Sparse grids for boundary integral equations

M. Griebel*, P. Oswald†, T. Schiekofer*

†Bell Laboratories, Lucent Technologies
600 Mountain Av., Murray Hill, NJ 07974, USA
poswald@research.bell-labs.com

*University Bonn, Institute for Applied Mathematics
Wegelerstr. 6, 53115 Bonn, Germany
{griebel,schiekof}@iam.uni-bonn.de

Dedicated to Olof B. Widlund on the occasion of his 60th birthday

Abstract

The potential of sparse grid discretizations for solving boundary integral equations is studied for the screen problem on a square in $\mathbb{R}^3$. Theoretical and numerical results on approximation rates, preconditioning, adaptivity and compression for piecewise constant and linear sparse grid spaces are obtained.

Classification: 45L10, 65N38, 65R20, 65Y20
Keywords: boundary element method, sparse grids, adaptivity, prewavelets, matrix compression

1 Introduction

This is a case study for some special boundary integral equations on a two-dimensional manifold Γ in $\mathbb{R}^3$ (screen problems). We will focus on the example of a two-dimensional unit square in $\mathbb{R}^2$ embedded into $\mathbb{R}^3$ where

$$\Gamma = \{x \,:\, (x_1, x_2) \in [0,1]^2,\; x_3 = 0\}\;. \tag{1}$$

In general, $d\Gamma_x$ stands for the surface Lebesgue measure with respect to the variable x, $|x|_2$ denotes the Euclidean norm of x, and n_x is the vector field of normal vectors associated with Γ. We specifically have in mind the *single layer potential equation*

$$\frac{1}{4\pi}\int_\Gamma \frac{f(y)}{|x-y|_2}\, d\Gamma_y = g(x)\;, \tag{2}$$

which we will write as operator equation $Vf = g$. The problem (2) is essentially equivalent to the Dirichlet problem for the Laplace equation in the exterior domain $\mathbb{R}^3\backslash\Gamma$, with Dirichlet data g on Γ. The order of the pseudodifferential operator V is -1, it maps $\tilde{H}^{-1/2}(\Gamma)$ onto $H^{1/2}(\Gamma)$, and leads to a symmetric $\tilde{H}^{-1/2}(\Gamma)$-elliptic variational problem. For the definition of the spaces and more information on the solution and regularity theory of the equation (2), see subsections 2.1 and 2.2.

Another case of interest is the *hypersingular equation*

$$\frac{1}{4\pi}\int_\Gamma \frac{\partial}{\partial n_x}\frac{\partial}{\partial n_y}\left(\frac{1}{|x-y|_2}\right) f(y)\, d\Gamma_y = g(x)\;. \tag{3}$$

Written as operator equation $Df = g$, one has that $D : \tilde{H}^{1/2}(\Gamma) \to H^{-1/2}(\Gamma)$. The pseudodifferential operator D is of order 1, and leads to a symmetric variational problem which is $\tilde{H}^{1/2}(\Gamma)$-elliptic.

We aim at investigating the potential of sparse grid spaces [50, 2] for the solution of such boundary integral equations and related elliptic problems in Sobolev spaces on a tensor product domain in $\mathbb{R}^d$. For simplicity, the exposition is given for the unit cube $\Omega = I^d$ ($I \equiv [0,1]$, $d \geq 2$). What we call sparse grid spaces, are specific tensor product constructions which start from a univariate orthogonal splitting

$$L_2(I) = \bigoplus_{j \geq 0} W_j \ , \tag{4}$$

where the $W_j \subset L_2(I)$ are finite-dimensional subspaces of finite dimension $m_j \equiv \dim W_j$. In each W_j, we fix a basis $\Psi_j = \{\psi_{j,i} \ : \ i = 0, \ldots, m_j - 1\}$ for which we assume L_2-stability

$$\| \sum_i c_{j,i} \psi_{j,i} \|_{L_2}^2 \approx \sum_i c_{j,i}^2 \ , \tag{5}$$

uniformly in $\{c_{j,i}\}$ and $j \geq 0$. Throughout the paper, $A \approx B$ stands for a two-sided inequality between the expressions A and B, i.e., $A \leq C \cdot B$ and $B \leq C \cdot A$, where $C > 0$ denotes a generic constant the value of which may change with each appearance. Dependencies of such constants C on parameters will be indicated (or should be clear from the context). To rephrase (5), we assume that

$$\Phi = \bigcup_{j \geq 0} \Psi_j = \{\psi_{j,i} \ : \ i = 0, \ldots, m_j - 1, \ j = 0, 1, \ldots\} \tag{6}$$

forms a *semi-orthogonal Riesz basis* in $L_2(I)$. We also introduce the notation

$$V_k = \bigoplus_{j=0}^{k} W_j \ , \quad \Phi_k = \bigcup_{j=0}^{k} \Psi_j \ , \quad k = 0, 1, \ldots \ , \tag{7}$$

where the V_k form now an increasing, dense sequence of subspaces in $L_2(I)$, with Riesz bases Φ_k.

With this at hand, multivariate systems and space decompositions for $L_2(I^d)$ and Sobolev spaces on I^d can be constructed by tensor product techniques (for more details and specific examples, see section 2 and subsection 3.1). The traditional *full grid spaces* are associated with rectangular (or square) index sets. Set

$$V_{\mathbf{k}} = V_{k_1} \otimes \ldots \otimes V_{k_d} = \bigoplus_{\mathbf{0} \leq \mathbf{j} \leq \mathbf{k}} W_{\mathbf{j}}$$

and specifically $V_k(I^d) = V_{(k,\ldots,k)}$. As usual, tensor products of univariate function spaces are identified with function spaces on the corresponding product domain. E.g., bases in the pairwise orthogonal subspaces

$$W_{\mathbf{j}} = W_{j_1} \otimes \ldots \otimes W_{j_d}$$

are given by

$$\Psi_{\mathbf{j}} = \{\psi_{\mathbf{j},\mathbf{i}}(x) = \psi_{j_1,i_1}(x_1) \cdot \ldots \cdot \psi_{j_d,i_d}(x_d) : \mathbf{0} \leq \mathbf{i} \leq \mathbf{m_j} - \mathbf{1}\} \quad (\mathbf{m_j} = (m_{j_1}, \ldots, m_{j_d})) \ .$$

The complete system

$$\Phi^d = \bigcup_{\mathbf{j} \geq \mathbf{0}} \Psi_{\mathbf{j}}$$

forms again a semi-orthogonal Riesz basis in $L_2(I^d)$. Obviously, the finite subset $\Phi^d_{\mathbf{k}} = \Phi^d \cap V_{\mathbf{k}}$ of Φ^d forms a basis in $V_{\mathbf{k}}$, $\mathbf{k} \in \mathbb{Z}^d_+$, analogously for $V_k(I^d)$.

We will work with univariate subspaces of exponentially growing dimension, i.e., we assume

$$m_j \approx a^j \ , \qquad j \to \infty \ , \tag{8}$$

for some $a > 1$, typically for $a = 2$ (the so called *dyadic refinement case*). This implies

$$\dim W_{\mathbf{j}} = m_{j_1} \ldots m_{j_d} \approx a^{|\mathbf{j}|} , \quad \dim V_{\mathbf{k}} \approx a^{|\mathbf{k}|} , \quad \dim V_k(I^d) \approx a^{dk} , \tag{9}$$

where $|\mathbf{j}| = j_1 + \ldots + j_d$, $\mathbf{j} \in \mathbb{Z}_+^d$, and shows that the dimension of full grid spaces grows exponentially with the dimension d of the coordinate space. This is known as the curse of dimension. An alternative are *sparse grid spaces* which, in the simplest version, are defined with respect to a triangular index set:

$$\hat{V}_k(I^d) = \bigoplus_{|\mathbf{j}| \le k} W_{\mathbf{j}} = \sum_{|\mathbf{k}| = k} V_{\mathbf{k}} . \tag{10}$$

The latter representation as a sum of special full grid spaces is not a direct one. Nevertheless, a simple calculation shows that

$$\dim \hat{V}_k(I^d) \approx k^{d-1} a^k , \tag{11}$$

which favorably compares with the dimension of the full grid space $V_k(I^d)$ since the approximation power of the two types of discretization spaces is often the same for certain classes of smooth functions. The interplay between dimension and approximation will be discussed in more detail in section 2.

Although approximation with sparse grid spaces has been tried under various names in several application fields (Smolyak quadratures, hyperbolic cross approximation, Boolean blending schemes), the use for the discretization of operator equations is more recent. Several papers have appeared on their use for the discretization of partial differential equations, mostly on the basis of low order finite element constructions for second order boundary value problems [2, 19, 21, 50]. To our knowledge, for the closely related class of boundary integral equations of potential theory, except of [15] where smooth kernels are treated, no papers on the sparse grid technique have appeared so far. The main bulk of recent activities in this area have concentrated on novel methods such as the panel clustering approach [26, 27, 16], the fast multipole method [17, 18, 41] or wavelet schemes including operator or matrix compression [1], as well as on the p- and hp-version of the boundary element method. Both these directions have their definite merits but also a common drawback: they require a complicated implementation. E.g., the main practical obstacle of the wavelet compression methods are the complicated rules in the matrix compression step without which the optimal work estimates of the wavelet algorithms would be lost. We refer to recent work by W. Dahmen, R. Schneider, C. Schwab, T. von Petersdorff et al. [8, 9, 37, 38, 44]. For a survey of recent activities on the p- and hp-method for boundary integral equations, see [28, 29, 46].

For example, for a screen problem on a square as described above, a wavelet discretization (as many other methods based on full grid spaces $V_k(I^2)$) would lead to a discrete problem of dimension $\approx a^{2k}$ described by a dense linear system with a matrix containing $\approx a^{4k}$ nonzero entries. Thus, any iterative method would cost $\ge ca^{4k}$ arithmetical operations per iteration which is prohibitive and implies the necessity of matrix compression (or similar strategies such as panel clustering). It is possible to sparsify the involved matrix-vector multiplications such that approximate solutions (of the same order of approximation $\mathrm{O}(a^{-kt})$, with t depending on the error norm and the regularity of the solution) can be found within $\mathrm{O}(k^r a^{2k})$ arithmetical operations. Here, r is an integer which in most of the papers containing practicable compression schemes varies between 1 and 4, even though theory predicts that one could set $r = 0$. I.e., theoretically, a complexity of $\mathrm{O}(a^{2k})$ arithmetical operations could be reached, provided that the compression and iteration method are optimally implemented.

Our starting point was the following observation: If we replace the full grid space of a wavelet method for I^2 by the corresponding sparse grid space, then under some regularity assumptions (which need to be investigated more thoroughly) we might reach an approximation order close to $\mathrm{O}(a^{-kt})$ by looking for the Galerkin solution in the sparse grid space. The size of the corresponding linear system would be $\approx ka^k$, with an overall of $\approx k^2 a^{2k}$ coefficients in the stiffness matrix. Although the economical computation of these coefficients within machine precision is an important aspect, we will assume for a moment that for a certain choice of a basis in $\hat{V}_k(I^2)$ the stiffness matrix has been computed in $\mathrm{O}(k^{\hat{r}} a^{2k})$ operations where $\hat{r} \ge 0$ is fixed. Multilevel preconditioners, providing uniformly bounded condition numbers for the preconditioned linear system and needing only $\mathrm{O}(ka^k)$ operations for each preconditioning step, can be constructed following [20, 21]. Thus, without implementing any ingenious compression rules, we might end up with an algorithm of almost the same arithmetical complexity for this dense system as the optimal wavelet compression scheme. This clearly requires a close look at all details and a thorough performance testing. E.g., so far our discussion was based on the assumption of a certain extra global regularity of the solutions which is in contrast to the appearance of corner and edge singularities for this kind of boundary integral equations.

The paper is organized as follows. Section 2 contains information on the regularity behavior of the solutions of the above model screen problem on a square, about various Sobolev classes, and their relationship to approximation spaces defined by decomposition norms with respect to the systems $\Phi(I^d)$. In subsection 2.4, an optimal preconditioning result for elliptic problems in subspaces of Sobolev spaces of order $-r < s < r$ is derived under rather general conditions on the systems Ψ_j (the parameter $r > 0$ depends on Ψ_j). On the other hand, (best) approximation estimates for these spaces show some significant differences in their behavior. These estimates depend on the ellipticity order s, the type of subspace (full or sparse grid spaces), and the regularity assumptions. It turns out that for approximation in Sobolev spaces of negative order (e.g., in the case $s = -1/2$ we are interested in when dealing with (2)), there appears a dimension-dependent penalty for using sparse grid spaces $\hat{V}_k(I^d)$ in comparison with the full grid spaces $V_k(I^d)$ which slightly damps our expectations. See subsection 2.3 for details.

In section 3 our numerical experiments are presented. The examples of univariate Riesz bases Φ used in this part are based on piecewise constant and linear spline functions. In subsection 3.1 we give details on these spline prewavelet systems. The exact computation of Galerkin stiffness matrices associated with (2) is outlined in subsection 3.2. Subsection 3.3 contains the results of the numerical experiments for capacity calculations (i.e., the solution of (2) with $g \equiv 1$) using sparse grid spaces and their comparison with results for the larger full grid spaces. In subsection 3.4 we construct adapted sparse grid spaces which, by adding more refinement only near the edges, significantly enhances the approximation properties of the scheme. Our preliminary testing allowed us to obtain capacity approximations of relative accuracy up to 10^{-4}, i.e., in the non-asymptotical range, with slightly less unknowns than reported for the hp-method in [28].

2 Theory: Approximation and Preconditioning

2.1 Definitions and preliminaries

Without further mentioning, we will use the notation given in the above introduction. Concerning the semi-orthogonal Riesz system Φ, we will assume for simplicity that (8) holds from now on. Other assumptions will be added later.

Given an index set $\mathcal{J} \subset \mathbb{Z}_+^d$, set

$$V_{\mathcal{J}} = \bigoplus_{\mathbf{j} \in \mathcal{J}} W_{\mathbf{j}} \, . \tag{12}$$

Partial cases are the index sets $\mathcal{J}_k = \{\mathbf{j} \in \mathbb{Z}_+^d \, : \, |\mathbf{j}|_\infty \le k\}$ and $\hat{\mathcal{J}}_k = \{\mathbf{j} \in \mathbb{Z}_+^d \, : \, |\mathbf{j}| \le k\}$ which yield the full grid spaces $V_k(I^d)$ and the sparse grid spaces $\hat{V}_k(I^d)$. To fix the notation, for multi-indices $\mathbf{j} \in \mathbb{Z}_+^d$ we set

$$|\mathbf{j}| = \sum_{l=1}^{d} j_l \quad \text{and} \quad |\mathbf{j}|_\infty = \max_{l=1,\ldots,d} j_l \, .$$

More general choices have appeared in the literature, see [2, 4]. We start with a simple lemma.

Lemma 1 *Under the assumption (8), the dimension $n_{\mathcal{J}}$ of $V_{\mathcal{J}}$ satisfies*

$$n_{\mathcal{J}} = \sum_{\mathbf{j} \in \mathcal{J}} m_{j_1} \ldots m_{j_d} \, .$$

In particular,

$$n_{k,d} \equiv \dim V_k(I^d) \approx a^{kd} , \qquad \hat{n}_{k,d} \equiv \dim \hat{V}_k(I^d) \approx k^{d-1} a^k .$$

The Sobolev spaces of interest for the study of integral equations and sparse grid approximations will be listed next. We assume that the reader is familiar with Sobolev spaces on $\mathbb{R}^d$, compare [30, 48]. Set

$$H^s(I^d) \equiv H^s(\mathbb{R}^d)|_{I^d} .$$

I.e., $g \in \mathcal{D}'(I^d)$ belongs to $H^s(I^d)$ if it is the restriction of some $f \in H^s(\mathbb{R}^d)$ to the cube I^d, and

$$\|g\|_{H^s(I^d)} = \inf_{g=f|_{I^d}} \|f\|_{H^s(\mathbb{R}^d)} .$$

We also need

$$\tilde{H}^s(I^d) = \{g = f|_{I^d} \ : \ f \in H^s(\mathbb{R}^d), \ \mathrm{supp}\, f \subset I^d\} ,$$

equipped with the norm $\|g\|_{\tilde{H}^s(I^d)} = \|f\|_{H^s(\mathbb{R}^d)}$. Properties of these *isotropic Sobolev spaces* are investigated in [48, 30]. E.g., the two scales are dual to each other:

$$(H^s(I^d))' = \tilde{H}^{-s}(I^d) , \quad (\tilde{H}^s(I^d))' = H^{-s}(I^d) , \quad -\infty < s < \infty . \tag{13}$$

The scales are closed under complex interpolation. The essential difference is that $\tilde{H}^s(I^d)$ mimics the case of homogeneous essential boundary conditions:

$$\tilde{H}^s(I^d) = H_0^s(I^d) \equiv \mathrm{clos}_{H^s(I^d)}\, C_0^\infty(I^d) , \ s > \frac{1}{2}, \ s \neq \frac{3}{2}, \frac{5}{2}, \dots .$$

In addition to isotropic Sobolev spaces, we need to consider also *Sobolev spaces of functions with dominating mixed derivative*, see [43]. Given a vector of smoothness exponents $\mathbf{s} = (s_1, \dots, s_d) \in \mathbb{R}^d$, these spaces are defined by the following tensor product construction:

$$H^{\mathbf{s}}_{\mathrm{mix}}(I^d) = H^{s_1}(I) \otimes \ldots \otimes H^{s_d}(I) . \tag{14}$$

Since we are exclusively dealing with tensor products of Hilbert spaces, the definitions for tensor product spaces can be given most directly in terms of orthogonal decompositions [49] which fits well with our semi-orthogonal splittings $\{W_j\}$ and systems Φ, respectively. Analogously, we could introduce $\tilde{H}^{\mathbf{s}}_{\mathrm{mix}}(I^d)$. In the isotropic case ($s_1 = \ldots = s_d = s$), we use the notation $H^s_{\mathrm{mix}}(I^d)$ and $\tilde{H}^s_{\mathrm{mix}}(I^d)$, respectively. These spaces play a key role in the theory of hyperbolic cross approximation in the periodic case [47] which is similar to sparse grid approximation as used in this paper. As a particular case, we mention that $H^2_{\mathrm{mix}}(I^2)$ possesses the explicit equivalent norm

$$\|f\|^2_{H^2_{\mathrm{mix}}(I^2)} \approx \|f\|^2_{L^2(I^2)} + \|D^{(2,0)}f\|^2_{L^2(I^2)} + \|D^{(0,2)}f\|^2_{L^2(I^2)} + \|D^{(2,2)}f\|^2_{L^2(I^2)} .$$

The last term involves a mixed fourth–order derivative and describes the additional smoothness requirement of functions from H^2_{mix} compared to the larger isotropic Sobolev space H^2. Equivalent norms for the above Sobolev spaces based on decomposition techniques will be discussed in subsection 2.3 below.

2.2 Solution and regularity theory

This subsection collects some known facts on the solution and regularity theory for the equation (2) on a square (similar results are known for (3) and more general polyhedral screens Γ, see [39, 40]). We will therefore identify $\Gamma = I^2$. The following proposition is cited from [28], compare also [7, 29, 39, 40, 45]. For a general introduction, see [24, Chapter 8].

Proposition 2 a) *For any $g \in H^{1/2}(I^2)$, there exists a unique solution $f \in \tilde{H}^{-1/2}(I^2)$ of the variational problem*

$$a_V(f,h) = \langle g,h\rangle_{H^{1/2}\times\tilde{H}^{-1/2}} \qquad \forall\, h \in \tilde{H}^{-1/2}(I^2) \tag{15}$$

associated with the single-layer potential equation (2). The bilinear form

$$a_V(f,h) \equiv \langle Vf,h\rangle_{H^{1/2}\times\tilde{H}^{-1/2}} \ , \qquad f,h \in \tilde{H}^{-1/2}(I^2) \ ,$$

generated by V is symmetric and $\tilde{H}^{-1/2}$-elliptic.
b) *If $g \in H^2(I^2)$, the following decomposition into regular and singular parts holds for the solution f of (2) and (15),respectively:*

$$f = f_0^{\mathrm{reg}} + \sum_{i=1}^{4} f_i^{\mathrm{sing}} \ , \tag{16}$$

where $f_0^{\mathrm{reg}} \in H^{1-\varepsilon}(I^2)$ is the regular part. The singular parts f_i^{sing} associated with the corners of the square are similar for all four corners, and split further into corner and corner-edge singularity components (multiplied by suitable cut-off functions). The corner singularity component associated with $(0,0)$, written in polar coordinates (r,θ), takes the form

$$f^{\mathrm{c}}(r,\theta) = b_1 r^{\gamma-1}\omega(\theta) \ , \quad \omega(\theta) \in H^{1-\varepsilon}[0,\pi/2] \ , \tag{17}$$

while the corner-edge singularity component associated with $(0,0)$ and with the edge along the axis $\theta = 0$ can be written as

$$f^{\mathrm{ce}}(r,\theta) = (b_2 r^{\gamma-1} + r^{-1}c(r))\sin^{-1/2}\theta \ , \quad r^{-1/2}c'(r),\ r^{-3/2}c(r) \in L_2(I) \ . \tag{18}$$

Here, b_1, b_2 are constants depending on g, $\gamma = 0.2966...$ is an absolute constant, while $\varepsilon > 0$ can be taken arbitrarily small. Analogous representations hold for the other corners and corner-edge combinations.

Due to the presence of edge and corner singularities as described by Proposition 2 b), solutions of the single layer potential equation (2) possess very low global regularity. Even for very smooth data g (e.g. $g = 1$ as considered in section 3) the solution f of (15) only belongs to $H^{-\varepsilon}(I^2)$, but not to $L_2(I^2)$. As is stated in [28], the results of [39, 40] which are the basis for the above proposition can be extended to higher regularity. E.g., if $g \in H^t(I^2)$ for $t > 2$ is assumed in Proposition 2 then (16) holds with $f_0^{\mathrm{reg}} \in \tilde{H}^{t-1-\varepsilon}$, on the expense of including more special terms into the singular parts. Since the new singularity components are necessarily included into $H^{1-\varepsilon}(I^2)$, the main singularity behavior remains the same as described by (17), (18). The decomposition (16) (more precisely, the leading singularity terms derived from (17), (18)) have been used to guide the correct mesh refinement in adaptive h- and hp-schemes based on graded meshes [28, 31]. This issue will be taken up in connection with adaptive sparse grid spaces in subsection 3.4.

2.3 Decomposition norms and approximation

We summarize some results from the theory of approximation spaces which are needed to produce stable splittings for the Sobolev spaces introduced in subsection 2.1. Our sources are [21, 34, 35, 36]. Since we are not interested in the most general results, we concentrate on assumptions which are easy to verify and cover our applications. Let us first assume that our univariate system Φ is not only a semi-orthogonal Riesz system in $L_2(I)$ but also provides approximation order m, i.e., satisfies a Jackson type estimate

$$\inf_{v_k \in V_k(I)} \|f - v_k\|_{L_2(I)} \le Ca^{-km}|f|_{H^m(I)} \qquad \forall\, f \in H^m(I) \ , \tag{19}$$

and possesses a Bernstein inequality

$$\|v_k\|_{H^r(I)} \le Ca^{kr}\|v_k\|_{L_2(I)} \qquad \forall\, v_k \in V_k(I) \ . \tag{20}$$

Here, $m \geq 1$ is an integer, and $r \in (0, m]$ is a real number. Examples will be discussed in subsection 3.1.

Under these assumptions, along the lines of [35, 36] one easily proves

Proposition 3 *Let Φ be as introduced in section 1, and assume that (19), (20) hold as stated above.*
a) *After suitable scaling, the system Φ forms a Riesz basis in the spaces $H^s(I)$ for $0 \leq s < r$. In particular, one has*

$$\|f\|^2_{H^s(I)} \approx \sum_{j=0}^{\infty} a^{2sj} \|w_j\|^2_{L_2(I)}\,, \quad 0 \leq s < r\,, \tag{21}$$

where

$$f = \sum_{j=0}^{\infty} w_j\,, \quad w_j \in W_j\,,\ j \geq 0\,, \tag{22}$$

is the unique decomposition of $f \in H^s(I)$ with respect to the L_2-orthogonal subspace system $\{W_j\}$.
b) *By a duality consideration, the appropriately scaled system Φ forms a Riesz basis for $\tilde{H}^s(I)$ and $-r < s < 0$:*

$$\|f\|^2_{\tilde{H}^s(I)} \approx \sum_{j=0}^{\infty} a^{2sj} \|w_j\|^2_{L_2(I)} \qquad \forall\ f \in \tilde{H}^s(I) \tag{23}$$

(here the L_2-convergence in (22) is to be replaced by distributional convergence).
c) *The lower estimate*

$$\sum_{j=0}^{\infty} a^{2tj} \|w_j\|^2_{L_2(I)} \leq C \|f\|^2_{H^t(I)} \qquad \forall\ f \in H^t(I) \tag{24}$$

in (21) can be extended to the enlarged parameter range $0 \leq t < m$.

Part b) is of importance for the preconditioning theory of subsection 2.4 in the case of the single layer potential equation, part c) will be used below for the approximation results.

The above formulation is tuned to the needs of this paper where we concentrate on the problem (2). To cover also the hypersingular equation, we would need the analog of part a) for the space $\tilde{H}^s(I)$ (specifically for $s = 1/2$). We will briefly outline the argument which yields such a result for the range $0 \leq s \leq 1$. Suppose that we have a semi-orthogonal system Φ which in addition to its Riesz basis property in $L_2(I)$ is (after scaling) also a Riesz basis in $H_0^1(I) = \tilde{H}^1(I)$. The usual way to achieve this is to start with a system as in Proposition 3 and to modify all those basis functions $\psi_{j,i} \in \Psi_j$ that do not vanish at the end-points such that the modified systems $\Psi_{j,0}$ belong to $H_0^1(I)$, that they remain Riesz bases for their spans $W_{j,0}$, and that this new subspace system provides again a semi-orthogonal decomposition of $L_2(I)$. For basis functions of local support, the perturbations can usually be controlled to yield the basis property in $H_0^1(I)$ for a scaled version of $\Phi_0 = \cup_{j\geq 0} \Psi_{j,0}$. Then one obtains for $0 \leq s \leq 1$ by interpolation

$$\|f\|^2_{\tilde{H}^s(I)} \approx \sum_{j=0}^{\infty} a^{2sj} \|w_{j,0}\|^2_{L_2(I)} \qquad \forall\ f \in \tilde{H}^s(I)\,, \tag{25}$$

where $w_{j,0} \in W_{j,0}$, $j \geq 0$. By duality, the result is also true for $H^s(I)$ and $-1 \leq s < 0$. For piecewise linear functions, this approach has already been demonstrated in [35, 36].

The above preparations almost immediately allow us to derive analogous results for multivariate Sobolev spaces on I^d, at least for the parameter ranges of interest. E.g., as in [21] we may use

$$H^s(I^d) = \bigcap_{l=1}^{d} L_2(I) \otimes \ldots \otimes H^s(I) \otimes \ldots \otimes L_2(I)\,, \quad s > 0\,. \tag{26}$$

In (26), the spaces on both sides coincide as sets and have equivalent norms. The definition of the Hilbert space in the right-hand side of (26) is explained in [21, section 2] where basic results on decomposition norms in tensor-product Hilbert spaces are stated. Together with Proposition 3, the trivial result for $s=0$, and duality arguments for the range $s<0$, we arrive at

Proposition 4 *Let the assumptions of Proposition 3 be satisfied. After suitable scaling, the system $\Phi(I^d)$ as introduced in section 1 forms a Riesz basis in the spaces $H^s(I^d)$ for $0 \le s < r$ and in the spaces $\tilde{H}^s(I^d)$ for $-r < s < 0$. In particular,*

$$||f||^2_{H^s(I^d)} \approx \sum_{\mathbf{j}\ge\mathbf{0}} a^{2s|\mathbf{j}|_\infty} ||w_{\mathbf{j}}||^2_{L_2(I^d)} \,, \quad 0 \le s < r \,, \tag{27}$$

where

$$f = \sum_{\mathbf{j}\ge\mathbf{0}} w_{\mathbf{j}} \,, \quad w_{\mathbf{j}} \in W_{\mathbf{j}} \,, \, \mathbf{j} \ge \mathbf{0} \,, \tag{28}$$

is the unique decomposition of $f \in H^s(I^d)$ with respect to the L_2-orthogonal subspace system $\{W_{\mathbf{j}}\}$. Analogously, for $-r < s < 0$:

$$||f||^2_{\tilde{H}^s(I^d)} \approx \sum_{\mathbf{j}\ge\mathbf{0}} a^{2s|\mathbf{j}|_\infty} ||w_{\mathbf{j}}||^2_{L_2(I^d)} \qquad \forall\, f \in \tilde{H}^s(I^d) \,. \tag{29}$$

Finally, the analog of (24) is valid:

$$\sum_{\mathbf{j}\ge\mathbf{0}} a^{2t|\mathbf{j}|_\infty} ||w_{\mathbf{j}}||^2_{L_2(I^d)} \le C||f||^2_{H^t(I^d)} \qquad \forall\, f \in H^t(I^d) \,,\ 0 \le t < m \,. \tag{30}$$

Extensions to the spaces $\tilde{H}^s(I^d)$ with $s>0$ are straightforward. We note that the decomposition norms used in Proposition 4 can be made the subject of a separate investigation: used as norms they lead to scales of *approximation spaces* (consult [34, 35, 36, 8] for results in this direction).

In the same way as Proposition 4 serves isotropic Sobolev spaces, decomposition norms can be found for the $H^{\mathbf{s}}_{\mathrm{mix}}$-spaces. We only formulate a partial result which immediately follows from (14), Proposition 3, and [21, Proposition 1].

Proposition 5 *Let the assumptions of Proposition 3 be satisfied. After suitable scaling, the system $\Phi(I^d)$ forms a Riesz basis in the spaces $H^{\mathbf{s}}_{\mathrm{mix}}$ if $0 \le s_l < r$, $l = 1, \ldots, d$. In particular, using the notation from (28),*

$$||f||^2_{H^{\mathbf{s}}_{\mathrm{mix}}(I^d)} \approx \sum_{\mathbf{j}\ge\mathbf{0}} a^{2\mathbf{s}\cdot\mathbf{j}} ||w_{\mathbf{j}}||^2_{L_2(I^d)} \quad \forall\, f \in H^{\mathbf{s}}_{\mathrm{mix}} \,. \tag{31}$$

Here, we could have even allowed for choosing different systems Φ in the different coordinate directions (accordingly, the condition on $\mathbf{s}$ would read $\mathbf{0} \le \mathbf{s} < \mathbf{r}$, where the vector $\mathbf{r}$ describes the possibly different exponents r in the Bernstein inequality (20)). These and other modifications (such as assuming a direction-dependent $\mathbf{a}$ instead of a single a) could also be applied in Proposition 4 giving the results more flexibility. Comparing (31) in the isotropic case $s_1 = \ldots = s_d = s$ with (27), the difference is in the exponents of the forefactors which are $2s|\mathbf{j}|$ for the mixed case and therefore larger than $2s|\mathbf{j}|_\infty$ for the usual Sobolev spaces.

We will next derive error estimates for the Galerkin approximations with respect to $V_{\mathcal{J}}$ for the variational problem (15): Find $u_{\mathcal{J}} \in V_{\mathcal{J}}$ such that

$$a_V(u_{\mathcal{J}}, v_{\mathcal{J}}) = \langle g, v_{\mathcal{J}} \rangle_{H^{1/2}\times\tilde{H}^{-1/2}} \qquad \forall\, v_{\mathcal{J}} \in V_{\mathcal{J}} \,. \tag{32}$$

In formulating (32), it is implicitly assumed that $V_{\mathcal{J}} \subset \tilde{H}^{-1/2}(I^2)$, which is a very weak requirement. By Proposition 2, the Lax-Milgram theorem guarantees solvability, and Cea's lemma leads to a two-sided estimate via best approximations in the $\tilde{H}^{-1/2}$-norm

$$\sqrt{a_V(f-u_{\mathcal{J}}, f-u_{\mathcal{J}})} \approx ||f-u_{\mathcal{J}}||_{\tilde{H}^{-1/2}} \approx \inf_{v_{\mathcal{J}}\in V_{\mathcal{J}}} ||f-v_{\mathcal{J}}||_{\tilde{H}^{-1/2}} \equiv e_{\mathcal{J}}(f)_{-1/2}$$

for the error between the solutions of the continuous problem (15) and the approximate problem (32). If $-m < s < 0$ or $0 \le s < r$, respectively, then for estimating the best approximations $e_{\mathcal{J}}(f)_s$ in the $\tilde{H}^s$- or H^s-norm, respectively, we can switch to the decomposition norms. We will carry out the estimation for the case $-m < s < 0$. First of all, by duality and (30) we have

$$\begin{aligned} \|f\|^2_{\tilde{H}^s} &= \sup_{v \in H^{-s}} \frac{(f,v)^2_{L_2}}{\|v\|^2_{H^{-s}}} = \sup_{\tilde{w}_{\mathbf{j}} \in W_{\mathbf{j}}\,:\, v=\sum_{\mathbf{j}} \tilde{w}_{\mathbf{j}} \in H^{-s}} \frac{(\sum_{\mathbf{j}} (w_{\mathbf{j}}, \tilde{w}_{\mathbf{j}})_{L_2})^2}{\|v\|^2_{H^{-s}}} \\ &\le C \sum_{\mathbf{j}} 2^{2|\mathbf{j}|_\infty s} \|w_{\mathbf{j}}\|^2_{L_2} \sup_{\tilde{w}_{\mathbf{j}} \in W_{\mathbf{j}}\,:\, v=\sum_{\mathbf{j}} \tilde{w}_{\mathbf{j}} \in H^{-s}} \frac{\sum_{\mathbf{j}} 2^{-2|\mathbf{j}|_\infty s} \|\tilde{w}_{\mathbf{j}}\|_{L_2}}{\|v\|^2_{H^{-s}}} \le C \sum_{\mathbf{j}} 2^{2|\mathbf{j}|_\infty s} \|w_{\mathbf{j}}\|^2_{L_2} \,. \end{aligned}$$

For $0 \le s < r$, the analogous result is contained in (27). Thus,

$$\begin{aligned} e_{\mathcal{J}}(f)_s &\le C \sum_{\mathbf{j} \notin \mathcal{J}} a^{2s|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|^2_{L_2} \le C (\max_{\mathbf{j} \notin \mathcal{J}} a^{2(s-t)|\mathbf{j}|_\infty}) \sum_{\mathbf{j} \notin \mathcal{J}} a^{2t|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|^2_{L_2} \\ &\le C (\max_{\mathbf{j} \notin \mathcal{J}} a^{2(s-t)|\mathbf{j}|_\infty}) \|f\|^2_{H^t} \,, \quad \max(-r, s) < t < m \,. \end{aligned}$$

The last step is covered by (30) if $0 \le t < m$ resp. by (29) if $-r < t < r$.

Now, let us specify the result for the two specific choices of $\mathcal{J}$ associated with the full and sparse grid spaces $V_k(I^d)$ and $\hat{V}_k(I^d)$. In the full grid case ($\mathcal{J} = \mathcal{J}_k$), we have the standard result:

$$e_k(f)_s \le C a^{(s-t)k} \|f\|_{H^t} \,, \quad -m < s < r \,,\ \max(-r, s) < t \le m \,. \tag{33}$$

Note that the saturation case $t = m$ can be derived directly from (19) and a tensor-product argument. When applied to the solution u_k of the discrete variational problem (32) on the full grid space $V_k(I^2)$ and to smooth solutions $f \in H^t(I^2)$ of (2), (33) leads to

$$\|f - u_k\|_{\tilde{H}^{-1/2}} \le C a^{(-1/2-t)k} \|f\|_{H^t} \,, \quad -1/2 < t \le m \,. \tag{34}$$

The estimate (33) should be compared with the following result for the sparse grid spaces ($\mathcal{J} = \hat{\mathcal{J}}_k$).

Proposition 6 *Let the univariate semi-orthogonal space splitting $\{W_j\}$ resp. the system Φ and its multi-dimensional counterparts be introduced as above (specifically, (19), (20) are assumed). Let $-m < s < r$ and $\max(-r, s) < t \le m$. Then, for the best approximations $\hat{e}_k(f)_s$ with respect to the sparse grid spaces $\hat{V}_k(I^d)$ (set $\mathcal{J} = \hat{\mathcal{J}}_k$ in the above definitions), the assumption $f \in H^t(I^d)$ only implies*

$$\hat{e}_k(f)_s \le C a^{(s-t)k/d} \|f\|_{H^t} \,. \tag{35}$$

On the other hand $f \in H^t_{\mathrm{mix}}(I^d)$, $0 \le t \le m$, leads to

$$\hat{e}_k(f)_s \le C \|f\|_{H^t_{\mathrm{mix}}} \begin{cases} a^{(s-t)k} & , \quad 0 \le s < r \,,\ s < t \le m \,, \\ a^{-tk} k^{d-1} & , \quad s = 0 < t \le m \,, \\ a^{(s/d-t)k} & , \quad -m < s < 0 \le t \le m \,. \end{cases} \tag{36}$$

In particular, if $d = 2$ then the error for the Galerkin solution $\hat{u}_k$ of the discrete variational problem (32) with respect to the sparse grid space $\hat{V}_k(I^2)$ satisfies

$$\|f - \hat{u}_k\|_{\tilde{H}^{-1/2}} \le C \begin{cases} a^{(-1/4-t/2)k} \|f\|_{H^t} \\ a^{(-1/4-t)k} \|f\|_{H^t_{\mathrm{mix}}} \end{cases} \tag{37}$$

depending on whether f belongs to $H^t(I^2)$, or $H^t_{\mathrm{mix}}(I^2)$ ($0 \le t \le m$).

Proof. To see (35), just compute the forefactor in the general estimate of $e_{\mathcal{J}}(f)_s$ given above (and its counterpart for $0 \le s < r$) for the specific index set $\hat{\mathcal{J}}_k$. Alternatively, observe that $V_{[k/d]} \subset \hat{V}_k$ and apply (33).

We come to the justification of (36). For $0 \le t < r$ ($-m < s < \min(t, r)$), we can proceed as in the above estimation for $e_{\mathcal{J}}(f)_s$ but use Proposition 5 instead of (30). This gives

$$\hat{e}_k(f)_s^2 \le C(\max_{\mathbf{j} \notin \hat{\mathcal{J}}_k} a^{2s|\mathbf{j}|_\infty - 2t|\mathbf{j}|}) \|f\|^2_{H^t_{\text{mix}}} .$$

Computing the maximum which is attained for, e.g., $\mathbf{j} = (k+1, 0, \ldots, 0)$ if $s \ge 0$ resp. for $\mathbf{j}$ with $j_l \approx k/d$, $l = 1, \ldots, d$, if $s < 0$, we obtain (36) for this range of t. Note that in this case the additional factor k^{d-1} can be dropped for $s = 0$ (a related, more general result for $s = 0$ is contained in [11]).

To cover the whole range of t, we need the auxiliary estimate

$$\|w_{\mathbf{j}}\|_{L_2} \le C a^{-|\mathbf{j}|t} \|f\|_{H^t_{\text{mix}}} , \quad 0 \le t \le m , \tag{38}$$

for the components of the decomposition (28) if $f \in H^t_{\text{mix}}(I^d)$. Here is the simple tensor-product argument to get (38): Let Q_j and $R_j (= Q_j - Q_{j-1})$ be the $L_2(I)$-orthogonal projections onto V_j and W_j, respectively. Then, by construction, the $L_2(I^d)$-orthoprojection $R_{\mathbf{j}}$ onto $W_{\mathbf{j}}$ is given by $R_{\mathbf{j}} = R_{j_1} \otimes \ldots \otimes R_{j_d}$ (for definitions and some properties of tensor products of operators, see [49, Section 3.4, 8.5]). Obviously, $\|R_j f\|_{L_2} \le C\|f\|_{L_2}$ and, as a consequence of (19),

$$\|R_j f\|_{L_2} \le \|f - Q_j f\|_{L_2} + \|f - Q_{j-1} f\|_{L_2} \le C a^{-mj} \|f\|_{H^m} , \quad f \in H^m(I) .$$

Interpolation theory of operators in Sobolev spaces yields

$$\|R_j\|_{H^t(I) \to L_2(I)} \le C a^{-tj} , \qquad 0 \le t \le m . \tag{39}$$

Now, for tensor products of operators $A_i : H_i \to \tilde{H}_i$, $i = 1, 2$, between separable Hilbert spaces there is an obvious norm estimate

$$\|A_1 \otimes A_2\|_{H_1 \otimes H_2 \to \tilde{H}_1 \otimes \tilde{H}_2} \le \|A_1\|_{H_1 \to \tilde{H}_1} \cdot \|A_2\|_{H_2 \to \tilde{H}_2}$$

which, iteratively applied, due to (39) yields

$$\|R_{\mathbf{j}}\|_{H^t_{\text{mix}}(I^d) \to L_2(I^d)} \le \prod_{l=1}^{d} \|R_{j_l}\|_{H^t(I) \to L_2(I)} \le C a^{-|\mathbf{j}|t} .$$

This establishes (38).

It remains to apply the inequality (38) componentwise. For $-r < s < r$, $0 \le t \le m$

$$\hat{e}_k(f)_s^2 \le C \sum_{|\mathbf{j}| > k} a^{2s|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|^2_{L_2} \le C\|f\|^2_{H^t_{\text{mix}}} \sum_{|\mathbf{j}| > k} a^{2s|\mathbf{j}|_\infty - 2t|\mathbf{j}|} .$$

Evaluating the last sum gives (36). Since (37) is a special case of previous estimates, the proof of Proposition 6 is completed.

It is important to mention that the above estimate for $e_{\mathcal{J}}(f)_s$ and the subsequent results for specific index sets seem to be sharp (possibly, except for the case $s = 0$ where partial improvements have been noted above). Here is the simple argument for the range $\max(0, s) \le t < r$. Pick the $\mathbf{j} \notin \mathcal{J}$ for which the maximum in the above estimates for $e_{\mathcal{J}}(f)_s$ is attained. Choose any function $w_{\mathbf{j}} \in W_{\mathbf{j}}$ with unit L_2-norm as candidate for the extremal f. By the norm equivalences from Proposition 4 and (31), we have

$$e_{\mathcal{J}}(w_{\mathbf{j}})_s \approx a^{s|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|_{L_2} = a^{s|\mathbf{j}|_\infty} , \quad \|w_{\mathbf{j}}\|_{H^t} \approx a^{t|\mathbf{j}|_\infty} , \quad \|w_{\mathbf{j}}\|_{H^t_{\text{mix}}} \approx a^{t|\mathbf{j}|} .$$

This gives the sharpness claim.

Proposition 6 exhibits an effect that we were not aware of when we started this research. For $s > 0$ and with the extra regularity for the solution expressed by the requirement $f \in H^t_{\mathrm{mix}}(I^d)$ (instead of $f \in H^t(I^d)$), according to (36) no loss of asymptotic approximation power occurs if the full grid space is replaced by a sparse grid space of less dimension. This is the well-known fact which was mentioned in the introduction, and which makes sparse grid spaces a good candidate for large-scale discretizations in higher-dimensional problems. For $s < 0$, this is no longer true. Although there is an improvement of the estimate when assuming $f \in H^t_{\mathrm{mix}}(I^d)$ instead of only $f \in H^t(I^d)$ $(0 < t \leq m)$, there still remains a gap of $-s(d-1)/d$ between the approximation orders achievable by the full and sparse grid spaces, respectively. A rough argument that explains the problems for $s < 0$ is as follows: While for $s \geq 0$ the approximation estimate essentially comes from the embeddings $H^t_{\mathrm{mix}}(I^d) \subset H^s_{\mathrm{mix}}(I^d) \subset H^s(I^d)$, for $s < 0$ the latter embedding is reversed: $H^s(I^d) \subset H^s_{\mathrm{mix}}(I^d)$. This means that for negative smoothness parameters the usual H^s-smoothness requirement is relaxed when changing to the spaces H^s_{mix}. The correct, best possible replacement is the embedding $H^{s/d}_{\mathrm{mix}}(I^d) \subset H^s(I^d)$ which explains the occurance of s/d instead of s.

We did not formulate the possible extension of (36) to the case $\max(-r, s/d) < t < 0$ which could be obtained from (31) by using duality arguments but does not have any immediate application. Note that the error estimates of this section need to be complemented by special handling of the singularity components in the solutions (see Proposition 2), otherwise they would apply only with very small t. E.g., for solutions of the single layer potential equation (2), the decomposition (16) in Proposition 2 points out the restriction $t < 0$ due to the corner-edge singularity component f^{ce} which contains the factor $\sin^{-1/2}\theta$ as the dominating singularity. See the discussion in subsection 3.4.

2.4 Preconditioners

Based on the results on decomposition norms presented in the previous subsection, the derivation of preconditioners is straightforward and can be given in great generality. Consider a generic discrete variational problem on some $V_{\mathcal{J}}$:

$$a(u_{\mathcal{J}}, v_{\mathcal{J}}) = G(v_{\mathcal{J}}) \qquad \forall\, v_{\mathcal{J}} \in V_{\mathcal{J}}\ , \tag{40}$$

where $a(\cdot,\cdot)$ is a symmetric elliptic bilinear form and G a linear functional on a Sobolev space $H^s(I^d)$ $(0 \leq s < r)$ or $\tilde{H}^s(I^d)$ $(-r < s < 0)$, respectively. In the following we will use all the time the notation $H^s(I^d)$ and $W_{\mathbf{j}}$ assuming that the reader is able to decide when a replacement by $\tilde{H}^s(I^d)$ or $W_{\mathbf{j},0}$ is appropriate. If we choose a basis in $V_{\mathcal{J}}$ we arrive at an equivalent linear system

$$A_{\mathcal{J}} x_{\mathcal{J}} = g_{\mathcal{J}} \tag{41}$$

for the coefficient vector $x_{\mathcal{J}}$ of the solution $u_{\mathcal{J}} \in V_{\mathcal{J}}$ of (40). The coefficient matrix $A_{\mathcal{J}}$ is symmetric positive definite and, at least in our applications, dense. Its condition number behavior depends on s and the basis and, therefore, possibly on $\mathcal{J}$, too. The main result in this subsection is

Proposition 7 *Under the above assumptions, if one chooses the scaled system*

$$\Phi_{\mathcal{J},s} \equiv \{a^{-s|\mathbf{j}|_\infty}\psi_{\mathbf{j},\mathbf{i}} \ :\ \mathbf{0} \leq \mathbf{i} \leq \mathbf{m_j} - \mathbf{1},\ \mathbf{j} \in \mathcal{J}\} \tag{42}$$

as the basis in $V_{\mathcal{J}}$ then the spectral condition number of the discretization matrix $A_{\mathcal{J}}$ in (41) is bounded, uniformly in $\mathcal{J}$. The bound depends on s and on the ellipticity constants (with respect to the Sobolev norm in $H^s(I^d)$) of the bilinear form $a(\cdot,\cdot)$ in (40).

Note that the scaling given in (42) assumes that (5) holds. Thus, the above statement is just a reformulation of the stability of the subspace splitting

$$\{V_{\mathcal{J}}, a(\cdot,\cdot)\} = \sum_{\mathbf{j}\in\mathcal{J}} \{W_{\mathbf{j}}, a^{2s|\mathbf{j}|_\infty}(\cdot,\cdot)_{L_2}\} \tag{43}$$

in conjunction with (5) and (42). For the connection of the notions of stable subspace splittings with additive Schwarz operators and multilevel preconditioners of hierarchical basis type, see [35]. Note that, under the assumptions made, the scaling factor used in (42) could be replaced by

$$a(\psi_{\mathbf{j},\mathbf{i}}, \psi_{\mathbf{j},\mathbf{i}})^{-1/2} \approx a^{-s|\mathbf{j}|_\infty} , \tag{44}$$

without changing the result. In matrix terminology, Proposition 7 together with the above remarks tells us that the matrix $A_{\mathcal{J}}$ associated with the tensor-product multilevel basis $\Phi_{\mathcal{J}} = \Phi_{\mathcal{J},0}$ and given by the entries $a(\psi_{\mathbf{j},\mathbf{i}}, \psi_{\mathbf{j}',\mathbf{i}'})$, where $\mathbf{0} \le \mathbf{i} \le \mathbf{m_j} - \mathbf{1}$, $\mathbf{0} \le \mathbf{i}' \le \mathbf{m_{j'}} - \mathbf{1}$, $\mathbf{j}, \mathbf{j}' \in \mathcal{J}$, is well-conditioned after diagonal scaling (with a condition number estimate that does not depend on $\mathcal{J}$). Even though there are alternative bases and representation systems in sparse grid spaces, we have run the numerical experiments described in the next section exclusively by precomputing the symmetric, uniformly well-conditioned matrices

$$B_{\mathcal{J}} = \mathrm{diag}(A_{\mathcal{J}})^{-1/2} A_{\mathcal{J}} \mathrm{diag}(A_{\mathcal{J}})^{-1/2} .$$

Moreover, for the adaptive sparse grid method described in subsection 3.4, we make also use of the fact that the well-conditioning result of Proposition 7 can be extended to discretization matrices associated with the linear span of an *arbitrary collection of scaled* $\psi_{\mathbf{j},\mathbf{i}}$. This allows to use adaptivity based on adding individual basis functions for better resolution of the singular parts along edges, without reconsidering the preconditioning aspect.

3 Numerical Tests

3.1 Examples of semi-orthogonal systems Φ

By now, many examples of univariate semi-orthogonal Riesz bases are available that fit the requirements of our framework. We do not aim at a general framework, see [5, 6, 10, 12] for recent work and references on multilevel systems of well-localized functions and their adaption to the case of a finite interval. Since our numerical testing concentrates on (2), we are looking for simple candidates that fit the assumptions of the $\tilde{H}^{-1/2}(I)$ case. E.g., our bases need not to fulfill special boundary conditions (in contrast, applications for the hypersingular equation (3) would require homogeneous boundary conditions at the endpoints of the interval). Our numerical experiments use piecewise constant and linear spline functions on uniform dyadic grids (with stepsize 2^{-j}). Thus, below we will always have $a = 2$.

Example 1: *Piecewise constant functions.* This case is used for its simplicity. Let V_j be the space of piecewise constant functions on a uniform partition of $[0, 1]$ into 2^j dyadic intervals. By $\phi_{j,i}$ we denote the characteristic function of the dyadic interval $[i2^{-j}, (i+1)2^{-j}) \subset I$, $i = 0, \ldots, 2^j - 1$. For later use, let us call the left endpoint $x = i2^{-j}$ of this interval *grid point associated with the nodal basis function* $\phi_{j,i}$ in V_j. The standard system Φ associated with this subspace sequence $\{V_j\}$ is the *Haar system* which is even orthonormal. For our convenience, we introduce the following enumeration of the Haar functions. We set $\psi_{0,0}(x) = \phi_{0,0}(x) \equiv 1$, $x \in I$ ($m_0 = 1$), for the only basis function in $W_0 = V_0$, and

$$\psi_{j,i}(x) = 2^{(j-1)/2}\psi(2^{j-1}x - i) , \quad x \in I , i = 0, \ldots, 2^{j-1} - 1 , \ j = 1, 2, \ldots ,$$

for the basis functions in $W_j = V_j \ominus_{L_2} V_{j-1}$, where

$$\psi(x) = \begin{cases} 1 & , \quad x \in [0, 1/2) \\ -1 & , \quad x \in [1/2, 1) \\ 0 & , \quad x \notin [0, 1) \end{cases}$$

is shown in Figure 1 a). Obviously, for $j > 0$ we have $m_j = 2^{j-1}$, and

$$\psi_{j,i} = 2^{(j-1)/2}(\phi_{j,2i} - \phi_{j,2i+1}) , \qquad i = 0, \ldots, 2^{j-1} - 1 .$$

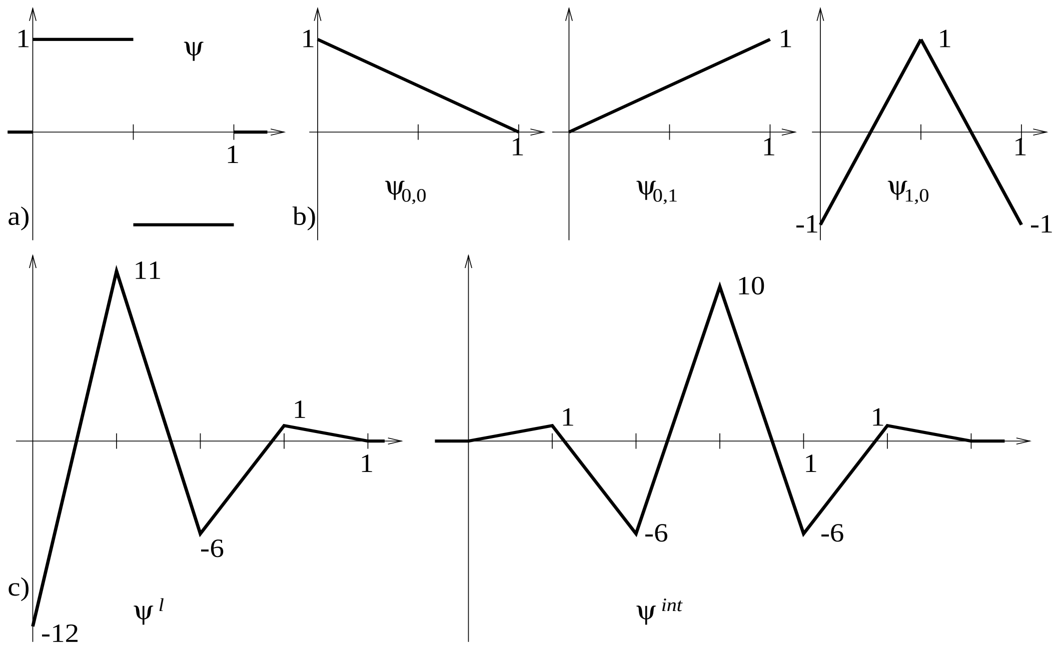


Figure 1: Generating functions for Examples 1 and 2.

It is then natural to call the midpoint of the support of $\psi_{j,i}$ (i.e., the grid point associated with $\phi_{j,2i+1}$) *grid point associated with the prewavelet function* $\psi_{j,i}$ in W_j (for $j = 0$, $x = 0$ is associated with the only function $\psi_{0,0}$ in W_0). This convention allows to identify the functions in Φ with the hierarchically enumerated dyadic points in the interval $[0, 1)$. Moreover, by taking cartesian products such bijections carry over to the multivariate systems, and can be used to visually represent the basis functions of the multivariate spaces $V_{\mathbf{j}}$, $W_{\mathbf{j}}$, V_k, $\hat{V}_k$, etc. See Figure 2 below for some illustrations.

Unfortunately, since for piecewise constant functions we have at best $r = 1/2$ in (20), the theory of section 2 is not fully applicable in the case of interest ($s = -1/2$). However, many results not covered can still be established (up to logarithmical terms, at least). E.g., sharp preconditioning results based on Haar type splittings have been obtained in [36].

Example 2: *Semi-orthogonal linear spline prewavelets.* This example goes back to Chui and Wang [6] who investigated semi-orthogonal spline prewavelets of arbitrary order. We only deal with the linear case. Let now $\{V_j\}$ be the spaces of linear splines with respect to the same sequence of uniform dyadic partitions of I. The standard nodal basis functions for V_j are again denoted by $\phi_{j,i}$, these are the usual hat functions associated with the dyadic grid points $i2^{-j}$, $i = 0, \ldots, 2^j$.

The bases for $W_0 = V_0$ ($m_0 = 2$) and W_1 ($m_1 = 1$), respectively, are shown in Figure 1 b). To generate a suitable basis Ψ_j for the orthogonal complement spaces $W_j = V_j \ominus_{L_2} V_{j-1}$, $j \geq 2$, one needs the scaled shifts of three functions ψ^{int}, ψ^l (these functions are shown in Figure 1 c)) and $\psi^r(x) = \psi^l(1-x)$ which serve the interior (int), the left (l) and right (r) boundary of the interval I. Again, we can associate dyadic points in $I = [0, 1]$ with the functions of the resulting system Φ in a natural way (the resulting visualizations are identical with that of Example 1, except for the additional right endpoint which occurs for $j = 0$, see Figure 2). More precisely, for general $j \geq 2$, the basis functions $\psi_{j,i}$ in Ψ_j could be introduced as

$$\psi_{j,i}(x) = \alpha_{j,i} \cdot \begin{cases} \psi^l(2^{j-1}x) & , \quad i = 0 \\ \psi^{int}(2^{j-1}x - i + 1) & , \quad i = 1, \ldots, 2^{j-1} - 2 \\ \psi^r(1 - 2^{j-1}(1-x)\} & , \quad i = 2^{j-1} - 1 \end{cases} ,$$

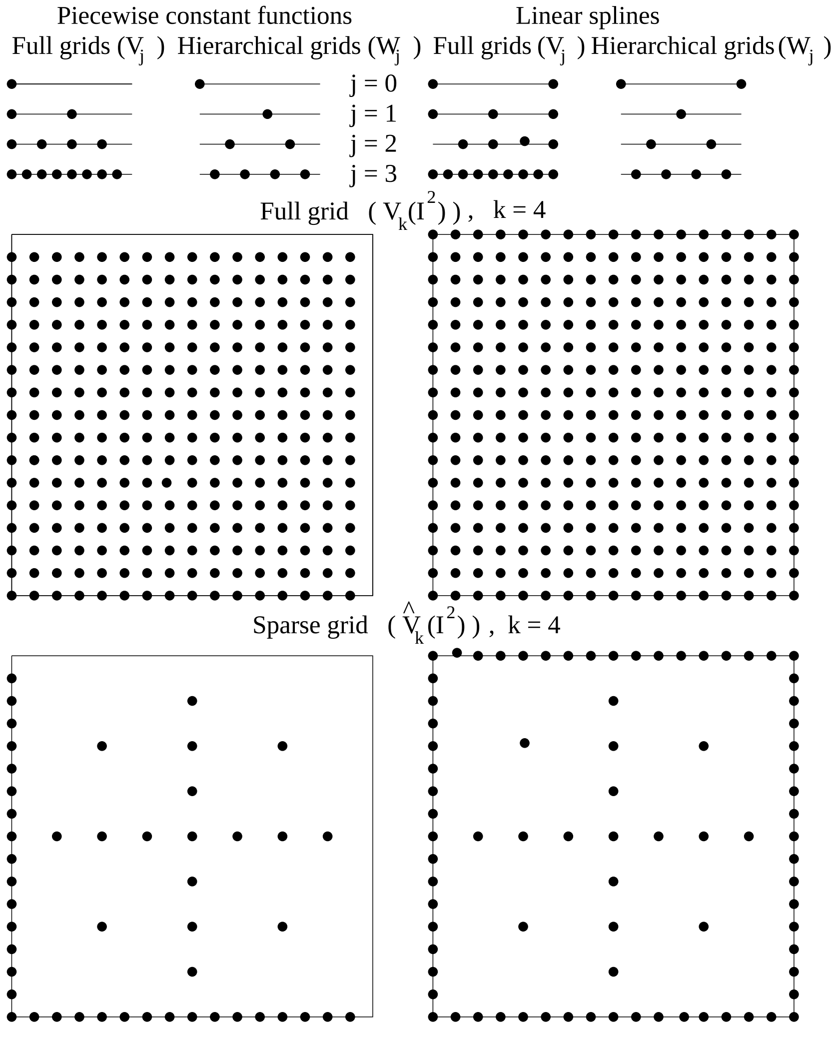


Figure 2: Grid point sets for Examples 1 and 2.

with appropriate scaling factors $\alpha_{j,i} \approx 2^{j/2}$ to ensure (5). This choice of basis functions in W_j is most economical since it realizes minimal support length (and therefore minimal mask length, see below) for the $\psi_{j,i}$. Note that the resulting system Φ is only semi-orthogonal, not fully orthogonal as in the case of the Haar system. The uniform L_2-stability (5) of the Ψ_j and the Riesz basis property of Φ in $L_2(I)$ have been shown in [6], respectively. According to Proposition 3, after appropriate scaling, this system serves as semi-orthogonal Riesz basis for $H^s(I)$ if $0 \leq s < 3/2$, and for $\tilde{H}^s(I)$ if $-3/2 < s < 0$ (for linear spline spaces, the relevant numbers for comparing with the general statements of subsection 2.3 are $r = 3/2$ and $m = 2$, see [35]). Thus, this example is fully covered by the approximation estimates and preconditioning results of section 2. A similar system (based on spaces incorporating homogeneous Dirichlet boundary conditions) has been used for constructing robust prewavelet solvers for anisotropic second–order elliptic boundary value problems in [21].

Table 1: Coefficient masks for expressing $\psi_{j,i}$ by nodal basis functions

Example 1		Example 2	
type	mask	type	mask
$\psi_{0,0}$	$[\,\underline{1}\,]$	$\psi_{0,0},\ \psi_{1,0}$	$[\,\underline{1}\,]$
$\psi_{j,i},\ j>0$	$[\,1 \quad \underline{-1}\,]$	$\psi_{1,0}$	$[\,-1 \quad \underline{1} \quad 1\,]$
		$\psi_{j,0},\ j\geq 2$	$[\,-12 \quad \underline{11} \quad -6 \quad 1\,]$
		$\psi_{j,i},\ i=1,\ldots,2^{j-1}-2,\ j\geq 2$	$[\,1 \quad -6 \quad \underline{10} \quad -6 \quad 1\,]$
		$\psi_{j,2^{j-1}-1},\ j\geq 2$	$[\,1 \quad -6 \quad \underline{11} \quad -12\,]$

3.2 Exact computation of the stiffness matrix entries

In this subsection, we indicate how to compute the entries of the stiffness matrices $A_{\mathcal{J}}$ arising from (32) when the corresponding section of $\Phi(I^d)$ is chosen as basis in $V_{\mathcal{J}}$, i.e., the quantities

$$a_{\mathbf{j},\mathbf{i};\mathbf{j}',\mathbf{i}'} = a_V(\psi_{\mathbf{j},\mathbf{i}}, \psi_{\mathbf{j}',\mathbf{i}'}) \;.$$

Due to the simplicity of the kernel function of the integral operator V, the tensor product structure of all constructions and the use of low order piecewise polynomial functions on dyadic partitions, these coefficients can be computed from *exact analytic expressions* following [28, 31]. To be self-contained, we state a few details of our implementation. First of all, the computation of the $a_{\mathbf{j},\mathbf{i};\mathbf{j}',\mathbf{i}'}$ can be reduced to the computation of the quantities

$$b_{\mathbf{j},\mathbf{i};\mathbf{j}',\mathbf{i}'} = a_V(\phi_{\mathbf{j},\mathbf{i}}, \phi_{\mathbf{j}',\mathbf{i}'}) \;,$$

since each $\psi_{\mathbf{j},\mathbf{i}}$ is a linear combination of $\phi_{\mathbf{j},\cdot}$, with coefficients expressable by a few tensor product masks. For the above examples, the underlying one-dimensional masks, i.e., the nonzero coefficients in the representations

$$\psi_{j,i} = \sum_{\hat{i}} a_{\hat{i}}^{(j,i)} \phi_{j,\hat{i}} \tag{45}$$

correspond to the few types of ψ-functions depicted in Figure 1 and are given for convenience in Table 1. We have neglected the scaling factors $2^{(j-1)/2}$ resp. $\alpha_{j,i}$ which are absorbed in the final diagonal scaling step, and show only the segment of nonzero coefficients $a_{\hat{i}}^{(j,i)}$ in (45). The underlined coefficient stands in front of the $\phi_{j,\hat{i}}$ which is associated with the grid point associated also with $\psi_{j,i}$.

Although the precomputation of the $b_{\mathbf{j},\mathbf{i};\mathbf{j}',\mathbf{i}'}$ is not necessary, it leads in practice to a substantial speedup. Note that in our particular situation further savings are possible by using the multilevel structure and symmetry considerations. Since all $\phi_{\mathbf{j},\mathbf{i}}$ are piecewise tensor products of polynomials of degree $\leq K$ ($K \leq 1$ in the above examples) over dyadic intervals, it is enough to know values at the corresponding dyadic points for the antiderivatives of the following type:

$$A_{klrs}(x,y,v,w) := \iiiint \frac{x^k y^l v^r w^s}{\sqrt{(x-v)^2+(y-w)^2}}\, dv\, dw\, dx\, dy \;, \quad 0 \leq k,l,r,s \leq K \;. \tag{46}$$

Note that the number of such antiderivatives grows dramatically with K. There exist recurrence formulae due to Maischak [31] that allow the exact computation (up to roundoff errors) of the above integrals A_{klrs} which can be used for small values of K. Note that computing $a_{\mathbf{j},\mathbf{i};\mathbf{j}',\mathbf{i}'}$ from the values of $A_{klrs}(x,y,v,w)$ at dyadic points corresponds to numerically calculating a high order difference of a smooth function which, in turn, requires special care in the implementation of these recursions. The amount of work per computed value of A_{klrs} is constant but grows excessively with K. For details, we refer to [31, 28] (we caution the reader that in [31, 28] and in all other papers available to us containing the recursions, some slight index errors are involved: $G_{lr}(\cdot)$ must be replaced by $G_{rl}(\cdot)$ in the recurrence formula for $G_{lrs}(\cdot)$).

It can be concluded from these brief remarks that for our examples and stiffness matrices corresponding to index sets $\mathcal{J} \subset \mathcal{J}_L$, $d = 2$,

a) the number of operations to compute all values of the integrals A_{klrs} at dyadic points that are necessary in step b) is bounded by $\mathrm{O}(2^{2L})$,

b) the number of operations to compute any individual $b_{\mathbf{j},\mathbf{i};\mathbf{j}',\mathbf{i}'}$ or $a_{\mathbf{j},\mathbf{i};\mathbf{j}',\mathbf{i}'}$ $(\mathbf{j},\mathbf{j}' \in \mathcal{J})$ from the values computed in step a) is bounded by an absolute constant, while

c) the number of *different* $b_{\mathbf{j},\mathbf{i};\mathbf{j}',\mathbf{i}'}$ or $a_{\mathbf{j},\mathbf{i};\mathbf{j}',\mathbf{i}'}$ to be computed and stored, is generally bounded by $\mathrm{O}(L^2 2^{2L})$ but depends on $\mathcal{J}$. A closer look at the procedure outlined above shows that for the sparse grid space $(\mathcal{J} = \hat{\mathcal{J}}_L)$ only $\mathrm{O}(2^{2L})$ entries of the stiffness matrix are different.

This substantiates our claim made in section 1 that a suboptimal complexity for the assembly of the stiffness matrices $A_{\hat{\mathcal{J}}_L}$ stemming from sparse grid multilevel discretizations can be achieved. However, especially in connection with adaptive methods (see subsection 3.4) and for more realistic applications, such a claim needs revision, and numerical integration of the entries of the stiffness matrices (using carefully tuned adaptive quadrature rules) cannot be avoided. In this paper, where we are only performing qualitative tests with the square screen problem, we do not touch this important issue further. However, comparing the costs for numerical integration and of the application of the recurrence formulae to compute the entries of the stiffness matrices reveals that, already for the square screen problem, numerical integration would be more reasonable. Additionally, due to round-off errors arising in the computations using the exact analytic formulae, numerical integration becomes of interest, especially for high levels L.

3.3 Full and sparse grid spaces: Numerical comparison

We have run comparative tests for both examples for Φ (constant and linear case) described in the previous section. All tests are for the capacity problem for a unit square screen which is described by (2), $\Gamma = I^2$, and a right-hand side $g \equiv 1$, i.e., we compute approximate solutions of the problem

$$Vf \equiv \frac{1}{4\pi} \int_{I^2} \frac{f(y)}{|x-y|}\, dy = 1 \; . \tag{47}$$

The capacity of the screen is defined as the average value of the solution f of (47):

$$\mathcal{C} = \frac{1}{4\pi} \int_{I^2} f(y)\, dy \; . \tag{48}$$

For the square screen under consideration, its numerical value is known from the literature as $\mathcal{C} = 0.366789...$ (cf. [14]). If we denote by u_L and $\hat{u}_L$, respectively, the Galerkin solution of (47) with respect to V_L and $\hat{V}_L$, respectively, we obtain approximate capacities $\mathcal{C}_L$ and $\hat{\mathcal{C}}_L$, respectively, as averages of these approximate solutions. Here and in the following, L denotes the discretization level number. It is well-known that

$$\begin{aligned} \delta_L &\equiv |\mathcal{C} - \mathcal{C}_L| = \mathcal{C} - \mathcal{C}_L = \tfrac{1}{4\pi} a_V(f-u_L, f-u_L) \approx \|f-u_L\|^2_{\tilde{H}^{-1/2}} \approx e_L(f)^2_{-1/2} \; , \\ \hat{\delta}_L &\equiv |\mathcal{C} - \hat{\mathcal{C}}_L| = \mathcal{C} - \hat{\mathcal{C}}_L = \tfrac{1}{4\pi} a_V(f-\hat{u}_L, f-\hat{u}_L) \approx \|f-\hat{u}_L\|^2_{\tilde{H}^{-1/2}} \approx \hat{e}_L(f)^2_{-1/2} \; . \end{aligned} \tag{49}$$

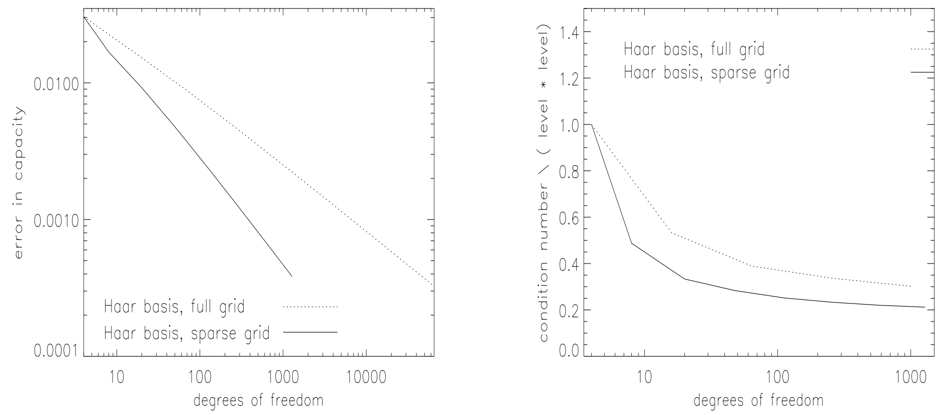
Therefore, it is enough to concentrate on capacity errors.

We start with the piecewise constant case. For the full grid spaces V_L and for the sparse grid spaces $\tilde{V}_L$, we obtain the results given in Table 2. The computations are performed by using the diagonally preconditioned discretization matrices $B_L \equiv B_{\mathcal{J}_L}$ and $\hat{B}_L \equiv B_{\hat{\mathcal{J}}_L}$, respectively, as introduced at the end of subsection 2.4. N denotes the number of grid points, the condition number of B_L is denoted by κ_L, It_L denotes the number of cg-iterations to reach a relative error reduction of 10^{-6} in the residual of the preconditioned system (from a zero starting vector). Finally, $q_L = \delta_{L-1}/\delta_L$ is the quotient of two successive errors in the capacity. Analogously we define $\hat{N}$, $\hat{\kappa_L}$, $\hat{It}_L$, and $\hat{q}_L$. The numbers in parentheses for full grid spaces are taken from [36, Table 3,4], where a different Haar preconditioner is used for the same problem. They are included to allow for a better comparison between full grid and sparse grid spaces in the asymptotical range. The numerical tests of the present paper are limited to problems of relatively small dimension, since we precompute and store the dense matrix $A_{\mathcal{J}}$ resp. $B_{\mathcal{J}}$. In Figure 3 the error in the capacity versus the number of degrees of freedom as well as the condition number behavior are displayed.

Table 2: Results for the tensor product Haar basis for full and sparse grids

	Full grid spaces					Sparse grid spaces				
L	N	κ_L	It_L	δ_L	q_L	$\hat{N}$	$\hat{\kappa}_L$	$\hat{It}_L$	$\hat{\delta}_L$	$\hat{q}_L$
1	4	1.00 (1.98)	1 (1)	0.030452	–	4	1.00	1	0.030452	–
2	16	2.13 (3.02)	3 (3)	0.016812	1.81	8	1.95	2	0.016831	1.81
3	64	3.50 (4.55)	8 (9)	0.009133	1.84	20	3.00	4	0.009209	1.83
4	256	5.39 (6.53)	12 (12)	0.004806	1.90	48	4.53	8	0.004915	1.87
5	1024	7.53 (8.78)	14 (15)	0.002483	1.94	112	6.29	12	0.002596	1.89
6	4096	(11.32)	(17)	(0.001268)	(1.96)	256	8.39	14	0.001369	1.90
7	16384	(14.14)	(20)	(0.000643)	(1.97)	576	10.79	17	0.000723	1.89
8	65536	(17.20)	(22)	(0.000325)	(1.98)	1280	13.56	20	0.000384	1.88

Figure 3: Error in capacity and condition number/L^2 vs. degrees of freedom



From these results we see that we obtain a similar convergence rates for the discrete capacities for both full and sparse grid spaces, even though the number of unknowns for the latter is only $O(L2^L)$ compared with 2^{2L} unknowns used for the full grid discretization. Note that for the full grid case, we are in agreement with the known error estimate of $\delta_L = O(2^{-L(1-\varepsilon)})$, where $\varepsilon > 0$ can be taken arbitrarily small (see, e.g., [28, section 7]). For the sparse grid spaces, a formal extrapolation of (37) in conjunction with Proposition 2 and (49) would only give $\hat{\delta}_L = O(2^{-L(1/2-\varepsilon)})$. Even though slightly better estimates can be proved due to the specific type of singularity functions involved (as should be expected from comparing with the numerical evidence given in Table 2), there is no hope for obtaining asymptotically the same approximation rate as for full grid spaces. More importantly, the ultimate goal should be adaptive methods since the presence of edge-corner singularities in the solution f of (47) leads to results that are far from the theoretical optimum. As can be extrapolated from (34) and (49), if the solution of (47) would be sufficiently smooth, the expected saturation order for capacity approximations using piecewise constant and piecewise linear functions would be $O(h^3)$ and $O(h^5)$, respectively (here,

h is the maximal element diameter of a generic partition). See [40] for approximation schemes using graded tensor-product meshes towards the edges that restore these rates asymptotically for the true, low-regularity solution in (2).

Furthermore, we see from Table 2 that the use of tensor product Haar functions results in still slightly growing condition numbers for the diagonally preconditioned stiffness matrices and for the cg-iteration count. This is in accordance with the results of [36] from which an asymptotical behavior

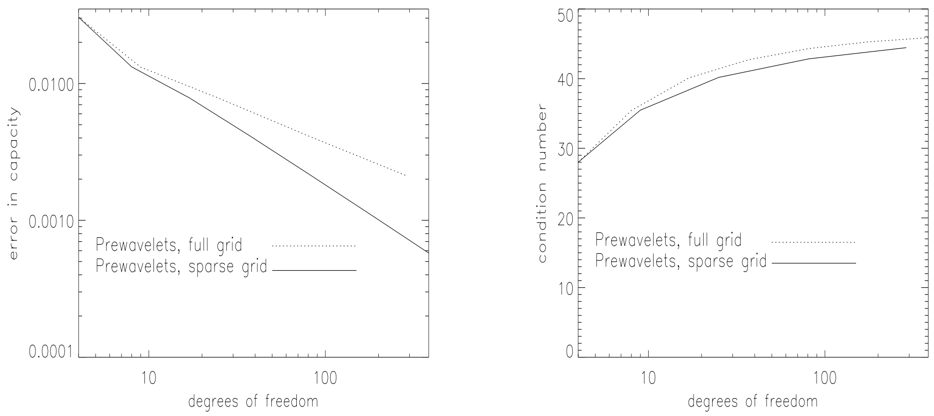
$$\kappa_L,\ \hat{\kappa}_L \approx L^2$$

can be concluded. The dependency of the condition number on L can also be seen in the right part of Figure 3. There, the quotient of condition number and L^2 that should asymptotically behave like a constant is plotted versus degrees of freedom. However, even though in this case the preconditioning effect is only suboptimal, the absolute iteration numbers are moderate and acceptable in practice.

Table 3: Results for linear spline prewavelets on full and sparse grids

	Full grid spaces					Sparse grid spaces				
L	N	κ_L	It_L	δ_L	q_L	$\hat{N}$	$\hat{\kappa}_L$	$\hat{It}_L$	$\hat{\delta}_L$	$\hat{q}_L$
0	4	1.00	1	0.030452	–	4	1.00	1	0.030452	–
1	9	35.47	3	0.013142	2.32	8	35.42	2	0.013216	2.30
2	25	40.18	6	0.007785	1.69	17	40.12	4	0.007839	1.69
3	81	42.83	9	0.004113	1.89	37	42.72	7	0.004173	1.88
4	289	44.45	16	0.002115	1.94	81	44.32	9	0.002169	1.92
5						177	45.29	10	0.001120	1.94
6						385	45.88	11	0.000577	1.94

Figure 4: Error in capacity and condition number vs. degrees of freedom



The results for the linear spline prewavelets can be seen in Table 3, and the corresponding graphical outputs are given in Figure 4. We have used the same notation and settings as described before. As predicted by Proposition 7 (see also [36]), the use of the semi-orthogonal prewavelet systems leads for both full and sparse grid linear spline spaces to condition numbers which are bounded from above by a constant independent of the number of levels L. Besides, this constant is fairly reasonable (even though it is larger than for the Haar case, at least in the range $L \leq 8$). On the other hand, there is no serious improvement in both approximation order and absolute error size (compared to the increased complexity of all computations). Again, the reason is the inadequate handling of the singularities in the solution of (47). This problem is addressed in the next subsection.

3.4 Adaptive sparse grid spaces

Since the solution of our model problem possesses a singular behavior along the boundary of $\Gamma = I^2$, adaptively refined grids instead of quasiuniform ones are surely more appropriate. This line of research is well-documented in the literature, see, e.g., [33, 32, 29]. Tensor-product grids which are refined towards the edges will generally lead to good results.

It is intriguing to examine the potential of *adaptive sparse grid spaces*. The main supporting argument is that sparse grid spaces are, by construction, well-suited for the approximation of *edge singularities* where refinement is required only towards the edge, and not in the direction along the edge. We have decided to follow some empirical arguments borrowed from [28]. First of all, since the *leading singularity components* of the solution f are essentially known (see Proposition 2), one can construct an adapted sparse grid space in the following way. Define suitably regularized functions that describe the behavior of the leading singularity. This can be done separately for the different corner/edge combinations. E.g., we have chosen the function $\Psi_\varepsilon(x,y)$ proposeded in [28, formula (28)]

$$\Psi_\varepsilon(x,y) = (r+\varepsilon^2)^{\gamma-1}\cdot\left(\sqrt{\sin(\theta)} + \frac{1-(\cos(\theta)+\varepsilon)^{\gamma-0.5}}{\sqrt{\sin(\theta)}+\varepsilon}\right) + \frac{(x+\varepsilon)^{\gamma-0.5}}{\sqrt{y+\varepsilon}}\,, \tag{50}$$

which serves the corner/edge singularity near $(0,0)$ and the edge described by $\theta = 0$ (compare Proposition 2 b), r, θ are polar coordinates, $x = r\cos\theta$, $y = r\sin\theta$, and $\gamma = 0.2966...$). Here, $\varepsilon > 0$ is a regularization parameter that ensures that $\Psi_\varepsilon(x,y)$ is smoothly defined everywhere on $[0,1]^2$. For our computations we choose $\varepsilon = 10^{-4}$. To obtain a function $\Phi_\varepsilon(x,y)$ that describes the singularities on the whole boundary we have defined $\Phi_\varepsilon(x,y) = \Psi_\varepsilon(x,y) + \Psi_\varepsilon(x,1-y) + \Psi_\varepsilon(1-x,y) + \Psi_\varepsilon(1-x,1-y)$.

This singularity function is now used to generate the adaptive sparse grid space. We describe the procedure for the linear spline prewavelet case. Since $\Phi_\varepsilon(x,y)$ is smooth on I^2, we can uniquely decompose it with respect to the *sparse hierarchical basis*, i.e., the tensor product of the well-known hierarchical basis on I consisting of dyadic hat functions. Starting with the four corners of the domain $[0,1]^2$ corresponding to the four basis functions in the space $W_{(0,0)} = V_0(I^2)$ on the coarsest level, a basis function $\psi_{\mathbf{j},\mathbf{i}}$ (or speaking in geometric terms, an associated *sparse grid point*) is introduced when the corresponding coefficient in the hierarchical basis decomposition is larger in absolute value than a given threshold $\eta > 0$. Furthermore, we set a maximum level $L_{\max}$ that is allowed in the resulting grid, i.e., we only pick basis functions as long as $|\mathbf{j}|_\infty \le L_{\max}$. We are aware of the fact that due to the available norm equivalences in $\tilde{H}^{-1/2}(I^2)$, see Proposition 4, looking directly at a decomposition of $\Phi_\varepsilon(x,y)$ with respect to $\Phi(I^d)$ would be much more reliable. However, the hierarchical basis decomposition is much simpler to obtain and, as it turned out, already leads to surprisingly good results.

With this approach, the *adaptive sparse grids* shown in Figure 5 have been obtained. The number of grid points N (i.e., the dimension of the linear span of all $\psi_{\mathbf{j},\mathbf{i}}$ which we will call *adaptive sparse grid space* and denote for brevity by $\hat{V}_N^{\mathrm{adapt}}$) is also indicated. Note that the space $\hat{V}_N^{\mathrm{adapt}}$ is not a space of bilinear functions on an adapted partition of I^2 into $\approx N$ rectangles but the linear span of a set of N tensor product prewavelet functions. The adaptive sparse grid only serves for visualizing the selected set of basis functions. The numerical results for adaptive sparse grid spaces using linear spline prewavelets are given in Table 4. As in the case of regular sparse grids, the condition number is bounded from above by a constant. This has been theoretically explained at the end of subsection 2.4. Note that when comparing the results in Tables 3 and 4, the comparison should be based on the dimensions of the corresponding spaces (i.e., on the number of grid points in the regular and adaptive sparse grids, respectively), and not on L. The maximum level $L_{\max}$ for an adaptive sparse grid is much higher than in a comparable regular sparse grid. This is also the reason for the slightly higher condition number in the adaptive case. What has not been done so far is to look at the complexity estimates for computing the stiffness matrix A_N^{adapt}.

We have run some analogous experiments for the piecewise constant case, see Table 5. Altogether, these results show the potential of adaptive sparse grid spaces for problems where *edge singularities* are dominating, as it is the case for the capacity problem. Note that our results compare well with

Figure 5: Adaptive sparse grids ($L_{\max} = 4,\ 12,\ 21,$ resp. $N = 57,\ 217,\ 1841$)

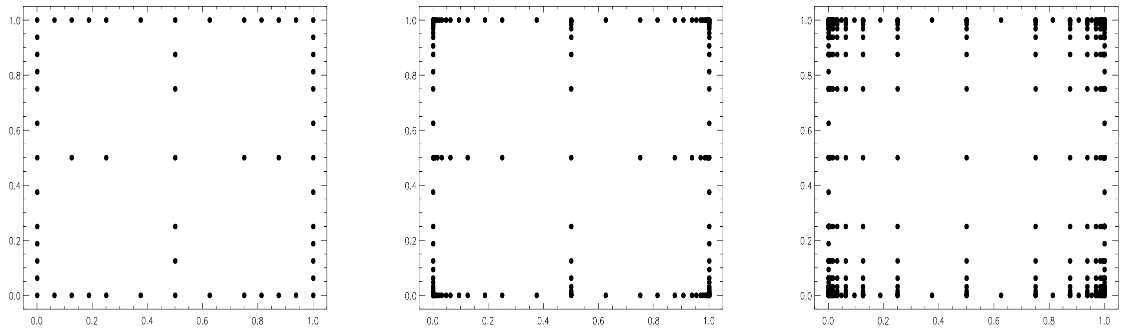


Table 4: Results for linear spline prewavelets on adaptive sparse grids

$L_{\max}$	N	$\hat{\kappa}_N^{\text{adapt}}$	$\hat{It}_N^{\text{adapt}}$	$\hat{\delta}_N^{\text{adapt}}$
2	17	40.11	4	0.007838
3	37	42.72	7	0.004173
4	57	44.26	8	0.002183
5	77	45.16	9	0.001151
6	97	45.68	9	0.000624
8	137	46.15	10	0.000223
10	177	46.30	12	0.000123

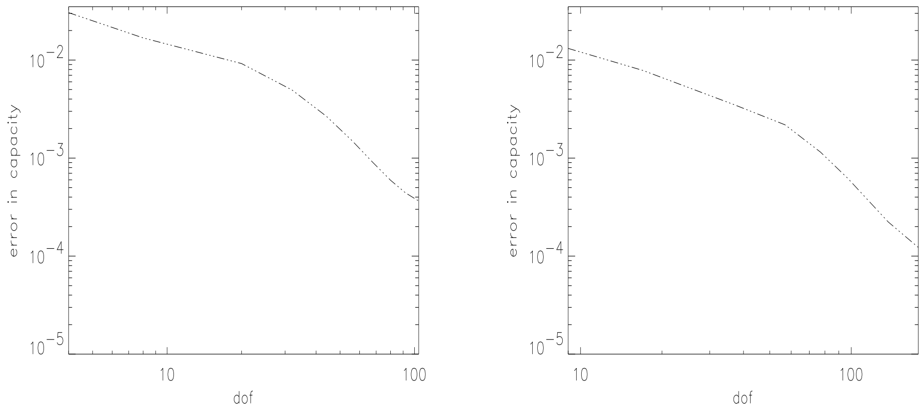
Table 5: Results for the Haar basis on adaptive sparse grids

$L_{\max}$	N	$\hat{\kappa}_N^{\text{adapt}}$	$\hat{It}_N^{\text{adapt}}$	$\hat{\delta}_N^{\text{adapt}}$
3	20	3.00	4	0.009209
4	32	3.24	6	0.004946
5	44	3.33	8	0.002676
6	56	3.35	8	0.001499
7	68	3.37	9	0.000899
8	80	3.40	9	0.000595
9	92	3.41	9	0.000442
10	104	3.42	9	0.000365

numerical results for the hp-method, see [28] (to make a realistic comparison, the values of $C - C_N$ given in the third column of Table 1 in [28] should be divided by a factor 2 since the authors there consider the square $[-1,1]^2$ instead of I^2). As a method that exhibits exponential convergence rates for this kind of application [28], the hp-version of the boundary element method is certainly more powerful in the asymptotic range. However, versions of adaptive sparse grid spaces that give good approximations within a certain range of accuracy at low implementational costs might provide an alternative, and serve as building blocks for efficient approximation schemes for composite manifolds Γ occuring in realistic applications. As this topic needs more theoretical understanding and practical testing, we will not go into further details here.

In a final remark, we will touch the following issue. The bottleneck of any type of integral equation solver are the cpu-time and amount of storage connected with the assembly of the stiffness matrix and, subsequently, the complexity of matrix-vector multiplications. Our main emphasis was on easing this

Figure 6: Errors in the capacity for the Haar basis and linear spline prewavelets



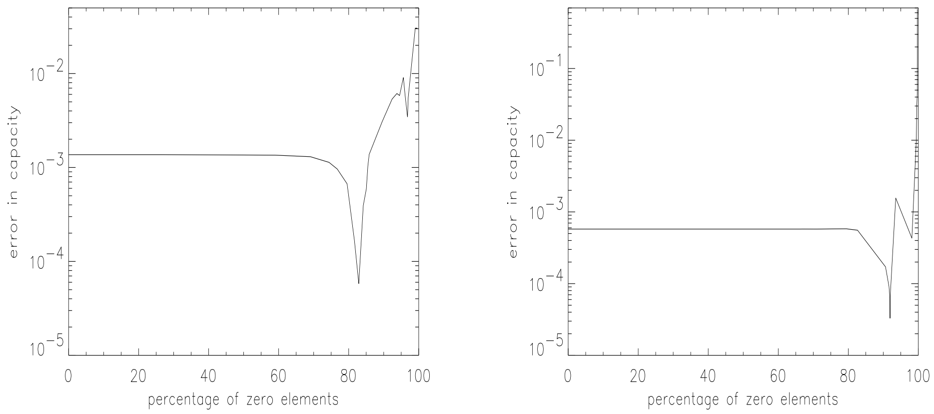
problem by *dimension reduction* of the ansatz space, by using (adaptive) sparse grid spaces. Alternative approaches use *matrix compression*, see [8] for an overview. Can compression be used on top of dimension reduction? As numerical testing shows this is possible but only to a certain extent. All our tests consistently demonstrated that up to 5-fold compression leaves the computed capacity value essentially unchanged. At the same time, compression did not have any serious impact on condition numbers and iteration count. In Figure 7 results for compressing the stiffness matrix $\hat{A}_L$ in the piecewise constant case (Haar basis) and for linear spline prewavelets are given for $L = 7$ and $L = 5$, respectively. The x-axis denotes the compression rate (i.e., percentage of zero entries in the compressed stiffness matrix) whereas the y-axis denotes the error in the capacity. This example is representative (for both element types, non-adaptive and adaptive sparse grid spaces, and a large range of L we essentially obtained the same curves). In any case, the potential amount of savings coming from additional matrix compression is within a constant factor, and does not significantly change with L or N. In the tests, we have simply set to zero all those entries of the diagonally preconditioned stiffness matrices which absolute values are below a given threshold ε. We admit that it is not a practical approach yet, since we first assembled the whole stiffness matrix and then compressed it. For wavelet discretizations, there exist techniques to prevent this difficulty, compare [44, 38]. There, decay estimates for the off-diagonal entries in the diagonally scaled stiffness matrix in the wavelet basis are used to avoid the assembly of a small entry a priori. See also [22, 23] for compression in the case of operators of arbitrary order in specially defined spaces including the cases of full and sparse grid spaces considered in the present paper.

4 Concluding remarks

In this paper, sparse grid spaces based on piecewise constant and linear functions are applied to the single layer potential equation on a square screen $\Gamma = I^2$. This equation is equivalent to a $\tilde{H}^{-1/2}$-elliptic variational problem, a case which has not been considered in connection with the sparse grid approach before. As is well-known for H^s-elliptic problems with $s > 0$, under special assumptions on higher order mixed derivatives (such as $f \in H^t_{\text{mix}}(I^d)$, instead of $f \in H^t(I^d)$, $t > s$), no loss of asymptotic approximation power occurs if the traditional full grid spaces are replaced by sparse grid spaces of much smaller dimension. Hence the sparse grid approach seems to be attractive for large scale discretizations in higher dimensional problems.

For $s < 0$, we showed in Proposition 6 that the approximation order in the sparse grid case slightly deteriorates with the dimension d if compared to the full grid case. This deterioration is still balanced with the substantially less number of grid points in the sparse grid case. Multilevel bases $\{\psi_{\mathbf{j},\mathbf{i}}\}$ built from tensor products of Haar functions resp. from linear spline prewavelets are used for both full and

Figure 7: Capacity error vs. compression rate, Haar basis ($L = 7$) and prewavelets ($L = 5$)



sparse grid spaces, and lead to well-conditioned representations for the arising stiffness matrices. The preconditioning result of Proposition 7 applies to conventional full and sparse grid spaces as well as to adaptive sparse grid spaces formed as the linear span of any suitable collection of $\psi_{\mathbf{j},\mathbf{i}}$.

We have focused in our numerical part on the case $\Gamma = I^2$ for the single layer potential equation. We show numerically that sparse grid spaces lead to Galerkin approximations in the $\tilde{H}^{-1/2}$-energy norm (measured by the error of capacity approximations) that are practically of almost the same order as obtained from the corresponding full grid spaces. The achieved dimension reduction can be further enhanced by using adaptive sparse grids constructed from the main singularity component. This eventually leads to low-dimensional ansatz spaces with impressive approximation potential in the non-asymptotical range. Although most of these findings are experimental, they are indicators for the potential of the sparse grid concept in applications which involve dense discretization matrices and/or solutions with strong edge singularities. In a certain sense, our approach to numerically solving this particular boundary integral equation can be interpreted as a method of *a priori compression* (by reducing the dimension of the ansatz space) and *a priori adaptivity* (by taking into account analytic information about the behavior of leading singularities). The efficient computation and further compression of discretization matrices and a posteriori solution adaptivity need still to be studied.

Due to the underlying tensor product construction, the sparse grid method as presented in this paper is directly applicable to polyhedral surfaces which consist of a few patches of rectangular shape. Generalizations to problems on curved surfaces Γ could be attempted using the ansatz made in [3, 13]. However, arbitrary manifolds can hardly be treated properly. Nevertheless, the numerical results obtained in section 3 might give some suggestions for the potential of anisotropic grid refinement, the construction of low-dimensional ansatz spaces, and the use of multilevel methods.

References

[1] G. Beylkin, R. Coifman, V. Rokhlin, Fast wavelet transforms and numerical algorithms I, Comm. Pure and Appl. Math. 44 (1991), 141-183.

[2] H.-J. Bungartz, Dünne Gitter und deren Anwendung bei der adaptiven Lösung der dreidimensionalen Poisson-Gleichung, Ph.D. Thesis, Technical University Munich (1992).

[3] H.-J. Bungartz and T. Dornseifer, Sparse grids: Recent developments for elliptic partial differential equations, Technical University Munich, SFB-Report 342/02/97 A, 1997.

[4] H.-J. Bungartz, M. Griebel, A note on the complexity of the Poisson equation and related elliptic equations for spaces of bounded mixed derivative, J. Complexity, submitted, 1997.

[5] C. K. Chui, J. A. Lian, A study of orthonormal multi-wavelets, J. Appl. Numer. Math. 20 (1996), 273–298.

[6] C. K. Chui, J. Z. Wang, On compactly supported spline wavelets, Trans. Amer. Math. Soc. 330 (1992), 903–915.

[7] M. Costabel, Boundary integral equations on Lipschitz domains: Elementary results, SIAM J. Numer. Anal. 19 (1988), 613–626.

[8] W. Dahmen, Wavelet and multiscale methods for operator equations, Acta Numerica 6 (1997), 55–228.

[9] W. Dahmen, B. Kleemann, S. Prössdorf, R. Schneider, A multiscale method for the double layer potential equation on a polyhedron, in Advances in Computational Mathematics (H. P. Dikshit, C. Micchelli, eds.), World Scientific, Singapore, 1994, 1–40.

[10] W. Dahmen, A. Kunoth, K. Urban, Wavelets in numerical analysis and their quantitative properties, in Proceedings of Chamonix 1996, A. Le Mèhautè, C. Rabut, and L. L. Schumaker (eds.), Vanderbild University Press, Nashville, TN, 1997

[11] R. A. DeVore, S. V. Konyagin, V. N. Temlyakov, Hyperbolic wavelet approximation, Constr. Approx. 14 (1998), 1–26.

[12] G. Donovan, J. S. Geronimo, D. P. Hardin, P. Massopust, Construction of orthonormal wavelets using fractal interpolation functions, SIAM J. Numer. Anal. 27 (1996), 1158–1192.

[13] T. Dornseifer, Diskretisierung allgemeiner elliptischer Differentialgleichungen in krummlinigen Koordinatensystemen auf dünnen Gittern, Ph.D. Thesis, Technical University Munich (1997).

[14] S. Abou El-Seoud, V. J. Ervin, E. P. Stephan, An improved boundary element method for the charge density of a thin electrified plate in $\mathbb{R}^3$, Math. Meth. Appl. Sci. 13 (1990), 291-303.

[15] K. Frank, Optimal numerical solution of multivariate integral equations, Ph.D. Thesis, University Kaiserslautern, Shaker Verlag, 1997.

[16] K. Giebermann, Schnelle Summationsverfahren zur numerischen Lösung von Integralgleichungen für Streuprobleme im $\mathbb{R}^3$, Ph.D. Thesis, University Karlsruhe (1997).

[17] L. F. Greengard, The Rapid Evaluation of Potential Fields in Particle Systems, MIT Press, Cambridge, Massachusetts, 1987.

[18] L. Greengard, V. Rokhlin, A new version of the fast multipole method for the Laplace equation in three dimensions, Acta Numerica 6 (1997), 229–269.

[19] M. Griebel, A domain decomposition method using sparse grids, in Proceedings DDM7, Contemporary Math. vol. 157, Amer. Math. Soc., Providence, 1994, 255–261.

[20] M. Griebel, P. Oswald, On additive Schwarz preconditioners for sparse grid discretizations, Numer. Math. 66 (1994), 449–463.

[21] M. Griebel, P. Oswald, Tensor product type subspace splittings and multilevel iterative methods for anisotropic problems, Adv. Comput. Math. 4 (1995), 171–206.

[22] M. Griebel, S. Knapek, Optimized approximation spaces for operator equations, University Bonn, SFB-256-Report, 1998.

[23] M. Griebel, S. Knapek, T. Schiekofer, Compression of Galerkin discretizations using wavelets of tensor-product type, in preparation.

[24] W. Hackbusch, Integral Equations, Theory and Numerical Treatment, Birkhäuser, Basel, 1995.

[25] W. Hackbusch, Iterative Solution of Large Sparse Systems of Equations, Springer, New York, 1994.

[26] W. Hackbusch, Z. P. Nowak, On the fast matrix multiplication in the boundary element method by panel clustering, Numerische Mathematik 54 (1989), 463-491.

[27] W. Hackbusch, C. Lage, S. A. Sauter, On the efficient realization of sparse matrix techniques for integral equations with focus on panel clustering, cubature and software design aspects, Berichtsreihe des Mathematischen Seminars Kiel, University Kiel, 1995

[28] N. Heuer, M. Maischak, E. P. Stephan, The hp-version of the boundary element method for screen problems, IFAM, University Hannover, Preprint ifam6, 1994 (submitted to Numer. Math.)

[29] H. Holm, M. Maischak, E. P. Stephan, The hp-version of the boundary element method for Helmholtz screen problems, Computing 57 (1996), 105–134.

[30] J.-L. Lions, E. Magenes, Non-Homogeneous Boundary Value Problems and Applications I, Springer, Berlin, 1972.

[31] M. Maischak, hp-Methoden für Randintegralgleichungen bei 3D-Problemen, Theorie und Implementierung, Ph. D. Thesis, IFAM, University Hannover, 1996.

[32] M. Maischak, P. Mund, E. P. Stephan, Adaptive multilevel BEM for acoustic scattering, Comput. Methods Appl.Mech.Engrg. 150 (1997) 351-367.

[33] P. Mund, E. P. Stephan, Adaptive two-level boundary element methods for the single layer potential in $\mathbb{R}^3$, IFAM, University Hannover, Preprint ifam14, 1997.

[34] G. Nießen, On the stability of sparse grid splittings, Report Nr. 115, IGPM, RWTH Aachen, 1995.

[35] P. Oswald, Multilevel Finite Element Approximation: Theory and Applications, Teubner, Stuttgart, 1994.

[36] P. Oswald, Multilevel-splitted norms for $H^{-1/2}$, Computing (1998), to appear (see also http://cm.bell-labs.com/who/poswald).

[37] T. von Petersdorff, R. Schneider, C. Schwab, Multiwavelets for second kind integral equations, SIAM J. Numer. Anal., Vol. 34 (1997), pp. 2212-2227.

[38] T. von Petersdorff, C. Schwab, Fully discrete multiscale Galerkin BEM, in Multiscale Wavelet Methods for PDEs (W. Dahmen, A. Kurdila, P. Oswald, eds.), Academic Press, San Diego, 1997, 287–346.

[39] T. von Petersdorff, E. P. Stephan, Decomposition in edge and corner singularities for the solution of the Dirichlet problem of the Laplacian in a polyhedron, Math. Nachr. 149 (1990), 71–104.

[40] T. von Petersdorff, E. P. Stephan, Regularity of mixed boundary value problems in $\mathbb{R}^3$ and boundary element methods on graded meshes, Math. Meth. Appl. Sci. 12 (1990), 229–249.

[41] V. Rokhlin, Rapid Solution of Integral Equations of Classical Potential Theory, Journal of Computational Physics 60(2) (1985), 187-207.

[42] S. Sauter, Über die effiziente Verwendung des Galerkinverfahrens zur Lösung Fredholmscher Integralgleichungen, Ph.D. Thesis, University Kiel (1992).

[43] H.-J. Schmeisser, H. Triebel, Topics in Fourier Analysis and Function Spaces, Geest& Portig, Leipzig, 1987; Wiley, Chichester, 1987.

[44] R. Schneider, Multiskalen- und Wavelet-Matrixkompression: Analysisbasierte Methoden zur effizienten Lösung grosser vollbesetzter Gleichungssysteme, Habilitation, Technical University Darmstadt, 1995.

[45] E. P. Stephan, Boundary integral equations for screen problems in $\mathbb{R}^3$, Int. Equations and Oper. Theory 10 (1987), 257–263.

[46] E. P. Stephan, The h-p boundary element method for solving 2- and 3-dimensional problems, IFAM, University Hannover, Preprint ifam8, 1995.

[47] V. N. Teml'yakov, Approximation of Periodic Functions, Nova Sci. Publ., New York, 1993.

[48] H. Triebel, Interpolation Theory, Function Spaces, Differential Operators, Dt. Verlag Wiss., Berlin 1978; North-Holland, Amsterdam, 1978.

[49] J. Weidmann, Linear Operators in Hilbert Spaces, Springer, New York, 1980.

[50] C. Zenger, Sparse grids, in Parallel Algorithms for PDE, Proc. 6th GAMM Seminar, Kiel (W. Hackbusch, ed.), Vieweg, Braunschweig, 1991, 241–251.