\Omega)$  (the dual of  $H_0^1(\Omega)$ ) and in  $H^{1/2}(\Gamma)$  (the image of  $H^1(\Omega)$  by the trace operator), respectively.

Recall that the trace operator  $T$  is defined on a function  $w \in C(\bar{\Omega})$  as the restriction of  $w$  to  $\Gamma$  and this definition is then generalized to functions in Sobolev spaces by a denseness argument. In particular, the trace operator is well defined on  $H^1(\Omega)$ . For any function  $v$  in  $H^1(\Omega)$  we denote by  $v_\Gamma$  its trace,

$$(1.2) \quad v_\Gamma := T(v) = v|_\Gamma, \quad v \in H^1(\Omega).$$

The Lax–Milgram analysis also yields the inequalities

$$(1.3) \quad c_0 \|v\|_{H^1(\Omega)} \leq \|\Delta v\|_{H^{-1}(\Omega)} + \|v_\Gamma\|_{H^{1/2}(\Gamma)} \leq c_1 \|v\|_{H^1(\Omega)}, \quad v \in H^1(\Omega).$$

Here the constants  $c_0, c_1$  depend on  $\Omega$  and on the particular choice of norms employed on  $H^1(\Omega)$  and  $H^{1/2}(\Gamma)$ .

Our interest centers on the question of how well we can numerically recover  $u$  in the  $H^1$  norm when we do not have sufficient knowledge to guarantee a unique solution to (1.1). There are many possible settings to which our techniques apply, but we shall focus on the following scenario:

- (i) We have a complete knowledge of  $f$  but we do not know  $g$ .
- (ii) The function  $g$  belongs to a known compact subset  $K^B$  of  $H^{\frac{1}{2}}(\Gamma)$ . Thus, membership in  $K^B$  describes our knowledge of the boundary data. The function  $u$  we wish to recover comes from the set

$$(1.4) \quad K := \{u : u \text{ solves (1.1) for some } g \in K^B\},$$

which is easily seen from (1.3) to be a compact subset of  $H^1(\Omega)$ .

- (iii) We have access to finitely many data observations of the unknown solution  $u$ , in terms of a vector

$$(1.5) \quad \lambda(u) := (\lambda_1(u), \dots, \lambda_m(u)) \in \mathbb{R}^m,$$

where the  $\lambda_j$  are fixed and known linear functionals defined on the functions from  $K$ .

Natural candidates for the compact set  $K^B$  are balls of Sobolev spaces that are compactly embedded in  $H^{\frac{1}{2}}(\Gamma)$ . We thus restrict our attention for the remainder of this paper to the case

$$(1.6) \quad K^B := U(H^s(\Gamma)),$$

for some  $s > \frac{1}{2}$ , where  $U(H^s(\Gamma))$  denotes the unit ball of a  $H^s(\Gamma)$  with respect to the norm  $\|\cdot\|_{H^s(\Gamma)}$ . The precise definition of  $H^s(\Gamma)$  and its norm  $\|\cdot\|_{H^s(\Gamma)}$  is described later. We have assumed that  $K^B$  is the unit ball only for convenience. The arguments given below hold in the case when  $K^B$  is a ball of  $H^s(\Gamma)$  centered at 0 of any radius  $R$ . The numerical algorithms proposed and analyzed do not require the knowledge of the radius  $R$  of that ball.

**1.2. The optimal recovery benchmark.** Let  $w_j := \lambda_j(u)$ ,  $j = 1, \dots, m$ , and let

$$(1.7) \quad w := (w_1, \dots, w_m) = \lambda(u) \in \mathbb{R}^m,$$

be the vector of data observations. Therefore, the totality of information we have about  $u$  is that it lies in the compact set

$$(1.8) \quad K_w := \{u \in K : \lambda(u) = w\}.$$

Our problem is to numerically find a function  $\hat{u} \in H^1(\Omega)$  which approximates simultaneously all the  $u \in K_w$ . This is a special case of the problem of optimal recovery from data (see [17, 29, 21]). The optimal recovery, i.e., the best choice for  $\hat{u}$ , has the following well known theoretical description. Let  $B(K_w)$  be a smallest ball in  $H^1(\Omega)$  which contains  $K_w$ , let  $R(K_w) := R(K_w)_{H^1(\Omega)}$  be its radius, and let  $\hat{u}(w)$  be its center. These are called the *Chebychev ball, radius, and center*, respectively. Note that the Chebyshev ball  $B(K_w)$  is unique in  $H^1(\Omega)$ ; see [6]. Then  $R(K_w)$  is the *optimal recovery error*, that is, the smallest error we can have for recovering  $u$  in the norm of  $H^1(\Omega)$ , and  $\hat{u}(w)$  is an *optimal recovery* of  $u$ .

We are interested in understanding how small  $R(K_w)$  is and what are the numerical algorithms which are *near optimal* in recovering  $u$  from the given data  $w$ . We say that an algorithm  $w \mapsto \hat{u} = \hat{u}(w)$  delivers *near optimal recovery* with constant  $C$  if

$$(1.9) \quad \|u - \hat{u}(w)\|_{H^1(\Omega)} \leq CR(K_w), \quad w \in \mathbb{R}^m.$$

Of course, we want  $C$  to be a reasonable constant independent of  $m$ . Our results actually deliver a recovery estimate of the form

$$(1.10) \quad \|u - \hat{u}(w)\|_{H^1(\Omega)} \leq R(K_w) + \varepsilon, \quad w \in \mathbb{R}^m,$$

where  $\varepsilon > 0$  can be made arbitrarily small at the price of higher computational cost. In this sense, the recovery is near optimal with constant  $C > 1$  in (1.9) that can be made arbitrarily close to 1.

**1.3. A connection with the recovery of harmonic functions.** There is a natural restatement of our recovery problem in terms of harmonic functions. Let  $f$  be the right side of (1.1), where  $f$  is a known fixed element of  $H^{-1}(\Omega)$ . Let  $u_0$  be the function in  $H^1(\Omega)$  which is the solution to (1.1) with  $g = 0$ . Then, we can write any function  $u \in K$  as

$$(1.11) \quad u = u_0 + u_{\mathcal{H}},$$

where  $u_{\mathcal{H}}$  is a harmonic function in  $H^1(\Omega)$  which has boundary value  $g = T(u_{\mathcal{H}})$  with  $g \in K^B$ . Recall our assumption that  $K^B$  is the unit ball of  $H^s(\Gamma)$  with  $s > \frac{1}{2}$ .

Let  $\mathcal{H}^s(\Omega)$  denote the set of harmonic functions  $v$  defined on  $\Omega$  for which  $v_{\Gamma} \in H^s(\Gamma)$ . We refer the reader to [3], where a detailed study of spaces like  $\mathcal{H}^s(\Omega)$  is presented. We define the norm on  $\mathcal{H}^s(\Omega)$  to be the one induced by the norm on  $H^s(\Gamma)$ , namely,

$$(1.12) \quad \|v\|_{\mathcal{H}^s(\Omega)} := \|v_{\Gamma}\|_{H^s(\Gamma)}, \quad v \in \mathcal{H}^s(\Omega).$$

There exist several equivalent definitions of norms on  $H^s(\Gamma)$ , as discussed later. For the moment, observe that from (1.3) it follows the existence of a constant  $C_s$  such that

$$(1.13) \quad \|v\|_{H^1(\Omega)} \leq C_s \|v\|_{\mathcal{H}^s(\Omega)}, \quad v \in \mathcal{H}^s(\Omega).$$

Indeed, the space  $\mathcal{H}^s(\Omega)$  is a Hilbert space that is compactly embedded in  $H^1(\Omega)$ , as a consequence of the compact embedding of  $H^s(\Gamma)$  in  $H^{1/2}(\Gamma)$ . We denote by  $K^{\mathcal{H}}$  the unit ball of  $\mathcal{H}^s(\Omega)$ ,

$$(1.14) \quad K^{\mathcal{H}} := U(\mathcal{H}^s(\Omega)).$$

Since the function  $u_0$  in (1.11) is fixed, it follows from (1.6) that

$$(1.15) \quad R(K_w) = R(K_{w'}^{\mathcal{H}})_{H^1(\Omega)}, \quad w' := \lambda(u_{\mathcal{H}}) = w - \lambda(u_0).$$

There are two conclusions that can be garnered from this reformulation. The first is that the optimal error in recovering  $u \in K_w$  is the same as that in recovering the harmonic function  $u_{\mathcal{H}} \in K_{w'}^{\mathcal{H}}$  in the  $H^1(\Omega)$  norm. The harmonic recovery problem does not involve  $f$  except in determining  $w'$ . The second point is that one possible numerical algorithm for our original problem is to first construct a sufficiently accurate approximation  $\hat{u}_0$  to  $u_0$  and then to numerically implement an optimal recovery of a harmonic function in  $K^{\mathcal{H}}$  from data observations. This numerical approach requires the computation of  $w'$ . In theory,  $u_0$  is known to us since we have a complete knowledge of  $f$ . However,  $u_0$  must be computed and any approximation  $\hat{u}_0$  will induce an error. Although this error can be made arbitrarily small, it means that we only know  $w'$  up to a certain numerical accuracy. One can thus view the harmonic reformulation as an optimal recovery problem with perturbed observations of  $w'$ . The numerical algorithm presented here follows this approach. Its central constituent, namely the recovery of harmonic functions from a finite number of noisy observations, can be readily employed as well in a number of different application scenarios described, e.g., in [11].

**1.4. Optimal recovery from centrally symmetric sets in a Hilbert space.** As noted above, the problem of recovering  $u \in K_w$  is directly related to the problem of recovering the harmonic component  $u_{\mathcal{H}} \in K^{\mathcal{H}}$  from the given data observations  $w'$ . Note that  $K^{\mathcal{H}}$  is the unit ball of the Hilbert space  $\mathcal{H}^s(\Omega)$ . There is a general approach for optimal recovery from data observations in this Hilbert space setting, as discussed, e.g., in [17, 16]. We first recollect the general principles of this technique which will be applied in section 3.1 to our specific setting.

Let  $\mathcal{H}$  be any Hilbert space and suppose that  $\lambda_1, \dots, \lambda_m \in \mathcal{H}^*$  are linearly independent functionals from  $\mathcal{H}^*$ . Let  $X$  be a Banach space such that  $\mathcal{H}$  is continuously embedded in  $X$ . We are interested in optimal recovery of a function  $v$  in the norm  $\|\cdot\|_X$ , knowing that  $v \in \mathcal{K} := U(\mathcal{H})$ , the unit ball of  $\mathcal{H}$ . If  $w \in \mathbb{R}^m$  is the vector of observations, we define the *minimal norm interpolant* as

$$(1.16) \quad v^*(w) = \operatorname{argmin}\{\|v\|_{\mathcal{H}} : v \in \mathcal{H} \text{ and } \lambda(v) = w\}.$$

*Remark 1.1.* If  $\mathcal{K}_w := \{u \in \mathcal{K} : \lambda(u) = w\}$  is nonempty, the minimum norm interpolant  $v^*(w)$  is a Chebyshev center of  $\mathcal{K}_w$  in  $X$ . That is, the minimal norm interpolant gives optimal recovery with constant  $C = 1$ . In other words, this Chebyshev center does not depend on  $X$ . The radius  $R(\mathcal{K}_w)_X$  will, however, generally depend on  $X$ .

To prove this remark, first note that any  $v \in \mathcal{K}_w$  may be written as  $v = v^*(w) + \eta$  where  $\eta$  belongs to the null space  $\mathcal{N}$  of  $\lambda$ . Because  $v^*(w)$  has minimal norm,  $\eta$  is orthogonal to  $v^*(w)$  and hence from the Pythagorean theorem

$$\|v - v^*(w)\|_{\mathcal{H}}^2 = \|v\|_{\mathcal{H}}^2 - \|v^*(w)\|_{\mathcal{H}}^2 \leq 1 - \|v^*(w)\|_{\mathcal{H}}^2 =: r^2,$$

because  $\|v\|_{\mathcal{H}} \leq 1$ . Notice that  $v^*(w) - \eta$  is also in  $\mathcal{K}_w$ . It follows that  $\mathcal{K}_w$  is precisely the ball in the affine space  $v^*(w) + \mathcal{N}$  centered at  $v^*(w)$  and of radius  $r$ . In particular,  $\mathcal{K}_w$  is centrally symmetric around  $v^*(w)$  and we now show that  $v^*(w)$  is a Chebyshev center of a Chebyshev ball in  $X$ . Recall that unless  $X$  is uniformly convex, these quantities might not be uniquely defined [6]. Let  $u \in \mathcal{K}_w$  be the furthest away from  $v^*(w)$ , i.e.,

$$u \in \arg \max_{v \in \mathcal{K}_w} \|v^*(w) - v\|_X,$$

and set  $\delta := \|v^*(w) - u\|_X$ . Since  $\mathcal{K}_w$  is centrally symmetric around  $v^*(w)$ ,  $u' := 2v^*(w) - u \in \mathcal{K}_w$  and  $\|u - u'\|_X = 2\delta$ . For any  $\bar{v} \in X$  with  $\bar{v} \neq v^*(w)$ , the triangle inequality yields

$$2\delta = \|u - u'\|_X \leq \|u - \bar{v}\|_X + \|u' - \bar{v}\|_X \leq 2 \max_{v \in \mathcal{K}_w} \|\bar{v} - v\|_X,$$

and so

$$\max_{v \in \mathcal{K}_w} \|v^*(w) - v\|_X \leq \max_{v \in \mathcal{K}_w} \|\bar{v} - v\|_X,$$

which shows that  $v^*(w)$  is indeed a Chebyshev center in  $X$ . In particular,

$$\|v - v^*(w)\|_X \leq R(\mathcal{K}_w)_X, \quad v \in \mathcal{K}_w.$$

Standard Hilbert space analysis shows that the mapping  $w \mapsto v^*(w)$  is a linear operator. More importantly, it has a natural expression that is useful for numerical computation. Namely, from the Riesz representation theorem each  $\lambda_j$  can be described as

$$\lambda_j(v) = \langle v, \phi_j \rangle_{\mathcal{H}}, \quad v \in \mathcal{H},$$

where  $\phi_j \in \mathcal{H}$  is called the Riesz representer of  $\lambda_j$ . The minimal norm interpolant has the representation

$$(1.17) \quad v^* = \sum_{j=1}^m a_j^* \phi_j,$$

where  $a^* = (a_1^*, \dots, a_m^*)$  solves the system of equations

$$G a^* = w, \quad G := (\langle \phi_i, \phi_j \rangle_{\mathcal{H}})_{i,j=1,\dots,m},$$

with  $G$  being the Gramian matrix associated to  $\phi_1, \dots, \phi_m$ .

*Remark 1.2.* Note that  $v^*(w)$  is exactly the  $\mathcal{H}$ -orthogonal projection of  $w$  onto the space spanned by the  $m$  Riesz representers.

*Remark 1.3.* In the case where  $\mathcal{H}$  is a more general Banach space, we are still ensured that the minimal norm interpolation is a near-optimal recovery with constant  $C = 2$ . However, its dependence on the data  $w$  is no longer linear and the above observation regarding its computation does not apply.

Our proposed numerical recovery scheme is based on approximately realizing (1.17).

**1.5. Objectives and outline.** The main goal of this paper is to create numerical algorithms which are guaranteed to produce a near optimal recovery  $\hat{u}$  from the given data  $w$  and to analyze their practical implementation. We begin in section 2 with some remarks on the definition of the space  $H^s(\Gamma)$  and its norm, which are of importance both in the accuracy analysis and the practical implementation of recovery algorithms.

Next, we turn to the description of our numerical algorithms. The algorithms we propose and analyze are based on the general approach for optimal recovery described in section 1.4 when this approach is applied to our particular PDE setting. We describe a solution algorithm which takes into consideration the effect of numerical perturbations. We first consider the case when the linear functionals  $\lambda_j$  are defined on all of  $H^1(\Omega)$  and then adapt this algorithm to the case when the linear functionals are point evaluations

$$(1.18) \quad \lambda_j(u) := u(x_j), \quad x_j \in \bar{\Omega}, \quad j = 1, \dots, m.$$

Point evaluations are not defined on all of  $H^1(\Omega)$  when  $d > 1$ ; however, they are defined on  $K$  when the smoothness order  $s$  is large enough.

The critical ingredient in our proposed algorithm is the numerical computation of the Riesz representers  $\phi_j$  of the restrictions of  $\lambda_j$  to the Hilbert space  $\mathcal{H}^s(\Omega)$ . Each of these Riesz representers is characterized as a solution to an elliptic problem and can be computed offline since it does not involve the measurement vector  $w$ . Our suggested numerical method for approximating  $\phi_j$  is based on finite element discretizations and is discussed in section 4. We establish quantitative error bounds for the numerical approximation in terms of the mesh size. Numerical illustrations of the optimal recovery algorithm are given in section 5.

Note that the optimal recovery error over the class  $K$  strongly depends on the choice of the linear functionals  $\lambda_j$ . For example, in the case of point evaluation, this error can be very large if the data sites  $\{x_j\}_{j=1}^m$  are poorly positioned, or small if they are optimally positioned. This points to the importance of the Gelfand widths and sampling numbers. They describe the optimal recovery error over  $K$  with optimal choice of functionals in the general case and the point evaluation case, respectively. The numerical behavior of these quantities in our specific setting is discussed in section 6.

**2. The spaces  $H^s(\Gamma)$  and  $\mathcal{H}^s(\Omega)$ .** In this section, we discuss the definition and basic properties of the spaces  $H^s(\Gamma)$  and  $\mathcal{H}^s(\Omega)$ . We refer to [1] for a general treatment of Sobolev spaces on domains  $D \subset \mathbb{R}^d$ . Recall that for fractional orders  $r > 0$ , the norm of  $H^r(D)$  is defined as

$$\|v\|_{H^r(D)}^2 := \|v\|_{H^k(D)}^2 + \sum_{|\alpha|=k} \int_{D \times D} \frac{|\partial^\alpha v(x) - \partial^\alpha v(y)|^2}{|x-y|^{d+2(r-k)}} dx dy,$$

where  $k$  is the integer such that  $k < r < k+1$ , and  $\|v\|_{H^k(D)}^2 := \sum_{|\alpha| \leq k} \|\partial^\alpha v\|_{L^2(D)}^2$  is the standard  $H^k$ -norm.

**2.1. Equivalent definitions of  $H^s(\Gamma)$ .** Let  $\Omega$  be any bounded Lipschitz domain in  $\mathbb{R}^d$ . We recall the trace operator  $T$  introduced in section 1.1. One first possible definition of the space  $H^s(\Gamma)$ , for any  $s \geq \frac{1}{2}$ , is as the restriction of  $H^{s+\frac{1}{2}}(\Omega)$  to  $\Gamma$ , that is,

$$H^s(\Gamma) = T(H^{s+\frac{1}{2}}(\Omega)),$$

with norm

$$(2.1) \quad \|g\|_{H^s(\Gamma)} := \min \left\{ \|v\|_{H^{s+\frac{1}{2}}(\Omega)} : v_\Gamma = g \right\}.$$

The resulting norm is referred to as the trace norm definition for  $H^s(\Gamma)$ .

There is a second, more intrinsic way to define  $H^s(\Gamma)$ , by properly adapting the notion of Sobolev smoothness to the boundary. This can be done by locally mapping the boundary onto domains of  $\mathbb{R}^{d-1}$  and requiring that the pullback of  $g$  by such transformation have  $H^s$  smoothness on such domains. We refer the reader to [12] and [19] for the complete intrinsic definition, where it is proved to be equivalent to the trace definition for a range of  $s$  that depends on the smoothness of the boundary  $\Gamma$ .

For small values of  $s$ , Sobolev norms for  $H^s(\Gamma)$  may also be equivalently defined without the help of local parameterizations, as contour integrals. For example, if  $0 < s < 1$  and  $\Omega$  is a Lipschitz domain, we define

$$\|g\|_{H^s(\Gamma)}^2 := \|g\|_{L^2(\Gamma)}^2 + \int_{\Gamma \times \Gamma} \frac{|g(x) - g(y)|^2}{|x - y|^{d-1+2s}} dx dy,$$

and if  $s = 1$  and  $\Omega$  is a polygonal domain, we define

$$(2.2) \quad \|g\|_{H^1(\Gamma)}^2 := \|g\|_{L^2(\Gamma)}^2 + \|\nabla_\Gamma g\|_{L^2(\Gamma)}^2,$$

where  $\nabla_\Gamma$  is the tangential gradient, and likewise

$$\|g\|_{H^s(\Gamma)}^2 := \|g\|_{H^1(\Gamma)}^2 + \int_{\Gamma \times \Gamma} \frac{|\nabla_\Gamma g(x) - \nabla_\Gamma g(y)|^2}{|x - y|^{d-1+2(s-1)}} dx dy$$

for  $1 < s < 2$ . In the numerical illustration given in section 5, we will specifically take the value  $s = 1$  and a square domain using the definition (2.2).

When  $\Omega$  has smooth boundary, it is known that the trace definition and intrinsic definition of the  $H^s(\Gamma)$  norms are equivalent for all  $s \geq 1/2$ . On the other hand, when  $\Omega$  does not have a smooth boundary, it is easily seen that the two definitions are not equivalent unless restrictions are made on  $s$ . Consider, for example, the case of polygonal domains of  $\mathbb{R}^2$ : it is easily seen that the trace  $v_\Gamma$  of a smooth function  $v \in C^\infty(\bar{\Omega})$  has a tangential gradient  $\nabla_\Gamma v_\Gamma$  that generally has jump discontinuities at the corner points and thus does not belong to  $H^{1/2}(\Gamma)$ . In turn, the equivalence between the trace and intrinsic norms only holds for  $s < \frac{3}{2}$  and in such case we limit the value of  $s$  to this range. The same restriction  $s < 3/2$  applies to a polyhedral domain in the case  $d = 3$ .

**2.2. The regularity of functions in  $\mathcal{H}^s(\Omega)$ .** We next give some remarks on the Sobolev smoothness of functions from the space  $\mathcal{H}^s(\Omega)$  when  $s > 1/2$ . Clearly, such harmonic functions are infinitely smooth inside  $\Omega$  and also belong to  $H^1(\Omega)$ , but one would like to know for which value of  $r$  they belong to  $H^r(\Omega)$ . To answer this question, we consider  $v \in \mathcal{H}^s(\Omega)$ . By the definition of  $\mathcal{H}^s(\Omega)$ ,  $v$  is harmonic in  $\Omega$  and  $v_\Gamma \in H^s(\Gamma)$ . Having assumed that  $s$  is in the admissible range where all above definitions of the  $H^s(\Gamma)$  norms are equivalent, and using the first one, we know that there exists a function  $\tilde{v} \in H^{s+\frac{1}{2}}(\Omega)$  such that  $\tilde{v}_\Gamma = v_\Gamma$

$$\|\tilde{v}\|_{H^{s+\frac{1}{2}}(\Omega)} = \|v_\Gamma\|_{H^s(\Gamma)} = \|v\|_{\mathcal{H}^s(\Omega)}.$$

We define  $\bar{v} := v - \tilde{v}$  so that  $v = \tilde{v} + \bar{v}$ . We are interested in the regularity of  $\bar{v}$  since it will give the regularity of  $v$ . Notice that  $\bar{v}_\Gamma = 0$  and

$$-\Delta \bar{v} = \bar{f} := \Delta \tilde{v}.$$

The function  $\bar{f}$  belongs to the Sobolev space  $H^{s-\frac{3}{2}}(\Omega)$  and we are left with the classical question of the regularizing effect in Sobolev scales when solving the Laplace equation with Dirichlet boundary conditions. Obviously, when  $\Omega$  is smooth, we find that  $\bar{v} \in H^{s+\frac{1}{2}}(\Omega)$  and so we have obtained the continuous embedding

$$\mathcal{H}^s(\Omega) \subset H^r(\Omega), \quad r = s + \frac{1}{2}.$$

For less smooth domains, the smoothing effect is limited (in particular, by the presence of singularities on the boundary of  $\Omega$ ), i.e.,  $\bar{v}$  is only guaranteed to be in  $H^r(\Omega)$  where  $r$  may be less than  $s + 1/2$ ; see [12]. More precisely

$$\mathcal{H}^s(\Omega) \subset H^r(\Omega),$$

where

$$(2.3) \quad r := \min \left\{ s + \frac{1}{2}, r^* \right\}.$$

Here,  $r^* = r^*(\Omega)$  is the limiting bound for the smoothing effect:

- (i) For smooth domains  $r^* = \infty$ .
- (ii) For convex domains  $r^* = 2$ .
- (iii) For nonconvex polygonal domains in  $\mathbb{R}^2$ , or a polyhedron in  $\mathbb{R}^3$ , one has  $3/2 < r^* < 2$  where the value of  $r^*$  depends on the re-entrant angles.
- (iv) In particular for polygons, we can take  $r^* = 1 + \frac{\pi}{\omega} - \varepsilon$  for any  $\varepsilon > 0$  where  $\omega$  is the largest inner angle.

Note that  $r^*$  could be strictly smaller than  $s + \frac{1}{2}$ .

In summary, for an admissible range of  $r > 1$  that depends on  $s$  and  $\Omega$  one has the continuous embedding  $\mathcal{H}^s(\Omega) \subset H^r(\Omega)$ , and so there exists a constant  $C_1$  that depends on  $(r, s)$  and  $\Omega$ , such that

$$(2.4) \quad \|v\|_{H^r(\Omega)} \leq C_1 \|v\|_{\mathcal{H}^s(\Omega)} = C_1 \|v_\Gamma\|_{H^s(\Gamma)}, \quad v \in \mathcal{H}^s(\Omega).$$

**3. A near optimal recovery algorithm.** In this section, we present a numerical algorithm for solving (1.1) when the information about the boundary value  $g$  is incomplete. We first work under the assumption that the  $\lambda_j$ 's are continuous over  $H^1(\Omega)$ , and assumed to be linearly independent (linear independence can be guaranteed by throwing away dependent functionals when necessary). We prove that the proposed numerical recovery algorithm is near optimal. We then adapt our approach to the case where the  $\lambda_j$ 's are point evaluations (see (1.18)), and therefore not continuous over  $H^1(\Omega)$  when  $d \geq 2$ .

**3.1. Minimum norm data fitting.** We now apply the general optimal recovery principles discussed in section 1.4 to our specific setting in which

$$\mathcal{H} = \mathcal{H}^s(\Omega) \quad \text{and} \quad X = H^1(\Omega).$$

Let  $\phi_j \in \mathcal{H}^s(\Omega)$  be the Riesz representer of the functional  $\lambda_j$  when viewed as a functional on  $\mathcal{H}^s(\Omega)$ . In other words

$$\lambda_j(v) = \langle v, \phi_j \rangle_{\mathcal{H}^s(\Omega)}, \quad v \in \mathcal{H}^s(\Omega).$$

We assume that the  $\lambda_j$  are linearly independent on  $\mathcal{H}^s(\Omega)$  and thus the Gramian matrix

$$G = (g_{i,j})_{i,j=1,\dots,m}, \quad g_{i,j} := \langle \phi_j, \phi_i \rangle_{\mathcal{H}^s(\Omega)} = \lambda_i(\phi_j),$$

is invertible.

Now, let  $u = u_0 + u_{\mathcal{H}}$ , with  $u_{\mathcal{H}} \in K^{\mathcal{H}} = U(\mathcal{H}^s(\Omega))$  be the function in  $K$  that gave rise to our data observation  $w$ . So, we have

$$w' = w - \lambda(u_0) = \lambda(u_{\mathcal{H}}).$$

If  $a^*$  is the vector in  $\mathbb{R}^m$  which satisfies  $Ga^* = w'$ , then  $u_{\mathcal{H}}^* := \sum_{j=1}^m a_j^* \phi_j$  is the function of minimum  $\mathcal{H}^s(\Omega)$  norm which satisfies the data  $w'$ , i.e.,  $\lambda(u_{\mathcal{H}}^*) = w'$ . We have seen that

$$\|u_{\mathcal{H}} - u_{\mathcal{H}}^*\|_{H^1(\Omega)} \leq R(K_{w'}^{\mathcal{H}})_{H^1(\Omega)},$$

namely,  $u_{\mathcal{H}}^*$  is the optimal recovery of the functions in  $K_{w'}^{\mathcal{H}}$ . Note that the recovery error is measured in  $H^1$  not in  $\mathcal{H}^s(\Omega)$ . In turn (see (1.15)), the function  $u^* := u_{\mathcal{H}}^* + u_0$  is the optimal recovery for functions in  $K_w$ :

$$\|u - u^*\|_{H^1(\Omega)} \leq R(K_w)_{H^1(\Omega)}.$$

The idea behind our proposed numerical method is to numerically construct a function  $\hat{u} \in H^1(\Omega)$  that approximates  $u^*$  well. If, for example, we have for  $\varepsilon > 0$  the bound

$$\|u^* - \hat{u}\|_{H^1(\Omega)} \leq \varepsilon,$$

then for any  $u \in K$ , we have by the triangle inequality

$$\|u - \hat{u}\|_{H^1(\Omega)} \leq R(K_w)_{H^1(\Omega)} + \varepsilon.$$

Given any  $C > 1$ , by taking  $\varepsilon$  small enough, we have that  $\hat{u}$  is a near best recovery of the functions in  $K_w$  with constant  $C$ .

**3.2. The numerical recovery algorithm for  $H^1$ -continuous functionals.** Motivated by the above analysis, we propose the following numerical algorithm for solving our recovery problem. The algorithm involves approximations of the function  $u_0$  and the Riesz representers  $\phi_j$ , typically computed by finite element discretizations, and the application of the linear functionals  $\lambda_j$  to these approximations. In order to avoid extra technicalities, here we make the assumption that the applications of the functionals to a known finite element function can be exactly computed.

We first work under the additional assumption that the linear functionals  $\lambda_j$  are not only defined on  $K$  but that they are continuous over  $H^1(\Omega)$ . We define  $\Lambda$  as the maximum of the norms of the  $\lambda_j$  on  $H^1(\Omega)$ . In this case

$$(3.1) \quad |\lambda_j(v)| \leq \Lambda \|v\|_{H^1(\Omega)}, \quad v \in H^1(\Omega).$$

In what follows, throughout this paper, we use the following weighted  $\ell_2$  norm on  $\mathbb{R}^m$ :

$$\|z\| := \left( \frac{1}{m} \sum_{j=1}^m |z_j|^2 \right)^{1/2} = m^{-1/2} \|z\|_{\ell_2}, \quad z = (z_1, \dots, z_m) \in \mathbb{R}^m.$$

In particular, we have

$$\|\lambda(v)\| \leq \Lambda \|v\|_{H^1(\Omega)}, \quad v \in H^1(\Omega).$$

Given a user prescribed accuracy  $\varepsilon > 0$ , our algorithm does the following four steps involving intermediate tolerances  $(\varepsilon_1, \varepsilon_2)$ .

*Step 1.* We numerically find an approximation  $\hat{u}_0$  to  $u_0$  which satisfies

$$(3.2) \quad \|u_0 - \hat{u}_0\|_{H^1(\Omega)} \leq \varepsilon_1.$$

To find such a  $\hat{u}_0$ , we use standard or adaptive FEM methods. Given that  $\hat{u}_0$  has been constructed, we define  $\hat{w} := w - \lambda(\hat{u}_0)$ . Then, for  $w' := w - \lambda(u_0)$  we have (see (3.2))

$$(3.3) \quad \|w' - \hat{w}\| \leq \Lambda \varepsilon_1.$$

On the other hand, since  $|\lambda_j(v)| \leq \Lambda \|v\|_{H^1(\Omega)} \leq \Lambda_s \|v\|_{\mathcal{H}^s(\Omega)} \leq \Lambda_s$ , where

$$\Lambda_s := C_s \Lambda$$

(see (3.1), (1.13), and (1.14)), we derive that

$$(3.4) \quad \|w'\| \leq \Lambda_s.$$

Thus by triangle inequality, we also find that

$$(3.5) \quad \|\hat{w}\| \leq \Lambda_s + \Lambda \varepsilon_1.$$

*Step 2.* For each  $j = 1, \dots, m$ , we numerically compute an approximation  $\hat{\phi}_j \in H^1(\Omega)$  to  $\phi_j$  which satisfies

$$(3.6) \quad \|\phi_j - \hat{\phi}_j\|_{H^1(\Omega)} \leq \varepsilon_2, \quad j = 1, \dots, m.$$

This numerical computation is crucial and is performed during the offline phase of the algorithm. We detail it in section 4. Note that the  $\hat{\phi}_j$ 's are not assumed to be in  $\mathcal{H}^s(\Omega)$ , and, in particular, not assumed to be harmonic functions.

*Step 3.* We define and compute the matrix

$$\hat{G} = (\hat{g}_{i,j})_{i,j=1,\dots,m}, \quad \hat{g}_{i,j} := \lambda_i(\hat{\phi}_j),$$

and thus  $|\hat{g}_{i,j} - g_{i,j}| \leq \Lambda \varepsilon_2$  for all  $i, j$ .

It follows that for the matrix  $R := G - \hat{G}$  we have

$$\|R\|_1 \leq m \Lambda \varepsilon_2,$$

where we use the shorthand notation  $\|\cdot\|_1 := \|\cdot\|_{\ell_1 \rightarrow \ell_1}$  for matrices. Since  $G$  is invertible, we are ensured that  $\hat{G}$  is also invertible for  $\varepsilon_2$  small enough. We define

$$M := \|G^{-1}\|_1, \quad \hat{M} := \|\hat{G}^{-1}\|_1.$$

While these two norms are finite, their size will depend on the nature and the positioning of the linear functionals  $\lambda_j$ ,  $j = 1, \dots, m$ , as it will be seen in the section on numerical experiments. These two numbers are close to one another when  $\varepsilon_2$  is small

since  $\hat{M}$  converges towards the unknown quantity  $M$  as  $\varepsilon_2 \rightarrow 0$ . In particular, we have

$$|M - \hat{M}| = ||G^{-1}||_1 - \|\hat{G}^{-1}\|_1 \leq \|G^{-1} - \hat{G}^{-1}\|_1 = \|\hat{G}^{-1}RG^{-1}\|_1 \leq M\hat{M}m\Lambda\varepsilon_2,$$

from which we obtain that

$$(3.7) \quad M \leq \frac{\hat{M}}{1 - m\hat{M}\Lambda\varepsilon_2} \quad \text{and} \quad \hat{M} \leq \frac{M}{1 - mM\Lambda\varepsilon_2},$$

provided that  $mM\Lambda\varepsilon_2 < 1$  and  $m\hat{M}\Lambda\varepsilon_2 < 1$ . We also have the bound

$$(3.8) \quad \|\hat{G}^{-1} - G^{-1}\|_1 \leq \frac{M^2}{1 - mM\Lambda\varepsilon_2}m\Lambda\varepsilon_2 =: \delta.$$

It is important to observe that  $\delta$  can be made arbitrarily small by diminishing  $\varepsilon_2$ .

*Step 4.* We numerically solve the  $m \times m$  algebraic system  $\hat{G}\hat{a} = \hat{w}$ , thereby finding a vector  $\hat{a} = (\hat{a}_1, \dots, \hat{a}_m)$ . We then define  $\hat{u}_{\mathcal{H}} := \sum_{j=1}^m \hat{a}_j \hat{\phi}_j$  and our recovery of  $u$  is  $\hat{u} := \hat{u}_0 + \hat{u}_{\mathcal{H}}$ . This step can be implemented by standard linear algebra solvers.

One major advantage of the above algorithm is that Steps 1–3 can be performed offline since they do not involve the data  $w$ . That is, we can compute  $\hat{u}_0$ , the approximate Riesz representers  $\hat{\phi}_j$  and the approximate Gramian  $\hat{G}$  and its inverse without knowing  $w$ . In this way, the computation of  $\hat{u}$  from given data  $w$  can be done fast online by Step 4 which only involves solving an  $m \times m$  linear system. This may be a significant advantage, for example, when having to process a large number of measurements for the same set of sensors.

**3.3. A near optimal recovery bound.** The following theorem shows that a near optimal recovery of  $u$  can be reached provided that the tolerances in the above described algorithm are chosen small enough.

**THEOREM 3.1.** *For any prescribed  $\varepsilon > 0$ , if the tolerances  $(\varepsilon_1, \varepsilon_2)$  are small enough such that  $mM\Lambda\varepsilon_2 < 1$  and*

$$(3.9) \quad \varepsilon_1 + mM\Lambda_s\varepsilon_2 + (C_0 + \varepsilon_2)(mM\Lambda\varepsilon_1 + m(\Lambda_s + \Lambda\varepsilon_1)\delta) \leq \varepsilon,$$

*where  $C_0 := \max_{j=1, \dots, m} \|\phi_j\|_{H^1(\Omega)}$  and  $\delta = \frac{M^2}{1 - mM\Lambda\varepsilon_2}m\Lambda\varepsilon_2$ , then the function  $\hat{u}$  generated by the above algorithm satisfies*

$$\|u - \hat{u}\|_{H^1(\Omega)} \leq R(K_w)_{H^1(\Omega)} + \varepsilon \quad \text{for every } u \in K_w.$$

*Thus, for any  $C > 1$  it is a near optimal recovery of  $u$  with constant  $C$  provided  $\varepsilon$  is taken sufficiently small.*

*Proof.* Let  $u = u_0 + v$  be our target function in  $K_w$ . We define  $w' = w - \lambda(u_0)$  and  $v^* := v^*(w')$  which is the Chebyshev center of  $K_{w'}^{\mathcal{H}}$ . We recall the algebraic system  $Ga^* = w'$  associated to the characterization of  $v^*$  (see (1.17)). We write

$$(3.10) \quad \begin{aligned} \|u_{\mathcal{H}}^* - \hat{u}_{\mathcal{H}}\|_{H^1(\Omega)} &\leq \left\| \sum_{j=1}^m a_j^*(\phi_j - \hat{\phi}_j) \right\|_{H^1(\Omega)} + \left\| \sum_{j=1}^m (a_j^* - \hat{a}_j) \hat{\phi}_j \right\|_{H^1(\Omega)} \\ &\leq \|a^*\|_{\ell_1} \varepsilon_2 + \|a^* - \hat{a}\|_{\ell_1} (C_0 + \varepsilon_2), \end{aligned}$$

where we have used (3.6) and the fact that

$$\|\hat{\phi}_j\|_{H^1(\Omega)} \leq \|\phi_j\|_{H^1(\Omega)} + \|\phi_j - \hat{\phi}_j\|_{H^1(\Omega)} \leq C_0 + \varepsilon_2.$$

Note that

$$(3.11) \quad \|a^*\|_{\ell_1} = \|G^{-1}w'\|_{\ell_1} \leq M\|w'\|_{\ell_1} \leq Mm\|w'\| \leq mM\Lambda_s,$$

where we have used that  $\|w'\|_{\ell_1} \leq m\|w'\|$  and inequality (3.4). Therefore, it follows from (3.10) and (3.11) that

$$(3.12) \quad \|u_{\mathcal{H}}^* - \hat{u}_{\mathcal{H}}\|_{H^1(\Omega)} \leq mM\Lambda_s \varepsilon_2 + \|a^* - \hat{a}\|_{\ell_1} (C_0 + \varepsilon_2).$$

For the estimation of  $\|a^* - \hat{a}\|_{\ell_1}$ , we introduce the intermediate vector  $\tilde{a} \in \mathbb{R}^m$ , which is the solution to the system  $G\tilde{a} = \hat{w}$ . Clearly,

$$\|\tilde{a} - a^*\|_{\ell_1} = \|G^{-1}(\hat{w} - w')\|_{\ell_1} \leq M\|\hat{w} - w'\|_{\ell_1} \leq Mm\|\hat{w} - w'\| \leq mM\Lambda\varepsilon_1,$$

where we invoked (3.3). On the other hand, in view of (3.8) and (3.5), we have

$$\|\tilde{a} - \hat{a}\|_{\ell_1} = \|(G^{-1} - \hat{G}^{-1})\hat{w}\|_{\ell_1} \leq \delta\|\hat{w}\|_{\ell_1} \leq m\delta\|\hat{w}\| \leq m(\Lambda_s + \Lambda\varepsilon_1)\delta.$$

Combining these two estimates, we find that

$$\|a^* - \hat{a}\|_{\ell_1} \leq mM\Lambda\varepsilon_1 + m(\Lambda_s + \Lambda\varepsilon_1)\delta.$$

We now insert this bound into (3.12) to obtain

$$\|u_{\mathcal{H}}^* - \hat{u}_{\mathcal{H}}\|_{H^1(\Omega)} \leq mM\Lambda_s \varepsilon_2 + (C_0 + \varepsilon_2)(mM\Lambda\varepsilon_1 + m(\Lambda_s + \Lambda\varepsilon_1)\delta).$$

Thus, for  $u^* := u_0 + u_{\mathcal{H}}^*$  and using (3.2), we have

$$(3.13) \quad \begin{aligned} \|u^* - \hat{u}\|_{H^1(\Omega)} &\leq \|u_0 - \hat{u}_0\|_{H^1(\Omega)} + \|u_{\mathcal{H}}^* - \hat{u}_{\mathcal{H}}\|_{H^1(\Omega)} \\ &\leq \varepsilon_1 + mM\Lambda_s \varepsilon_2 + (C_0 + \varepsilon_2)(mM\Lambda\varepsilon_1 + m(\Lambda_s + \Lambda\varepsilon_1)\delta) \leq \varepsilon. \end{aligned}$$

Since  $u = u_0 + u_{\mathcal{H}}$ , we have

$$\|u - u^*\|_{H^1(\Omega)} = \|u_{\mathcal{H}} - u_{\mathcal{H}}^*\|_{H^1(\Omega)} \leq R(K_{w'}^{\mathcal{H}})_{H^1(\Omega)} = R(K_w)_{H^1(\Omega)},$$

and the statement of the theorem follows from this inequality and (3.13).  $\square$

*Remark 3.2.* We point out that  $M$ , the norm of the matrix  $G^{-1}$ , grows potentially fast as  $m$  gets larger indicating that the Riesz representers become closer to be linearly dependent. This leads to numerical difficulties when computing the estimator which, as noted in Remark 1.2, amounts to the  $\mathcal{H}$ -orthogonal projection onto the space spanned by the representers. Therefore, regularization strategies might also be needed to compete with the ill-conditioning of  $G$  in the asymptotic regime. One typical strategy is to set to 0 the smallest eigenvalues of  $G$  according to a given threshold and apply the pseudo-inverse. We refer, for example, to [2] where this regularized projection strategy is reviewed and analyzed in detail.

*Remark 3.3.* Note that in numerical computations the quantity  $\hat{M}$  is available while  $M$  is unknown. Thus, in practice, in order to achieve the prescribed accuracy  $\varepsilon$ , we can first impose that  $\varepsilon_2 < (2m\hat{M}\Lambda)^{-1}$  and derive the inequalities (see (3.7))

$$M \leq \frac{\hat{M}}{1 - m\hat{M}\Lambda\varepsilon_2} \leq 2\hat{M}, \quad \|G^{-1} - \hat{G}^{-1}\|_1 \leq \frac{\hat{M}^2}{1 - m\hat{M}\Lambda\varepsilon_2} m\Lambda\varepsilon_2 \leq 2\hat{M}^2 m\Lambda\varepsilon_2 =: \hat{\delta},$$

where the last inequality is proven in a similar fashion to (3.8). If we then follow the proof of Theorem 3.1, the requirement in (3.13) can be substituted by

$$\varepsilon_1 + 2m\hat{M}\Lambda_s\varepsilon_2 + (C_0 + \varepsilon_2)(2m\hat{M}\Lambda\varepsilon_1 + m(\Lambda_s + \Lambda\varepsilon_1)\hat{\delta}) \leq \varepsilon,$$

and thus all participating quantities are computable.

*Remark 3.4.* The result in Theorem 3.1 can easily be extended to the case of noisy data, that is, to the case when the observations

$$\tilde{w} = w + \eta,$$

where  $\eta$  is a noise vector of norm  $\|\eta\| \leq \kappa$ . Indeed, the application of the algorithm to this noisy data leads to finding in Step 1 the vector  $\hat{w} := w + \eta - \lambda(\hat{u}_0)$  that satisfies

$$\|w' - \hat{w}\| \leq \Lambda\varepsilon_1 + \kappa \quad \text{and} \quad \|\hat{w}\| \leq \Lambda_s + \varepsilon_1\Lambda + \kappa,$$

where  $w' = w - \lambda(u_0)$ . Inspection of the above proof shows that under the same assumption as in Theorem 3.1, one has the recovery bound

$$\|u - \hat{u}\|_{H^1(\Omega)} \leq R(K_w)_{H^1(\Omega)} + \varepsilon + C\kappa \quad \text{for every } u \in K_w,$$

where  $C := (M + \delta)m(C_0 + \varepsilon_2)$ .

*Remark 3.5.* For simplicity, we did not introduce in the above analysis the possible errors in the application of the  $\lambda_i$  to the approximations  $\hat{u}_0$  and  $\hat{\phi}_j$ , and in the numerical solution to the system  $\hat{G}\hat{a} = \hat{w}$ , which would simply result in similar conditions involving the extra tolerance parameters.

**3.4. Point evaluation data.** We now want to extend the numerical algorithm and its analysis to the case when the data functionals  $\lambda_j$ ,  $j = 1, \dots, m$ , are point evaluations

$$\lambda_j(h) := h(x_j), \quad x_j \in \overline{\Omega}, \quad j = 1, \dots, m.$$

Of course these functionals are not defined for general functions  $h$  from  $H^1(\Omega)$ . However, we can formulate the recovery problem whenever the functionals  $\lambda_j$  are well defined on  $K$ . We now discuss settings when this is possible.

Recall that any  $u \in K$  can be written as  $u = u_0 + u_{\mathcal{H}}$ , where  $u_0$  is the solution to (1.1) with right side  $f$  and  $g = 0$  and  $u_{\mathcal{H}} \in \mathcal{H}^s(\Omega)$ . Point evaluation is well defined for the harmonic functions  $u_{\mathcal{H}} \in \mathcal{H}^s(\Omega)$ , provided the points are in  $\Omega$ . In addition, they are well defined for points on the boundary  $\Gamma$  if the space  $\mathcal{H}^s(\Omega)$  continuously embeds into  $C(\overline{\Omega})$ . For  $d = 2$ , this is the case when  $s > 1/2$  and when  $d = 3$ , this is the case when  $s > 1$ .

Concerning  $u_0$ , we will need some additional assumption to guarantee that point evaluation of  $u_0$  makes sense at the data sites  $x_j$ ,  $j = 1, \dots, m$ . For example, it is enough to assume that  $u_0$  is globally continuous or at least in a neighborhood of each of these points. This can be guaranteed by assuming an appropriate regularity of  $f$ . In this section, we assume that one of these settings holds. We then write

$$w'_j := u_{\mathcal{H}}(x_j) = w_j - u_0(x_j), \quad j = 1, \dots, m,$$

and follow the algorithm of the previous section with the following simple modifications:

*Modified Step 1.* We numerically find an approximation  $\hat{u}_0$  to  $u_0$ , which, in addition to

$$\|u_0 - \hat{u}_0\|_{H^1(\Omega)} \leq \varepsilon_1,$$

satisfies the requirement

$$(3.14) \quad \max_{i=1,\dots,m} |u_0(x_i) - \hat{u}_0(x_i)| \leq \varepsilon_1.$$

To find such a  $\hat{u}_0$  we use standard or adaptive FEM methods. Given that  $\hat{u}_0$  has been constructed, we define  $\hat{w}_j := w_j - \hat{u}_0(x_j)$ ,  $j = 1, \dots, m$ , and thus, using (3.14), we have  $\|w' - \hat{w}\| \leq \varepsilon_1$ .

*Modified Step 2.* For each  $j = 1, \dots, m$ , we numerically compute an approximation  $\hat{\phi}_j$  to  $\phi_j$ , which, in addition to

$$\|\phi_j - \hat{\phi}_j\|_{H^1(\Omega)} \leq \varepsilon_2, \quad j = 1, \dots, m,$$

satisfies the condition

$$(3.15) \quad \max_{i=1,\dots,m} |\phi_j(x_i) - \hat{\phi}_j(x_i)| \leq \varepsilon_2, \quad i, j = 1, \dots, m.$$

Condition (3.15) ensures that in Step 3 we can choose the entries  $\hat{g}_{i,j}$  of the matrix  $\hat{G}$  as

$$\hat{g}_{i,j} = \hat{\phi}_j(x_i), \quad i, j = 1, \dots, m.$$

Steps 3 and 4 of our algorithm remain the same as in the previous section.

**THEOREM 3.6.** *With the above modifications, Theorem 3.1 holds with the exact same statement in this point evaluation setting.*

*Proof.* The proof is the same as that of Theorem 3.1. □

**4. Finite element approximations of the Riesz representers.** The computation of an approximation  $\hat{u}_0$  to  $u_0$ , required in Step 1 of the algorithm, can be carried out by standard finite element Galerkin schemes. Depending on our knowledge on  $f$  one can resort to known a priori estimates for  $\varepsilon_1$ , or may employ standard a posteriori estimates to ensure that the underlying discretization provides a desired target accuracy. Therefore, in the remainder of this section, we focus on a numerical implementation of Step 2 of the proposed algorithm.

Our proposed numerical algorithm for Step 2 is to use finite element methods to generate the approximations  $\hat{\phi}_j$  of the Riesz representers  $\phi_j$ . Note that each of the functions  $\phi_j$  is harmonic on  $\Omega$  but we do not require that the sought after numerical approximation  $\hat{\phi}_j$  is itself harmonic but only that it provides an accurate  $H^1(\Omega)$  approximation to  $\phi_j$ . This allows us to use finite element approximations which are themselves not harmonic. However, the  $\hat{\phi}_j$  will necessarily have to be close to being harmonic since they approximate a harmonic function in the  $H^1(\Omega)$  norm.

Our numerical approach to constructing a  $\hat{\phi}_j$ , discussed in section 4.1, is to use discretely harmonic finite elements. Here,  $\hat{\phi}_j$  is a discrete harmonic extension of a finite element approximation to the trace  $\psi_j = T(\phi_j)$  computed by solving a Galerkin problem. In order to reduce computational cost (see Remark 4.2), we incorporate discrete harmonicity as constraints and introduce in section 4.2 an *equivalent saddle*

*point formulation* that has the same solution  $\hat{\phi}_j$ , and which is the one that we practically employ in the numerical experiments given in section 5. We give in section 4.4 an a priori analysis with error bounds for  $\|\phi_j - \hat{\phi}_j\|_{H^1(\Omega)}$  in terms of the finite element mesh size, in the case where the measurement functionals are continuous on  $H^1(\Omega)$ . These error bounds can in turn be used to ensure the prescribed accuracy  $\varepsilon_2$  in Step 2. We finally discuss in section 4.5 the extensions to the point value case where pointwise error bounds on  $|\hat{\phi}_j(x_i) - \hat{\phi}_j(x_i)|$  are also needed.

In order to simplify notation, we describe these procedures for finding an approximation  $\hat{\phi}$  to the Riesz representer  $\phi \in \mathcal{H}^s = \mathcal{H}^s(\Omega)$  of a given linear functional  $\nu$  on  $\mathcal{H}^s$ . This numerical procedure is then applied with  $\nu = \lambda_j$  to find the numerical approximations  $\hat{\phi}_j$  to the Riesz representer  $\phi_j$ .

For simplicity, throughout this section, we work under the assumption that  $\Omega$  is a polygonal domain of  $\mathbb{R}^2$  or polyhedral domain of  $\mathbb{R}^3$ . This allows us to define finite element spaces based on triangular or simplicial partitions of  $\Omega$  that in turn induce similar partitions on the boundary. We assume that  $\frac{1}{2} < s < \frac{3}{2}$ , which is the relevant range for such domains, as explained in section 2. Our analysis can be extended to more general domains with smooth or piecewise smooth boundaries, for example by using isoparametric elements near the boundary, however at the price of considerably higher technicalities.

**4.1. A Galerkin formulation.** Let  $s > 1/2$  be fixed and assume that  $\nu$  is any linear form continuous on  $\mathcal{H}^s(\Omega)$  with norm

$$(4.1) \quad C_s := \max\{\nu(v) : \|v\|_{\mathcal{H}^s(\Omega)} = 1\}.$$

In view of the definition of the  $\mathcal{H}^s$  norm, the representer  $\phi \in \mathcal{H}^s(\Omega)$  of  $\nu$  for the corresponding inner product can be defined as

$$\phi = E\psi,$$

where  $E$  is the harmonic extension operator of (4.3) below and where  $\psi \in H^s(\Gamma)$  is the solution to the following variational problem:

$$(4.2) \quad \langle \psi, \eta \rangle_{H^s(\Gamma)} = \mu(\eta) := \nu(E\eta), \quad \eta \in H^s(\Gamma).$$

Note that this problem admits a unique solution and we have

$$\|\psi\|_{H^s(\Gamma)} = \|\phi\|_{\mathcal{H}^s(\Omega)} = C_s.$$

Recall that

$$(4.3) \quad Eg := \operatorname{argmin}\{\|\nabla v\|_{L^2(\Omega)} : v_\Gamma = g\}.$$

The function  $Eg$  is characterized by  $T(Eg) = g$  and

$$\int_{\Omega} \nabla Eg \cdot \nabla v = 0, \quad v \in H_0^1(\Omega).$$

From the left inequality in (1.3), one has

$$(4.4) \quad \|Eg\|_{H^1(\Omega)} \leq C_E \|g\|_{H^{1/2}(\Gamma)}, \quad g \in H^{1/2}(\Gamma),$$

where  $C_E$  can be taken to be the inverse of the constant  $c_0$  in (1.3).

Therefore, one approach to discretizing this problem is the following: consider finite element spaces  $\mathbb{V}_h$  associated to a family of meshes  $\{\mathcal{T}_h\}_{h>0}$  of  $\Omega$ , where as usual  $h$  denotes the maximum meshsize. We define  $\mathbb{T}_h$  to be the space obtained by restriction of  $\mathbb{V}_h$  on the boundary  $\Gamma$ , that is,

$$\mathbb{T}_h = T(\mathbb{V}_h).$$

Since we have assumed that  $\Omega$  is a polygonal or polyhedral domain, the space  $\mathbb{T}_h$  is a standard finite element space for the boundary mesh. Having also assumed that  $s < 3/2$ , when using standard  $H^1$  conforming finite elements such as  $\mathbb{P}_k$ -Lagrange finite elements, we are ensured that  $\mathbb{T}_h \subset H^s(\Gamma)$ . We denote by

$$\mathbb{W}_h := \{v_h \in \mathbb{V}_h : T(v_h) = 0\},$$

the finite element space with homogeneous boundary conditions.

We define the discrete harmonic extension operator  $E_h$  associated to  $\mathbb{V}_h$  as follows: for  $g_h \in \mathbb{T}_h$ ,

$$E_h g_h := \operatorname{argmin}\{\|\nabla v_h\|_{L^2(\Omega)} : v_h \in \mathbb{V}_h, T(v_h) = g_h\}.$$

Note that  $E_h g_h$  is not harmonic. Similar to  $E$ , the function  $E_h g_h$  is characterized by  $T(E_h g_h) = g_h$  and

$$\int_{\Omega} \nabla E_h g_h \cdot \nabla v_h = 0, \quad v_h \in \mathbb{W}_h.$$

Then, we define the approximation  $\phi_h \in \mathbb{V}_h$  to  $\phi$  as

$$\phi_h = E_h \psi_h,$$

where  $\psi_h \in \mathbb{T}_h$  is the solution to the following variational problem:

$$(4.5) \quad \langle \psi_h, g_h \rangle_{H^s(\Gamma)} = \mu_h(g_h) := \nu(E_h g_h), \quad g_h \in \mathbb{T}_h.$$

Here we are assuming that, in addition to be defined on  $\mathcal{H}^s(\Omega)$ , the functional  $\nu$  is also well defined on the space  $\mathbb{V}_h$ . We shall further consider separately two instances where this is the case: (i)  $\nu$  is a continuous functional on  $H^1(\Omega)$  and (ii)  $\nu$  is a point evaluation functional.

Note that (4.5) is *not* the straightforward Galerkin approximation of (4.2), since  $\mu_h$  differs from  $\mu$ . This complicates somewhat the further conducted convergence analysis. The numerical method we employ for computing  $\phi_h$  is to numerically solve an equivalent saddle point problem described below.

We apply the strategy (4.5) to  $\nu := \lambda_j$  for each  $j$  and thereby obtain the corresponding approximations  $\hat{\phi}_j := \phi_h \in \mathbb{V}_h$ . Since Step 2 requires that we guarantee the error  $\|\phi_j - \hat{\phi}_j\|_{H^1} \leq \varepsilon_2$ , our main goal in this section is to establish a quantitative convergence bound for  $\|\phi - \phi_h\|_{H^1}$ . We also need to establish a pointwise convergence bound for  $|\phi(x) - \phi_h(x)|$  when considering the modified version of Step 2 in the case that the measurements are point values.

Similar to  $E$ , it will be important in our analysis to control the stability of  $E_h$  in the sense of a bound

$$(4.6) \quad \|E_h g_h\|_{H^1(\Omega)} \leq D_E \|g_h\|_{H^{1/2}(\Gamma)}, \quad g_h \in \mathbb{T}_h,$$

with a constant  $D_E$  that is independent of  $h$ . However, such a uniform bound is not readily inherited from the stability of  $E$ . As observed in [8], its validity is known to depend on the existence of uniformly  $H^1$ -stable linear projections onto  $\mathbb{V}_h$  preserving the homogeneous boundary condition, that is, projectors  $P_h$  onto  $\mathbb{V}_h$  that satisfy

$$(4.7) \quad P_h(H_0^1(\Omega)) = \mathbb{W}_h \quad \text{and} \quad \|P_h v\|_{H^1(\Omega)} \leq B \|v\|_{H^1(\Omega)}, \quad v \in H^1(\Omega),$$

for some  $B$  independent of  $h$ . One straightforward consequence of this is that if  $v \in H^1(\Omega)$  with  $v|_\Gamma \in \mathbb{T}_h$ , then  $P_h(v)|_\Gamma = v|_\Gamma$ .

We next show that the existence of such projectors is sufficient to guarantee the stability of  $E_h$ . For this, suppose (4.7) holds and  $g_h \in \mathbb{T}_h$ . Then  $P_h E g_h \in \mathbb{V}_h$  and the trace of  $P_h E g_h$  is equal to  $g_h$ . It follows that

$$\begin{aligned} \|E_h g_h - P_h E g_h\|_{H^1(\Omega)} &\leq C_P \|\nabla E_h g_h - \nabla P_h E g_h\|_{L^2(\Omega)} \\ &\leq C_P \|\nabla E_h g_h\|_{L_2(\Omega)} + C_P \|\nabla P_h E g_h\|_{L^2(\Omega)}, \\ &\leq 2C_P \|P_h E g_h\|_{H^1(\Omega)}, \end{aligned}$$

where  $C_P$  is the Poincaré constant for  $\Omega$ . Here, the last inequality follows from the minimizing property of  $E_h g_h$ . Thus, by triangle inequality, one has

$$\begin{aligned} \|E_h g_h\|_{H^1(\Omega)} &\leq (1 + 2C_P) \|P_h E g_h\|_{H^1(\Omega)} \leq (1 + 2C_P) B \|E g_h\|_{H^1(\Omega)} \\ &\leq (1 + 2C_P) B C_E \|g_h\|_{H^{1/2}(\Gamma)}, \end{aligned}$$

which is (4.6) with  $D_E = (1 + 2C_P) B C_E$ .

The requirement of uniformly stable projectors  $P_h$  with the property (4.7) is satisfied by projectors of Scott–Zhang type [28] when the family of meshes  $\{\mathcal{T}_h\}_{h>0}$  is shape regular, that is, when all elements  $T$  have a uniformly bounded ratio between their diameters  $h(T)$  and the diameter  $\rho(T)$  of their inner circle. In other words, the shape parameter

$$(4.8) \quad \sigma = \sigma(\{\mathcal{T}_h\}_{h>0}) := \sup_{h>0} \max_{T \in \mathcal{T}_h} \frac{h(T)}{\rho(T)}$$

is finite. In all that follows in the present paper, we work under such an assumption on the meshes  $\mathcal{T}_h$ . Therefore, (4.6) holds when  $\mathbb{V}_h$  is subordinate to such partitions.

**4.2. A saddle point formulation.** Before attacking the convergence analysis, we need to stress an important computational variant of the above described Galerkin method, that leads to the same solution  $\phi_h$ . It is based on imposing harmonicity via a Lagrange multiplier. For this purpose, we introduce the Hilbert space  $X^s(\Omega)$  that consists of all  $v \in H^1(\Omega)$  such that  $v|_\Gamma \in H^s(\Gamma)$ , and equip it with the norm

$$\|v\|_{X^s(\Omega)} := \left( \|v|_\Gamma\|_{H^s(\Gamma)}^2 + \|\nabla v\|_{L^2(\Omega)}^2 \right)^{1/2}.$$

Then, the Riesz representer  $\phi$  is equivalently determined as the solution of the saddle point problem: find  $(\phi, \pi) \in X^s(\Omega) \times H_0^1(\Omega)$  such that

$$(4.9) \quad \begin{aligned} a(\phi, v) + b(v, \pi) &= \nu(v), & v \in X^s(\Omega), \\ b(\phi, z) &= 0, & z \in H_0^1(\Omega), \end{aligned}$$

where the bilinear forms are given by

$$a(\phi, v) := \langle \phi|_\Gamma, v|_\Gamma \rangle_{H^s(\Gamma)} \quad \text{and} \quad b(v, \pi) := \langle \nabla v, \nabla \pi \rangle_{L^2(\Omega)}.$$

Clearly the second equation in (4.9) means that  $\phi$  is harmonic and testing the first equation with a  $v \in \mathcal{H}^s(\Omega)$  shows that  $\phi$  is the Riesz representer of  $\nu$ .

This saddle point formulation is well-posed: the bilinear forms  $a$  and  $b$  obviously satisfy the continuity properties

$$a(\phi, v) \leq \|\phi_\Gamma\|_{H^s(\Gamma)} \|v_\Gamma\|_{H^s(\Gamma)} \leq \|\phi\|_{X^s(\Omega)} \|v\|_{X^s(\Omega)}, \quad \phi, v \in X^s(\Omega),$$

and for the standard norm  $\|v\|_{H_0^1(\Omega)} = \|\nabla v\|_{L^2(\Omega)}$ ,

$$b(v, \pi) \leq \|\nabla v\|_{L^2(\Omega)} \|\nabla \pi\|_{L^2(\Omega)} \leq \|v\|_{X^s(\Omega)} \|\pi\|_{H_0^1(\Omega)}, \quad v \in X^s(\Omega), \pi \in H_0^1(\Omega).$$

In addition, for all  $v \in \mathcal{H}^s(\Omega)$ , one has

$$\|v\|_{X^s(\Omega)}^2 \leq \|v_\Gamma\|_{H^s(\Gamma)}^2 + \|v\|_{H^1(\Omega)}^2 \leq \|v_\Gamma\|_{H^s(\Gamma)}^2 + C_E^2 \|v\|_{H^{1/2}(\Gamma)}^2 \leq (1 + C_E^2) a(v, v),$$

which shows that  $a$  is coercive on the null space of  $b$  in  $X^s(\Omega)$ . Finally, the bilinear form  $b$  satisfies the inf-sup condition

$$\inf_{\pi \in H_0^1(\Omega)} \sup_{v \in X^s(\Omega)} \frac{b(v, \pi)}{\|v\|_{X^s(\Omega)} \|\pi\|_{H_0^1(\Omega)}} \geq \inf_{\pi \in H_0^1(\Omega)} \frac{b(\pi, \pi)}{\|\pi\|_{X^s(\Omega)} \|\pi\|_{H_0^1(\Omega)}} = 1.$$

Therefore, the standard LBB theory ensures existence and uniqueness of the solution pair  $(\phi, \pi)$ .

We now discretize the saddle point problem by searching for  $(\phi_h, \pi_h) \in \mathbb{V}_h \times \mathbb{W}_h$  such that

$$\begin{aligned} a(\phi_h, v_h) + b(v_h, \pi_h) &= \nu(v_h), & v_h \in \mathbb{V}_h, \\ b(\phi_h, z_h) &= 0, & z_h \in \mathbb{W}_h. \end{aligned}$$

*Remark 4.1.* The equivalence with the previous derivation of  $\phi_h$  by the Galerkin approach is easily checked: the second equation tells us that the solution  $\phi_h$  is discretely harmonic, and therefore equal to  $E_h \psi_h$  for some  $\psi_h \in \mathbb{T}_h$ . Then taking  $v_h$  of the form  $E_h g_h$  for  $g_h \in \mathbb{T}_h$  gives us exactly the Galerkin formulation (4.5).

This discrete saddle point problem is uniformly well-posed when we equip the space  $\mathbb{W}_h$  with the  $H_0^1$  norm, and the space  $\mathbb{V}_h$  with the  $X^s$  norm. The continuity of  $a$  and  $b$ , and the inf-sup condition for  $b$  follow by the exact same arguments applied to the finite element spaces, with the same constants. On the other hand, we need to check the uniform ellipticity of  $a$  in the space  $\mathbb{V}_h^{\mathcal{H}} \subset \mathbb{V}_h$  of discretely harmonic functions, which can be defined as

$$\mathbb{V}_h^{\mathcal{H}} := \{v_h \in \mathbb{V}_h : b(v_h, z_h) = 0, z_h \in \mathbb{W}_h\},$$

or equivalently as the image of  $\mathbb{T}_h$  by the operator  $E_h$ . For all  $v_h \in \mathbb{V}_h^{\mathcal{H}}$  and  $g_h = T(v_h)$ , we write

$$\|v_h\|_{X^s(\Omega)}^2 \leq \|g_h\|_{H^s(\Gamma)}^2 + \|v_h\|_{H^1(\Omega)}^2 \leq \|g_h\|_{H^s(\Gamma)}^2 + D_E^2 \|g_h\|_{H^{1/2}(\Gamma)}^2 \leq (1 + D_E^2) a(v_h, v_h),$$

where we have used the discrete stability of  $E_h$ .

*Remark 4.2.* In practice, we use this discrete saddle point formulation for the computation of  $\phi_h$  rather than the equivalent Galerkin formulation (4.5) for the following reason. Let  $N_h := \dim \mathbb{V}_h$ ,  $M_h := \dim \mathbb{W}_h$ , and  $P_h := \dim \mathbb{T}_h = N_h - M_h$ . Computing the right-hand side load vector in (4.2) requires computing discretely harmonic

extensions of  $P_h$  basis functions, which means solving  $P_h$  linear systems of dimension  $M_h$ . In addition, one has to solve the sparse linear system (4.5) of size  $P_h$  followed by another system of size  $M_h$  to compute  $\phi_h = E_h \psi_h$ . Using optimal iterative solvers of linear complexity the minimum amount of work needed to compute one representer scales then like

$$P_h M_h \sim N_h^{1 + \frac{d-1}{d}}$$

while solving the saddle point problem requires the order of  $N_h + M_h \sim N_h$  operations. On the other hand the characterization of  $\phi_h$  through (4.5) appears to be more convenient when deriving error bounds for  $\|\phi - \phi_h\|_{H^1(\Omega)}$ . This is the objective of the next sections.

**4.3. Preparatory results.** In the derivation of error bounds for  $\|\phi - \phi_h\|_{H^1(\Omega)}$ , we will need several ingredients.

The first is the following lemma that quantifies the perturbation induced by using  $E_h$  in place of  $E$ .

LEMMA 4.3. *For any  $g_h \in \mathbb{T}_h$ , one has*

$$(4.10) \quad \|(E - E_h)g_h\|_{H^1(\Omega)} \leq C_2 h^{r-1} \|g_h\|_{H^s(\Gamma)},$$

where  $C_2$  depends on  $r$  and  $s$ , the shape-parameter  $\sigma$ , and on the geometry of  $\Omega$ .

*Proof.* From the properties of  $E$  and  $E_h$ , one has

$$\langle \nabla(E - E_h)g_h, \nabla v_h \rangle = 0, \quad v_h \in \mathbb{W}_h.$$

This orthogonality property shows that

$$\|\nabla(Eg_h - E_h g_h)\|_{L^2(\Omega)} \leq \|\nabla(Eg_h - E_h g_h - v_h)\|_{L^2(\Omega)}, \quad v_h \in \mathbb{W}_h,$$

and therefore,

$$\begin{aligned} \|\nabla(Eg_h - E_h g_h)\|_{L^2(\Omega)} &\leq \min_{v_h \in \mathbb{V}_h, T(v_h) = g_h} \|\nabla(Eg_h - v_h)\|_{L^2(\Omega)} \\ &\leq \|\nabla(Eg_h - P_h Eg_h)\|_{L^2(\Omega)}, \end{aligned}$$

where  $P_h$  is the stable projector that preserves homogeneous boundary condition; see (4.7). It follows that

$$\|\nabla(Eg_h - E_h g_h)\|_{L^2(\Omega)} \leq (1 + B) \min_{v_h \in \mathbb{V}_h} \|Eg_h - v_h\|_{H^1(\Omega)},$$

where  $B$  is the uniform  $H^1$ -stability bound on  $P_h$ . By standard finite element approximation estimates and (2.4), we have

$$\min_{v_h \in \mathbb{V}_h} \|Eg_h - v_h\|_{H^1(\Omega)} \leq Ch^{r-1} \|Eg_h\|_{H^r(\Omega)} \leq CC_1 h^{r-1} \|g_h\|_{H^s(\Gamma)},$$

where the constant  $C$  depends on  $r$  and on the shape parameter  $\sigma$ . The estimate (4.10) follows by Poincaré inequality since  $Eg_h - E_h g_h \in H_0^1(\Omega)$ .  $\square$

The second ingredient concerns the regularity of the solution to the variational problem

$$(4.11) \quad \langle \kappa, v \rangle_{H^s(\Gamma)} = \gamma(v), \quad v \in H^s(\Gamma).$$

For a general linear functional  $\gamma \in H^{-s}(\Gamma)$ , that is, continuous on  $H^s(\Gamma)$ , we are only ensured that the solution  $\kappa$  is bounded in  $H^s(\Gamma)$ , with  $\|\kappa\|_{H^s(\Gamma)} = \|\gamma\|_{H^{-s}(\Gamma)}$ . However, if  $\gamma$  has some extra regularity, this then translates into additional regularity of  $\kappa$ .

As a simple example, consider the case where  $\gamma$  is, in addition, continuous on  $L^2(\Gamma)$ , that is

$$(4.12) \quad \gamma(v) = \langle g, v \rangle_{L^2(\Gamma)}$$

for some  $g \in L^2(\Gamma)$ , and assume that we work with  $s = 1$  and a polygonal domain. Then the variational problem has a solution  $\kappa \in H^1(\Gamma)$  and, in addition,  $\kappa \in H^2(E)$  for each edge  $E$  with weak second derivative given by

$$-\kappa'' = g - \kappa \in L^2(\Gamma).$$

In turn, standard finite element approximation estimates yield

$$\min_{\kappa_h \in \mathbb{T}_h} \|\kappa - \kappa_h\|_{H^1(\Gamma)} \leq Ch\|g\|_{L^2(\Gamma)},$$

with a constant  $C$  that depends on the shape parameter  $\sigma$ .

Of course, gain of regularity theorems for elliptic problems are known in various contexts. However, we have not found a general treatment of gain of regularity that addresses the setting of this paper. In going forward, we do not wish to systematically explore this gain in regularity and approximability for more general values of  $s$  and smoothness of  $\gamma$  since this would significantly enlarge the scope of this paper. Instead, we state it as the following general assumption.

**Assumption R.** For  $s > \frac{1}{2}$  and  $\delta > 0$ , there exists  $r(s, \delta) > 0$  such that if  $\gamma \in H^{-s+\delta}(\Gamma)$  for some  $\delta > 0$ , then the solution  $\kappa$  to (4.11) satisfies

$$(4.13) \quad \min_{\kappa_h \in \mathbb{T}_h} \|\kappa - \kappa_h\|_{H^s(\Gamma)} \leq Ch^{r(s, \delta)}\|\gamma\|_{H^{-s+\delta}(\Gamma)},$$

with a constant  $C$  that depends on  $s$ ,  $\delta$ , and on the shape parameter  $\sigma$ .

The above example shows that  $r(1, 1) = 1$  for a polygonal domain. We expect that this assumption always holds for the range  $\frac{1}{2} < s < \frac{3}{2}$  that is considered here.

**4.4. An a priori error estimate for  $\|\phi - \phi_h\|_{H^1(\Omega)}$ .** In this section, we work under the assumption that the linear functional  $\nu$  is continuous on  $H^1(\Omega)$  with norm

$$C_\nu := \max\{\nu(v) : \|v\|_{H^1(\Omega)} = 1\}.$$

Let us first check that this assumption implies a uniform a priori bound on  $\|\psi_h\|_{H^s(\Gamma)}$ . Indeed, we may write

$$\|\psi_h\|_{H^s(\Gamma)}^2 = \langle \psi_h, \psi_h \rangle_{H^s(\Gamma)} = \nu(E_h \psi_h) \leq C_\nu D_E \|\psi_h\|_{H^{1/2}(\Gamma)} \leq C_\nu D_E \|\psi_h\|_{H^s(\Gamma)},$$

where the first inequality used (4.6). Therefore,

$$(4.14) \quad \|\psi_h\|_{H^s(\Gamma)} \leq C_\nu D_E.$$

We have seen in section 2 that the function  $\phi$  belongs to the standard Sobolev space  $H^r(\Omega)$  for  $r$  defined in (2.3). We use this  $r$  throughout this section. From (2.4), there exists a constant  $C_1$  such that

$$(4.15) \quad \|Ew\|_{H^r(\Omega)} \leq C_1 \|w\|_{H^s(\Gamma)}, \quad w \in H^s(\Gamma).$$

As noted in section 2, the amount of smoothness  $r$  depends both on  $s$  and on the geometry of  $\Omega$ . What is important for us is that since  $s > 1/2$ , we have shown in section 2 that  $r > 1$ . For example, for smooth domains it is  $r = s + \frac{1}{2}$ . The fact that  $\phi \in H^r(\Omega)$  hints that the finite element approximation  $\phi_h$  to  $\phi$  should converge with a certain rate.

This is indeed the case as given in the following result.

**THEOREM 4.4.** *Under Assumption R, we have*

$$(4.16) \quad \|\phi - \phi_h\|_{H^1(\Omega)} \leq CC_\nu h^t,$$

where  $t = \min\{r - 1, r(s, s + \frac{1}{2}) + r(s, s - \frac{1}{2})\}$ . The constant  $C$  depends on  $s$  and on the geometry of  $\Omega$ , and on the family of meshes through the shape parameter  $\sigma$ .

*Proof.* We use the decomposition

$$(4.17) \quad \phi - \phi_h = E\psi - E_h\psi_h = E(\psi - \psi_h) + (E - E_h)\psi_h.$$

The second term can be estimated with the help of Lemma 4.3 applied to  $g_h = \psi_h$  which gives

$$\|(E - E_h)\psi_h\|_{H^1(\Omega)} \leq C_2 h^{r-1} \|\psi_h\|_{H^s(\Gamma)} \leq C_2 D_E C_\nu h^{r-1}$$

from the a priori estimate (4.14) for  $\psi_h$ . We thus have obtained a bound in  $\mathcal{O}(h^{r-1})$  for the  $H^1$  norm of the second term in (4.17).

For the first term, we know that

$$\|E(\psi - \psi_h)\|_{H^1(\Omega)} \leq C_E \|\psi - \psi_h\|_{H^{1/2}(\Gamma)},$$

and so we are led to estimate  $\psi - \psi_h$  in the  $H^{1/2}(\Gamma)$  norm. For this purpose, we introduce the intermediate solution  $\bar{\psi}_h \in \mathbb{T}_h$  to the problem

$$\langle \bar{\psi}_h, g_h \rangle_{H^s(\Gamma)} = \mu(g_h) = \nu(Eg_h), \quad g_h \in \mathbb{T}_h,$$

and we use the decomposition

$$(4.18) \quad \psi - \psi_h = (\psi - \bar{\psi}_h) + (\bar{\psi}_h - \psi_h).$$

We estimate the second term in (4.18) by noting that for any  $g_h \in \mathbb{T}_h$ ,

$$\langle \bar{\psi}_h - \psi_h, g_h \rangle_{H^s(\Gamma)} = \nu((E - E_h)g_h) \leq C_\nu \|(E - E_h)g_h\|_{H^1(\Omega)} \leq C_\nu C_2 h^{r-1} \|g_h\|_{H^s(\Gamma)},$$

where we have again used Lemma 4.3. Taking  $g_h = \bar{\psi}_h - \psi_h$  we obtain a bound  $\mathcal{O}(h^{r-1})$  for its  $H^s(\Gamma)$  norm, and in turn for its  $H^{1/2}(\Gamma)$  norm.

It remains to estimate  $\|\psi - \bar{\psi}_h\|_{H^{1/2}(\Gamma)}$ . Note that  $\bar{\psi}_h$  is exactly the Galerkin approximation of  $\psi$  since we use the same linear form  $\mu$  in both problems. In fact, we have

$$\langle \psi - \bar{\psi}_h, g_h \rangle_{H^s(\Gamma)} = 0, \quad g_h \in \mathbb{T}_h,$$

that is  $\bar{\psi}_h$  is the  $H^s$ -orthogonal projection of  $\psi$  onto  $\mathbb{T}_h$ , and therefore,

$$\|\psi - \bar{\psi}_h\|_{H^s(\Gamma)} = \min_{\kappa_h \in \mathbb{T}_h} \|\psi - \kappa_h\|_{H^s(\Gamma)}.$$

Since the linear form  $\mu$  satisfies

$$|\mu(g)| = |\nu(Eg)| \leq C_\nu \|Eg\|_{H^1(\Omega)} \leq C_\nu C_E \|g\|_{H^{1/2}(\Gamma)},$$

and thus belongs to  $H^{-1/2}(\Gamma)$ , we may apply the estimate (4.13) to  $\gamma = \nu$ ,  $\kappa = \psi$ ,  $\delta = s - \frac{1}{2} > 0$ , to reach

$$(4.19) \quad \|\psi - \bar{\psi}_h\|_{H^{1/2}(\Gamma)} \leq \|\psi - \bar{\psi}_h\|_{H^s(\Gamma)} \leq CC_\nu C_E h^{r(s, s - \frac{1}{2})}.$$

This proves the theorem for the value  $t = \min\{r - 1, r(s, s - \frac{1}{2})\} > 0$ .

We finally improve the value of  $t$  by using a standard Aubin–Nitsche duality argument as follows. We now take  $\kappa$  to be the solution of (4.11) with

$$\gamma(v) = \langle \psi - \bar{\psi}_h, v \rangle_{H^{1/2}(\Gamma)}, \quad v \in H^{1/2}(\Gamma),$$

where  $\langle \cdot, \cdot \rangle_{H^{1/2}(\Gamma)}$  stands for the  $H^{1/2}$  scalar product associated with the norm  $\|\cdot\|_{H^{1/2}(\Gamma)}$ . We then write

$$\|\psi - \bar{\psi}_h\|_{H^{1/2}(\Gamma)}^2 = \langle \psi - \bar{\psi}_h, \psi - \bar{\psi}_h \rangle_{H^{1/2}(\Gamma)} = \langle \kappa, \psi - \bar{\psi}_h \rangle_{H^s(\Gamma)} = \langle \kappa - \kappa_h, \psi - \bar{\psi}_h \rangle_{H^s(\Gamma)},$$

where the last equality comes from Galerkin orthogonality. It follows that

$$\begin{aligned} \|\psi - \bar{\psi}_h\|_{H^{1/2}(\Gamma)}^2 &\leq \|\kappa - \kappa_h\|_{H^s(\Gamma)} \|\psi - \bar{\psi}_h\|_{H^s(\Gamma)} \\ &\leq Ch^{r(s, s + \frac{1}{2})} \|\psi - \bar{\psi}_h\|_{H^{1/2}(\Gamma)} \|\psi - \bar{\psi}_h\|_{H^s(\Gamma)}, \end{aligned}$$

where we have again used (4.13) now with  $\delta = s + \frac{1}{2}$ . Using the already established estimate (4.19), it follows that

$$\|\psi - \bar{\psi}_h\|_{H^{1/2}(\Gamma)} \leq CC_E C_\nu h^{\tilde{t}},$$

with  $\tilde{t} := r(s, s + \frac{1}{2}) + r(s, s - \frac{1}{2})$ . With all such estimates, the desired convergence bound follows with  $t := \min\{r - 1, \tilde{t}\}$ .  $\square$

*Remark 4.5.* In the case of a polygonal domain and  $s = 1$  which is further considered in our numerical experiments, we know that  $r = \frac{3}{2}$  and  $r(1, 1) = 1$  so that  $\tilde{t} \geq r(1, \frac{3}{2}) \geq 1$ . In turn, the convergence bound is established with  $t = r - 1 = \frac{1}{2}$ .

#### 4.5. The case of point value evaluations.

We discuss now the case where

$$\nu(v) = \delta_z(v) = v(z)$$

for some  $z \in \overline{\Omega}$ . In order to guarantee that point evaluation is a continuous functional on  $\mathcal{H}^s$ , we assume that

$$s > \frac{d-1}{2},$$

that is  $s > \frac{1}{2}$  for  $d = 2$ , and  $s > 1$  for  $d = 3$ . We want to find the Riesz representer of such a point evaluation functional on  $\mathcal{H}^s$ . Note that our assumption on  $s$  ensures the continuous embeddings

$$H^s(\Gamma) \subset C(\Gamma),$$

as well as

$$\mathcal{H}^s(\Omega) \subset H^r(\Omega) \subset C(\bar{\Omega}),$$

since in view of (2.3)

$$r = \min \left\{ s + \frac{1}{2}, r^* \right\} > \frac{d}{2},$$

where in the inequality we recall that  $r^* > \frac{3}{2}$  for polygonal domains.

The point evaluation functional  $\nu$  is thus continuous on  $\mathcal{H}^s(\Omega)$  with norm  $C_s$  bounded independently of the position of  $z$ . Of course, the Galerkin scheme analyzed above for  $\nu \in H^1(\Omega)^*$  continues to make sense since  $\nu$  is well defined on the space  $\mathbb{V}_h$ .

As explained in section 3.4, the prescriptions in Step 2 of the recovery algorithm need to be strengthened in the point evaluation setting. Thus, we are interested in bounding the pointwise error  $|\phi(x) - \phi_h(x)|$  at the measurement points, in addition to the  $H^1$ -error  $\|\phi - \phi_h\|_{H^1(\Omega)}$ . In what follows, we establish a modified version of Theorem 4.4 in the point value setting that gives a convergence rate for  $\|\phi - \phi_h\|_{H^1(\Omega)}$ , and, in addition, for  $\|\phi - \phi_h\|_{L^\infty(\Omega)}$  ensuring the pointwise error control. We stress that the numerical method remains unchanged, that is,  $\phi_h$  is defined in the exact same way as previously. The new ingredients that are needed in our investigation are two classical results on the behavior of the finite element method with respect to the  $L^\infty$  norm.

The first one is the so-called weak discrete maximum principle which states that there exists a constant  $C_{\max}$  such that, for all  $h > 0$ ,

$$(4.20) \quad \|E_h g_h\|_{L^\infty(\Omega)} \leq C_{\max} \|g_h\|_{L^\infty(\Gamma)}, \quad g_h \in \mathbb{T}_h.$$

This result was first established in [5] with constant  $C_{\max} = 1$  for piecewise linear Lagrange finite elements under acuteness assumptions on the angles of the simplices. The above version with  $C_{\max} \geq 1$  is established in [27] for Lagrange finite elements of any degree on  $2d$  polygonal domains, under the more general assumption that the meshes  $\{\mathcal{T}_h\}_{h>0}$  are quasi-uniform (in addition to shape regularity, all elements of  $\mathcal{T}_h$  have diameters of order  $h$ ). A similar result is established in [15] on  $3d$  convex polyhedrons.

The second ingredient we need is a stability property in the  $L^\infty$  norm of the Galerkin projection  $R_h : H_0^1(\Omega) \rightarrow \mathbb{W}_h$  where  $R_h v$ ,  $v \in H_0^1(\Omega)$ , is defined by

$$\int_{\Omega} \nabla R_h v \cdot \nabla v_h = \int_{\Omega} \nabla v \cdot \nabla v_h, \quad v_h \in \mathbb{W}_h.$$

Specifically, this result states that there exists a constant  $C_{\text{gal}}$  and exponent  $a \geq 0$  such that, for all  $h > 0$ ,

$$(4.21) \quad \|R_h v\|_{L^\infty(\Omega)} \leq C_{\text{gal}} (1 + |\ln(h)|)^a \|v\|_{L^\infty(\Omega)}, \quad v \in L^\infty(\Omega) \cap H_0^1(\Omega),$$

that is, the Ritz projection is stable and quasi-optimal, uniformly in  $h$ , up to a logarithmic factor. This result is established in [27] for Lagrange finite elements on  $2d$  polygonal domains and quasi-uniform partitions, with  $a = 1$  in the case of piecewise linear elements and  $a = 0$  for higher order elements. A similar result is established in [15] with  $a = 0$  for convex polygons and polyhedrons. Going further, we assume that the choice of finite element meshes ensures the validity of (4.20) and (4.21).

We begin our analysis with the observation that under the additional mesh assumptions, Lemma 4.3 can be adapted to obtain an estimate on  $\|(E - E_h)g_h\|_{L^\infty(\Omega)}$ .

LEMMA 4.6. *For any  $g_h \in \mathbb{T}_h$ , one has*

$$(4.22) \quad \|(E - E_h)g_h\|_{L^\infty(\Omega)} \leq C_3(1 + |\ln(h)|)^a h^{r - \frac{d}{2}} \|g_h\|_{H^s(\Gamma)},$$

where  $C_3$  depends on  $(r, s)$ , the geometry of  $\Omega$ , and the family of meshes through  $C_{\text{gal}}$ .

*Proof.* For any  $v_h \in \mathbb{V}_h$  such that  $T(v_h) = g_h$ , we write

$$\|(E - E_h)g_h\|_{L^\infty(\Omega)} \leq \|Eg_h - v_h\|_{L^\infty(\Omega)} + \|E_hg_h - v_h\|_{L^\infty(\Omega)}.$$

It is readily seen that  $E_hg_h - v_h = R_h(E_hg_h - v_h) = R_h(Eg_h - v_h)$ . Indeed,  $R_hE_hg_h - R_hEg_h \in \mathbb{W}_h$  and  $\int_\Omega \nabla(R_h(E_hg_h - Eg_h)) \cdot \nabla v_h = \int_\Omega \nabla(E_hg_h - Eg_h) \cdot \nabla v_h = 0$  for all  $v_h \in \mathbb{W}_h$ . Therefore, by (4.21), we obtain

$$\|(E - E_h)g_h\|_{L^\infty(\Omega)} \leq (1 + C_{\text{gal}}(1 + |\ln(h)|)^a) \min_{v_h \in \mathbb{V}_h, T(v_h) = g_h} \|Eg_h - v_h\|_{L^\infty(\Omega)}.$$

On the other hand, we are ensured that  $Eg_h$  belongs to  $H^r(\Omega)$  where  $r > \frac{d}{2}$ , and therefore has Hölder smoothness of order  $r - \frac{d}{2} > 0$  with

$$\|Eg_h\|_{C^{r-\frac{d}{2}}(\Omega)} \leq C_e \|Eg_h\|_{H^r(\Omega)} \leq C_e C_1 \|g_h\|_{H^s(\Gamma)},$$

where  $C_e$  is the relevant continuous embedding constant. By standard finite element approximation theory,

$$\min_{v_h \in \mathbb{V}_h, T(v_h) = g_h} \|Eg_h - v_h\|_{L^\infty(\Omega)} \leq Ch^{r - \frac{d}{2}} \|Eg_h\|_{C^{r-\frac{d}{2}}(\Omega)},$$

where  $C$  depends on  $r$  and the shape-parameter  $\sigma$ , and therefore we obtain (4.22).  $\square$

We are now in position to give an adaptation of Theorem 4.4 to the point value setting.

**THEOREM 4.7.** *Under Assumption R, for any  $t_1 < \min\{r - \frac{d}{2}, r(s, s + \frac{1}{2}) + r(s, s - \frac{1}{2})\}$ , one has*

$$(4.23) \quad \|\phi - \phi_h\|_{H^1(\Omega)} \leq Ch^{t_1},$$

and for any  $t_2 < \min\{r - \frac{d}{2}, 2r(s, s - \frac{d-1}{2})\}$ , one has

$$(4.24) \quad \|\phi - \phi_h\|_{L^\infty(\Omega)} \leq Ch^{t_2}.$$

The constant  $C$  depends in both cases on  $s$ ,  $t_1$ , and  $t_2$ , on the geometry of  $\Omega$ , as well as on the family of meshes through the constants  $C_{\text{max}}$  and  $C_{\text{gal}}$ , and the shape parameter  $\sigma$ .

*Proof.* We estimate  $\|\phi - \phi_h\|_{H^1(\Omega)}$  by adapting certain steps in the proof of Theorem 4.4. The first change lies in the a priori estimate of the  $H^s(\Gamma)$  norm of  $\psi_h$  that was previously given by (4.14) which is not valid anymore since  $C_\nu = \infty$ . Instead, we write

$$\begin{aligned} \|\psi_h\|_{H^s(\Gamma)}^2 &= \langle \psi_h, \psi_h \rangle_{H^s(\Gamma)} = \nu(E_h\psi_h) \leq \|E_h\psi_h\|_{L^\infty(\Omega)} \leq C_{\text{max}} \|\psi_h\|_{L^\infty(\Gamma)} \\ &\leq C_{\text{max}} B_s \|\psi_h\|_{H^s(\Gamma)}, \end{aligned}$$

where we have used (4.20) and where  $B_s$  is the continuous embedding constant between  $H^s(\Gamma)$  and  $L_\infty(\Gamma)$ . In turn, we find that

$$(4.25) \quad \|\psi_h\|_{H^s(\Gamma)} \leq C_{\text{max}} B_s,$$

which results in the slightly modified estimate

$$\|(E - E_h)\psi_h\|_{H^1(\Omega)} \leq C_2 C_{\max} B_s h^{r-1},$$

for the second term of (4.17).

For the first term  $E(\psi - \psi_h)$ , we proceed in a similar manner to the proof of Theorem 4.4. Namely, we estimate the  $H^{1/2}(\Gamma)$  norms of two summands in (4.18). The estimate of  $\|\bar{\psi}_h - \psi_h\|_{H^{1/2}(\Gamma)}$  is modified as follows. We note that for any  $g_h \in \mathbb{T}_h$ ,

$$\begin{aligned} \langle \bar{\psi}_h - \psi_h, g_h \rangle_{H^s(\Gamma)} &= \nu((E - E_h)g_h) \leq \|(E - E_h)g_h\|_{L^\infty(\Omega)} \\ &\leq C_3(1 + |\ln(h)|)^a h^{r-\frac{d}{2}} \|g_h\|_{H^s(\Gamma)}, \end{aligned}$$

where we have now used Lemma 4.6. Taking  $g_h = \bar{\psi}_h - \psi_h$  we obtain a bound of order  $\mathcal{O}(h^{r-\frac{d}{2}})$  up to logarithmic factors for its  $H^s$  norm, and in turn for its  $H^{1/2}$  norm. The estimate of  $\|\psi - \bar{\psi}_h\|_{H^{1/2}(\Gamma)}$  is left unchanged and of order  $\mathcal{O}(h^{\tilde{t}})$ . Combining these various estimates, we have established (4.23) for any  $t_1 < \min\{r - \frac{d}{2}, \tilde{t}\}$ , with  $\tilde{t} := r(s, s + \frac{1}{2}) + r(s, s - \frac{1}{2})$ .

We next estimate  $\|\phi - \phi_h\|_{L^\infty(\Omega)}$  by the following adaptation of the proof of Theorem 4.4. For the first term  $(E - E_h)\psi_h$  of (4.17) we use Lemma 4.6 combined with the estimate (4.25) of  $\psi_h$  which give us

$$\|(E - E_h)\psi_h\|_{L^\infty(\Omega)} \leq C_{\max} B_s C_3(1 + |\ln(h)|)^a h^{r-\frac{d}{2}}.$$

For the second term  $E(\psi - \psi_h)$ , we use the continuous maximum principle to obtain

$$\|E(\psi - \psi_h)\|_{L^\infty(\Omega)} \leq \|\psi - \psi_h\|_{L^\infty(\Gamma)} \leq \|\psi_h - \bar{\psi}_h\|_{L^\infty(\Gamma)} + \|\psi - \bar{\psi}_h\|_{L^\infty(\Gamma)}$$

For the first summand, we write

$$\|\psi_h - \bar{\psi}_h\|_{L^\infty(\Gamma)} \leq C_e \|\psi_h - \bar{\psi}_h\|_{H^s(\Gamma)},$$

where  $C_e$  is the relevant continuous embedding constant, and we have already observed that  $\|\psi_h - \bar{\psi}_h\|_{H^s(\Gamma)}$  satisfies a bound in  $\mathcal{O}(h^{r-\frac{d}{2}})$  up to logarithmic factors. For the second summand, we may write

$$\|\psi - \bar{\psi}_h\|_{L^\infty(\Gamma)} \leq C_e \|\psi - \bar{\psi}_h\|_{H^s(\Gamma)},$$

where  $C_e$  is the relevant continuous embedding constant. Since  $\nu$  belongs to  $H^{-s+\delta}(\Gamma)$  for all  $\delta < s - \frac{d-1}{2}$ , we can apply the estimate (4.13) to reach a convergence bound

$$\|\psi - \bar{\psi}_h\|_{H^s(\Gamma)} \leq C h^{r(s, \delta)},$$

where  $C$  depends on the closeness of  $\delta$  to  $s - \frac{d-1}{2}$ , and on the family of meshes through the shape parameter  $\sigma$ . Combining these estimates then gives (4.24) for any  $t_2 < \min\{r - \frac{d}{2}, \tilde{t}\}$  where  $\tilde{t} = r(s, s - \frac{d-1}{2})$ , since  $\delta$  can be picked arbitrarily close to  $s - \frac{d-1}{2}$ .

We can improve the range of  $t_2$  as follows: pick any  $\bar{s}$  such that  $\frac{d-1}{2} < \bar{s} < s$  and write

$$\|\psi - \bar{\psi}_h\|_{L^\infty(\Gamma)} \leq C_e \|\psi - \bar{\psi}_h\|_{H^{\bar{s}}(\Gamma)},$$

where  $C_e$  is the relevant continuous embedding constant. We then apply a similar Aubin–Nitsche argument to derive an estimate

$$\|\psi - \bar{\psi}_h\|_{H^{\bar{s}}(\Gamma)} \leq C h^{r(s, \delta) + r(s, s - \bar{s})}.$$

Combining these estimates gives (4.24) for any  $t_2 < \min\{r - \frac{d}{2}, \bar{t}\}$ , where  $\bar{t} := 2r(s, s - \frac{d-1}{2})$  since  $\bar{s}$  can be picked arbitrarily close to  $\frac{d-1}{2}$  and  $\delta$  arbitrarily close to  $s - \frac{d-1}{2}$ .  $\square$

**5. Numerical illustrations.** In this section, we implement some examples of our numerical method. For this, we have to specify the domain  $\Omega$ , the functionals  $\lambda_j$ , and a function  $u \in H^1(\Omega)$  which gives rise to the data vector  $w = \lambda(u)$ . While our numerical method can be applied to general choices for these quantities, in our illustrations we make these choices so that the computations are not too involved but yet allow us the flexibility to illustrate certain features of our algorithm. The specific choices we make for our numerical example are the following.

**The domain.** In order to simplify the presentation, we restrict ourselves when  $\Omega = (0, 1)^2$  but point out again that the algorithm can be extended to more general domains.

**The function  $u$ .** For the function  $u$  we choose the harmonic function  $u = u_{\mathcal{H}}$  where

$$(5.1) \quad u_{\mathcal{H}}(x, y) = e^x \cos(y), \quad (x, y) \in \Omega := (0, 1)^2.$$

This choice means that  $u_0 = 0$  and therefore allows us not to deal with the computation of  $\hat{u}_0$ . This choice corresponds to the right side  $f = 0$ . Note that the trace of  $u_{\mathcal{H}}$  on the boundary  $\Gamma$  is piecewise smooth and continuous. Therefore, we have  $T(u_{\mathcal{H}}) \in H^1(\Gamma)$ . We take  $s = 1$  as our assumption on the value of  $s$ . This means that we shall seek Riesz representor for the functionals given below when viewed as acting on  $\mathcal{H}^1(\Omega)$ .

**5.1. The case of linear functionals defined on  $H^1(\Omega)$ .** In this section, we consider numerical experiments for linear functionals defined on  $H^1(\Omega)$ . In our illustrative example, we relabel these functionals by double indices associated with a regular square grid. More precisely,

$$(5.2) \quad \lambda_{i,j}(v) := \frac{1}{\sqrt{2\pi r^2}} \int_{\Omega} v(z) e^{-\frac{|z-z_{i,j}|^2}{r^2}} dz, \quad v \in H^1(\Omega), \quad i, j = 1, \dots, \sqrt{m}.$$

Here, we assume that  $m$  is a square integer and  $r = 0.1$  in our simulations. The centers  $z_{i,j} \in \Omega$  are uniformly distributed

$$(5.3) \quad z_{i,j} := \frac{1}{\sqrt{m} + 1} (i, j), \quad i, j = 1, \dots, \sqrt{m}.$$

Recall that our numerical algorithm as described in section 3.2 is based on finite element methods. Specifically, we use the finite element spaces

$$\mathbb{V}_h := \{v_h \in C^0(\bar{\Omega}) : v_h|_T \in \mathcal{Q}^1, \quad T \in \mathcal{T}_h\},$$

where  $\mathcal{T}_h$  are subdivisions of  $\Omega$  made of squares of equal side length  $h$  and  $\mathcal{Q}^1$  denotes the space of polynomials of degree at most 1 in each direction. In order to study the effect of the mesh-size we specifically consider

$$h = h_n := 2^{-n}, \quad n = 4, \dots, 9,$$

that is, bilinear elements on uniformly refined meshes with mesh-size  $2^{-n}$ .

We display in Table 5.1 the results of our numerical recovery algorithm. The entries in the table are the recovery errors

$$e(m, n) := \|u_{\mathcal{H}} - \hat{u}_{\mathcal{H}}\|_{H^1(\Omega)},$$

where  $\hat{u}_{\mathcal{H}} \in \mathbb{V}_{h_n}$  is the recovery for the particular values of  $m$  and  $n$ .

We have proven in this paper that our numerical recovery algorithm is near optimal with constant  $C$  that can be made arbitrarily close to one by choosing  $n$  sufficiently large. This means that the error  $e(m, n)$  satisfies  $e(m, n) \leq CR(K_w^{\mathcal{H}})_{H^1(\Omega)}$

TABLE 5.1

Recovery error  $e(m, n)$  for different amounts of Gaussian measurements  $m$  and finite element refinements  $n$ .

$n \backslash m$	4	9	16	25	36
4	0.7	0.28	0.2	141.73	49.43
5	0.7	0.28	0.18	16.0	16.31
6	0.7	0.28	0.18	0.2	1.79
7	0.7	0.28	0.18	0.16	0.11
8	0.7	0.28	0.18	0.09	0.06
9	0.7	0.28	0.18	0.09	0.06

for  $n$  sufficiently large. Increasing the number  $m$  of measurements is expected to decrease this Chebyshev radius. While one is tempted to think that the entries in each column of the table provides an upper bound for the Chebyshev radius of  $K_w^{\mathcal{H}}$  for these measurements, this is not guaranteed since we are only measuring the error for one function from  $K_w$ , namely  $u_{\mathcal{H}}$ , and not all possible functions from  $K_w$ . However, the entries in any given column provide a lower bound for the Chebyshev radius of  $K_w^{\mathcal{H}}$  provided  $n$  is sufficiently large.

Increasing the number  $m$  of measurements requires a finer resolution, i.e., increasing  $n$ , of the finite element discretization until the perturbation  $\varepsilon$  in Theorem 3.1 is sufficiently small. This is indeed confirmed by the results in Table 5.1 where stagnating error bounds (in each fixed column) indicate the corresponding tip-over point. We notice, in particular, that for small values of  $n$ , the error becomes very large as  $m$  grows. This is explained by the fact that the Gramian matrix  $G$  becomes severely ill-conditioned, and in turn the prescriptions on  $\|G - \hat{G}\|_1$  cannot be fulfilled when using finite element approximation of the Riesz representers on too coarse meshes. An overall convergence of the recovery error to zero can, of course, only take place when both  $m$  and  $n$  increase.

**5.2. The case of point value measurements.** In this section, we describe our numerical experiments in the case where the linear functionals  $\lambda_{i,j}$  are point evaluations at points from  $\bar{\Omega}$ . Recall that while the  $\lambda_{i,j}$  are not defined for general functions in  $H^1(\Omega)$  they are defined for functions in the model class  $K^{\mathcal{H}} := U(\mathcal{H}^s(\Omega))$  provided  $s$  is sufficiently large ( $s > 1/2$  for  $d = 2$  and  $s > 1$  for  $d = 3$ ). This means that the optimal recovery problem is well posed in such a case. We have given in section 3.4 sufficient conditions on a numerical algorithm to give near optimal recovery and then we have gone on to show in section 4.5 that our proposed numerical algorithm based on discrete harmonics converges to a near optimal recovery with any constant  $C > 1$  provided that the finite element spaces are discretized fine enough.

In the numerical experiments of this section, we again take  $\Omega = (0, 1)^2$ ,  $s = 1$ , and the data to be the point values of the harmonic function  $u_{\mathcal{H}}$  defined in (5.1). We choose the evaluation points to be the  $z_{i,j}$  of (5.3). We now provide in Table 5.2 the recovery error  $e(m, n)$ . The observed behavior is similar to the case of Gaussian averages; see Table 5.1.

**5.3. Additional comments on the approximation of Riesz representers.** We provide a little more information on the computation of the Riesz representers that may be of interest to the reader. We work in the same setting as in the previous

TABLE 5.2

Recovery error  $e(m, n)$  for different amounts of point evaluation measurements  $m$  and refinements  $n$ .

n \ m	4	9	16	25	36
4	0.70	0.28	0.19	14.43	15.49
5	0.70	0.28	0.18	32.56	8.02
6	0.70	0.28	0.18	1.51	2.27
7	0.70	0.28	0.18	0.53	0.89
8	0.70	0.28	0.18	0.20	0.14
9	0.70	0.28	0.18	0.14	0.11

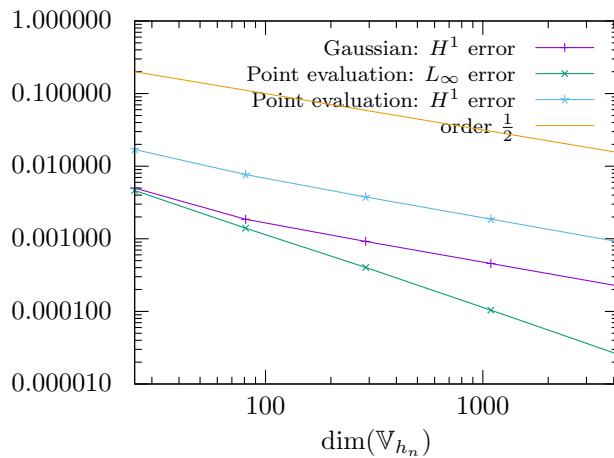


FIG. 5.1. Approximation errors for the Riesz representers of the Gaussian and point evaluation functionals.

sections. Let us begin with the rate of convergence of our numerical approximations to the Riesz representers.

We first consider the computation of the Riesz representer for the Gaussian measurement functional centered at  $z = z_{i,j} := (0.75, 0.5)$ . Let  $\phi_n \in \mathbb{V}_{h_n}$  be the approximation to the Riesz representer  $\phi$  produced by the finite element computation. Figure 5.1 shows the error  $\|\phi_n - \phi\|_{H^1(\Omega)}$ ,  $n = 2, \dots, 6$ . This graph indicates an error decay  $Cn^{-1/2} = Ch_n$  (Theorem 4.4 only guarantees  $Ch_n^{1/2}$ ; see also Remark 4.5).

Next, consider the computation of the Riesz representer for point evaluation at the same  $z$ . Figure 5.1 reports the numerical computations of error in both the  $H^1(\Omega)$  and  $L^\infty(\Omega)$  norms. Again, the graph indicates an error decay  $Ch_n$  for the  $H^1(\Omega)$  norm and a decay rate closer to  $Ch_n^2$  for the  $L^\infty(\Omega)$  norm which are better than the rate guaranteed by Theorems 4.4 and 4.7.

**5.4. Convergence of the estimator.** We conclude the analysis of our algorithm with some remarks on the convergence of the estimator provided by our algorithm as  $m, n \rightarrow +\infty$ . We continue with the case where  $K = K^\mathcal{H} = U(\mathcal{H}^s(\Omega))$  with  $s > 1/2$  and  $X$  is a Banach space for which  $\mathcal{H}^s(\Omega)$  embeds into  $X$ . For each  $m = 1, 2, \dots$ , let

$$\Lambda_m := \{\lambda_{1,m}, \dots, \lambda_{m,m}\}$$

be the set of measurement functionals. Recall that we are either dealing with functionals in  $H^1(\Omega)^*$  or with point evaluations which belong to  $\mathcal{H}^s(\Omega)^*$  by our assumption on  $s$ . We assume that the data vectors

$$(5.4) \quad w_m := w_m(u) := (\lambda_{1,m}(u), \dots, \lambda_{m,m}(u)), \quad m = 1, 2, \dots,$$

are observations of a fixed function  $u \in K$ . We are interested in what conditions on the sets  $\Lambda_m$ ,  $m = 1, 2, \dots$ , ensure that the functions  $\hat{u}_{m,n}$  produced by our algorithm converge in  $\|\cdot\|_X$  to  $u$  as  $m, n \rightarrow \infty$  and whether this convergence is uniform over  $u \in K$ .

Theorem 3.1 says that

$$(5.5) \quad \|u - \hat{u}_{m,n}\|_X \leq R(K_{w_m(u)})_X + \varepsilon_{n,m},$$

where for each fixed  $m$ , the error  $\varepsilon_{n,m} \rightarrow 0$  as  $n \rightarrow \infty$ . Therefore, if  $R(K_{w_m(u)})_X \rightarrow 0$  as  $m \rightarrow \infty$ , then we know that given any error tolerance  $\eta > 0$ , the error  $\|u - \hat{u}_{m,n}\|_X$  will not exceed  $\eta$  provided we take  $m$  sufficiently large and then  $n$  sufficiently large (depending on  $m$ ). Thus, we are left with the question of whether  $R(K_{w_m(u)})_X \rightarrow 0$ , uniformly over  $u \in K$ , as the number  $m$  of measurements tends to  $+\infty$ . In other words, we would like to know whether

$$(5.6) \quad R_m(K^{\mathcal{H}}) := \max_{u \in K} R(K_{w_m(u)})_X \rightarrow 0, \quad m \rightarrow \infty.$$

This is equivalent to asking when does

$$(5.7) \quad \tilde{R}_m(K^{\mathcal{H}}) := \max\{\|v\|_X : v \in K^{\mathcal{H}} \cap \mathcal{N}_m\} \rightarrow 0 \quad \text{as } m \rightarrow \infty.$$

where  $\mathcal{N}_m$  is the null space of  $\lambda_{1,m}, \dots, \lambda_{m,m}$ . Indeed,  $\tilde{R}_m(K^{\mathcal{H}}) \leq 2R_m(K^{\mathcal{H}}) \leq 2\tilde{R}_m(K^{\mathcal{H}})$ . Obviously, the validity of (5.7) requires some density assumption on the sets  $\Lambda_m$  in  $X^*$  as  $m \rightarrow \infty$ .

We illustrate how (5.7) is verified by considering the two specific cases studied in our numerical experiments, i.e., the cases of Gaussian and pointwise measurements on the Cartesian mesh points  $z_{i,j}$  defined in (5.3). We limit our discussion, as was done in our numerical examples, to the case when  $X = H^1(\Omega)$ ,  $\Omega = (0, 1)^2$ , and  $K = K^{\mathcal{H}}$ , where  $\mathcal{H} = \mathcal{H}^1(\Omega)$ . It follows that the functions in  $K$  are in  $U(H^{3/2}(\Omega))$  and hence (by the Sobolev embedding theorem) are not only continuous but in a ball of  $\text{Lip } \alpha$  for each  $\alpha < 1/2$  with

$$\|f\|_{\text{Lip } \alpha} := \max \left\{ \|f\|_{L^\infty(\Omega)}, \sup_{\substack{x, y \in \Omega \\ x \neq y}} \frac{|f(x) - f(y)|}{|x - y|^\alpha} \right\}.$$

We recall that the measurement functionals in  $\Lambda_m$  are associated with a grid of points  $z_{i,j}$  where  $m = k^2$ ,  $k \geq 1$ , an integer. We denote by  $\Omega_m = (h, 1-h) \times (h, 1-h) \subset \mathbb{R}^2$  the convex hull of these grid points.

Before going further, let us mention the following remark, which we will use next.

*Remark 5.1.* Let  $R := (a, b) \times (c, d) \subset \mathbb{R}^2$  be a rectangle in  $\mathbb{R}^2$ , and let  $R_x := R \setminus (b-h, b) \times [c, d]$ . Let  $g \in U(H^{1/2}(R))$  such that  $\|g\|_{L^2(R_x)} \leq \varepsilon$ . Note that

$\|g\|_{B_\infty^{1/2}(L^2(R))} \leq \|g\|_{H^{1/2}(R)} \leq 1$  due to the embedding of these spaces, and thus if we define  $g_h : R_x \rightarrow \mathbb{R}$  by  $g_h(x, y) := g(x + h, y)$ , we have

$$\|g_h - g\|_{L^2(R_x)} \leq h^{1/2}.$$

Then, the following bound on the norm of  $g$  on the whole rectangle  $R$  holds:

$$(5.8) \quad \|g\|_{L^2(R)} \leq \sqrt{2}(\varepsilon + h^{1/2}),$$

since

$$\begin{aligned} \|g\|_{L^2(R)}^2 &\leq \|g\|_{L^2(R_x)}^2 + \|g_h\|_{L^2(R_x)}^2 \leq \|g\|_{L^2(R_x)}^2 + [\|g_h - g\|_{L^2(R_x)} + \|g\|_{L^2(R_x)}]^2 \\ &\leq \varepsilon^2 + [h^{1/2} + \varepsilon]^2 < 2[h^{1/2} + \varepsilon]^2. \end{aligned}$$

Similar statements hold if  $R_x$  is replaced by  $R \setminus (a + h, a) \times (c, d)$ ,  $R \setminus (a, b) \times (d - h, d)$ , or  $R \setminus (a, b) \times (c, c + h)$ .

Now, we consider first the case of point evaluation. In this case, any function  $v$  appearing in the set of (5.7) is in  $U(H^{3/2}(\Omega))$  and vanishes on a grid of points with spacing  $h = m^{-1/2}$ . Standard Finite Element estimation shows that

$$(5.9) \quad \|v\|_{H^1(\Omega_m)} \leq Cm^{-1/4},$$

with  $C$  an absolute constant. To extend the above estimate to  $\Omega$ , we use (5.8) applied to appropriate rectangles obtained from  $\Omega_m$  and with  $h = m^{-1/2}$ ,  $\varepsilon = Cm^{-1/4}$  (see (5.9)), and  $g = v$ ,  $g = \partial_x v$ , or  $g = \partial_y v$ . Whence, we deduce that

$$\|v\|_{H^1(\Omega)} \leq Cm^{-1/4}$$

for a different absolute constant  $C$ .

We next consider the more intricate case of Gaussian measurements. To prove (5.7), we assume that (5.7) does not hold and derive a contradiction. So, assume there is a sequence of functions  $v_m \in U(\mathcal{H}^1(\Omega)) \cap \mathcal{N}_m$ ,  $m = 1, 2, \dots$ , that does not tend to 0 in  $H^1(\Omega)$ . This means that we can extract an increasing sequence  $(m_j)$  for which

$$(5.10) \quad \|v_{m_j}\|_{H^1(\Omega)} \geq \delta > 0, \quad j = 1, 2, \dots,$$

for some  $\delta > 0$ . Since  $\mathcal{H}^1(\Omega)$  is continuously embedded into  $H^{3/2}(\Omega)$ , and  $H^{3/2}(\Omega)$  is compactly embedded in  $H^s(\Omega)$  for each  $1 < s < \frac{3}{2}$ , we can further extract a subsequence (which we continue to denote by  $(m_j)$ ) such that  $v_{m_j}$  strongly converges towards a  $v^*$  in  $H^s(\Omega)$ . Note, in particular, that  $v^*$  is a continuous function, in fact a Lip  $s - 1$  function on  $\Omega$ , and we have

$$(5.11) \quad \|v^* - v_{m_j}\|_{C(\Omega)}, \quad \|v^* - v_{m_j}\|_{H^1(\Omega)} \leq A\|v^* - v_{m_j}\|_{H^s(\Omega)} \rightarrow 0, \quad j \rightarrow \infty,$$

with  $A$  an absolute constant. Hence,

$$(5.12) \quad \|v^*\|_{H^1(\Omega)} \geq \delta.$$

Let us now examine how the measurement functionals act on any continuous function  $v \in C(\Omega)$ . For any such  $v$ , we define

$$(5.13) \quad \tilde{v} := g_r * \mathcal{E}v,$$

where

$$g_r(z) := \frac{1}{\sqrt{2\pi r^2}} e^{-\frac{1}{2} \frac{|z|^2}{r^2}}$$

is the Gaussian function used for the measurements and  $\mathcal{E}$  is the extension operator by 0 outside of  $\Omega$ . The function  $\tilde{v}$  is a continuous function on  $\mathbb{R}^2$  and for any of our measurement functionals we have

$$(5.14) \quad \lambda_{i,j}(v) = \tilde{v}(z_{i,j}), \quad 1 \leq i \leq j.$$

We claim that  $\tilde{v}^*$  is identically zero on  $\Omega$ . Indeed, we have

$$(5.15) \quad |\tilde{v}^*(x)| \leq |\tilde{v}^*(x) - \tilde{v}_{m_j}(x)| + |\tilde{v}_{m_j}(x) - \tilde{v}_{m_j}(z)| + |\tilde{v}_{m_j}(z)|, \quad x, z \in \Omega.$$

Given any  $x \in \Omega$  and  $\varepsilon > 0$ , because of (5.11), we can make the first term in (5.15) smaller than  $\varepsilon/2$  for any sufficiently large  $j$ . Then by again taking  $j$  sufficiently large, there is a grid point  $z$  for this  $m_j$  that is sufficiently close to  $x$  so that the second term in (5.15) is also smaller than  $\varepsilon/2$ . Since  $\tilde{v}_{m_j}(z) = 0$  we see that  $|\tilde{v}^*(x)| < \varepsilon$ . This proves that  $\tilde{v}^* = 0$ .

Now that we know that  $\tilde{v}^*$  is identically zero on  $\Omega$ , our final step is to show that this implies that  $v^*$  is identically zero on  $\Omega$ . Denoting by  $\hat{f}$  the Fourier transform of  $f$ , it follows that

$$0 = \int_{\Omega} \tilde{v}^* v^* = \int_{\mathbb{R}^2} \tilde{v}^* \mathcal{E} v^* = (2\pi)^{-2} \int_{\mathbb{R}^2} \hat{g}_r |\widehat{\mathcal{E} v^*}|^2.$$

The positivity of  $\hat{g}_r$  implies that  $\mathcal{E} v^* = 0$  and therefore  $v^* = 0$  on  $\Omega$ . This contradicts (5.12) and is our desired contradiction.

Returning now to the general setting where  $X$  and the measurement sets  $\Lambda_m$  are general, we will have that (5.7) holds only if the sets  $\Lambda_m$  become sufficiently dense in  $X^*$  as  $m$  gets large. The precise rate of convergence of  $R(K^{\mathcal{H}})$  towards 0 will depend on the prior class  $\mathcal{H}$  and on the choice of the  $\Lambda_m$ ,  $m = 1, 2, \dots$ . In the next section we discuss in more detail the choice of  $\Lambda_m$ ,  $m = 1, 2, \dots$ , in order to obtain the best achievable rate.

**6. Optimal data sites: Gelfand widths and sampling numbers.** In this section, we make some comments on the number of measurements  $m$  that are needed to guarantee a prescribed error in the recovery of  $u$ . Bounds on  $m$  are known to be governed by the Gelfand width for the case of general linear functionals and by sampling numbers when the functionals are required to be point evaluations. We explain what is known about these quantities for our specific model classes. As we shall see, these issues are not completely settled for the model classes studied in this paper. The problem of finding the best choice of functionals, respectively point evaluations, is in need of further research.

We have seen that the accuracy of the optimal recovery of  $u \in K_w$  is given by the Chebyshev radius  $R(K_w) := R(K_w)_{H^1(\Omega)}$  or, equivalently,  $R(K_w^{\mathcal{H}}) := R(K_w^{\mathcal{H}})_{H^1(\Omega)}$  for the harmonic component. The worst case recovery error  $R(K)$  over the class  $K$  is defined by

$$(6.1) \quad R(K)_{H^1(\Omega)} := \sup_{w \in \mathbb{R}^m} R(K_w)_{H^1(\Omega)}.$$

Notice that this worst case error fixes the measurement functionals but allows the measurements  $w$  to come from any function in  $K$ . Both the individual error  $R(K_w)$

and the worst case error  $R(K)$  are very dependent on the choice of the data functionals  $\lambda_j$ . For example, in the case that these functionals are point evaluations at points  $z_1, \dots, z_m \in \bar{\Omega}$ , then  $R(K_w)$  and  $R(K)$  will depend very much on the positioning of these points in  $\bar{\Omega}$ .

In the case of general linear functionals, one may fix  $m$  and then search for the  $\lambda_1, \dots, \lambda_m$  that minimize the worst case recovery error over the class  $K$ . This minimal worst case error is called the *Gelfand width* of  $K$ . If we restrict the linear functionals to be given by point evaluation, we could correspondingly search for the sampling points  $x_1, \dots, x_m$  minimizing the worst case recovery error. This minimal error is called the *deterministic sampling number* of  $K$ .

The goal of this section is not to provide new results on Gelfand widths and sampling numbers, since we regard this as a separate issue in need of a systematic study, but to discuss what is known about them in our setting and refer the reader to the relevant papers. Let us recall that  $R(K_w)$  is equivalent to  $R(K_w^{\mathcal{H}})_{H^1}$  and so we restrict our discussion in what follows to sampling of harmonic functions.

**6.1. Optimal choice of functionals.** Suppose we fix the number  $m$  of observation to be allowed and ask what is the optimal choice for the  $\lambda_j$ ,  $j = 1, \dots, m$ , and what is the optimal error of recovery for this choice. The answer to the second question is given by the Gelfand width of  $K$ . Given a compact set  $K$  of a Banach space  $X$ , we define the *Gelfand width* of  $K$  in  $X$  by

$$(6.2) \quad d^m(K)_X := \inf_{\lambda_1, \dots, \lambda_m} R(K)_X,$$

where the infimum is taken over the linear functionals defined on  $X$ . Let us mention that this definition differs from that employed in the classical literature [23] where  $d^m(K)_X$  is defined as the infimum over all spaces  $F$  of codimension  $n$  of  $\max\{\|v\|_X : v \in K \cap F\}$ . The two definitions are equivalent in the case where  $K$  is a centrally symmetric set such that  $K + K \subset CK$  for some constant  $C \geq 1$ .

Any set of functionals which attains the infimum in (6.2) would be optimal. The Gelfand width is often used as a benchmark for performance since it says that no matter how the  $m$  functionals  $\lambda_1, \dots, \lambda_m$  are chosen, the error of recovery of  $u \in K$  cannot be better than  $d^m(K)_X$ .

When  $X$  is a Hilbert space and  $K$  is the ball of a Hilbert space  $Y$  with compact embedding in  $X$ , it is known that the Gelfand width coincides with the Kolmogorov width, that is

$$d^m(K)_X = d_m(K)_X := \inf_{\dim(E)=m} \text{dist}(K, E)_X = \inf_{\dim(E)=m} \max\{\|v - P_E v\|_X : v \in K\},$$

where the infimum is taken over all linear spaces  $E$  of dimension  $m$ . This is precisely our setting as discussed in section 3: taking  $X = H^1 := H^1(\Omega)$  and  $K$  as in (1.4), we have

$$(6.3) \quad d^m(K)_{H^1(\Omega)} = d^m(K^{\mathcal{H}})_{H^1(\Omega)} = d_m(K^{\mathcal{H}})_{H^1(\Omega)} \sim d_m(K^B)_{H^{1/2}(\Gamma)} = d^m(K^B)_{H^{1/2}(\Gamma)},$$

where the equivalence follows from (1.3). Upper and lower bounds for the Gelfand width of  $K^B$  in  $L_2(\Gamma)$  are explicitly given in [22].

We can estimate the rate of decay of the Kolmogorov and Gelfand width of  $K^B$  in  $H^{1/2}(\Gamma)$  by the following general argument: as explained in section 2.1, for the admissible range of smoothness, the Sobolev spaces  $H^s(\Gamma)$  have an intrinsic description

by locally mapping the boundary onto domains of  $\mathbb{R}^{d-1}$ . More precisely, in [19] and [12], the  $H^s(\Gamma)$  norm of  $g$  is defined as

$$(6.4) \quad \|g\|_{H^s(\Gamma)} := \left( \sum_{j=1}^J \|g_j\|_{H^s(R_j)}^2 \right)^{1/2},$$

where the  $R_j$  are open bounded rectangles of  $\mathbb{R}^{d-1}$  that are mapped by transforms  $\gamma_j$  into portions  $\Gamma_j$  that constitute a covering of  $\Gamma$ , and  $g_j = g \circ \gamma_j$  are the local pullbacks.

From this it readily follows that the Gelfand and Kolmogorov  $m$ -width of the unit ball of  $H^s(\Gamma)$  in the norm  $H^t(\Gamma)$ , with  $0 \leq t < s$  behaves similar to that of the unit ball of  $H^s(R)$  in the norm  $H^t(R)$  where  $R$  is a bounded rectangle of  $\mathbb{R}^{d-1}$ . The latter is known to behave like  $m^{-\frac{s-t}{d-1}}$ . Therefore, for  $K^{\mathcal{H}} = U(\mathcal{H}^s)$  with  $s > \frac{1}{2}$  in the admissible range allowed by the boundary smoothness, one has

$$(6.5) \quad cm^{-\frac{s-1/2}{d-1}} \leq d^m(K^{\mathcal{H}})_{H^1(\Omega)} \leq Cm^{-\frac{s-1/2}{d-1}}, \quad m \geq 1,$$

where  $c$  and  $C$  are positive constants depending only on  $\Omega$  and  $s$ .

*Remark 6.1.* We have already observed in section 2 that the space  $\mathcal{H}^s(\Omega)$  is continuously embedded in the Sobolev space  $H^r(\Omega)$  with  $r := \max\{s + \frac{1}{2}, r^*\}$  and, in particular,  $r = s + \frac{1}{2}$  for smooth domains. However, the Gelfand and Kolmogorov widths of the unit ball of  $H^r(\Omega)$  in  $H^1(\Omega)$  have the slower decay rate  $m^{-\frac{r-1}{d}} = m^{-\frac{s-1/2}{d}}$  compared to (6.5) for  $\mathcal{H}^s(\Omega)$ . This improvement reflects the fact that the functions from  $\mathcal{H}^s(\Omega)$  have  $d$  variables but are, in fact, determined by functions of  $d-1$  variables. The reduction in dimension from  $d$  to  $d-1$  is related to the fact that in our formulation of our problem we have complete knowledge of  $f$ .

**6.2. Optimal choice of sampling points.** We turn to the particular setting where the  $\lambda_j$  are point evaluations functionals,

$$\lambda_j(v) = v(x_j)$$

at  $m$  points  $x_j \in \overline{\Omega}$ . Similar to the Gelfand width, the *deterministic sampling numbers* are defined as

$$(6.6) \quad \rho^m(K)_X := \inf_{x_1, \dots, x_m} R(K)_X.$$

A variant of this is to measure the worst case expected recovery error when the  $m$  points are chosen at random according to a probability distribution and search for the distribution that minimizes this error, leading to the *randomized sampling number* of  $K$ . Obviously, one has

$$(6.7) \quad \rho^m(K)_X \geq d^m(K)_X.$$

In the majority of the literature, deterministic and randomized sampling numbers are studied with error measured in the  $L_2(\Omega)$  norm. In this setting, concrete strategies for optimal deterministic and randomized point design have been given when  $K$  is the unit ball of a reproducing kernel Hilbert space  $H$  defined on  $\Omega$ . In particular, the recent results in [18, 14, 20, 7] show that under the assumption

$$\sum_{m>0} |d^m(K)_{L^2(\Omega)}|^2 < \infty,$$

then, for all  $t > \frac{1}{2}$ ,

$$\sup_{m \geq 1} m^t d^m(K)_{L^2(\Omega)} < \infty \implies \sup_{m \geq 1} m^t \rho^m(K)_{L^2(\Omega)} < \infty.$$

In words, under the above assumptions, optimal recovery in  $L_2(\Omega)$  has the same algebraic convergence rate when using optimally chosen point values compared to an optimal choice of general linear functionals.

While similar general results have not been established for Gelfand width and sampling numbers in the  $H^1$  norm, we argue that they hold in our particular setting where  $H = \mathcal{H}^s(\Omega)$ . For simplicity, as in section 4, we consider a domain that is either a polygon when  $d = 2$  or polyhedron when  $d = 3$ , and thus consider the range  $\frac{d-1}{2} < s < \frac{3}{2}$  where the restriction from below ensures that  $\mathcal{H}^s(\Omega) \subset C(\bar{\Omega})$ . Recalling the finite element spaces  $\mathbb{V}_h$  and their traces  $\mathbb{T}_h$  on the boundary, based on quasi-uniform meshes  $\{\mathcal{T}_h\}_{h>0}$ , we consider for a given  $h > 0$  the measurement points  $x_1, \dots, x_m$  that are the mesh vertices located on  $\Gamma$ . By the quasi-uniformity property the number  $m = m(h)$  of these points satisfies

$$ch^{1-d} \leq m \leq Ch^{1-d}$$

for some  $c, C > 0$  independent of  $h$ . If  $v \in \mathcal{H}^s(\Omega)$ , its trace  $v_\Gamma$  belongs to  $H^s(\Gamma)$ . Then, denoting by  $I_h$  the piecewise linear interpolant on the boundary, standard finite element approximation theory ensures the estimate

$$\|v_\Gamma - I_h v_\Gamma\|_{H^{1/2}(\Gamma)} \leq Ch^{s-\frac{1}{2}} \|v_\Gamma\|_{H^s(\Gamma)} = Ch^{s-\frac{1}{2}} \|v\|_{\mathcal{H}^s(\Omega)}$$

for some  $C$  that only depends on  $s$ . Therefore, introducing  $\tilde{v} := EI_h v$ , one has

$$\|v - \tilde{v}\|_{H^1(\Omega)} \leq C_E \|v_\Gamma - I_h v_\Gamma\|_{H^{1/2}(\Gamma)} \leq CD_E m^{-\frac{s-1/2}{d-1}} \|v\|_{\mathcal{H}^s(\Omega)}.$$

Since  $\tilde{v}$  only depends on the value of  $v$  at the points  $x_1, \dots, x_m$ , we have thus proved an upper bound of order  $m^{-\frac{s-1/2}{d-1}}$  for  $\rho^m(K^H)_{H^1(\Omega)}$ , and in turn for  $\rho^m(K)_{H^1(\Omega)}$ . In view of (6.7) and (6.5), a lower bound of the same order must hold. In summary, similar to the Gelfand widths, the sampling numbers satisfy

$$(6.8) \quad \tilde{c} m^{-\frac{s-1/2}{d-1}} \leq \rho^m(K)_{H^1(\Omega)} \leq \tilde{C} m^{-\frac{s-1/2}{d-1}}, \quad m \geq 1,$$

where  $\tilde{c}$  and  $\tilde{C}$  are positive constants depending only on  $\Omega$  and  $s$ .

## REFERENCES

- [1] R. A. ADAM AND J. F. FOURNIER, *Sobolev Spaces*, Elsevier, Amsterdam, 2003.
- [2] B. ADCOCK AND D. HUYBRECHS, *Frames and numerical approximation*, SIAM Rev., 61 (2019), pp. 443–473, <https://doi.org/10.1137/17M1114697>.
- [3] G. AUCHMUTY, *Reproducing kernels for Hilbert spaces of real harmonic functions*, SIAM J. Numer. Anal., 41 (2009), pp. 1994–2009, <https://doi.org/10.1137/080739628>.
- [4] S. L. BRUNTON, B. R. NOACK, AND P. KOUMOUTSAKOS, *Machine learning for fluid mechanics*, Annu. Rev. Fluid Mech., 36 (2020), pp. 477–508.
- [5] P. G. CIARLET AND P. A. RAVIART, *Maximum principle and uniform convergence for the finite element method*, Comput. Methods Appl. Mech. Engrg., 2 (1973), pp. 17–31.
- [6] R. DEVORE, G. PETROVA, AND P. WOJTASZCZYK, *Data Assimilation in Banach Spaces*, preprint, <https://arxiv.org/abs/1602.06342>, 2016.
- [7] M. DOLBEAULT, D. KRIEG, AND M. ULLRICH, *A Sharp Upper Bound for Sampling Numbers in  $L^2$* , preprint, <https://arxiv.org/abs/2204.12621>, 2022.

- [8] V. DOMÍNGUEZ AND F.-J. SAYAS, *Stability of discrete liftings*, C. R. Acad. Sci. Paris I, 337 (2003), pp. 885–898.
- [9] K. DURAISAMY, G. IACCARINO, AND H. XIAO, *Turbulence modeling in the age of data*, Annu. Rev. Fluid Mech., 51 (2019), pp. 357–377.
- [10] L. FORMAGGIA, J.-F. GERBEAU, F. NOBILE, AND A. QUARTERONI, *Numerical treatment of defective boundary conditions for the Navier–Stokes equations*, SIAM J. Numer. Anal., 40 (2002), pp. 376–401, <https://doi.org/10.1137/S003614290038296X>.
- [11] A. N. GALYBIN AND J. IRSIA, *On reconstruction of three-dimensional harmonic functions from discrete data*, Proc. Roy. Soc. A, 466 (2010), pp. 1935–1955.
- [12] P. GRISVARD, *Elliptic Problems in Non-smooth Domains*, Pitman, Boston, 1985.
- [13] L. GRINBERGAND AND G. E. KARNIADAKIS, *Outflow boundary conditions for arterial networks with multiple outlets*, Ann. Biomed. Eng., 51 (2008), pp. 1496–1514.
- [14] D. KRIEG AND M. ULLRICH, *Function values are enough for  $L_2$ -approximation*, Found. Comput. Math., 21 (2021), pp. 1141–1151.
- [15] D. LEYKEKHMAN AND B. LI, *Weak discrete maximum principle of finite element methods in convex polyhedra*, Math. Comp., 90 (2021), pp. 1–18.
- [16] A. MELKMAN AND C. MICCHELLI, *Optimal estimation of linear operators in Hilbert spaces from inaccurate data*, SIAM J. Numer. Anal., 16 (1979), pp. 87–105, <https://doi.org/10.1137/0716007>.
- [17] C. MICCHELLI AND T. RIVLIN, *A survey of optimal recovery*, in Optimal Estimation in Approximation Theory, C. A. Micchelli and T. J. Rivlin, eds., Plenum Press, New York, 1977, pp. 1–54.
- [18] M. MOELLER AND T. ULLRICH,  *$L_2$ -norm sampling discretization and recovery of functions from RKHS with finite trace*, Sampl. Theory Signal Process. Data Anal., 19 (2021), 13.
- [19] J. NECAS, *Direct Methods in the Theory of Elliptic Equations*, English translation of the 1967 monograph, Springer Monogr. Math., Springer, Heidelberg, 2012.
- [20] N. NAGEL, M. SCHÄFER, AND T. ULLRICH, *A new upper bound for sampling numbers*, Found. Comput. Math., 22 (2022), pp. 445–468.
- [21] E. NOVAK AND H. WOZNIAKOWSKI, *Tractability of Multivariate Problems, Vol. I: Linear Information*, EMS Tracts in Mathematics 6, Eur. Math. Soc. Publ. House, Zürich, 2008.
- [22] I. PESENSON, *Estimates of Kolmogorov, Gelfand, and linear widths on compact Riemannian manifolds*, Proc. AMS, 144 (2016), pp. 2985–2998.
- [23] A. PINKUS, *N-Widths in Approximation Theory*, Ergeb. Math. Grenzgeb. (3) 7, Springer, Berlin, Heidelberg, 2012.
- [24] K. RAJAGOPAL, *On boundary conditions for fluids of the differential type*, in Stokes Equations and Related Non-linear Problems, A. Sequeira, ed., Plenum Press, New York, 1995, pp. 273–278.
- [25] P. J. RICHARDS AND S. E. NORRIS, *Appropriate boundary conditions for computational wind engineering models revisited*, J. Wind Eng. Ind. Aerodyn., 99 (2011), pp. 257–266.
- [26] P. J. RICHARDS AND S. E. NORRIS, *Appropriate boundary conditions for computational wind engineering: Still an issue after 25 years*, J. Wind Eng. Ind. Aerodyn., 190 (2019), pp. 245–255.
- [27] A. H. SCHATZ, *A weak discrete maximum principle and stability of the finite element method in  $L_\infty$  on plane polygonal domains*, I, Math. Comp., 34 (1980), pp. 77–91.
- [28] L. R. SCOTT AND SH. ZHANG, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.
- [29] J. TRAUB AND H. WOZNIAKOWSKI, *A General Theory of Optimal Algorithms*, Academic Press, 1980.
- [30] H. XU, W. ZHANG, AND Y. WANG, *Explore missing flow dynamics by physics-informed deep learning: The parameterized governing systems*, Phys. Fluids, 33 (2021), 095116.
- [31] K. YOSIDA, *Functional Analysis*, Springer-Verlag, Berlin, 2012.