

Best N -term capacitance approximation on sparse grids

Peter Oswald

Bell Labs, Lucent Technologies

600 Mountain Av., Rm. 2C-403

Murray Hill, NJ 07974-0636, USA

e-mail: poswald@research.bell-labs.com

Abstract

In [5], adaptive sparse grid spaces spanned by a finite number of tensor-product L_2 -orthogonal Haar functions have been applied to capacitance calculations on a unit screen. In this note, we state asymptotically optimal approximation rates for this problem when choosing the best possible adaptive sparse grid space of a given dimension N . We also compare the results with other recent approaches to efficiently solve this problem and comment on some numerical tests. Details of the proofs and a discussion of the approximation-theoretical aspects will appear in [9].

1 Introduction

In [5], we obtained some theoretical and numerical results on the use of *sparse grid spaces* for the solution of the single layer potential equation in \mathbb{R}^3 . Specifically, for a flat square screen $I^2 \equiv [0, 1]^2$, the *single layer potential equation*

$$\frac{1}{4\pi} \int_{I^2} \frac{f(y)}{|x - y|_2} dy = g(x), \quad x \in I^2, \quad (1)$$

has been studied. As this problem can be cast in variational form and leads to a symmetric $H^{-1/2}$ -elliptic problem, Galerkin methods can be set up and allow for a straightforward analysis. E.g., convergence and error estimates in Sobolev norms (most naturally in the $H^{-1/2}$ -related energy norm) can be obtained for many natural discretization spaces.

There are two obstacles that trigger further investigations. First, one is interested in as small as possible computational subspaces since the discretization leads to *dense matrices* which is in contrast to the situation in finite element or finite difference methods for partial differential equations. Several approaches are under investigation (see [5] for a brief discussion) to overcome this problem. We only mention adaptive wavelet

compression schemes [2, 11] and the hp-version of the boundary element method [13] which will be used for comparison below. These methods also deal with the second obstacle: solutions of problems such as (1) exhibit very low global Sobolev smoothness due to dominant corner and edge singularities. For the important special case $g(x) \equiv 1$, the so-called *capacitance problem*

$$\frac{1}{4\pi} \int_{I^2} \frac{f(y)}{|x-y|_2} dy = 1, \quad x \in I^2, \quad (2)$$

the variational solution $f \in H^{-1/2}(I^2)$ does not even belong to $L_2(I^2)$. This leads to very slow convergence rates of any standard Galerkin method, both theoretically and practically. However, in analogy to elliptic problems in polyhedral domains, the ‘bad’ behavior of a solution f of (1) for smooth data g can be separated into a few singularity components associated with the edges and corners of I^2 , i.e., one can write

$$f = f^{sing} + f^{reg}, \quad (3)$$

where the *singular part* f^{sing} is a finite linear combination of specific, prescribed singularity functions (usually composed of terms of the form $\text{dist}(x, F)^\alpha$ and $\log(\text{dist}(x, F))$, where F is an edge or an vertex of I^2) while the *regular part* can be as smooth as wanted (limits are set by the smoothness class of g). See [10, 12] for details on the singularity decomposition (3) for (1) and similar screen problems. Thus, to obtain improved rates of convergence it would be enough to adapt the computational subspace such that it approximates the singularity functions in f^{sing} as well as the smooth part f^{reg} . In practical algorithms, this basic idea is implemented a priori (e.g., by using *graded meshes* in h- and hp-version boundary element methods [6]) or by using *feedback adaptivity schemes*, e.g., based on *a posteriori error estimators*, as suggested by different authors [6, 2].

Without becoming too detailed, let us mention some theoretical approximation results for the h- and hp-version of the boundary element method, on the one hand, and the wavelet schemes, on the other. Throughout the paper, wavelets are *semi-orthogonal spline wavelets* of low order m , even though results for this class of ansatz spaces are valid under much more general assumptions [2]. The boundary element spaces for the h-version are piecewise polynomials or splines of order m on certain sequences of partitions of I^2 (both quasi-uniform and adaptively refined ones) while in an hp-method, in addition, the polynomial degree may vary in each element of the underlying partition. Subsequently, we will specialize to the simplest case $m = 1$ of piecewise constant approximation. Our model problem will be (2). From [12] it follows that the singular part f^{sing} of the solution f of (2) is representable as a sum of singularity functions, with the leading singularities of the form $\sim \text{dist}(x, e)^{-1/2}$ near the interior part of any edge e of I^2 , and $\sim \text{dist}(x, P)^{\gamma-1}$, $\gamma \approx 0.2966\dots$, if x approaches a vertex P of I^2 . This leads to $f \in H^{-\varepsilon}(I^2)$ for any $\varepsilon > 0$, a result which brakes down (due to the edge singularity) for $\varepsilon = 0$. Throughout the paper, the notation ε stands for an arbitrarily small positive parameter.

To make a fair comparison between different approximation methods, we will relate error quantities to the dimension $N = \dim V_N$ of the computational subspace V_N from which the Galerkin solution is determined, and not to a meshsize parameter of the underlying partition or to the level number of a space in a wavelet multiresolution analysis. We admit that this way of comparison is still disputable since computational work and storage limitations may be quite different for subspaces with the same N . All estimates are given for the best approximations in the $H^{-1/2}(I^2)$ norm,

$$e_N(f)_{-1/2} = \inf_{v_N \in V_N} \|f - v_N\|_{H^{-1/2}},$$

which is equivalent to estimating the error in energy norm between the Galerkin solution $f_N \in V_N$ and the solution f of (2). Moreover, capacitance errors $\delta_N \equiv |\mathcal{C} - \mathcal{C}_N|$, where

$$\mathcal{C} = \frac{1}{4\pi} \int_{I^2} f(y) dy, \quad \mathcal{C}_N = \frac{1}{4\pi} \int_{I^2} f_N(y) dy, \quad (4)$$

are covered, too, since

$$\delta_N = \mathcal{C} - \mathcal{C}_N \asymp \|f - f_N\|_{H^{-1/2}}^2 \asymp e_N(f)_{-1/2}^2. \quad (5)$$

- *h-version with quasi-uniform partitions and fixed polynomial degree resp. non-adaptive wavelet spaces.* Here, standard estimates

$$e_N(u)_{-1/2} \leq C h_N^{t+1/2} \|u\|_{H^t}, \quad u \in H^t(I^2), \quad -1/2 < t \leq m,$$

hold with a mesh parameter $h_N \approx N^{-1/2}$, and lead in conjunction with the above regularity result for f to the estimate

$$e_N(f)_{-1/2} = O(N^{-(1/4-\varepsilon)}), \quad N \rightarrow \infty, \quad (6)$$

The asymptotic behavior in (6) is independent of m , and much worse than the saturation order $O(N^{-(m/2+1/4)})$ valid for approximating smooth functions from $H^m(I^2)$ with respect to the same spaces V_N .

- *h-version with graded meshes.* The estimate (6) can be improved if graded meshes are allowed for partitioning I^2 , see [12]. For I^2 , these are based on tensor-product partitions where the univariate partitions have $n \asymp \sqrt{N}$ grid points $\xi_i \in (0, 1)$ which behave like $\sim (i/n)^\beta$ near the left endpoint (analogous refinement is assumed at the right endpoint of $[0, 1]$). For appropriate β , the above mentioned saturation order can be reached:

$$e_N(f)_{-1/2} = O(N^{-(1/4+m/2)}), \quad N \rightarrow \infty, \quad (7)$$

for the associated spaces of piecewise polynomials or splines of order m on the above partitions, see [10, 12]. This improvement is achieved by allowing high aspect ratios of the rectangles (anisotropic refinement) near the edges.

- *hp-version on geometric meshes.* The best asymptotic estimates are known for the hp-method and a geometric tensor-product mesh (now the univariate meshes are given by $\xi_i \sim \sigma^{n/2-i}$, $\sigma < 1$, near the left endpoint of $[0, 1]$). The result for the particular case under consideration (see [6, 13]) is

$$e_N(f)_{-1/2} = O(e^{-cN^{1/4}}), \quad N \rightarrow \infty. \quad (8)$$

- *Adaptive wavelet approximation.* The basic idea is to determine a wavelet space V_N as the linear span of N carefully selected wavelets ψ_λ from different levels of the underlying multiresolution analysis. Theoretically, assuming that $f = \sum_\lambda c_\lambda \psi_\lambda$ is decomposed into a wavelet series, and that $\Psi = \{\psi_\lambda\}$ forms a Riesz basis in $H^{-1/2}$, the best one can do is to select the terms with the N largest $|c_\lambda|$. An algorithm which uses this basic idea has been described in [2]. The supporting approximation-theoretical result behind it has been known for some years [4, 7]. It is now referred to under the name *nonlinear N -term approximation* and has found important applications to image compression and adaptive algorithms, see [3]. The bad news, however, is that for our f this only leads to an estimate of

$$e_N(f)_{-1/2} = O(N^{-(1/2-\varepsilon)}), \quad N \rightarrow \infty, \quad (9)$$

again independently of m . This is a slight improvement over (6) but even the h-version on optimally chosen graded meshes with piecewise constants does asymptotically better than any adaptive wavelet space. The main reason is that for the wavelet bases considered in [2, 3] (and in most of the literature on solving boundary integral equations by wavelet methods), the nonlinear N -term approximation models *optimal isotropic local h -refinement*. Thus, for resolving the dominating edge singularities in the solution of (2), too many wavelet functions are necessary to improve the resolution along edges. This effect does not occur for point singularities and is practically invisible for edge singularities that are weaker than those exhibited by the solutions of screen problems (compare [1]).

Clearly, from the above one would prefer graded resp. geometric meshes (combined with h- resp. hp-methods) over wavelet type methods for the application under consideration. The exponential convergence of the hp-method is hard to beat in the asymptotic range. However, since the implementation of an hp-method for integral equations is by no means trivial, simpler and less optimal methods may still have a chance. E.g., well-understood adaptivity and compression strategies, preconditioning, and canonical data structures are some advantages of wavelet methods that one might wish to explore.

Improving upon the relatively weak approximation potential for solutions of screen problems while still working in a wavelet multiresolution analysis is suggested by the results on *adaptive sparse grid spaces* in [5]. In the present note we describe the approximation rates obtainable from these spaces in more quantitative terms. Roughly speaking, our general claim is that under the same assumptions on f , by changing from the

traditional, isotropic wavelet constructions on I^2 to tensor-product, anisotropic wavelet systems Ψ^* , the unsatisfactory rates of (9) can be replaced by

$$e_N^*(f)_{-1/2} = O(N^{-(1/4+m)}) , \quad N \rightarrow \infty , \quad (10)$$

where $e_N^*(f)_{-1/2}$ describes now the best N -term approximation with respect to the new wavelet system Ψ^* . Our point is that, even without going to graded meshes, we can expect good results if standard wavelet systems are replaced by tensor-product wavelet systems. We give precise statements for the case $m = 1$ (piecewise constant approximation) in section 2. Numerical experiments are presented in section 3.

2 N -term approximation by Haar functions

Let us give the definition of the Haar-wavelet systems ($m = 1$) under consideration. The characteristic function of a set Ω will be denoted by χ_Ω . Let \mathcal{D}_j be the system of dyadic intervals Δ of length $|\Delta| = 2^{-j}$, $j \geq 0$, of $I \equiv [0, 1]$. Any $\Delta \in \mathcal{D}_j$ uniquely splits into left (Δ^+) and right (Δ^-) half-intervals from \mathcal{D}_{j+1} . Set

$$\phi_\Delta = |\Delta|^{-1/2} \chi_\Delta , \quad \psi_\Delta = |\Delta|^{-1/2} (\chi_{\Delta^+} - \chi_{\Delta^-}) , \quad \Delta \in \mathcal{D} = \cup_{j \geq 0} \mathcal{D}_j ,$$

for the univariate scaled box functions and Haar functions, respectively. The *standard bivariate Haar system* is given by

$$\Psi_H = \cup_{j \geq 0} \Psi_j ,$$

where Ψ_0 consists of the only function χ_{I^2} , and

$$\Psi_j = \{ \psi_\Delta(x_1) \phi_{\Delta'}(x_2), \phi_\Delta(x_1) \psi_{\Delta'}(x_2), \psi_\Delta(x_1) \psi_{\Delta'}(x_2), \Delta, \Delta' \in \mathcal{D}_{j-1} \} , \quad j \geq 1 .$$

The supports of Haar functions from Ψ_j are dyadic squares of sidelength 2^{-j+1} , $j \geq 1$. In contrast, the Haar functions in the tensor-product bivariate Haar system

$$\Psi_H^* = \cup_{j_1, j_2 \geq 0} \Psi_{j_1, j_2}^* ,$$

where

$$\Psi_{j_1, j_2}^* = \{ \psi_\Delta(x_1) \psi_{\Delta'}(x_2), \Delta \in \mathcal{D}_{j_1-1}, \Delta' \in \mathcal{D}_{j_2-1} \} ,$$

possess rectangular support. For notational convenience, we defined $\mathcal{D}_{-1} = \{[0, 2]\}$ and $\psi_{[0,2]} = \phi_I$. Obviously, both systems are complete orthonormal systems in $L_2(I^2)$.

We are interested in the behavior of best N -term approximations with respect to Ψ_H^*

$$e_N^*(f)_s = \inf_{\Psi_N^* \equiv \{ \psi_1, \dots, \psi_N \} \subset \Psi_H^*} \inf_{v_N \in V_N^* \equiv \text{span } \Psi_N^*} \|f - v_N\|_{H^s} , \quad N \geq 1 , \quad (11)$$

in the $H^s(I^2)$ -norm. Due to the approximation and smoothness properties of piecewise constant functions, only the range $-1 < s < 1/2$ is of interest. Two main theorems

are established (for a detailed exposition and proofs, we refer to [9]). The first theorem serves functions from spaces of functions with dominating mixed derivatives which can be defined as tensor products of univariate Sobolev spaces:

$$H_{\text{mix}}^t(I^2) = H^t(I) \otimes H^t(I), \quad -\infty < t < \infty.$$

For $t = 0$, we have $H_{\text{mix}}^0(I^2) \cong L_2(I^2)$ while $f \in H_{\text{mix}}^1(I^2)$ if f belongs to $H^1(I^2)$ and additionally possesses a weak mixed derivative $\partial_{11}f \in L_2(I^2)$.

Theorem 1 *Let $f \in H_{\text{mix}}^t(I^2)$ for some $-1/2 < t \leq 1$. Then its best N -term approximations with respect to Ψ_H^* in $H^s(I^2)$, $N \geq 1$, satisfy*

$$e_N^*(f)_s \leq C \|f\|_{H_{\text{mix}}^t} \begin{cases} N^{-(t-s)}, & 0 < s < 1/2, \quad s < t \leq 1, \\ N^{-t}(1 + \log N)^t, & s = 0 < t < 1, \\ N^{-1}(1 + \log N)^{3/2}, & s = 0, \quad t = 1, \\ N^{-(t-s/2)}, & -1 < s < 0, \quad s/2 < t \leq 1. \end{cases} \quad (12)$$

In particular, if $f \in H_{\text{mix}}^1(I^2)$ then

$$e_N^*(f)_{-1/2} \leq CN^{-5/4} \|f\|_{H_{\text{mix}}^1}, \quad N \rightarrow \infty. \quad (13)$$

This estimate is applicable to the smooth part f^{reg} of solutions to (1), to achieve the $O(N^{-5/4})$ error bound in practice, one can, e.g., take subspaces spanned by the following subset of $\asymp 2^J$ Haar functions:

$$\Psi_{H,J}^* = \cup_{j_1, j_2 \geq 0 : j_1 + j_2 + \frac{1}{4} \max(j_1, j_2) \leq \frac{9}{8} J} \Psi_{j_1, j_2}^*. \quad (14)$$

Note that $\Psi_{H,J}^*$ spans a subspace of the standard sparse grid space of level J .

The second result covers certain types of *singularity functions*. We call $f \in L_1(I^2)$ an *edge singularity function with exponent* $\alpha \in [0, 1)$ if it has a continuous derivative $\partial_{11}f$ in the open square $(0, 1)^2$ and satisfies

$$|\partial_{kl}f(x_1, x_2)| \leq C(\min(x_1, 1 - x_1))^{-\alpha-k}(\min(x_2, 1 - x_2))^{-\alpha-l}, \quad (x_1, x_2) \in (0, 1)^2.$$

E.g., the singular part f^{sing} in (3) of the solution f of (2) possesses this property with $\alpha = 0.7034\dots$ (a more detailed analysis shows that for the capacitance problem better representations of f^{sing} can be found which would lead to edge singularity functions with $\alpha = 1/2$ as the appropriate value).

Theorem 2 *Let $-1 < s < 1/2$, and f be an edge singularity function with exponent α , where $0 \leq \alpha < \min(1/2 - s, 1/2 - s/2)$. Then $f \in H^s(I^2)$ and*

$$e_N^*(f)_s \leq C \begin{cases} N^{-(1-s)}, & 0 < s < 1/2, \\ N^{-1}(\log N)^{3/2}, & s = 0, \\ N^{-(1-s/2)}, & -1 < s < 0, \end{cases} \quad N \rightarrow \infty. \quad (15)$$

Roughly speaking, by optimally choosing N Haar functions from $\Psi_H^*(I^2)$, an edge singularity function with exponent α satisfying the above condition possesses the same asymptotic N -term approximation rate as smooth functions from $H_{\text{mix}}^1(I^2)$. For the case $s = -1/2$, we can have $0 \leq \alpha < 3/4$ which leads according to our above remarks to

$$e_N^*(f^{\text{sing}})_{-1/2} \leq CN^{-5/4}, \quad N \rightarrow \infty, \quad (16)$$

for the singular part of the solution f of the capacitance problem (2). Since, at the same time, we can assume that $f^{\text{reg}} \in H_{\text{mix}}^1(I^2)$ in (3), the two estimates (13) and (16) yield an analogous estimate for f itself. Finally, from (4) we see that the capacitance \mathcal{C} of the unit square screen can be approximated at a rate of $O(N^{-5/2})$ if optimal selections of N Haar functions from $\Psi_H^*(I^2)$ are used to build discretization spaces.

3 Numerical tests

This part complements the numerical experiments of [5, section 3.3-4]. There, capacitance approximations have been computed for full grid, sparse grid and adaptive sparse grid spaces. To reach a relative error of approximately 10^{-3} , subspaces of dimension $N = 65536$, $N = 1280$, and $N = 104$, respectively, were needed.

The proofs of Theorem 1 and 2 suggest that, in order to achieve the asymptotical error estimate $O(N^{-5/2})$, it is sufficient to take the union of the set $\Psi_{J_0, H}^*$ defined in (14) which serves the regular part f^{reg} , and a set Ψ_N^* consisting of $N \asymp 2^{J_0}$ functions from $\Psi_H^*(I^2)$ producing the N largest contributions to the upper bound

$$\|f^{\text{sing}}\|_{H^{-1/2}}^2 \leq C \sum_{j_1, j_2} \sum_{\psi \in \Psi_{j_1, j_2}^*} 2^{-\max(j_1, j_2)} |c_\psi(f^{\text{sing}})|^2 \quad (17)$$

for the singular part f^{sing} from (3). The proof of Theorem 2 also shows that the unknown Haar-Fourier coefficients $c_\psi(f^{\text{sing}})$ can be replaced by computable upper bounds. These, in turn, can be obtained from using appropriate majorants for f^{sing} (such as appearing in the definition of edge singularity functions with exponent α or obtained directly from the available singularity decompositions, see [6, 5]). Tuning N , J_0 , and choosing different majorants may lead to further improvement. In our experiments, the sets Ψ_N^* have been obtained from thresholding the sequence $\{2^{-\max(j_1, j_2)} |c_\psi(f^\alpha)|^2\}$ for the function $f^\alpha(x_1, x_2) = x_2^{-\alpha}$ (which mimics a univariate singularity along the edge $x_2 = 0$ of the unit square) and a straightforward symmetrization step (note that for the solution of (2) satisfies $f(x_1, x_2) = f(x_1, 1 - x_2) = f(1 - x_1, x_2)$). Using the values $\alpha = 1/2$, $J_0 = 3$, we found that the above-mentioned relative error of 10^{-3} can be reached by using 41 ansatz functions (the constant function from $\Psi_{0,0}^*$ and four functions with support along the edges from each of the sets $\Psi_{0,j}^* \cup \Psi_{j,0}^*$, $j = 2, \dots, 11$). This hints at the importance of dealing with the edge singularities adequately, and in the first place.

Let us give the numerical results in more detail. We start with giving values for the relative capacitance errors

$$\delta_N^{\text{rel}} = |\mathcal{C} - \mathcal{C}_N|/\mathcal{C} \quad (\mathcal{C} = 0.366789\dots)$$

obtained from discretizing (2) with respect to *full grid (fg) spaces* of level J ,

$$V_J(I^2) = \text{span } \cup_{0 \leq j_1, j_2 \leq J} \Psi_{j_1, j_2}^* , \quad \dim V_J(I^2) = 2^{2J} ,$$

sparse grid (sg) spaces of level J ,

$$V_J^*(I^2) = \text{span } \Psi_J^*(I^2) , \quad \Psi_J^*(I^2) \equiv \cup_{0 \leq j_1 \leq j_1 + j_2 \leq J} \Psi_{j_1, j_2}^* , \quad \dim V_J^*(I^2) = (J + 2)2^{J-1} ,$$

$1 \leq J \leq 8$, and two sequences of *adaptive sparse grid (asg) spaces* $V_N^* = \text{span } \Psi_N^*$ spanned by certain sets of N Haar functions $\psi \in \Psi_H^*(I^2)$, respectively. The numbers for the first three columns in Table 1 are taken from [5], where also the construction of the corresponding Ψ_N^* is described. In the fourth column, new adaptive sparse grid spaces associated with Theorems 1 and 2 (see below for details) are used. The numbers demonstrate the potential of best N -term approximation with respect to the tensor-product Haar basis $\Psi_H^*(I^2)$ for the problem at hand.

Table 1: Capacitance errors for various V_N

fg-spaces		sg-spaces		asg-spaces [5]		new asg-spaces	
N	δ_N^{rel}	N	δ_N^{rel}	N	δ_N^{rel}	N	δ_N^{rel}
4	0.08302	3	0.08302	20	0.02511	9	0.02516
16	0.04584	8	0.04589	32	0.01348	17	0.00738
64	0.02490	20	0.02511	44	0.00730	25	0.00254
256	0.01310	48	0.01340	56	0.00409	33	0.00130
1024	0.00677	112	0.00708	68	0.00245	41	0.00098
4096	0.00346	256	0.00373	80	0.00162	61	0.00069
16384	0.00175	576	0.00197	92	0.00121	81	0.00051
65536	0.00089	1280	0.00105	104	0.00100	101	0.00036

The last column in Table 1 represents the final result of a series of other tests. First, we ran tests with various J_0 , α , and threshold values $\theta > 0$, to determine suitable sets Ψ_N^* , and to study the sensitivity with respect to various parameters. More precisely, we have set

$$\Psi_N^* \equiv \Psi^*(\alpha, J_0, \theta) = \Psi_{J_0}^*(I^2) \cup (\cup_{k=1}^4 \Psi_k^*(\alpha, \theta)) \quad (18)$$

where

$$\Psi_1^*(\alpha, \theta) = \{\psi \in \Psi_{0, j}^* : 2^{-j/2} |c_\psi(f^\alpha)| \leq \theta, j \geq 0\} ,$$

corresponds to the above univariate singularity function f^α associated with the edge $x_2 = 0$ (the sets $\Psi_k^*(\alpha, \theta)$, $k = 2, 3, 4$, are defined from $\Psi_1^*(\alpha, \theta)$ by symmetry arguments and correspond to the other three edges). The derivation of explicit formulas for the Haar-Fourier coefficients $c_\psi(f^\alpha)$ represents no difficulty. Note that the behavior of f^{sing} near the vertices of the square is not taken into account by our choices in (18), it can only be absorbed by the ansatz functions from $\Psi_{J_0}^*$ and for sufficiently large J_0 . We have also decided to neglect the differences between $\Psi_{J_0, H}^*$ (see (14)) and $\Psi_{J_0}^*(I^2)$ which show

up only for larger J_0 . The symmetry properties of f also show that $c_\psi(f) = 0$ for all $\psi \in \cup_{j \geq 0} (\Psi_{j,1}^* \cup \Psi_{1,j}^*)$. We have neglected all these functions in Ψ_N^* before determining N for the tables below. In Table 2, we keep $\alpha = 1/2$ and vary J_0 and θ . The second column gives the value $j_{\max} = \max(j : \Psi_1^*(1/2, \theta) \cap \Psi_{0,j}^* \neq \emptyset)$. The relative capacitance errors shown indicate that an increase of J_0 is only helpful if a relative error $\leq 10^{-3}$ is needed. Also, the increase of J_0 leads to a significant growth of N while only a relatively small improvement of capacitance approximations can be observed. A possible reason is that the remaining significant error contributions are associated with corner singularities not yet resolved by the functions in $\Psi^*(\alpha, J_0, \theta)$, $J_0 \leq 7$.

Table 2: Results for different J_0 ($\alpha = 0.5$)

θ	j_{\max}	$J_0 = 3$		$J_0 = 4$		$J_0 = 5$		$J_0 = 6$		$J_0 = 7$	
		N	δ_N^{rel}	N	δ_N^{rel}	N	δ_N^{rel}	N	δ_N^{rel}	N	δ_N^{rel}
0.110	5	13	0.01356	17	0.01346	-	-	-	-	-	-
0.060	7	21	0.00418	25	0.00406	57	0.00390	133	0.00375	-	-
0.030	9	29	0.00171	33	0.00160	65	0.00143	141	0.00128	329	0.00115
0.015	11	41	0.00107	45	0.00097	73	0.00081	149	0.00065	337	0.00052
0.008	13	57	0.00086	61	0.00077	85	0.00063	157	0.00049	345	0.00036
0.004	15	81	0.00080	85	0.00070	105	0.00057	173	0.00044	357	0.00032
0.002	17	113	0.00078	117	0.00068	133	0.00055	197	0.00042	377	0.00030
0.001	19	157	0.00077	161	0.00068	177	0.00055	233	0.00041	405	0.00030

Table 3 documents the influence of the parameter α . This time, $J_0 = 4$ has been fixed, and values $\alpha = 0.3, 0.4, 0.6, 0.7$ are tried (recall that $\alpha = 0.5$ is covered by the second test set in Table 2). Naturally, j_{\max} (which is not shown) increases with α , for each fixed threshold θ . Since we have fixed J_0 , errors cannot be reduced indefinitely (they stagnate at a value of about 0.00067). Away from this error values, the results are insensitive to α . We conclude that an exact estimation of edge singularity exponents is only necessary to obtain close-to-optimal results in the asymptotic range $\delta_N^{rel} \leq 10^{-3}$.

Table 3: Results for different α ($J_0 = 4$)

θ	$\alpha = 0.3$		$\alpha = 0.4$		$\alpha = 0.6$		$\alpha = 0.7$	
	N	δ_N^{rel}	N	δ_N^{rel}	N	δ_N^{rel}	N	δ_N^{rel}
0.110	17	0.02503	17	0.02503	29	0.00406	41	0.00118
0.060	-	-	25	0.00727	37	0.00160	53	0.00081
0.030	25	0.00727	29	0.00406	49	0.00087	65	0.00077
0.015	29	0.00406	37	0.00160	57	0.00079	77	0.00076
0.008	33	0.00243	45	0.00115	81	0.00071	101	0.00070
0.004	49	0.00113	65	0.00080	105	0.00069	141	0.00068
0.002	69	0.00090	89	0.00071	153	0.00068	-	-
0.001	97	0.00073	121	0.00069	201	0.00067	-	-

Further insight can be gained from the following a posteriori analysis. Assume that we have computed the Galerkin solution f_{N^*} of (2) with respect to a subspace generated by a certain set Ψ^* containing N^* Haar functions $\psi \in \Psi_H^*(I^2)$. We can write f_{N^*} as a Haar polynomial

$$f_{N^*} = \sum_{\psi \in \Psi^*} c_\psi^* \psi .$$

Consider the sequence $d_\psi = 2^{-j_\psi/2} |c_\psi^*|$, where $j_\psi = \max(j_1, j_2)$ if $\psi \in \Psi_{j_1, j_2}^* \cap \Psi^*$, enumerate the functions in Ψ^* with respect to decreasing values of d_ψ :

$$\Psi^* = \{ \psi_k : d_{\psi_k} \geq d_{\psi_l}, 1 \leq k \leq l \leq N^* \} .$$

We look for smaller sets $\Psi_N^* \subset \Psi^*$ ($N \leq N^*$) to generate discretization spaces (this is equivalent to a restricted version of best N -term approximation: only subsets of Ψ^* are allowed). Let f_N denote the Galerkin solution with respect to such a space $V_N = \text{span } \Psi_N^*$. Then

$$\|f - f_N\|_{H^{-1/2}} \leq C_0 \inf_{v_N \in V_N} \|f - v_N\|_{H^{-1/2}} \leq C_0 (\|f - f_{N^*}\|_{H^{-1/2}} + \inf_{v_N \in V_N} \|f_{N^*} - v_N\|_{H^{-1/2}}) .$$

By [8, Theorem 2], we have

$$\frac{C_1}{(J^*)^2} \sum_{\psi \in \Psi^*} 2^{-j_\psi} |c_\psi|^2 \leq \left\| \sum_{\psi \in \Psi^*} c_\psi \psi \right\|_{H^{-1/2}}^2 \leq C_2 \sum_{\psi \in \Psi^*} 2^{-j_\psi} |c_\psi|^2 , \quad (19)$$

where $J^* = \max_{\psi \in \Psi^*} j_\psi$ and $0 < C_1, C_2 < \infty$ represent absolute constants. Using the upper estimate of (19), we get

$$\|f - f_N\|_{H^{-1/2}} \leq C_0 \left(\left(C_2 \sum_{\psi \in \Psi^* \setminus \Psi_N^*} d_\psi^2 \right)^{1/2} + \|f - f_{N^*}\|_{H^{-1/2}} \right) ,$$

Using the lower bound, one gets

$$\|f - f_N\|_{H^{-1/2}} \geq C_3 \left(\left(\frac{C_1}{(J^*)^2} \sum_{\psi \in \Psi^* \setminus \Psi_N^*} d_\psi^2 \right)^{1/2} - \|f - f_{N^*}\|_{H^{-1/2}} \right) .$$

Provided that N is small compared to N^* and neglecting $\|f - f_{N^*}\|_{H^{-1/2}}$, we see that

$$\Psi_N^* = \{ \psi_k : k = 1, \dots, N \} \quad (20)$$

is probably a good choice (it minimizes the upper bound for $\|f - f_N\|_{H^{-1/2}}$).

Table 4 shows results for this kind of a posteriori analysis. In the first experiment, we have used

$$\Psi^* = \Psi^*(0.5, 1, 0.001) \cup \left(\bigcup_{2 \leq j_1, j_2, j_1 + j_2 \leq 6} \Psi_{j_1, j_2}^* \right) \quad (N^* = 277) ,$$

which is a slightly smaller subset of $\Psi^*(0.5, 6, 0.001)$. In a second experiment more Haar functions from Ψ_{j_1, j_2}^* with $j_1, j_2 \geq 2$ and $7 \leq j_1 + j_2 \leq 12$ with supports near the

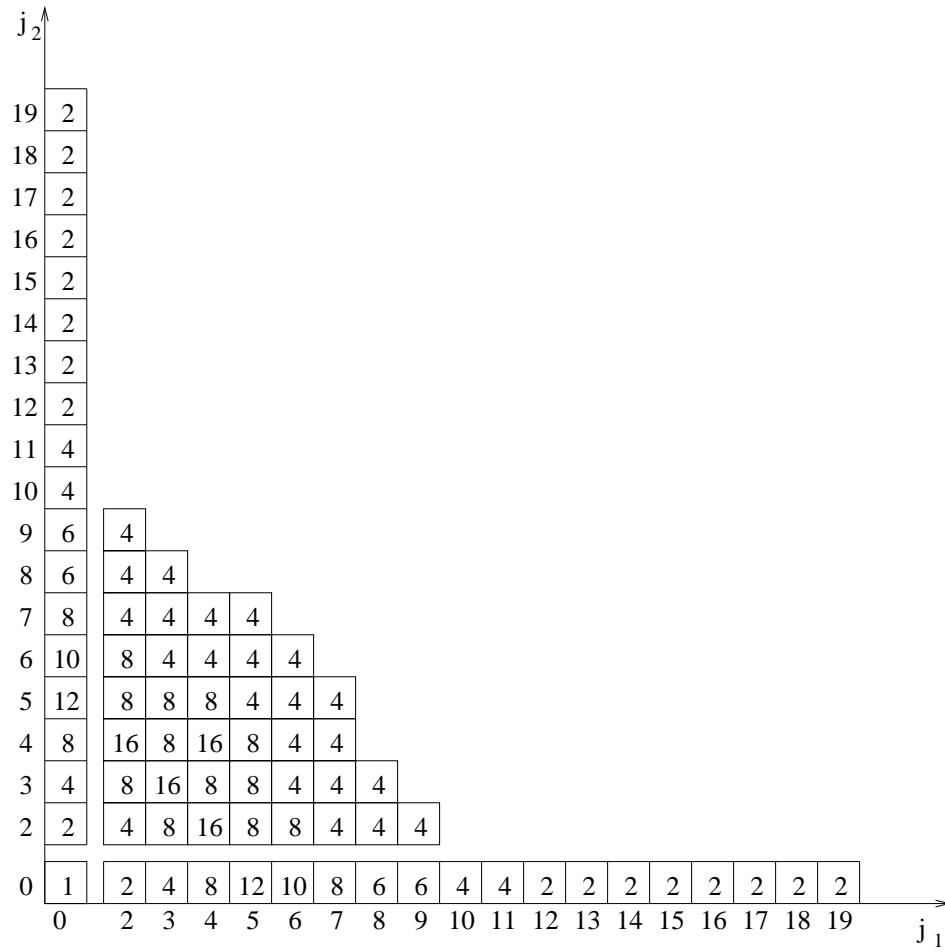


Figure 1: Distribution of $\psi \in \Psi^*$ with respect to $\Psi^*_{j_1, j_2}$

vertices have been included. This resulted in a set Ψ^* of size $N^* = 409$. The number of Haar functions taken from each Ψ_{j_1, j_2}^* is shown in Figure 1, they are symmetrically associated to the vertices/edges of the square to ensure better resolution of f^{sing} . Recall that no functions from $\cup_{j \geq 0} (\Psi_{j,1}^* \cup \Psi_{1,j}^*)$ are needed. We give values of δ_N^{rel} in the range $\leq N^*/2$ resp. until error saturation ($\delta_N^{rel} \asymp \delta_{N^*}^{rel}$) is achieved. In addition, the quantity $C(N) = N^{5/2} \delta_N$ is shown. Since the results differ only for $N > 75$, where further, optimal error reduction seems to necessitate better treatment of the vertex singularities and the smooth part, values for lesser N are not repeated for the second test. Choosing N to be of the form $N = 4n + 1$ is dictated by the symmetry properties of the solution f of (2).

Table 4: A posteriori construction of Ψ_N^*

Test 1			Test 2		
N	δ_N^{rel}	$C(N)$	N	δ_N^{rel}	$C(N)$
13	0.01356	3.03	77	0.000525	10.02
21	0.00418	3.10	85	0.000477	11.65
29	0.00171	2.85	93	0.000419	12.83
37	0.00109	3.33	101	0.000364	13.70
45	0.00089	4.44	109	0.000324	14.76
53	0.00081	6.06	117	0.000297	16.14
61	0.00069	7.40	125	0.000263	16.84
69	0.00060	8.68	141	0.000213	18.48
77	0.00053	10.02	157	0.000177	20.10
85	0.00048	11.71	173	0.000152	21.93
93	0.00046	13.93	189	0.000128	23.00
101	0.00044	16.65	205	0.000111	24.52
277	0.00041	-	409	0.000058	-

Figure 2 shows the distribution of the Haar functions in Ψ_N^* with respect to the Ψ_{j_1, j_2}^* for $N = 41, 77$, and 117 for the second test (gray color is used for the supports of the corresponding ψ_k with $k \leq 41$, white for $42 \leq k \leq 77$, and black for $78 \leq k \leq 117$). In each $\Psi_{j,0}^*$ resp. $\Psi_{0,j}^*$, $j = 7, \dots, 13$, exactly two functions associated with opposite edges are included into Ψ_{117}^* .

Looking at the numbers $C(N)$ given in Table 4, there remains some doubt about the constant in the asymptotic $O(N^{-5/2})$ estimate for δ_N^{rel} following from the results of section 2 (what we should expect is $C(N) \leq C$, at least asymptotically). Possibly, our a posteriori selection rule outlined above needs some improvement (note that $J^* = 19$ in both tests). Numerical instabilities in the computation of entries of the Galerkin discretization matrices for ψ with large j_ψ might be another reason for reduced approximation rates, see the discussion in [5, section 3.2]. However, even though it would be of some theoretical interest to numerically document this error behavior, from a more practical point of view it is more important to realize that the error reduction through

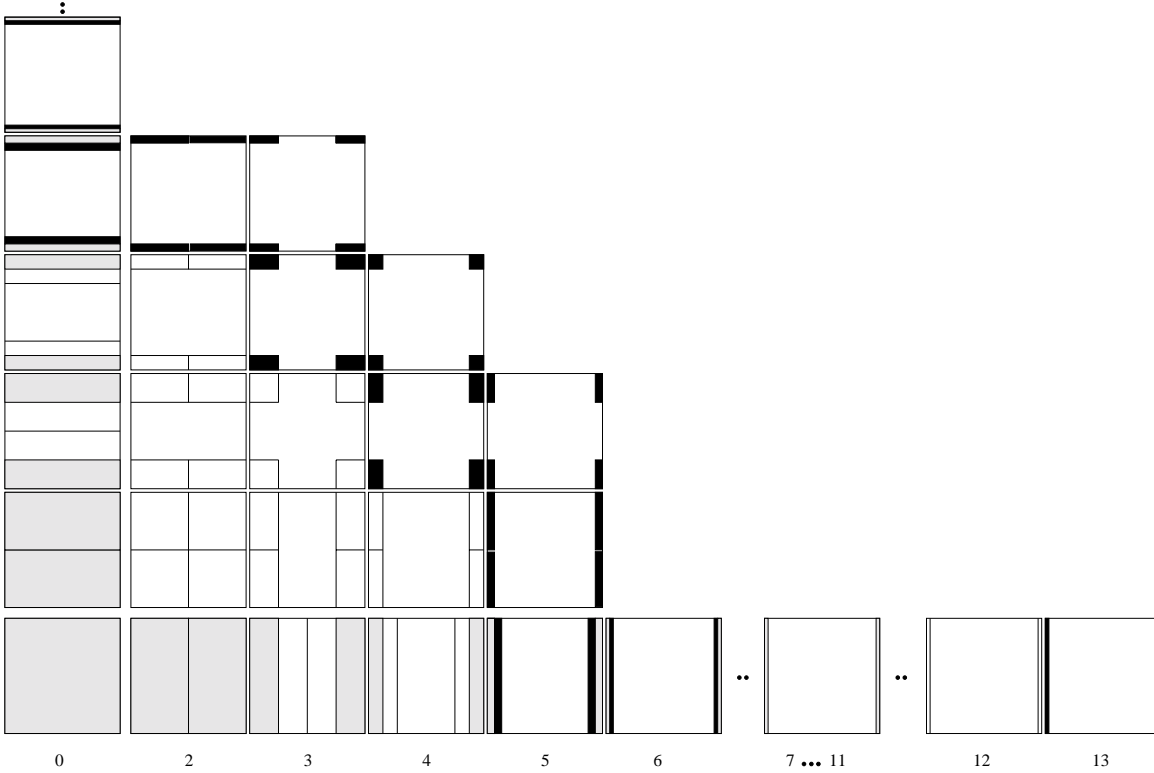


Figure 2: Distribution of $\psi \in \Psi_N^*$ with respect to Ψ_{j_1, j_2}^* , $N = 41, 77, 117$.

best N -term approximation techniques is most impressive in the range of small $N \leq 50$, where $\delta_{41}^{rel} < 10^{-3}$ is achieved by a pure 'edge refinement'. This observation should be helpful also for more realistic applications of boundary integral equations: anisotropic refinement along (sharp) edges of the surface Γ under consideration might improve solution quality considerably. In contrast, to guarantee relative errors of 10^{-4} and less, one seemingly needs $N \geq 200$ Haar basis functions. For this range, methods based on higher order polynomials ($m \geq 2$) might be more adequate.

References

- [1] S. Dahlke, R. A. DeVore, Besov regularity for elliptic boundary value problems, *Commun. Partial Diff. Eqns.* 22 (1997), 1–16.
- [2] W. Dahmen, Wavelet and multiscale methods for operator equations, *Acta Numerica* 6 (1997), 55–228.
- [3] R. A. DeVore, Nonlinear approximation, *Acta Numerica* 7 (1998), 51–150.

- [4] R. A. DeVore, B. Jawerth, V. Popov, Compression of wavelet decompositions, *Amer. J. Math.* 114 (1992), 737–785.
- [5] M. Griebel, P. Oswald, T. Schiekofer, Sparse grids for boundary integral equations, *Numer. Math.* (submitted).
- [6] N. Heuer, M. Maischak, E. P. Stephan, The hp-version of the boundary element method for screen problems, *Numer. Math.*, to appear (Preprint ifam6, IFAM, University Hannover, 1994).
- [7] P. Oswald, On the degree of nonlinear spline approximation in Besov-Sobolev spaces, *J. Approx. Th.* 61 (1990), 131-157.
- [8] P. Oswald, Multilevel norms for $H^{-1/2}$, *Computing* (1998), to appear.
- [9] P. Oswald, Best N -term approximation by Haar functions in H^s -norms, in *Approximation and Fourier Series* (S. M. Nikolskij, B. S. Kashin, A. Izaak, eds.), AFC, Russian Academy of Sciences, 1998, submitted.
- [10] T. von Petersdorff, Randwertprobleme der Elastizitätstheorie für Polyeder-Singularitäten und Approximation mit Randelementmethoden, PhD Thesis, TH Darmstadt, 1989.
- [11] 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.
- [12] 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.
- [13] E. P. Stephan, The h-p boundary element method for solving 2- and 3-dimensional problems, Preprint ifam8, IFAM, University Hannover, 1996.