Compression of boundary integral operators discretized by anisotropic wavelet bases

Helmut Harbrecht, Remo von Rickenbach

Departement Mathematik und Informatik
Fachbereich Mathematik
Universität Basel
CH-4051 Basel

Preprint No. 2023-04
April 2023

dmi.unibas.ch

COMPRESSION OF BOUNDARY INTEGRAL OPERATORS DISCRETIZED BY ANISOTROPIC WAVELET BASES

HELMUT HARBRECHT AND REMO VON RICKENBACH

Abstract. The present article is devoted to wavelet matrix compression for boundary integral equations when using anisotropic wavelet bases for the discretization. We provide a compression scheme which amounts to only $\mathcal{O}(N)$ relevant matrix coefficients in the system matrix without deteriorating the accuracy offered by the underlying Galerkin scheme. Here, N denotes the degrees of freedom in the related trial spaces. By numerical results we validate our theoretical findings.

1. Introduction

Many problems from engineering result in partial differential equations, which can often be solved efficiently by using the finite element method [3, 4]. However, if the loading of the equation is zero, under some circumstances, one may transform the partial differential equation in a domain into an integral equation on its boundary. Such an integral equation can then be solved using the boundary element method [23, 26]. There are many practical problems which can be treated with the boundary element method, such as for example the Laplacian or linear elasticity problems [22, 26], scattering problems [6], and homogenization problems [1, 5, 19].

A huge advantage of the boundary element method is that the integral domain under consideration is reduced from an n-dimensional surface to a $(n-1)$-dimensional hypersurface. This brings us a significant reduction in the number of the degrees of freedom, but since the integral kernels are, in general, nonlocal, also densely populated matrices.

To overcome the dense matrices, fast boundary element methods have been developed, such as *adaptive cross adaptation* [2], the *fast multipole method* [12], or the *wavelet matrix compression* [8, 24]. A comparison of these methods with respect to their advantages and disadvantages can be found in [14] for example. The computational cost of all these methods scale linearly or loglinearly in the number of degrees of freedom. Indeed, the wavelet matrix compression has been shown to have linear cost complexity, compare [8]. Moreover, the boundary integral operator is $s^\star$*-compressible* [27], with the consequence of a quasi-optimal convergence for the adaptive wavelet boundary element method [9, 11, 17, 28].

However, all these works consider isotropic wavelets, meaning that the mesh of the underlying multiscale hierarchy consists of isotropic elements. Therefore, only anisotropic singularities cannot be resolved properly. This leads to a loss in the convergence rate if the solution of the boundary integral equation exhibits such singularites. Anisotropic singularities, however, appear if the boundary under under consideration contains edges as this is the case, for example, for *Fichera's corner* [17]. This gives rise to considering anisotropic tensor product wavelets, which are allowed to refine in one coordinate direction whilst staying coarse in the other coordinate direction. With such wavelet functions, the disadvantage of the isotropic wavelets might be overcome.

Anisotropic tensor wavelets for boundary integral equations have been considered first in [13] in the context of sparse tensor product approximations. In [20], for both isotropic and nonisotropic boundary integral operators which are discretized with respect to sparse tensor product spaces, a compression scheme has been developed. This scheme which leads to an essentially linear number of the degrees of freedom therein, provided the underlying integro-differential operator is of the order $2q > \frac{1}{2}(\sqrt{5}-1) > 0$. We, on the other hand, will construct in the present article a *linearly* scaling compression scheme for integral operators of arbitrary order which are discretized with respect to the full tensor product space. In particular, our compression estimates can be used to improve the results from [20]. Note, however, that the computation of the matrix entries of the compressed system matrix is not a topic of the present article. This can be done by using the techniques and results of [15, 29].

The rest of the article is structured as follows. In Section 2, we introduce the boundary integral equation to be solved. Then, in Section 3, we define the anisotropic wavelet basis we shall use for the discretization on the unit square. Estimates on the size of the entries of the respective Galerkin matrix with respect to the unit square are derived in Section 4. The wavelet matrix compression is proposed in Section 5. The number of the remaining nonzero matrix entries is counted in Section 6. In Section 7, we generalize the wavelet matrix compression to the boundary of a Lipschitz domain. Consistency and convergence of the wavelet matrix compression is proven in Section 8. In Section 9, we provide numerical experiments to validate the results derived. Finally, in Section 10, we state concluding remarks.

Throughout the article, let us replace generic constants by the notation $A \lesssim B$, which means that A is bounded by a constant multiple of B, and, similarly we define $A \gtrsim B$ if and only if $B \lesssim A$. If $A \lesssim B$ and $B \lesssim A$, we write $A \sim B$. Moreover, if $\mathbf{j}, \mathbf{j}' \in \mathbb{N}_0^2$ are given multiindices, the inequality $\mathbf{j} \leq \mathbf{j}'$ is understood componentwise. Especially, the notion $\mathbf{j} < \mathbf{j}'$ means that $\mathbf{j} \leq \mathbf{j}'$ and $\mathbf{j} \neq \mathbf{j}'$. Finally, we set $\mathbf{1} := (1,1)$.

2. Problem Formulation

2.1. **Parametrization.** Throughout this article, we consider a bounded, piecewise smooth domain $\Omega \subseteq \mathbb{R}^3$ with Lipschitz boundary $\Gamma := \partial\Omega$. We us assume that Γ can be decomposed into r four-sided, smooth patches Γ_i, $i = 1, \ldots, r$, such that

$$\Gamma = \bigcup_{i=1}^{r} \Gamma_i.$$

This decomposition needs to be admissible, meaning that the intersection $\Gamma_i \cap \Gamma_j$ is for $i \neq j$ either empty, a common vertex, or a common edge of both Γ_i and Γ_j, cf. Figure 1. We next choose smooth diffeomorphisms $\boldsymbol{\gamma}_i : \square := [0,1]^2 \to \Gamma_i$ such that there exist constants c_i and C_i with

$$0 < c_i \leq \sqrt{\det\left(\mathbf{D}\boldsymbol{\gamma}_i(\boldsymbol{s})^\intercal \mathbf{D}\boldsymbol{\gamma}_i(\boldsymbol{s})\right)} \leq C_i < \infty, \quad \boldsymbol{s} \in \square. \tag{1}$$

This parametrization should fulfill the matching condition that $\boldsymbol{\gamma}_i$ and $\boldsymbol{\gamma}_j$ coincide up to orientation at a common edge of two neighbouring patches Γ_i and Γ_j.

2.2. **Boundary integral equation.** In the following, we intend to calculate the solution u of the boundary integral equation

$$\mathcal{A}u(\boldsymbol{x}) := \int_\Gamma K(\boldsymbol{x}, \boldsymbol{y}) u(\boldsymbol{y}) \, \mathrm{d}S_{\boldsymbol{y}} = g(\boldsymbol{x}), \quad \boldsymbol{x} \in \Gamma, \tag{2}$$

where $\mathcal{A} : H^q(\Gamma) \to H^{-q}(\Gamma)$ is a given boundary integral operator. Typically, the kernel k is asymptotically smooth of the order $2q$, that is, K is singular only on the

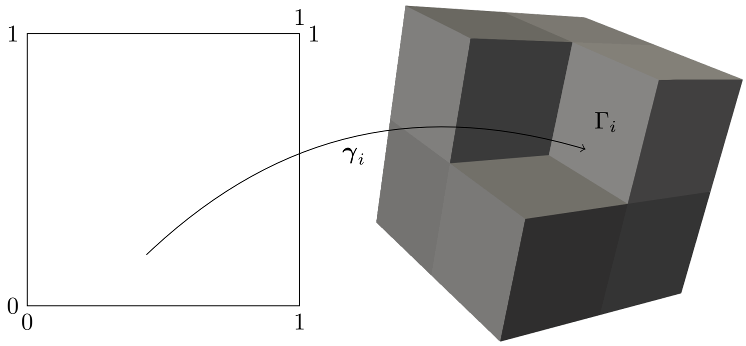

FIGURE 1. A parametrization of Fichera's vertex. The different shadings represent the different $r = 24$ patches Γ_j.

diagonal $\{\boldsymbol{x} = \boldsymbol{y}\}$ and smooth apart from it in terms of

$$\left|\partial_{\boldsymbol{x}}^{\boldsymbol{\alpha}} \partial_{\boldsymbol{y}}^{\boldsymbol{\beta}} K(\boldsymbol{x}, \boldsymbol{y})\right| \leq C_{\boldsymbol{\alpha},\boldsymbol{\beta}} \|\boldsymbol{x} - \boldsymbol{y}\|^{-(2+2q+|\boldsymbol{\alpha}|+|\boldsymbol{\beta}|)}, \tag{3}$$

provided that $2 + 2q + |\boldsymbol{\alpha}| + |\boldsymbol{\beta}| > 0$. We assume that $\mathcal{A}$ is bounded and strongly elliptic on $H^q(\Gamma)$, meaning that there exists a uniform constant $c > 0$, such that for any $u \in H^q(\Gamma)$, we have

$$\big\langle u, (\mathcal{A} + \mathcal{A}^\star) u \big\rangle_\Gamma \geq c\|u\|_{H^q(\Gamma)}^2. \tag{4}$$

Furthermore, for the sake of convenience the operator $\mathcal{A}$ is assumed to be injective. If this is not the case, but if its kernel is finite-dimensional and known in advance, then one can consider $\mathcal{A}$ as on operator

$$\mathcal{A}: \quad H^q(\Gamma)/\ker \mathcal{A} \quad \to \quad \big(H^q(\Gamma)/\ker \mathcal{A}\big)',$$

and the presented approach is still valid, which is the case for example for any interior Neumann problem, where the kernel consists of all constant functions.

A practical example, which can be written as a boundary integral equation, is the Laplace problem with homogeneous Dirichlet boundary data in three spatial dimensions for given boundary data $g \in H^{\frac{1}{2}}(\Gamma)$, i.e.,

$$\Delta v = 0 \text{ in } \Omega, \qquad v = g \text{ on } \Gamma.$$

It is well known that this problem is uniquely solvable. As described in detail in [23, 26], for example, we may write $v \in H^1(\Omega)$ as a layer potential of an unknown density $u \in H^q(\Gamma)$, that is $v = \mathcal{P}u$, where $\mathcal{P}$ is a linear and continuous boundary potential operator from $H^q(\Gamma)$ to $H^1(\Omega)$. By taking the trace of the equation $v = \mathcal{P}u$, we arrive at a boundary integral equation

$$\mathcal{A}u := \operatorname{tr}\big(\mathcal{P}u\big) = g. \tag{5}$$

Especially, in the case of the single layer and the double layer potential, the kernels are given by

$$K_s(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{4\pi\|\boldsymbol{x} - \boldsymbol{y}\|}, \qquad K_d(\boldsymbol{x}, \boldsymbol{y}) = \frac{\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{n}_{\boldsymbol{y}} \rangle}{4\pi\|\boldsymbol{x} - \boldsymbol{y}\|^3}.$$

It can easily be seen that k_s and k_d are asymptotically smooth kernels of the orders $2q = -1$ and $2q = 0$, respectively.

2.3. **Galerkin scheme.** By multiplying (2) with a test function $\phi \in H^q(\Gamma)$, we derive the variational formulation of the boundary integral equation under consideration:

$$\text{find } u \in H^q(\Gamma) \text{ such that } \langle \mathcal{A}u, \phi\rangle_\Gamma = \langle g, \phi\rangle_\Gamma \quad \text{for any } \phi \in H^q(\Gamma). \tag{6}$$

Similar to [8], we are considering a sequence of nested trial spaces $V_J \subseteq V_{J+1} \subseteq \ldots \subseteq H^q(\Gamma)$, which is asymptotically dense in $H^q(\Gamma)$. For any fixed level J (reflecting a mesh width of size $\sim 2^{-J}$), we restrict the variational formulation (6) to V_J to obtain the Galerkin problem

$$\text{find } u_J \in V_J \text{ such that } \langle \mathcal{A}u_J, \phi\rangle_\Gamma = \langle g, \phi\rangle_\Gamma \quad \text{for any } \phi \in V_J. \tag{7}$$

If $V_J = \operatorname{span}\{\psi_1, \ldots, \psi_{N_J}\} \subseteq H^q(\Gamma)$, the Galerkin problem (7) is equivalent to the linear system of equations

$$\mathbf{A}_J\mathbf{u}_J = \mathbf{g}_J, \quad \mathbf{A}_J = \big[\langle \mathcal{A}\psi_j, \psi_i\rangle_\Gamma\big]_{i,j=1}^{N_J}, \quad \mathbf{g}_J = \big[\langle g, \psi_j\rangle_\Gamma\big]_{j=1}^{N_J}, \quad \mathbf{u}_J = \big[u_j\big]_{j=1}^{N_J},$$

where $u_J(\boldsymbol{x}) = \sum_{j=1}^{N_J} u_j\psi_j(\boldsymbol{x})$. Especially, by means of Cea's Lemma, the solution $u_J \in V_J$ satisfies an estimate of the form

$$\|u - u_J\|_{H^q(\Gamma)} \lesssim \inf_{v_J \in V_J} \|u - v_J\|_{H^q(\Gamma)}.$$

Herein, the right-hand side can estimated further by imposing more knowledge on the trial spaces V_J.

3. Discretization

3.1. **Single-scale bases.** A natural choice of trial functions are piecewise polynomial functions, defined on the unit interval, tensorized, and then transported onto a surface patch Γ_i. We postpone this transportation to Section 7 and consider the unit square first. To this end, we first have to consider the unit interval $I = [0, 1]$. Given a level j, we want to construct a space V_j, with $\dim V_j \sim 2^j$, which consists of piecewise polynomial functions on the dyadic intervals $[k2^{-j}, (k+1)2^{-j}]$, $k = 0, 1, \ldots, 2^j - 1$. This is possible by choosing a suitable function ϕ and then rescaling it according to

$$\phi_{j,k}(x) := 2^{\frac{j}{2}}\phi\big(2^jx - k\big), \quad k \in \Delta_j,$$

where Δ_j is a suitable index set. We remark that this scaling implies $\|\phi_{j,k}\|_{L^2([0,1])} \sim 1$.

Outgoing from this construction, we can define any ansiotropic tensor product function and the corresponding tensor product spaces. In particular, for $\mathbf{j} = (j_1, j_2) \in \mathbb{N}_0^2$, and $\mathbf{k} = (k_1, k_2) \in \Delta_{\mathbf{j}} := \Delta_{j_1} \times \Delta_{j_2}$, we define the tensor product function

$$\phi_{\mathbf{j},\mathbf{k}}(\boldsymbol{x}) := \big(\phi_{j_1,k_1} \otimes \phi_{j_2,k_2}\big)(\boldsymbol{x}) = \phi_{j_1,k_1}(x_1)\phi_{j_2,k_2}(x_2).$$

With these functions, we define the trial space

$$V_j := \operatorname{span}\big\{\phi_{\mathbf{j},\mathbf{k}} : \mathbf{j} = (j, j),\, \mathbf{k} \in \Delta_{\mathbf{j}}\big\}.$$

The spaces V_j are said to have the approximation order $d \in \mathbb{N}$ given by

$$d = \sup\left\{s \in \mathbb{R} \colon \inf_{v_j \in V_j} \big\|v - v_j\big\|_{L^2(\square)} \lesssim 2^{-js}\|v\|_{H^s(\square)} \text{ for any } v \in H^s(\square)\right\} \tag{8}$$

and the regularity γ given by

$$\gamma = \sup\left\{s \in \mathbb{R} \colon V_j \subseteq H^s(\square)\right\}.$$

In the simplest case, we take piecewise constant scaling functions which are defined by $\phi := \mathbb{1}_{[0,1]}$. Then, for any fixed $j \in \mathbb{N}$, we define the local trial functions

$$\phi_{j,k} = 2^{\frac{j}{2}}\mathbb{1}_{[k2^{-j},(k+1)2^{-j}]} = 2^{\frac{j}{2}}\phi(2^j\cdot - k), \quad k \in \Delta_j := \{0, 1, \ldots, 2^j - 1\}.$$

This yields the well-known approximation spaces $V_j := \text{span}\{\phi_{j,k} : k \in \Delta_j\}$, having the parameters $\gamma = \frac{1}{2}$ and $d = 1$.

In general, piecewise polynomial functions of the order r result in an approximation order $d = r$. The regularity γ is, however, limited by the global smoothness of the trial functions. There holds $\gamma = \frac{1}{2}$ if they are discontinuous while there is $\gamma = \frac{3}{2}$ if they are continuous.

3.2. **Wavelet bases.** Although the above method is very intuitive, we have a lot of difficulties to deal with. As the boundary integral operators under consideration are not local, the Galerkin problem results in fully populated matrices. This drawback can, up to logarithmic terms, be overcome with fast boundary element methods like the fast multipole method [12]. An alternative approach is to consider specific, linear combinations of piecewise polynomial trial functions, which are called wavelets. For a full introduction into this topic, see for example [10, 16, 24].

The general idea is to discretize the complement of V_{j-1} in V_j. Roughly speaking, given a function u_j in V_j, the projection $\mathbb{P}_{j-1}u_j \in V_{j-1}$ is a good estimation on u_j, and the difference $u_j - \mathbb{P}_{j-1}u_j$ can be expressed in terms of complementary basis functions. To this end, we fix a minimal level j_0 and introduce complement spaces W_j for all $j > j_0$, satisfying

$$V_j = V_{j-1} \oplus W_j.$$

Similar as before, W_j is spanned by basis functions of the form

$$\psi_{j,k} := 2^{\frac{j-1}{2}}\psi\big(2^{j-1}\cdot - k\big), \quad k \in \nabla_j := \{0, \ldots, 2^{j-1} - 1\}, \quad j > j_0. \tag{9}$$

The function ψ is the so-called mother wavelet. Also here, we note that $\|\psi_{j,k}\|_{L^2(\Gamma)} \sim 1$. However, it can be shown that $\big\{\psi_{j,k}\big\}_{j,k}$ is a Riesz basis of multiple Sobolev spaces if properly scaled, compare e.g. [24]. We remark that the identity $V_j = V_{j-1} \oplus W_j$ implies that

$$V_J = V_{j_0} \oplus W_{j_0+1} \oplus W_{j_0+2} \oplus \cdots \oplus W_J. \tag{10}$$

For the sake of notational convenience, we set $W_{j_0} := V_{j_0}$ and denote $\psi_{j_0,k} := \phi_{j_0,k}$ for all $k \in \nabla j_0 := \Delta_{j_0}$.

By tensorising (10) with itself, we arrive at

$$V_J \otimes V_J = \bigoplus_{|\mathbf{j}|_\infty \le J} W_{j_1} \otimes W_{j_2}.$$

Hence, in view of (9), we can write

$$V_J \otimes V_J = \text{span}\,\Psi := \text{span}\,\{\psi_{\mathbf{j},\mathbf{k}} \colon |\mathbf{j}|_\infty \le J,\ \mathbf{k} \in \nabla_{\mathbf{j}}\} \tag{11}$$

with the tensor product wavelets $\psi_{\mathbf{j},\mathbf{k}} = \psi_{j_1,k_1} \otimes \psi_{j_2,k_2}$ and

$$\nabla_{\mathbf{j}} := \nabla_{j_1} \times \nabla_{j_2}, \quad \nabla_j := \big\{0, \ldots, \max\{0, 2^{j-1} - 1\}\big\}. \tag{12}$$

3.3. **Notation.** Let us define some notation which we will use throughout the remainder of this article. First, we define the support of a wavelet as

$$\Omega_{j,k} := \text{supp}\,\psi_{j,k}$$

and, accordingly,

$$\Omega_{\mathbf{j},\mathbf{k}} := \text{supp}\,\psi_{\mathbf{j},\mathbf{k}}.$$

Similarly, we define the singular support, i.e., the points at which a wavelet is not smooth, as

$$\Omega^\sigma_{j,k} := \text{singsupp}\,\psi_{j,k}, \quad \Omega^\sigma_{\mathbf{j},\mathbf{k}} := \text{singsupp}\,\psi_{\mathbf{j},\mathbf{k}}.$$

For a pair of wavelets $\psi_{\mathbf{j},\mathbf{k}}$ and $\psi_{\mathbf{j}',\mathbf{k}'}$, we let

$$\delta_{x_i} := \text{dist}\,\big(\Omega_{j_i,k_i}, \Omega_{j'_i,k'_i}\big), \quad \delta_{\text{tot}} := \text{dist}\,\big(\Omega_{\mathbf{j},\mathbf{k}}, \Omega_{\mathbf{j}',\mathbf{k}'}\big).$$

Moreover, we also define

$$\sigma_{x_i} := \begin{cases} \operatorname{dist}\left(\Omega_{j_i,k_i}, \Omega^{\sigma}_{j_i',k_i'}\right), & j_i \geq j_i', \\ \operatorname{dist}\left(\Omega^{\sigma}_{j_i,k_i}, \Omega_{j_i',k_i'}\right), & \text{otherwise.} \end{cases}$$

Finally, given a wavelet $\psi_{j,k}$, we say that $\psi_{j',k'}$ is located in the *far-field* of $\psi_{j,k}$ if there holds $\operatorname{dist}(\Omega_{j,k}, \Omega_{j',k'}) \gtrsim 2^{-\min\{j,j'\}}$, otherwise, we say that $\psi_{j',k'}$ is located in the *near-field* of $\psi_{j,k}$. For the tensorized wavelets, this threshold is the maximal support length, which amounts to $2^{-\min\{j_1,j_2,j_1',j_2'\}}$.

3.4. **Some important wavelet properties.** Wavelet functions have some very nice properties, see e.g. [7, 10, 24] for the full range of expressions. In this section, we will restrict ourselves on the most important ones, which are needed in the remainder of this article.

First, it is well-known (see e.g. [7, 10]) that a one-dimensional wavelet basis on the interval possesses a unique, biorthogonal dual basis. By tensorising this dual basis with itself, we get a biorthogonal dual basis on $\square$, which we denote by $\widetilde{\psi}$. This dual basis then provides the approximation order $\widetilde{d}$ and the regularity $\widetilde{\gamma} > 0$.

As already stated in [13, 24], the set $\Psi := \{\psi_{\mathbf{j},\mathbf{k}} \colon \mathbf{j} \geq j_0,\ \mathbf{k} \in \nabla_{\mathbf{j}}\}$ forms a Riesz basis of $L^2(\square)$, meaning that

$$\left\| \sum_{\mathbf{j},\mathbf{k}} c_{\mathbf{j},\mathbf{k}} \psi_{\mathbf{j},\mathbf{k}} \right\|^2_{L^2(\square)} \sim \sum_{\mathbf{j},\mathbf{k}} \left| c_{\mathbf{j},\mathbf{k}} \right|^2.$$

Moreover, in accordance with [7, 13], they satisfy the norm equivalences,

$$\|u\|^2_{H^s_\circ(\square)} \sim \sum_{\mathbf{j},\mathbf{k}} 2^{2s|\mathbf{j}|_\infty} \left| \langle \widetilde{\psi}_{\mathbf{j},\mathbf{k}}, u \rangle_\square \right|^2, \quad -\widetilde{\gamma} < s < \gamma, \tag{13}$$

where $H^s_\circ := H^s$ if $s \geq 0$ and $H^s_\circ := \widetilde{H}^s$ if $s < 0$.

For a multiindex $\mathbf{j}$, let us define the (non-orthogonal) projection onto $V_{\mathbf{j}}$ as

$$Q_{\mathbf{j}} u := \sum_{\mathbf{k} \in \nabla_{\mathbf{j}}} \left\langle \widetilde{\psi}_{\mathbf{j},\mathbf{k}}, u \right\rangle_\Gamma \psi_{\mathbf{j},\mathbf{k}},$$

whereas for $J \in \mathbb{N}$, we define the projection onto V_J as

$$Q_J u := \sum_{|\mathbf{j}|_\infty \leq J} Q_{\mathbf{j}} u = \sum_{|\mathbf{j}|_\infty \leq J} \sum_{\mathbf{k} \in \nabla_{\mathbf{j}}} \left\langle \widetilde{\psi}_{\mathbf{j},\mathbf{k}}, u \right\rangle_\Gamma \psi_{\mathbf{j},\mathbf{k}}. \tag{14}$$

By using the a tensor argument, the duality and the biorthogonality, the one-dimensional approximation property, which is derived e.g. in [24], generalizes to

$$\|u - Q_J u\|_{H^s_\circ(\square)} \lesssim 2^{J(s-t)} \|u\|_{H^t_\circ(\square)}, \quad s \leq t, \quad -\widetilde{d} < s < \gamma, \quad -\widetilde{\gamma} < t \leq d. \tag{15}$$

Moreover, there holds Bernstein's inequality

$$\|Q_{\mathbf{j}} u\|_{H^s_\circ(\square)} \lesssim 2^{|\mathbf{j}|_\infty (s-t)} \|Q_{\mathbf{j}} u\|_{H^t_\circ(\square)}, \quad t \leq s < \gamma,$$

and, regarding $Q_J u \in V_J$, also

$$\|Q_J u\|_{H^s_\circ(\square)} \lesssim 2^{J(s-t)} \|Q_J u\|_{H^t_\circ(\square)}, \quad t \leq s < \gamma. \tag{16}$$

Perhaps the most important property of wavelets for the present article is that they have *vanishing moments*, also called *cancellation property*, which is induced from the approximation order of the dual basis. Namely, in accordance with [7], there holds

$$\left| \left\langle \psi_{j,k}, u \right\rangle_{[0,1]} \right| \lesssim 2^{-(\widetilde{d}+\frac{1}{2})j} |u|_{W^{\widetilde{d},\infty}(\Omega_{j,k})}.$$

By explicitly enrolling the tensor product structure of the wavelet $\psi_{\mathbf{j},\mathbf{k}}$, we can immediately deduce that

$$\left|\left\langle\psi_{\mathbf{j},\mathbf{k}}, u\right\rangle_{\square}\right| \lesssim 2^{-(\widetilde{d}+\frac{1}{2})|\mathbf{j}|_1} |u|_{W^{2\widetilde{d},\infty}(\Omega_{\mathbf{j},\mathbf{k}})}. \tag{17}$$

Remark 3.1. *Due to the tensor product structure of the wavelets, we must tensorize scaling functions on the coarsest level with wavelets on a finer level. This means that we cannot use $\widetilde{d}$ vanishing moments in both directions. However, if $\mathcal{I} \subseteq \{j_1, j_2\}$ denotes the subset of indices corresponding to univariate wavelets with $\widetilde{d}$ vanishing moments, we have the estimate*

$$\left|\left\langle\psi_{\mathbf{j},\mathbf{k}}, u\right\rangle_{\square}\right| \lesssim 2^{-\frac{1}{2}|\mathbf{j}|_1 - \widetilde{d}\sum_{j\in\mathcal{I}} j} |u|_{W^{|\mathcal{I}|\widetilde{d},\infty}(\Omega_{\mathbf{j},\mathbf{k}})}.$$

4. Matrix entry estimates

In order to develop a compression scheme for the operator $\mathcal{A}$ with respect to the wavelet basis Ψ, we need to estimate the matrix entries in the Galerkin matrix

$$\mathbf{A}_J = \left[\left\langle\psi_{i',\mathbf{j}',\mathbf{k}'}, \mathcal{A}\psi_{i,\mathbf{j},\mathbf{k}}\right\rangle_{\Gamma}\right], \quad \text{where} \begin{cases} 1 \le i, i' \le r, \\ \mathbf{j}, \mathbf{j}' \ge \mathbf{j}_0, \\ \mathbf{k} \in \nabla_{\mathbf{j}},\ \mathbf{k}' \in \nabla_{\mathbf{j}'}, \end{cases}$$

where the wavelet function $\psi_{i,\mathbf{j},\mathbf{k}}$ is the lifting of the function $\psi_{\mathbf{j},\mathbf{k}}$ onto the patch Γ_i, i.e.,

$$\psi_{i,\mathbf{j},\mathbf{k}}(\boldsymbol{x}) := \psi_{\mathbf{j},\mathbf{k}}\big(\boldsymbol{\gamma}_i^{-1}(\boldsymbol{x})\big).$$

For now, let us consider the situation $r = 1$, where the only patch present is the unit square $\square$, in which case we can assume that $\boldsymbol{\gamma}_i = \mathrm{id}$. The discussion of the situation on a Lipschitz manifold is postponed to Section 7.

4.1. **Far-field estimates.** For the remainder of Section 4.1, we assume that $\delta_{\text{tot}} > 0$, which means that the *first compression* [8, 16, 24] applies. There exist estimations for the entries by Reich [20, 21], which make use of the vanishing moments in the one dimensional wavelets with the smallest corresponding support length. This is especially useful when considering thin, long wavelets with a small distance. We quote the following result.

Theorem 4.1 ([20, Theorem 2.1.9])**.** *For $\mathbf{j}, \mathbf{j}' \ge \mathbf{j}_0 + \mathbf{1}$, there holds*

$$\left|\left\langle\psi_{\mathbf{j}',\mathbf{k}'}, \mathcal{A}\psi_{\mathbf{j},\mathbf{k}}\right\rangle_{\square}\right| \lesssim 2^{-\frac{1}{2}(|\mathbf{j}|_1+|\mathbf{j}'|_1)} 2^{-\widetilde{d}(j^{(1)}+j^{(2)})} \operatorname{dist}\left(\Omega_{\mathbf{j},\mathbf{k}}, \Omega_{\mathbf{j}',\mathbf{k}'}\right)^{-(2+2q+2\widetilde{d})}.$$

Here, $\left\{j^{(1)}, j^{(2)}\right\} \subseteq \{j_1, j_1', j_2, j_2'\} \cap [j_0+1, \infty)$ can be any two distinct indices, the best behaviour is obtained by choosing the two largest indices.

Let us next derive an estimate which makes use of the vanishing moments in every one-dimensional wavelet, which is beneficial if the supports of the wavelets $\psi_{\mathbf{j},\mathbf{k}}$ and $\psi_{\mathbf{j}',\mathbf{k}'}$ are small.

Theorem 4.2. *For $\mathbf{j}, \mathbf{j}' \ge \mathbf{j}_0 + \mathbf{1}$, there holds*

$$\left|\left\langle\psi_{\mathbf{j}',\mathbf{k}'}, \mathcal{A}\psi_{\mathbf{j},\mathbf{k}}\right\rangle_{\square}\right| \lesssim 2^{-(\widetilde{d}+\frac{1}{2})(|\mathbf{j}|_1+|\mathbf{j}'|_1)} \delta_{\text{tot}}^{-(2+2q+4\widetilde{d})}. \tag{18}$$

Proof. By explicitly enrolling the tensor product structure of the wavelets, we can write

$$\left|\left\langle\psi_{\mathbf{j}',\mathbf{k}'}, \mathcal{A}\psi_{\mathbf{j},\mathbf{k}}\right\rangle_{\square}\right| \sim \int_0^1 \int_0^1 \int_{\square} K(\boldsymbol{x}, \boldsymbol{x}')\psi_{\mathbf{j},\mathbf{k}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\ \psi_{j_1',k_1'}(x_1')\,\mathrm{d}x_1'\ \psi_{j_2',k_2'}(x_2')\,\mathrm{d}x_2'.$$

We can use the vanishing moments of $\psi_{j_2',k_2'}$ to deduce that

$$\begin{aligned}&\big|\big\langle\psi_{\mathbf{j}',\mathbf{k}'},\mathcal{A}\psi_{\mathbf{j},\mathbf{k}}\big\rangle_\square\big|\\&\quad\lesssim 2^{-(\widetilde{d}+\frac{1}{2})j_2'}\left\|\int_0^1\int_\square \partial_{x_2'}^{\widetilde{d}}K\left(\boldsymbol{x},\left[\begin{smallmatrix}x_1'\\ \cdot\end{smallmatrix}\right]\right)\psi_{\mathbf{j},\mathbf{k}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\psi_{j_1',k_1'}(x_1')\,\mathrm{d}x_1'\right\|_{L^\infty(\Omega_{j_2,k_2})}.\end{aligned}$$

Note that we can differentiate under the integral because the kernel k is smooth and bounded on $\Omega_{\mathbf{j},\mathbf{k}}\times\Omega_{\mathbf{j}',\mathbf{k}'}$. The vanishing moments of $\psi_{j_1',k_1'}$ then allow us proceed with the estimate to

$$\begin{aligned}&\big|\big\langle\psi_{\mathbf{j}',\mathbf{k}'},\mathcal{A}\psi_{\mathbf{j},\mathbf{k}}\big\rangle_\square\big|\\&\quad\lesssim 2^{-(\widetilde{d}+\frac{1}{2})|\mathbf{j}'|_1}\left\|\int_0^1\int_0^1 \partial_{x_1'}^{\widetilde{d}}\partial_{x_2'}^{\widetilde{d}}K\left(\left[\begin{smallmatrix}x_1\\ x_2\end{smallmatrix}\right],\cdot\right)\psi_{j_1,k_1}(x_1)\,\mathrm{d}x_1\psi_{j_2,k_2}(x_2)\,\mathrm{d}x_2\right\|_{L^\infty(\Omega_{\mathbf{j}',\mathbf{k}'})}.\end{aligned}$$

By subsequently using the vanishing moments of ψ_{j_2,k_2}, and ψ_{j_1,k_1} as well, we finally arrive at

$$\big|\big\langle\psi_{\mathbf{j}',\mathbf{k}'},\mathcal{A}\psi_{\mathbf{j},\mathbf{k}}\big\rangle_\square\big|\lesssim 2^{-(\widetilde{d}+\frac{1}{2})(|\mathbf{j}|_1+|\mathbf{j}'|_1)}\big\|\partial_{x_1}^{\widetilde{d}}\partial_{x_2}^{\widetilde{d}}\partial_{x_1'}^{\widetilde{d}}\partial_{x_2'}^{\widetilde{d}}K(\boldsymbol{x},\boldsymbol{x}')\big\|_{L^\infty\left(\Omega_{\mathbf{j},\mathbf{k}}\times\Omega_{\mathbf{j}',\mathbf{k}'}\right)}.$$

If we remember the fact that the kernel k is asymptotically smooth of order $2q$, compare (3), we can deduce (18). □

4.2. **Near-field estimates.** As we will see in Section 5, we may use the previous estimates only if a wavelet pair is in the far-field, meaning that the supports are sufficiently far away. For the near field, we need to derive different estimates. In this case, we explicitly enroll the tensor product structure of the wavelets again. We will use an approach which is similar to the one created in [20].

To this end, we define the dimensionally reduced kernel

$$K_1(x,x'):=\int_0^1\int_0^1 K\left(\left[\begin{smallmatrix}x\\ y\end{smallmatrix}\right],\left[\begin{smallmatrix}x'\\ y'\end{smallmatrix}\right]\right)\psi_{j_2,k_2}(y)\psi_{j_2',k_2'}(y')\,\mathrm{d}y\,\mathrm{d}y' \tag{19}$$

and the operator $\mathcal{A}_1$ as the integral operator with the kernel K_1. By definition, the kernel K_1 depends on the wavelets ψ_{j_2,k_2} and $\psi_{j_2',k_2'}$, but the context will always clarify this relation.

Due to the tensor product structure, the dimensionally reduced operator obviously satisfies $\langle\psi_{\mathbf{j}',\mathbf{k}'},\mathcal{A}\psi_{\mathbf{j},\mathbf{k}}\rangle_\square=\langle\psi_{j_1',k_1'},\mathcal{A}_1\psi_{j_1,k_1}\rangle_{[0,1]}$. Moreover, similar as shown in [20, Lemma 2.1.5], the estimate

$$\left|\partial_x^\alpha\partial_{x'}^{\alpha'}K_1(x,x')\right|\lesssim 2^{-\frac{1}{2}(j_2+j_2')}|x-x'|^{-(2+2q+\alpha+\alpha')} \tag{20}$$

holds. However, there are also vanishing moments of the wavelets hidden in the kernel K_1, which can be used to improve the estimate (20) and hence also Theorem 2.1.7 in [20]:

Theorem 4.3. *Assume that* $0<\sigma_{x_1}\lesssim 2^{-\min\{j_1,j_1'\}}$, *and* $\max\{j_1,j_1'\},\max\{j_2,j_2'\}>j_0$. *Then, we have*

$$\left.\begin{aligned}&\big|\big\langle\psi_{j_1',k_1'},\mathcal{A}_1\psi_{j_1,k_1}\big\rangle_{[0,1]}\big|\\&\big|\big\langle\psi_{j_1,k_1},\mathcal{A}_1\psi_{j_1',k_1'}\big\rangle_{[0,1]}\big|\end{aligned}\right\}$$
$$\lesssim 2^{-\frac{1}{2}(j_2+j_2')}2^{-\widetilde{d}\max\{j_2,j_2'\}}2^{-\frac{1}{2}|j_1-j_1'|}2^{-\widetilde{d}\max\{j_1,j_1'\}}\sigma_{x_1}^{-(1+2q+2\widetilde{d})}.$$

Proof. We will simply derive the appropriate estimate for the kernel k_1 similar to (20). Then, the rest of the proof may be completed by simply following the arguments of [20].

If $x \neq x'$, then the function under the integral in (19) is bounded, so we may directly differentiate under the integral. Moreover, let us without loss of the generality assume that $j_2 > j_2'$. Then,

$$\begin{aligned}
\left|\partial_x^\alpha \partial_{x'}^{\alpha'} K_1(x,x')\right| &= \left|\int_{\Omega_{j_2,k_2}} \psi_{j_2,k_2}(y) \int_{\Omega_{j_2',k_2'}} \psi_{j_2',k_2'}(y')\, \partial_x^\alpha \partial_{x'}^{\alpha'}\, K\left(\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right], \left[\begin{smallmatrix} x' \\ y' \end{smallmatrix}\right]\right)\, \mathrm{d}y\, \mathrm{d}y'\right| \\
&\lesssim 2^{-(\widetilde{d}+\frac{1}{2})j_2} \sup_{y\in\Omega_{j_2,k_2}} \left|\int_{\Omega_{j_2',k_2'}} \psi_{j_2',k_2'}(y') \partial_y^{\widetilde{d}} \partial_x^\alpha \partial_{x'}^{\alpha'}\, K\left(\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right], \left[\begin{smallmatrix} x' \\ y' \end{smallmatrix}\right]\right)\, \mathrm{d}y'\right| \\
&\lesssim 2^{-(\widetilde{d}+\frac{1}{2})j_2} \int_{\Omega_{j_2',k_2'}} \underbrace{\left|\psi_{j_2',k_2'}(y_2')\right|}_{\lesssim 2^{j_2/2}} \underbrace{\left\|\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] - \left[\begin{smallmatrix} x' \\ y' \end{smallmatrix}\right]\right\|}_{\geq |x-x'|}^{-(2+2q+\widetilde{d}+\alpha+\alpha')} \mathrm{d}y_2' \\
&\lesssim 2^{-\frac{1}{2}(j_2+j_2')} 2^{-\widetilde{d}j_2} |x-x'|^{-(2+2q+\widetilde{d}+\alpha+\alpha')},
\end{aligned}$$

since $|\Omega_{j_2,k_2}| \lesssim 2^{-j_2}$.

As the remainder of the proof is based on the ideas of [8, Section 6], we just sketch it. Without loss of the generality, we may assume that $j_1' \leq j_1$. In this case, ψ_{j_1,k_1} is located on a smooth part of the wavelet $\psi_{j_1',k_1'}$, so we may decompose $\psi_{j_1',k_1'} = \tilde{f} + \bar{f}$ such that $\tilde{f}$ is a smooth function satisfying

$$\tilde{f}\big|_{\Omega_{j_1,k_1}} = \psi_{j_1',k_1'}\big|_{\Omega_{j_1,k_1}}.$$

This can be realized by Calderón's extension theorem [25] with $\|\tilde{f}\|_{H^s([0,1])} \lesssim 2^{sj_1'}$. Hence, we have

$$\left|\left\langle \psi_{j_1,k_1}, \mathcal{A}_1 \psi_{j_1',k_1'}\right\rangle_{[0,1]}\right| \leq \left|\left\langle \psi_{j_1,k_1}, \mathcal{A}_1 \tilde{f}\right\rangle_{[0,1]}\right| + \left|\left\langle \psi_{j_1,k_1}, \mathcal{A}_1 \bar{f}\right\rangle_{[0,1]}\right|.$$

The estimate for $\bar{f}$ follows directly from [20, Lemma 2.1.1].

For the function $\tilde{f}$, we define the operator

$$\mathcal{A}_1^\sharp \tilde{f}(x) := \int_{\mathbb{R}} \chi(x)\chi(x') K_1(x,x') \tilde{f}(x')\, \mathrm{d}x'$$

by employing a smooth cutoff function χ satisfying $\chi|_{[0,1]} = 1$. Then, $\mathcal{A}_1^\sharp$ is a pseudo-differential operator of the order $m = 1 + 2q + \widetilde{d}$, cf. [18]. Remarking that in our case, we have $c(j_2, j_2') = 2^{-\frac{1}{2}(j_2+j_2')} 2^{-\widetilde{d}j_2}$, we may apply [20, Lemma 2.1.4] and the fact that $\sigma_{x_1} \lesssim 2^{-j_1'}$ to conclude. □

Remark 4.4. *Up to now, we have just considered a reduction to the first coordinate direction. Nevertheless, as also done in* [20]*, a reduction to the second coordinate is possible by using a similar definition for the operator $\mathcal{A}_2$, and the same estimates hold with exchanged indices.*

5. Matrix compression scheme

To keep the number of the degrees of freedom small, we need to introduce a compression scheme, according to which many matrix entries do not have to be calculated, whilst obtaining convergence with the full rate offered by the underlying Galerkin scheme. We differ between the first compression and the second compression, but for either case, we require a matrix block error which is controlled by a level dependent parameter $\sigma_{\mathbf{j},\mathbf{j}'}$ given by

$$\sigma_{\mathbf{j},\mathbf{j}'} := 2J(d'-q) - d'(|\mathbf{j}|_\infty + |\mathbf{j}'|_\infty) + \kappa(2|\mathbf{j}|_\infty + 2|\mathbf{j}'|_\infty - |\mathbf{j}|_1 - |\mathbf{j}'|_1). \tag{21}$$

Here, $d' > d$ and $\kappa > 0$ are sufficiently small, but fixed real numbers, which are introduced in order to avoid logarithmic terms in the consistency estimates.

5.1. **Far-field: First compression.** In the case of the first compression, we consider a pair of wavelets $\psi_{\mathbf{j},\mathbf{k}}$ and $\psi_{\mathbf{j}',\mathbf{k}'}$, whose supports are located sufficiently far away from each other. As we will see, we need to estimate a sum of matrix coefficients by an integral, which requires that the minimal distance between the respective wavelets' supports is large enough. In two dimensions, we must have a minimal distance, which is at least as wide the largest face of the included supports, namely $2^{-\min\{j_1,j_2,j_1',j_2'\}}$.

If this is not the case, however, we can make use of the tensor product structure and estimate the sum only in the coordinate direction of x_i, which results in a minimal distance of $2^{-\min\{j_i,j_i'\}}$. This procedure basically follows [20], but is adapted here to the setting on the full tensor product space.

5.1.1. *Compression in the $\boldsymbol{x}$- and $\boldsymbol{y}$-coordinate.* For a fixed maximal level J, we define the compressed matrix for the first compression $\mathbf{A}_J^{c_1,1}$ as

$$\big[\mathbf{A}_J^{c_1,1}\big]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')} := \begin{cases} 0, & \mathbf{j},\mathbf{j}' \geq \mathbf{j}_0 + \mathbf{1}, \quad \delta_{\text{tot}} > \mathcal{B}_{\mathbf{j},\mathbf{j}'}, \\ \big[\mathbf{A}_J\big]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}, & \text{otherwise}, \end{cases} \tag{22}$$

for $\mathbf{k} \in \nabla_{\mathbf{j}}$, $\mathbf{k}' \in \nabla_{\mathbf{j}'}$, and $|\mathbf{j}|_\infty, |\mathbf{j}'|_\infty \leq J$. Herein, the cutoff parameter $\mathcal{B}_{\mathbf{j},\mathbf{j}'}$ is given as

$$\mathcal{B}_{\mathbf{j},\mathbf{j}'} := a \max\left\{2^{-\min\{j_1,j_2,j_1',j_2'\}}, 2^{\frac{\sigma_{\mathbf{j},\mathbf{j}'} - \tilde{d}(|\mathbf{j}|_1 + |\mathbf{j}'|_1)}{2q+4\tilde{d}}}\right\}, \tag{23}$$

where $a > 0$ is a fixed real number.

Theorem 5.1. *Let $\mathbf{R}_J := \mathbf{A}_J - \mathbf{A}_J^{c_1,1}$. Then, for the matrix block $\mathbf{R}_{\mathbf{j},\mathbf{j}'}$, we have the estimate*

$$\|\mathbf{R}_{\mathbf{j},\mathbf{j}'}\|_2 \lesssim a^{-(2q+4\tilde{d})} 2^{-\sigma_{\mathbf{j},\mathbf{j}'}}$$

with a generic constant that is independent of the refinement level J.

Proof. We advance similar as in [8]. First, we define the set

$$\nabla_{\mathbf{j}}^{\mathcal{B}} := \{\mathbf{k} \in \nabla_{\mathbf{j}} \colon \delta_{\text{tot}} > \mathcal{B}_{\mathbf{j},\mathbf{j}'}\}.$$

Then, we estimate the column sum of the block $\mathbf{R}_{\mathbf{j},\mathbf{j}'}$ by

$$\sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} \big|r_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}\big| = \sum_{\mathbf{k}\in\nabla_{\mathbf{j}}^{\mathcal{B}}} \big|\big\langle \psi_{\mathbf{j}',\mathbf{k}'}, \mathcal{A}\psi_{\mathbf{j},\mathbf{k}}\big\rangle_\square\big| \lesssim 2^{-(\tilde{d}+\frac{1}{2})(|\mathbf{j}|_1+|\mathbf{j}'|_1)} \sum_{\mathbf{k}\in\nabla_{\mathbf{j}}^{\mathcal{B}}} \delta_{\text{tot}}^{-(2+2q+4\tilde{d})},$$

where the last inequality is due to Theorem 4.2. By the compression rule (22), we have the relation $\delta_{\text{tot}} \geq \mathcal{B}_{\mathbf{j},\mathbf{j}'}$, and, since also $\mathcal{B}_{\mathbf{j},\mathbf{j}'} \gtrsim 2^{-j_1}, 2^{-j_2}$, we can estimate the sum by an integral, yielding

$$\begin{aligned}\sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} \big|r_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}\big| &\lesssim 2^{-(\tilde{d}+\frac{1}{2})(|\mathbf{j}|_1+|\mathbf{j}'|_1)} 2^{|\mathbf{j}|_1} \int_{\|\boldsymbol{x}\|\geq\mathcal{B}_{\mathbf{j},\mathbf{j}'}} \|\boldsymbol{x}\|^{-(2+2q+4\tilde{d})} \,\mathrm{d}\boldsymbol{x} \\ &\lesssim 2^{-(\tilde{d}+\frac{1}{2})(|\mathbf{j}|_1+|\mathbf{j}'|_1)} 2^{|\mathbf{j}|_1} \mathcal{B}_{\mathbf{j},\mathbf{j}'}^{-(2q+4\tilde{d})}.\end{aligned}$$

As we also have $\mathcal{B}_{\mathbf{j},\mathbf{j}'} \geq a 2^{\frac{\sigma_{\mathbf{j},\mathbf{j}'} - \tilde{d}(|\mathbf{j}|_1 + |\mathbf{j}'|_1)}{2q+4\tilde{d}}}$, we obtain

$$\sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} \big|r_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}\big| \lesssim a^{-(2q+4\tilde{d})} 2^{\frac{1}{2}(|\mathbf{j}|_1 - |\mathbf{j}'|_1)} 2^{-\sigma_{\mathbf{j},\mathbf{j}'}}. \tag{24}$$

Using exactly the same arguments, we can likewise derive the estimate for the row sums

$$\sum_{\mathbf{k}'\in\nabla_{\mathbf{j}'}} \big|r_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}\big| \lesssim a^{-(2q+4\tilde{d})} 2^{\frac{1}{2}(|\mathbf{j}'|_1 - |\mathbf{j}|_1)} 2^{-\sigma_{\mathbf{j},\mathbf{j}'}}. \tag{25}$$

Similar to [8], we now use the estimate for the operator norm of a matrix

$$\|\mathbf{R}_{\mathbf{j},\mathbf{j}'}\|_2^2 \le \|\mathbf{R}_{\mathbf{j},\mathbf{j}'}\|_1 \|\mathbf{R}_{\mathbf{j},\mathbf{j}'}\|_\infty = \|c\mathbf{R}_{\mathbf{j},\mathbf{j}'}\|_1 \|c^{-1}\mathbf{R}_{\mathbf{j},\mathbf{j}'}\|_\infty, \quad c > 0, \tag{26}$$

which gives us, together with (24), (25), the desired result

$$\begin{aligned}\left\|\mathbf{R}_{\mathbf{j},\mathbf{j}'}\right\|_2^2 &\le \left(\max_{\mathbf{k}'\in\nabla_{\mathbf{j}'}} \sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} 2^{\frac{|\mathbf{j}'|_1-|\mathbf{j}|_1}{2}} \left|r_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}\right|\right)\left(\max_{\mathbf{k}\in\nabla_{\mathbf{j}}} \sum_{\mathbf{k}'\in\nabla_{\mathbf{j}'}} 2^{\frac{|\mathbf{j}|_1-|\mathbf{j}'|_1}{2}} \left|r_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}\right|\right) \\ &\lesssim a^{-2(2q+4\tilde{d})} 2^{-2\sigma_{\mathbf{j},\mathbf{j}'}}. \end{aligned}$$

□

Remark 5.2. *Similar to* [20], *using Theorem 4.1 and the cutoff parameter*

$$\tilde{\mathcal{B}}_{\mathbf{j},\mathbf{j}'} := a \max\left\{2^{-\min\{j_1,j_2,j_1',j_2'\}},\ 2^{\frac{\sigma_{\mathbf{j},\mathbf{j}'}-\tilde{d}(j^{(1)}+j^{(2)})}{2q+2\tilde{d}}}\right\},$$

where $\{j^{(1)}, j^{(2)}\} \subseteq \{j_1, j_2, j_1', j_2'\}$*, we have a compression scheme*

$$\left[\mathbf{A}_J^{c_1,2}\right]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')} := \begin{cases} 0, & j^{(1)}, j^{(2)} > j_0, \quad \delta_{tot} > \tilde{\mathcal{B}}_{\mathbf{j},\mathbf{j}'}, \\ \left[\mathbf{A}_J\right]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}, & \textit{otherwise.} \end{cases} \tag{27}$$

The requirement $j^{(1)}, j^{(2)} > 0$ *is necessary to ensure the validity of Theorem 4.1.*

By modifying the appropriate calculations, we get that the corresponding difference matrix $\mathbf{R} = \mathbf{A} - \mathbf{A}_J^{c_1,2}$ *satisfies*

$$\|\mathbf{R}_{\mathbf{j},\mathbf{j}'}\| \lesssim a^{-(2q+2\tilde{d})} 2^{-\sigma_{\mathbf{j},\mathbf{j}'}}.$$

This will be important when we consider the complexity since we can also compress matrix blocks where scaling functions are involved in at most two coordinate directions.

5.1.2. *Compression in only one coordinate direction.* As remarked earlier, we need at least that $\delta_{\text{tot}} \gtrsim 2^{-\min\{j_1,j_2,j_1',j_2'\}}$ in the proof of Theorem 5.1 to estimate the row and column sums of the matrix blocks by an integral. If this is not the case, we may estimate the sum by an integral in just *one* coordinate direction x_i. This leads to restrictions on the distance in only this coordinate direction. Especially when the term $2^{-\min\{j_1,j_2,j_1',j_2'\}}$ in (23) is too large, this approach is beneficial. As all the derivations can be found in [20], we just quote the results. We also remark that we have exchanged the $|\cdot|_1$-norms from [20] with $|\cdot|_\infty$-norms in (21), since we are not working on a sparse tensor product space but on the full tensor product space.

Let us define the parameters

$$\begin{aligned}\mathcal{D}^{x_1}_{\mathbf{j},\mathbf{j}'} &:= a \max\left\{2^{-\min\{j_1,j_1'\}}, 2^{\frac{\sigma_{\mathbf{j},\mathbf{j}'}-\tilde{d}(j^{(1)}+j^{(2)})-\min\{j_2,j_2'\}}{1+2q+2\tilde{d}}}\right\}, \\ \mathcal{D}^{x_2}_{\mathbf{j},\mathbf{j}'} &:= a \max\left\{2^{-\min\{j_2,j_2'\}}, 2^{\frac{\sigma_{\mathbf{j},\mathbf{j}'}-\tilde{d}(j^{(1)}+j^{(2)})-\min\{j_1,j_1'\}}{1+2q+2\tilde{d}}}\right\}.\end{aligned}$$

We can then define the compressed value

$$v_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')} := \begin{cases} 0, \quad \text{if} \begin{cases}\delta_{x_1} > \mathcal{D}^{x_1}_{\mathbf{j},\mathbf{j}'}, \\ \delta_{x_2} \le a 2^{-\min\{j_2,j_2'\}},\end{cases} \\ 0, \quad \text{if} \begin{cases}\delta_{x_2} > \mathcal{D}^{x_2}_{\mathbf{j},\mathbf{j}'}, \\ \delta_{x_1} \le a 2^{-\min\{j_1,j_1'\}},\end{cases} \\ \left[\mathbf{A}_J\right]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}, \quad \text{otherwise,} \end{cases}$$

and then the compressed matrix by the rule

$$\big[\mathbf{A}_J^{c_1,3}\big]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')} := \begin{cases} v_{(\mathbf{j},\mathbf{k}),(\mathbf{k}',\mathbf{k}')}, & \text{if } j^{(1)}, j^{(2)} > j_0, \\ \big[\mathbf{A}_J\big]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}, & \text{otherwise.} \end{cases} \tag{28}$$

The latter definition is just a restriction of the compression to the matrix blocks $\mathbf{A}_{\mathbf{j},\mathbf{j}'}$, for which $j^{(1)}, j^{(2)} > 0$, meaning that we can use Theorem 4.1 to estimate the corresponding matrix entries.

With these definitions, out of the proof of Theorem 2.3.1 in [20], one immediately obtains the following result:

Theorem 5.3. *Let* $\mathbf{R}_J := \mathbf{A}_J - \mathbf{A}_J^{c_1,3}$. *Then, the compressed matrix blocks satisfy the estimate*

$$\big\|\mathbf{R}_{\mathbf{j},\mathbf{j}'}\big\|_2 \lesssim a^{-(1+2q+2\widetilde{d})} 2^{-\sigma_{\mathbf{j},\mathbf{j}'}}$$

with a generic constant that is independent of the refinement level J.

By combining (22), (27), and (28), we can define the *first compression of the matrix* by

$$\big[\mathbf{A}_J^{c_1}\big]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')} := \begin{cases} 0, & [\mathbf{A}_J^{c_1,\ell}]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')} = 0 \text{ for some } \ell \in \{1,2,3\}, \\ [\mathbf{A}_J]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}, & \text{otherwise.} \end{cases} \tag{29}$$

This compression affects the far-field of the system matrix in wavelet coordinates.

5.2. **Near-field: Second compression.** Up to now, we have considered wavelets with disjoint and distant supports. As we will see, we can also discard many entries if the supports of the wavelet pairs are close or even if they overlap, where a strict requirement is that the distance of the support of the smaller wavelet to the singular support of the larger wavelet is sufficiently big.

We will only use one direction for the second compression as done by Reich in [20, 21], but with improved parameters. We define

$$\begin{aligned} \mathcal{E}_{\mathbf{j},\mathbf{j}'} &:= a2^{\frac{\sigma_{\mathbf{j},\mathbf{j}'} - \widetilde{d}(\max\{j_1,j_1'\}+\max\{j_2,j_2'\}) - \min\{j_1,j_1',j_2,j_2'\}}{1+2q+2\widetilde{d}}}, \\ \mathcal{F}^{x_1}_{\mathbf{j},\mathbf{j}'} &:= a2^{\frac{\sigma_{\mathbf{j},\mathbf{j}'} - \widetilde{d}(\max\{j_1,j_1'\}+\max\{j_2,j_2'\}) - \min\{j_2,j_2'\}}{1+2q+2\widetilde{d}}}, \\ \mathcal{F}^{x_2}_{\mathbf{j},\mathbf{j}'} &:= a2^{\frac{\sigma_{\mathbf{j},\mathbf{j}'} - \widetilde{d}(\max\{j_1,j_1'\}+\max\{j_2,j_2'\}) - \min\{j_1,j_1'\}}{1+2q+2\widetilde{d}}}. \end{aligned}$$

Then, the compressed values are given by

$$w^{(1)}_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')} := \begin{cases} 0, \quad \text{if } \begin{cases} \sigma_{x_1} > \mathcal{E}_{\mathbf{j},\mathbf{j}'}, \\ \delta_{x_1} \le a2^{-\min\{j_1,j_1'\}}, \\ a2^{-\min\{j_1,j_2,j_1',j_2'\}} > \delta_{x_2} > a2^{-\min\{j_2,j_2'\}}, \end{cases} \\ 0, \quad \text{if } \begin{cases} \sigma_{x_2} > \mathcal{E}_{\mathbf{j},\mathbf{j}'}, \\ \delta_{x_2} \le a2^{-\min\{j_2,j_2'\}}, \\ a2^{-\min\{j_1,j_2,j_1',j_2'\}} > \delta_{x_1} > a2^{-\min\{j_1,j_1'\}}, \end{cases} \\ \big[\mathbf{A}_J\big]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}, \quad \text{otherwise,} \end{cases}$$

$$w^{(2)}_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')} := \begin{cases} 0, \quad \text{if} \begin{cases} \sigma_{x_1} > \mathcal{F}^{x_1}_{\mathbf{j},\mathbf{j}'}, \\ \delta_{x_1} \leq a2^{-\min\{j_1,j_1'\}}, \\ \delta_{x_2} \leq a2^{-\min\{j_2,j_2'\}}, \end{cases} \\ 0, \quad \text{if} \begin{cases} \sigma_{x_2} > \mathcal{F}^{x_2}_{\mathbf{j},\mathbf{j}'}, \\ \delta_{x_2} \leq a2^{-\min\{j_2,j_2'\}}, \\ \delta_{x_1} \leq a2^{-\min\{j_1,j_1'\}}, \end{cases} \\ \big[\mathbf{A}_J\big]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}, \quad \text{otherwise.} \end{cases}$$

Similar as in the first compression, we define the corresponding compressed matrices as

$$\big[\mathbf{A}^{c_2,1}_J\big]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')} := \begin{cases} w^{(1)}_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}, & \text{if } \max\{j_1,j_1'\}, \max\{j_2,j_2'\} > j_0, \\ \big[\mathbf{A}_J\big]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}, & \text{otherwise,} \end{cases} \tag{30}$$

$$\big[\mathbf{A}^{c_2,2}_J\big]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')} := \begin{cases} w^{(2)}_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}, & \text{if } \max\{j_1,j_1'\}, \max\{j_2,j_2'\} > j_0, \\ \big[\mathbf{A}_J\big]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}, & \text{otherwise.} \end{cases} \tag{31}$$

Combining these two compression schemes leads to the second compressed matrix

$$\big[\mathbf{A}^{c_2}_J\big]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')} := \begin{cases} 0, & \big[\mathbf{A}^{c_2,\ell}_J\big]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')} = 0 \text{ for some } \ell \in \{1,2\}, \\ \big[\mathbf{A}_J\big]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}, & \text{otherwise.} \end{cases} \tag{32}$$

Remark 5.4. *It suffices to compress only the entries with* $a2^{-\min\{j_1,j_2,j_1',j_2'\}} \geq \delta_{x_1}, \delta_{x_2}$*. Otherwise, we have*

$$\delta_{\text{tot}} \geq \max\{\delta_{x_1},\, \delta_{x_2}\} > a2^{-\min\{j_1,j_2,j_1',j_2'\}}$$

and the first compression applies, meaning that either the entries are zero, or that, as we will see, there are only $\mathcal{O}(4^J)$ *such entries.*

For the remainder of this section, let us without loss of the generality assume that $j_1' \leq j_1$. The following estimate holds:

Theorem 5.5. *The matrix blocks of the perturbed matrix* $\mathbf{R}_J := \mathbf{A}_J - \mathbf{A}^{c_2}_J$ *satisfy the estimate*

$$\big\|\mathbf{R}_{\mathbf{j},\mathbf{j}'}\big\| \lesssim a^{-(1+2q+2\widetilde{d})} 2^{-\sigma_{\mathbf{j},\mathbf{j}'}}$$

with a generic constant that is independent of the refinement level J*.*

Proof. It suffices to consider the two coordinate directions separately. For $\mathbf{A}^{c_2,1}_J$, we first consider the case where $\delta_{x_1} \leq a2^{-\min\{j_1,j_1'\}}$ and $\delta_{x_2} > a2^{-\min\{j_2,j_2'\}}$.

First, assume that $j_1' \neq \min\{j_1, j_2, j_1', j_2'\}$. Since we assumed that $j_1' \leq j_1$, this means that either $j_2 < j_1'$, or $j_2' < j_1'$, so we have $\min\{j_2, j_2'\} = \min\{j_1, j_2, j_1', j_2'\}$, resulting in

$$\delta_{x_2} > a2^{-\min\{j_2,j_2'\}} = a2^{-\min\{j_1,j_2,j_1',j_2'\}}.$$

Hence, according to (30), we do not compress such entries here and therefore, they do not contribute to the block error.

Let us therefore consider the case where $j_1' = \min\{j_1, j_2, j_1', j_2'\}$. For the sake of comfortability, let us define the index sets

$$\mathcal{I}_{\mathbf{j},\mathbf{k}} := \left\{ \mathbf{k}' \in \nabla_{\mathbf{j}'} : \substack{\sigma_{x_1} > \mathcal{E}_{\mathbf{j},\mathbf{j}'} \\ \delta_{x_1} \leq a2^{-j_1'} \\ a2^{-j_1'} \geq \delta_{x_2} \geq a2^{-\min\{j_2,j_2'\}}} \right\},$$

and likewise,

$$\mathcal{I}_{\mathbf{j}',\mathbf{k}'} := \left\{ \mathbf{k} \in \nabla_{\mathbf{j}} : \substack{\sigma_{x_1} > \mathcal{E}_{\mathbf{j},\mathbf{j}'} \\ \delta_{x_1} \leq a2^{-j_1'} \\ a2^{-j_1'} \geq \delta_{x_2} \geq a2^{-\min\{j_2,j_2'\}}} \right\}.$$

As one readily verifies, the cardinality of these sets is bounded by

$$\left|\mathcal{I}_{\mathbf{j},\mathbf{k}}\right| \lesssim 2^{j_2'-j_1'}, \qquad \left|\mathcal{I}_{\mathbf{j}',\mathbf{k}'}\right| \lesssim 2^{j_1-j_1'}2^{j_2-j_1'}.$$

Next, we recall that, according to Theorem 4.3, we have the estimate

$$\left|r_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}\right| \lesssim 2^{\frac{1}{2}(j_1'-j_1)}2^{-\frac{1}{2}(j_2+j_2')}2^{-\widetilde{d}(j_1+\max\{j_2,j_2'\})}\sigma_{x_1}^{-(1+2q+2\widetilde{d})}.$$

This allows us to estimate the column sums by

$$\begin{aligned}\sum_{\mathbf{k}\in\nabla_{\mathbf{j}}}\left|r_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}\right| &\lesssim \sum_{\mathbf{k}\in\mathcal{I}_{\mathbf{j}',\mathbf{k}'}} 2^{\frac{1}{2}(j_1'-j_1)}2^{-\frac{1}{2}(j_2+j_2')}2^{-\widetilde{d}(j_1+\max\{j_2,j_2'\})}\sigma_{x_1}^{-(1+2q+2\widetilde{d})}\\ &\lesssim 2^{j_1-j_1'}2^{j_2-j_1'}2^{\frac{1}{2}(j_1'-j_1)}2^{-\frac{1}{2}(j_2+j_2')}2^{-\widetilde{d}(j_1+\max\{j_2,j_2'\})}\mathcal{E}_{\mathbf{j},\mathbf{j}'}^{-(1+2q+2\widetilde{d})}\\ &\lesssim a^{-(1+2q+2\widetilde{d})}2^{\frac{1}{2}(|\mathbf{j}|_1-|\mathbf{j}'|_1)}2^{-\sigma_{\mathbf{j},\mathbf{j}'}}.\end{aligned}$$

Similarly, we may estimate the row sums by

$$\begin{aligned}\sum_{\mathbf{k}'\in\nabla_{\mathbf{j}'}}\left|r_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}\right| &\lesssim \sum_{\mathbf{k}'\in\mathcal{I}_{\mathbf{j},\mathbf{k}}} 2^{\frac{1}{2}(j_1'-j_1)}2^{-\frac{1}{2}(j_2+j_2')}2^{-\widetilde{d}(j_1+\max\{j_2,j_2'\})}\sigma_{x_1}^{-(1+2q+2\widetilde{d})}\\ &\lesssim 2^{j_2'-j_1'}2^{\frac{1}{2}(j_1'-j_1)}2^{-\frac{1}{2}(j_2+j_2')}2^{-\widetilde{d}(j_1+\max\{j_2,j_2'\})}\mathcal{E}_{\mathbf{j},\mathbf{j}'}^{-(1+2q+2\widetilde{d})}\\ &\lesssim a^{-(1+2q+2\widetilde{d})}2^{\frac{1}{2}(|\mathbf{j}'|_1-|\mathbf{j}|_1)}2^{-\sigma_{\mathbf{j},\mathbf{j}'}}.\end{aligned}$$

Hence, we can argue in complete analogy to the proof of Theorem 5.1.

By using exactly the same arguments, but with interchanging the coordinate directions, we may also control the compression error of the entries, for which $\delta_{x_1} > a2^{-\min\{j_1,j_1'\}}$ and $\delta_{x_2} \leq a2^{-\min\{j_2,j_2'\}}$. This implies the control of the error for the whole matrix $\mathbf{A}_J^{c_2,1}$.

For the matrix $\mathbf{A}_J^{c_2,2}$, we may use exactly the same arguments as in the proof of Theorem 2.3.2 in [20], with the only adaptation being that we have to use Theorem 4.3 to estimate the matrix entries instead of Theorem 2.1.7 in [20]. □

Finally, by using an additive argument, we can pose the main theorem of this section.

Theorem 5.6. *Consider the compressed matrix*

$$\left[\mathbf{A}_J^c\right]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')} := \begin{cases}0, & [\mathbf{A}_J^{c_\ell}]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')} = 0 \textit{ for some } \ell\in\{1,2\},\\ [\mathbf{A}_J]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}, & \textit{otherwise.}\end{cases} \tag{33}$$

Then, the block error is controlled by

$$\left\|\mathbf{A}_J - \mathbf{A}_J^c\right\|_2 \lesssim \varepsilon\ 2^{-\sigma_{\mathbf{j},\mathbf{j}'}},$$

with

$$\varepsilon := \max\left\{a^{-(2q+4\widetilde{d})},\, a^{-(2q+2\widetilde{d})},\, a^{-(1+2q+2\widetilde{d})}\right\}.$$

Remark 5.7. *We have improved both the parameters $\mathcal{E}_{\mathbf{j},\mathbf{j}'}$ and $\mathcal{F}_{\mathbf{j},\mathbf{j}'}^{x_i}$ in contrast to* [20]*: For both $\mathcal{E}_{\mathbf{j},\mathbf{j}'}$ and $\mathcal{F}_{\mathbf{j},\mathbf{j}'}^{x_i}$, we gain an additional factor*

$$2^{-\widetilde{d}\max\{j_\ell,j_\ell'\}}, \quad \ell\neq i,$$

in the above error estimates, but we have to pay another $\widetilde{d}$ in the denominator. This is not mandatory to ensure linear complexity, but it reduces the number of required vanishing moments, for example in the case of piecewise constant wavelets for the single layer operator to as few as three, as it is the case if an isotropic wavelet basis is used.

Much more important, we win the additional factor

$$2^{-\frac{\min\{j_1,j_2,j_1',j_2'\}}{1+2q+2\widetilde{d}}}$$

for $\mathcal{E}_{\mathbf{j},\mathbf{j}'}$ in the above error estimates. This is strictly necessary to ensure a linear complexity.

6. Complexity

We are now going to count the number of nonzero matrix coefficients of the compression matrix $\mathbf{A}_J^c$ and show that this number is asymptotically bounded by $N_J = 4^J$. In the arguments below, it is crucial that the exponents are positive to estimate the sum asymptotically by the largest term. To this end, for the sake of simplicity, we will for the remainder of this section assume that $\kappa > 0$ is sufficiently small. This is not problematic since we only have a bounded number of restrictions on κ, depending only on the uniform constants d', $\widetilde{d}$, and the order of the operator $2q$. Moreover, as we will see, we need to require the inequalities

$$d < d' < \min\Big\{\widetilde{d}+q,\ \widetilde{d}+2q,\ 2+4q+3\widetilde{d},\ \frac{1}{2}+3q+2\widetilde{d},\ \frac{1}{2}+4q+2\widetilde{d},\ \frac{1}{2}+q+\widetilde{d},\ \frac{1}{2}+2q+\widetilde{d}\Big\}. \tag{34}$$

Since this inequality is strict, there will especially always be space for a small $\kappa > 0$, which can be inserted between d' and the minimum over all these terms.

First of all, we note that the restriction of the compression to the appropriate matrix blocks in (27), (30), and (31) never causes a problem. Indeed, if we can not compress a matrix block, then at least two indexes in the set $\{j_1, j_2, j_1', j_2'\}$ are equal to j_0. In particular, as $\dim V_{j_0} \sim 2^{j_0}$, there are only $\mathcal{O}(2^{2j_0}) = \mathcal{O}(1)$ rows and columns corresponding to such situation. As every row and column contains at $\mathcal{O}(N_J)$ entries, these are at most $\mathcal{O}(N_J)$ entries in total.

For the compressed blocks, we organize the proof in the following steps. First, we split up the unit square in at most nine regions, compare Figure 2, corresponding to the distance in each coordinate direction being either big or small. These nine regions correspond to the four possible cases

$$\text{(I)}\begin{cases}\delta_{x_1} > a2^{-\min\{j_1,j_1'\}},\\ \delta_{x_2} > a2^{-\min\{j_2,j_2'\}},\end{cases} \qquad \text{(II)}\begin{cases}\delta_{x_1} \le a2^{-\min\{j_1,j_1'\}},\\ \delta_{x_2} > a2^{-\min\{j_2,j_2'\}},\end{cases}$$
$$\text{(III)}\begin{cases}\delta_{x_1} > a2^{-\min\{j_1,j_1'\}},\\ \delta_{x_2} \le a2^{-\min\{j_2,j_2'\}},\end{cases} \qquad \text{(IV)}\begin{cases}\delta_{x_1} \le a2^{-\min\{j_1,j_1'\}},\\ \delta_{x_2} \le a2^{-\min\{j_2,j_2'\}}.\end{cases} \tag{35}$$

In Section 6.1, we will show that already the first compression gives a linear complexity when there holds

$$\delta_{\text{tot}} > a2^{-\min\{j_1,j_2,j_1',j_2'\}}.$$

Then, in Section 6.2, we consider the wavelet pairs whose supports are closer together than $a2^{-\min\{j_1,j_2,j_1',j_2'\}}$ and we will show the linear complexity for those regions as well.

6.1. Complexity of the first compression.

We advance by the type of the compression which is performed on the matrix entries. First, we count all the nontrivial entries remaining from the compression scheme (22) in the case when

$$\mathcal{B}_{\mathbf{j},\mathbf{j}'} \sim 2^{\frac{\sigma_{\mathbf{j},\mathbf{j}'} - \widetilde{d}(|\mathbf{j}|_1 + |\mathbf{j}'|_1)}{2q+4\widetilde{d}}}. \tag{36}$$

Theorem 6.1. *Assume that we set all matrix entries to zero where the underlying wavelets satisfy $\delta_{\text{tot}} > \mathcal{B}_{\mathbf{j},\mathbf{j}'}$ with $\mathcal{B}_{\mathbf{j},\mathbf{j}'}$ given by (36). Then, only $\mathcal{O}(4^J)$ nontrivial entries remain.*

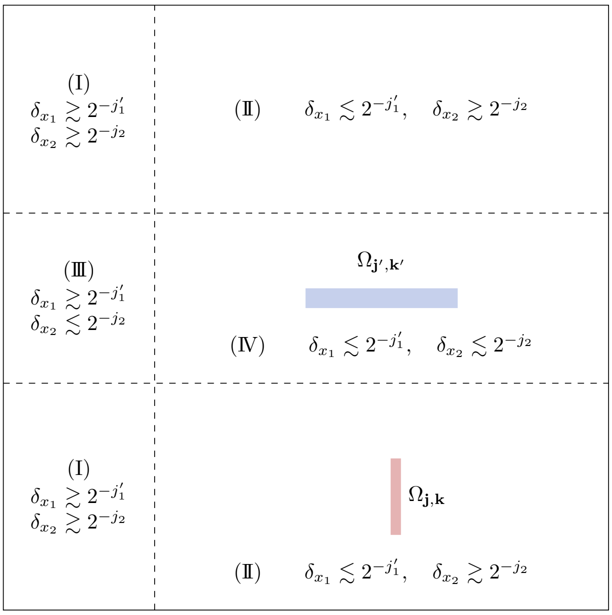

FIGURE 2. Graphical illustration of the regions described in (35). Note that there are at most nine different regions.

Proof. In any column of a block $\mathbf{A}_{\mathbf{j},\mathbf{j}'}$, we find $\mathcal{O}(2^{|\mathbf{j}'|_1}\mathcal{B}^2_{\mathbf{j},\mathbf{j}'})$ entries, for which the distance of the supports is bounded by $\mathcal{B}_{\mathbf{j},\mathbf{j}'}$. Since there are $\mathcal{O}(2^{|\mathbf{j}|_1})$ columns in such a block, there are at most $\mathcal{O}(2^{|\mathbf{j}|_1+|\mathbf{j}'|_1}\mathcal{B}^2_{\mathbf{j},\mathbf{j}'})$ nonzero entries per block. Hence, the total complexity for the whole matrix is given by

$$\begin{aligned}
\mathcal{C} &\lesssim \sum_{\substack{|\mathbf{j}|_\infty \le J\\ |\mathbf{j}'|_\infty \le J}} 2^{|\mathbf{j}|_1+|\mathbf{j}'|_1} 2^{\frac{\sigma_{\mathbf{j},\mathbf{j}'}-\tilde{d}(|\mathbf{j}|_1+|\mathbf{j}'|_1)}{q+2\tilde{d}}} \\
&\lesssim \sum_{\substack{|\mathbf{j}|_\infty \le J\\ |\mathbf{j}'|_\infty \le J}} 2^{\frac{2J(d'-q)-(d'-2\kappa)(|\mathbf{j}|_\infty+|\mathbf{j}'|_\infty)+(\tilde{d}+q-\kappa)(|\mathbf{j}|_1+|\mathbf{j}'|_1)}{q+2\tilde{d}}} \\
&= 2^{2J\frac{d'-q}{q+2\tilde{d}}}\left(\sum_{|\mathbf{j}|_\infty\le J} 2^{\frac{-(d'-2\kappa)|\mathbf{j}|_\infty+(\tilde{d}+q-\kappa)|\mathbf{j}|_1}{q+2\tilde{d}}}\right)^2. \qquad (37)
\end{aligned}$$

To calculate the sum, we explicitly enrol the indices j_1 and j_2. Then, after calculating the two sums over j_2, and shifting the index in the second sum, we obtain

$$\begin{aligned}
&\sum_{j_1=0}^{J}\left(\sum_{j_2=0}^{j_1-1} 2^{\frac{-(d'-2\kappa)j_1+(\tilde{d}+q-\kappa)(j_1+j_2)}{q+2\tilde{d}}} + \sum_{j_2=j_1}^{J} 2^{\frac{-(d'-2\kappa)j_2+(\tilde{d}+q-\kappa)(j_1+j_2)}{q+2\tilde{d}}}\right) \qquad (38)\\
&\quad\lesssim \sum_{j_1=0}^{J}\left(2^{\frac{\tilde{d}+q+\kappa-d'}{q+2\tilde{d}}j_1}2^{\frac{\tilde{d}+q-\kappa}{q+2\tilde{d}}j_1} + 2^{\frac{\tilde{d}+q-\kappa}{q+2\tilde{d}}j_1}2^{\frac{\tilde{d}+q+\kappa-d'}{q+2\tilde{d}}j_1}\sum_{j_2=0}^{J-j_1}2^{\frac{\tilde{d}+q+\kappa-d'}{q+2\tilde{d}}j_2}\right)\\
&\quad\lesssim \sum_{j_1=0}^{J}\left(2^{\frac{2\tilde{d}+2q-d'}{q+2\tilde{d}}j_1} + 2^{\frac{\tilde{d}+q+\kappa-d'}{q+2\tilde{d}}J}2^{\frac{\tilde{d}+q-\kappa}{q+2\tilde{d}}j_1}\right)\\
&\quad\lesssim 2^{J\frac{2\tilde{d}+2q-d'}{q+2\tilde{d}}}.
\end{aligned}$$

Inserting this result back into (37), we obtain

$$\mathcal{C} \lesssim 2^{2J\frac{q+2\tilde{d}}{q+2\tilde{d}}} = \mathcal{O}(4^J),$$

which is what we wanted to show. □

Remark 6.2. *Theorem 6.1 implies that there are only $\mathcal{O}(4^J)$ entries in the region* (I). *Indeed, in view of* (23)*, if $\mathcal{B}_{\mathbf{j},\mathbf{j}'} \sim a2^{-\min\{j_1,j_2,j_1',j_2'\}}$, then the entries in the region* (I) *are all set to zero. Otherwise, if* (36) *holds, then there are at most $\mathcal{O}(4^J)$ nontrivial entries by Theorem 6.1.*

Remark 6.3. *If we tensorize scaling functions on the coarsest level with wavelets, we must use the cutoff parameter from Remark 5.2. However, by using very similar arguments as in the proof above, one concludes that there are only $\mathcal{O}(4^J)$ nontrivial entries in this case. If we enrol the expression explicitly, we may assume that $j^{(1)} = |\mathbf{j}|_\infty$ and $j^{(2)} = |\mathbf{j}'|_\infty$. Since $q < d < d'$, the exponent in the second sum in* (38) *will then be negative, but thus the sum can be estimated by the lower limit in the exponent, which is j_1.*

6.2. **Complexity of the second compression.** In the next step, we want to cover the regions (Ⅱ) and (Ⅲ) in (35). Due to the symmetry in the problem, these two regions yield the same complexity, so we may only consider the region (Ⅱ), that is, $\delta_{x_1} \le 2^{-\min\{j_1,j_1'\}}$ and $\delta_{x_2} > 2^{-\min\{j_2,j_2'\}}$. For the sake of simplicity, let us for the remainder of this section assume that $j_1' \le j_1$, as the case $j_1' > j_1$ follows directly by exchanging $\mathbf{j}$ and $\mathbf{j}'$.

Lemma 6.4. *Consider all matrix entries in $\mathbf{A}_J^{\mathrm{c}}$ such that the underlying wavelet pairs $\psi_{\mathbf{j},\mathbf{k}}$ and $\psi_{\mathbf{j}',\mathbf{k}'}$ satisfy*

$$\delta_{x_1} \le a2^{-j_1'}, \quad \delta_{x_2} > a2^{-\min\{j_2,j_2'\}}. \tag{39}$$

Then, after the combination of the compression schemes (29) *and* (30)*, these are at most $\mathcal{O}(4^J)$ nontrivial entries.*

Proof. We may without loss of the generality assume that $j_1' = \min\{j_1,j_2,j_1',j_2'\}$. If not, then we must have either $j_1' > j_2$ or $j_1' > j_2'$ since we assumed $j_1' \le j_1$. Therefore, we would have

$$\delta_{\mathrm{tot}} \ge \delta_{x_2} > 2^{-\min\{j_2,j_2'\}} = 2^{-\min\{j_1,j_2,j_1',j_2'\}}.$$

Hence, these entries are either trivial or there are only $\mathcal{O}(4^J)$ of them due to Theorem 6.1. As a consequence, we may especially assume in the following $j_1' \le \min\{j_2,j_2'\}$.

We remark that we only have to consider the situation

$$\mathcal{D}^{x_2}_{\mathbf{j},\mathbf{j}'}\mathcal{E}_{\mathbf{j},\mathbf{j}'} \sim 2^{\frac{2\sigma_{\mathbf{j},\mathbf{j}'} - \tilde{d}\left(j^{(1)}+j^{(2)}+j_1+\max\{j_2,j_2'\}\right)-2j_1'}{1+2q+2\tilde{d}}},$$

since if $\mathcal{D}^{x_2}_{\mathbf{j},\mathbf{j}'} = a2^{-\min\{j_2,j_2'\}}$, in view of (39), all entries are compressed.

In order to estimate the entries, we need to consider four different cases: First, if $j_2 \le j_2'$ and $|\mathbf{j}|_\infty = j_1$, we have $j_1' \le j_2 \le j_1, j_2'$. Hence, by using $j^{(1)} := j_1$ and

$j^{(2)} := j_2'$, we conclude

$$\begin{aligned}
\mathcal{C} &\lesssim \sum_{\substack{j_1'\le j_1\\ j_2\le j_2'}} 2^{|\mathbf{j}|_1+|\mathbf{j}'|_1} 2^{\frac{2\sigma_{\mathbf{j},\mathbf{j}'}-\tilde{d}(j_1+j_2'+j_1+j_2')-2j_1'}{1+2q+2\tilde{d}}}\\
&\lesssim 2^{J\frac{4d'-4q}{1+2q+2\tilde{d}}} \sum_{\substack{j_1'\le j_1\\ j_2\le j_2'}} 2^{j_1\left(1-\frac{2d'+2\tilde{d}-2\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_2\left(1-\frac{2\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_1'\left(1-\frac{2+2\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_2'\left(1-\frac{2d'+2\tilde{d}-2\kappa}{1+2q+2\tilde{d}}\right)}\\
&\lesssim 2^{J\frac{4d'-4q}{1+2q+2\tilde{d}}} \sum_{j_1,j_2'\le J} 2^{j_1\left(2-\frac{2+2d'+2\tilde{d}}{1+2q+2\tilde{d}}\right)} 2^{j_2'\left(2-\frac{2d'+2\tilde{d}}{1+2q+2\tilde{d}}\right)}\\
&\lesssim 2^{J\left(\frac{4d'-4q}{1+2q+2\tilde{d}}+4-\frac{2+4d'+4\tilde{d}}{1+2q+2\tilde{d}}\right)}\\
&= 4^J.
\end{aligned}$$

Second, if $j_2 \le j_2'$ and $|\mathbf{j}|_\infty = j_2$, then we have the order $j_1' \le j_1 \le j_2 \le j_2'$, so with $j^{(1)} := j_2'$, $j^{(2)} := j_2$, we may directly sum up

$$\begin{aligned}
\mathcal{C} &\lesssim 2^{J\frac{4d'-4q}{1+2q+2\tilde{d}}} \sum_{j_1'\le j_1\le j_2\le j_2'} 2^{j_1\left(1-\frac{\tilde{d}+2\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_2\left(1-\frac{2d'+\tilde{d}-2\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_1'\left(1-\frac{2+2\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_2'\left(1-\frac{2d'+2\tilde{d}-2\kappa}{1+2q+2\tilde{d}}\right)}\\
&\lesssim 2^{J\frac{4d'-4q}{1+2q+2\tilde{d}}} \sum_{j_1\le j_2\le j_2'} 2^{j_1\left(2-\frac{2+\tilde{d}+4\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_2\left(1-\frac{2d'+\tilde{d}-2\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_2'\left(1-\frac{2d'+2\tilde{d}-2\kappa}{1+2q+2\tilde{d}}\right)}\\
&\lesssim 2^{J\frac{4d'-4q}{1+2q+2\tilde{d}}} \sum_{j_2\le j_2'} 2^{j_2\left(3-\frac{2+2d'+2\tilde{d}+2\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_2'\left(1-\frac{2d'+2\tilde{d}-2\kappa}{1+2q+2\tilde{d}}\right)}\\
&\lesssim 2^{J\frac{4d'-4q}{1+2q+2\tilde{d}}} \sum_{j_2'\le J} 2^{j_2'\left(4-\frac{2+4d'+4\tilde{d}}{1+2q+2\tilde{d}}\right)}\\
&\lesssim 4^J.
\end{aligned}$$

Third, if $j_2' \le j_2$ and $|\mathbf{j}|_\infty = j_1$, we have $j_1' \le j_2' \le j_2 \le j_1$, therefore the choice $j^{(1)} := j_1$, $j^{(2)} := j_2$ yields

$$\begin{aligned}
\mathcal{C} &\lesssim 2^{J\frac{4d'-4q}{1+2q+2\tilde{d}}} \sum_{j_1'\le j_2'\le j_2\le j_1} 2^{j_1\left(1-\frac{2d'+2\tilde{d}-2\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_2\left(1-\frac{2\tilde{d}+2\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_1'\left(1-\frac{2+2\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_2'\left(1-\frac{2d'-2\kappa}{1+2q+2\tilde{d}}\right)}\\
&\lesssim 2^{J\frac{4d'-4q}{1+2q+2\tilde{d}}} \sum_{j_2'\le j_2\le j_1} 2^{j_1\left(1-\frac{2d'+2\tilde{d}-2\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_2\left(1-\frac{2\tilde{d}+2\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_2'\left(2-\frac{2+2d'}{1+2q+2\tilde{d}}\right)}\\
&\lesssim 2^{J\frac{4d'-4q}{1+2q+2\tilde{d}}} \sum_{j_2\le j_1} 2^{j_1\left(1-\frac{2d'+2\tilde{d}-2\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_2\left(3-\frac{2+2d'+2\tilde{d}+2\kappa}{1+2q+2\tilde{d}}\right)}\\
&\lesssim 2^{J\frac{4d'-4q}{1+2q+2\tilde{d}}} \sum_{j_1\le J} 2^{j_1\left(4-\frac{2+4d'+4\tilde{d}}{1+2q+2\tilde{d}}\right)}\\
&\lesssim 4^J.
\end{aligned}$$

Finally, if $j_2' \le j_2$ and $|\mathbf{j}|_\infty = j_2$, we also have $j_1' \le j_1 \le j_2$. If we choose $j^{(1)} := j_2$ and $j^{(2)} := j_1$, then the complexity reads as

$$\begin{aligned}
\mathcal{C} &\lesssim 2^{J\frac{4d'-4q}{1+2q+2\tilde{d}}} \sum_{\substack{j_2'\le j_2\\ j_1'\le j_1\le j_2}} 2^{j_1\left(1-\frac{2\tilde{d}+2\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_2\left(1-\frac{2d'+2\tilde{d}-2\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_1'\left(1-\frac{2+2\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_2'\left(1-\frac{2d'-2\kappa}{1+2q+2\tilde{d}}\right)}\\
&\lesssim 2^{J\frac{4d'-4q}{1+2q+2\tilde{d}}} \sum_{j_1\le j_2} 2^{j_1\left(2-\frac{2+2\tilde{d}+4\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_2\left(2-\frac{4d'+2\tilde{d}-4\kappa}{1+2q+2\tilde{d}}\right)}\\
&\lesssim 2^{J\frac{4d'-4q}{1+2q+2\tilde{d}}} \sum_{j_2\le J} 2^{j_2\left(4-\frac{2+4d'+4\tilde{d}}{1+2q+2\tilde{d}}\right)}\\
&\lesssim 4^J.
\end{aligned}$$

□

Let us now consider the last possible situation, that is, we suppose that

$$\delta_{x_1} \le 2^{-\min\{j_1,j_1'\}}, \quad \delta_{x_2} \le 2^{-\min\{j_2,j_2'\}},$$

which describes the region (IV) in (35), or, the near-field region in Figure 2. In this case, the compression scheme (31) applies. Remember that we still assume that $j_1' \le j_1$.

Lemma 6.5. *Consider all matrix entries in* $\mathbf{A}_J^c$ *such that the underlying wavelet pairs* $\psi_{\mathbf{j},\mathbf{k}}$ *and* $\psi_{\mathbf{j}',\mathbf{k}'}$ *satisfy*

$$\delta_{x_1} \le a2^{-j_1'}, \qquad \delta_{x_2} \le a2^{-\min\{j_2,j_2'\}}.$$

Then, after the compression scheme (31)*, these are at most* $\mathcal{O}(4^J)$ *nontrivial entries.*

Proof. It suffices to count the respective matrix entries of $\mathbf{A}_J^{c_2,2}$. Suppose first that $j_2' \ge j_2$. In this case, the number of nontrivial entries in the matrix block $[\mathbf{A}_J^{c_2,2}]_{\mathbf{j},\mathbf{j}'}$ can be estimated by

$$N_{\mathbf{j},\mathbf{j}'} \lesssim 2^{|\mathbf{j}|_1+|\mathbf{j}'|_1} \mathcal{F}_{\mathbf{j},\mathbf{j}'}^{x_1} \mathcal{F}_{\mathbf{j},\mathbf{j}'}^{x_2} \sim 2^{|\mathbf{j}|_1+|\mathbf{j}'|_1} 2^{\frac{2\sigma_{\mathbf{j},\mathbf{j}'}-2\tilde{d}(j_1+j_2')-(j_1'+j_2)}{1+2q+2\tilde{d}}}. \tag{40}$$

Remarking that for $\kappa > 0$ sufficiently small, we have $d' - 2\kappa > 0$, so the properties $|\mathbf{j}|_\infty \ge j_1$ and $|\mathbf{j}'|_\infty \ge j_2'$ imply that the number of entries can be estimated by

$$\begin{aligned}
\mathcal{C}^{(1)} \lesssim \sum_{\substack{0\le j_1\le J\\ 0\le j_1'\le j_1}} \sum_{\substack{0\le j_2'\le J\\ 0\le j_2\le j_2'}} N_{\mathbf{j},\mathbf{j}'} \lesssim 2^{J\frac{4d'-4q}{1+2q+2\tilde{d}}} &\left(\sum_{\substack{0\le j_1\le J\\ 0\le j_1'\le j_1}} 2^{j_1\left(1-\frac{2d'-2\kappa+2\tilde{d}}{1+2q+2\tilde{d}}\right)} 2^{j_1'\left(1-\frac{1+2\kappa}{1+2q+2\tilde{d}}\right)} \right)\\
&\left(\sum_{\substack{0\le j_2'\le J\\ 0\le j_2\le j_2'}} 2^{j_2\left(1-\frac{1+2\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_2'\left(1-\frac{2d'-2\kappa+2\tilde{d}}{1+2q+2\tilde{d}}\right)} \right).
\end{aligned}$$

As the indices can be exchanged, the two sums are equal and can be estimated by

$$\sum_{j_1=0}^{J} 2^{j_1\left(1-\frac{2d'-2\kappa+2\tilde{d}}{1+2q+2\tilde{d}}\right)} \sum_{j_1'=0}^{j_1} 2^{j_1'\left(1-\frac{1+2\kappa}{1+2q+2\tilde{d}}\right)} \lesssim \sum_{j_1=0}^{J} 2^{j_1\left(2-\frac{1+2d'+2\tilde{d}}{1+2q+2\tilde{d}}\right)} \lesssim 2^{J\left(2-\frac{1+2d'+2\tilde{d}}{1+2q+2\tilde{d}}\right)}.$$

Putting everything together, we obtain that

$$\mathcal{C}^{(1)} \lesssim 2^{J\left(\frac{4d'-4q}{1+2q+2\tilde{d}}+4-\frac{2+4d'+4\tilde{d}}{1+2q+2\tilde{d}}\right)} = 4^J.$$

In the case where $j_2 \ge j_2'$, we need to argue in a slightly different way. By the assumption $j_1' \le j_1$, we have at least $\mathcal{O}(2^{|\mathbf{j}'|_1})$ nontrivial entries in the matrix block $[\mathbf{A}_J^{c_2,2}]_{\mathbf{j},\mathbf{j}'}$, as at every point at which the singular supports $\Omega_{j_1',k_1'}^\sigma$ and $\Omega_{j_2',k_2'}^\sigma$

intersect, there is at least one smaller wavelet $\psi_{\mathbf{j},\mathbf{k}}$ touching it. On the other hand, every nontrivial entry satisfies $\sigma_{x_1} \leq \mathcal{F}^{x_1}_{\mathbf{j},\mathbf{j}'}$ and $\sigma_{x_2} \leq \mathcal{F}^{x_2}_{\mathbf{j},\mathbf{j}'}$, so the number of nontrivial entries in a block $[\mathbf{A}^{c_2,2}_J]_{\mathbf{j},\mathbf{j}'}$ can be estimated by

$$N_{\mathbf{j},\mathbf{j}'} \lesssim 2^{|\mathbf{j}'|_1} \max\big\{1,\ 2^{|\mathbf{j}|_1}\mathcal{F}^{x_1}_{\mathbf{j},\mathbf{j}'}\mathcal{F}^{x_2}_{\mathbf{j},\mathbf{j}'}\big\}. \tag{41}$$

If the maximum in (41) is equal to 1, then it is easy to see that

$$\sum_{\mathbf{j}'\leq\mathbf{j}\leq\mathbf{J}} N_{\mathbf{j},\mathbf{j}'} \lesssim \sum_{j_1'\leq j_1\leq J}\ \sum_{j_2'\leq j_2\leq J} 2^{j_1'+j_2'} \lesssim \sum_{j_1\leq J}\sum_{j_2\leq J} 2^{j_1+j_2} \lesssim 4^J.$$

If, however, the maximum in (41) is not equal to 1, we have

$$N_{\mathbf{j},\mathbf{j}'} \lesssim 2^{|\mathbf{j}|_1+|\mathbf{j}'|_1} 2^{\frac{2\sigma_{\mathbf{j},\mathbf{j}'}-2\tilde{d}(j_1+j_2)-(j_1'+j_2')}{1+2q+2\tilde{d}}}.$$

Using again that $d'-2\kappa>0$, $|\mathbf{j}|_\infty \geq j_1$, $|\mathbf{j}'|_\infty \geq j_2'$, one obtaines

$$\mathcal{C}^{(2)} \lesssim 2^{J\frac{4d'-4q}{1+2q+2\tilde{d}}} \left(\sum_{\substack{0\leq j_1\leq J\\ 0\leq j_1'\leq j_1}} 2^{j_1\left(1-\frac{2d'-2\kappa+2\tilde{d}}{1+2q+2\tilde{d}}\right)} 2^{j_1'\left(1-\frac{1+2\kappa}{1+2q+2\tilde{d}}\right)}\right)$$
$$\left(\sum_{\substack{0\leq j_2\leq J\\ 0\leq j_2'\leq j_2}} 2^{j_2\left(1-\frac{2\tilde{d}+2\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_2'\left(1-\frac{2d'-2\kappa+1}{1+2q+2\tilde{d}}\right)}\right).$$

The first sum can be treated as in the previous case, whereas for the second sum, there holds

$$\sum_{\substack{0\leq j_2\leq J\\ 0\leq j_2'\leq j_2}} 2^{j_2\left(1-\frac{2\tilde{d}+2\kappa}{1+2q+2\tilde{d}}\right)} 2^{j_2'\left(1-\frac{2d'-2\kappa+1}{1+2q+2\tilde{d}}\right)} \lesssim \sum_{0\leq j_2\leq J} 2^{j_2\left(2-\frac{1+2d'+2\tilde{d}}{1+2q+2\tilde{d}}\right)} \lesssim 2^{J\left(2-\frac{1+2d'+2\tilde{d}}{1+2q+2\tilde{d}}\right)}.$$

Hence, we can conclude that also $\mathcal{C}^{(2)} \lesssim 4^J$. □

With the preceeding three lemmata and Theorem 6.1 at hand, we conclude the main result of this section.

Theorem 6.6. *Assume that* $-q < d \leq \widetilde{d}$ *and that* (34) *holds. Then, compressed matrix* $\mathbf{A}^c_J$ *arising from* (33) *contains at most* $\mathcal{O}(4^J)$ *nontrivial entries.*

Remark 6.7. *For the piecewise constant* ($d = 1$) *and piecewise bilinear* ($d = 2$) *wavelets, used most often in practice, one needs at least the following number of vanishing moments:*

	$d=1$	$d=2$
$2q=-1$	3	4
$2q=0$	2	3
$2q=1$	–	2

If $2q = 1$, *using piecewise constant wavelets is mathematically not meaningful because the energy space* $H^{\frac{1}{2}}(\Gamma)$ *cannot be discretized by discontinuous trial functions. Note that these are the same values as in the setting of isotropic wavelet bases, compare* [8].

7. The situation on a Lipschitz manifold

Up to now, we have only considered the situation on the unit square. As stated in Section 2, we are however interested in the solution of a boundary integral equation posed on the boundary Γ of a Lipschitz domain $D \subseteq \mathbb{R}^3$. Recall that the boundary Γ is given as the union of the patches Γ_i, which can be smoothly parametrized by $\boldsymbol{\gamma}_i : \square \to \Gamma_i$.

7.1. **Sobolev spaces on manifolds.** To properly define the Sobolev space $H^s(\Gamma)$, we choose for each patch Γ_i an extension $\widetilde{\Gamma}_i$ such that $\Gamma_i \Subset \widetilde{\Gamma}_i \Subset \Gamma$. Then, the family $\{\widetilde{\Gamma}_i\}_i$ is an overlapping decomposition of Γ and, by the Lipschitz continuity of Γ, every parametrization $\boldsymbol{\gamma}_i$ can be used to construct a piecewise smooth, globally Lipschitz continuous parametrization $\widetilde{\boldsymbol{\gamma}}_i : \widetilde{\square} \to \widetilde{\Gamma}_i$, where $\widetilde{\square}$ is a suitable superset of $\square$.

Next, we choose a partition of the unity $\{\chi_i\}_{i=1}^r$, consisting of nonnegative, smooth functions $\chi_i \in C_0^\infty(\mathbb{R}^3)$ such that

$$\sum_{i=0}^{r} \chi_i(\boldsymbol{x}) = 1 \text{ for all } \boldsymbol{x} \in \Gamma, \quad \chi_i(\boldsymbol{x}) = 0 \text{ for all } \boldsymbol{x} \in \Gamma \setminus \widetilde{\Gamma}_i.$$

We then define the Sobolev space $H^s(\Gamma)$ for $s \leq 1$ as the space of functions $u : \Gamma \to \mathbb{R}$, for which the norm

$$\|u\|_{H^s(\Gamma)} := \left(\sum_{i=1}^{r} \big\| (u \circ \widetilde{\boldsymbol{\gamma}}_i) \cdot (\chi_i \circ \widetilde{\boldsymbol{\gamma}}_i) \big\|_{H^s(\widetilde{\square})}^2 \right)^{\frac{1}{2}}$$

is finite. Obviously, this norm is induced by the inner product

$$\langle u, v \rangle_{H^s(\Gamma)} := \sum_{i=1}^{r} \big\langle (u \circ \widetilde{\boldsymbol{\gamma}}_i)(\chi_i \circ \widetilde{\boldsymbol{\gamma}}_i), (v \circ \widetilde{\boldsymbol{\gamma}}_i)(\chi_i \circ \widetilde{\boldsymbol{\gamma}}_i) \big\rangle_{H^s(\widetilde{\square})}.$$

For $-1 \leq s < 0$, we define $H^s(\Gamma)$ as usual as the dual of $H^{-s}(\Gamma)$ with respect to the $L^2(\Gamma)$-inner product.

Remark 7.1. *The above definition depends on the parametrizations $\boldsymbol{\gamma}_i$ and the chosen partition of the unity. However, it can be shown that the space $H^s(\Gamma)$ consists of the same functions, regardless which parametrizations and which partition of the unity is chosen* [30]. *Moreover, the requirement $|s| \leq 1$ amounts from the Lipschitz continuity of $\widetilde{\boldsymbol{\gamma}}_i$. For a $C^{k,\alpha}$-surface, it is possible to define these spaces up to $|s| \leq k + \alpha$.*

7.2. **Patchwise smooth wavelets.** Similar to the existing literature, see e.g. [8, 24], we discretize the energy space $H^q(\Gamma)$ with $q \leq 1$ by transporting the wavelet functions from the unit square onto Γ. Precisely, we define a basis function as

$$\psi_{i,\mathbf{j},\mathbf{k}} := \psi_{\mathbf{j},\mathbf{k}} \circ \boldsymbol{\gamma}_i^{-1},$$

and then consider the basis set

$$\Psi := \big\{ \psi_{i,\mathbf{j},\mathbf{k}} : 1 \leq i \leq r, \, \mathbf{j} \in \mathbb{N}_0^2, \, \mathbf{k} \in \nabla_{\mathbf{j}} \big\}.$$

To construct a trial space on the level J, we proceed as on the unit square, meaning that we cut the basis off at the level J, obtaining

$$V_J := \operatorname{span} \Psi_J, \quad \Psi_J := \big\{ \psi_{i,\mathbf{j},\mathbf{k}} : 1 \leq i \leq r, \, |\mathbf{j}|_\infty \leq J, \, \mathbf{k} \in \nabla_{\mathbf{j}} \big\}.$$

Note that the above wavelets are supported on a single patch Γ_i only. In general, they are not continuous over the patch boundaries, so they only attain a regularity of $\gamma = \frac{1}{2}$, regardless how smooth the wavelets are piecewise. It is possible to construct wavelets which are continuous over the patch boundaries, see e.g. [24], but $\gamma = \frac{1}{2}$ is sufficient if we want to discretize an operator of nonpositive order.

Because the wavelets are supported on a single patch, and each parametrization and cutoff function is smooth, one can generalize all the wavelet properties stated in Section 3.4. For the same reason, the Lipschitz continuity of the surface is no restriction in using $\widetilde{d} > 1$ vanishing moments for the cancellation property (17), since we do not have to consider the behaviour of the test function over the patch boundaries, and the expressions $|\cdot|_{W^{\tilde{d},\infty}(\Omega_{i,\mathbf{j},\mathbf{k}})}$ are well-defined.

Finally, as the spaces V_J coincide with the space spanned by all lifted dyadic indicator functions on the level J, we can directly quote the following lemma, compare [18].

Lemma 7.2. *For a continuous, strongly elliptic and injective operator $\mathcal{A} : H^q(\Gamma) \to H^{-q}(\Gamma)$, the Galerkin discretization is stable, meaning that*

$$\langle v_J, (\mathcal{A} + \mathcal{A}^\star) v_J \rangle_\Gamma \gtrsim \|v_J\|^2_{H^q(\Gamma)}, \quad v_J \in V_J,$$

for any sufficiently large J, and

$$\left|\langle v_J, \mathcal{A} w_J \rangle_\Gamma\right| \lesssim \|v_J\|_{H^q(\Gamma)} \|w_J\|_{H^q(\Gamma)}, \quad v_J, w_J \in V_J.$$

Furthermore, let u be the solution of (5) *and u_J the solution of* (7)*. Then, we have the convergence*

$$\|u - u_J\|_{H^t(\Gamma)} \lesssim 2^{-J(s-t)} \|u\|_{H^s(\Gamma)}, \quad 2q - d \le t \le q, \quad q \le s \le d,$$

provided that Γ is sufficiently regular.

Remark 7.3. *The condition that $\mathcal{A}$ is injective is, as already stated, not strictly necessary. It suffices if the kernel is finite-dimensional and known in advance.*

7.3. **Matrix estimates.** As we will see, all the matrix estimates on a given surface can be concluded from the matrix estimates on the unit square. Depending on the situation of the two patches on which the wavelets are supported, we need to differ between several cases. To this end, let

$$\hat{K}_{i,i'}(\hat{\boldsymbol{x}}, \hat{\boldsymbol{x}}') := K\big(\boldsymbol{\gamma}_i(\hat{\boldsymbol{x}}), \boldsymbol{\gamma}_{i'}(\hat{\boldsymbol{x}}')\big) \sqrt{\det\big(\mathbf{D}\boldsymbol{\gamma}_i^\intercal \mathbf{D}\boldsymbol{\gamma}_i\big)(\hat{\boldsymbol{x}})} \sqrt{\det\big(\mathbf{D}\boldsymbol{\gamma}_{i'}^\intercal \mathbf{D}\boldsymbol{\gamma}_{i'}\big)(\hat{\boldsymbol{x}}')} \tag{42}$$

with $1 \le i, i' \le r$ denote the transported kernel function. With the transported kernel function at hand, we find that

$$\begin{aligned} \langle \psi_{i,\mathbf{j}',\mathbf{k}'}, \mathcal{A}\psi_{i',\mathbf{j},\mathbf{k}} \rangle_\Gamma &= \int_\Gamma \int_\Gamma K(\boldsymbol{x}, \boldsymbol{x}') \psi_{\mathbf{j},\mathbf{k}}\big(\boldsymbol{\gamma}_i^{-1}(\boldsymbol{x})\big) \psi_{\mathbf{j}',\mathbf{k}'}\big(\boldsymbol{\gamma}_{i'}^{-1}(\boldsymbol{x}')\big) \,\mathrm{d}S_{\boldsymbol{x}'} \,\mathrm{d}S_{\boldsymbol{x}} \\ &= \int_\square \int_\square \hat{K}_{i,i'}(\hat{\boldsymbol{x}}, \hat{\boldsymbol{x}}') \psi_{\mathbf{j},\mathbf{k}}(\hat{\boldsymbol{x}}) \psi_{\mathbf{j}',\mathbf{k}'}(\hat{\boldsymbol{x}}') \,\mathrm{d}\hat{\boldsymbol{x}}' \,\mathrm{d}\hat{\boldsymbol{x}} \\ &= \langle \psi_{\mathbf{j}',\mathbf{k}'}, \hat{\mathcal{A}}_{i,i'} \psi_{\mathbf{j},\mathbf{k}} \rangle_\square, \end{aligned} \tag{43}$$

where we define $\hat{\mathcal{A}}_{i,i'}$ as the integral operator with the transported kernel $\hat{K}_{i,i'}$. Since that local parametrizations $\boldsymbol{\gamma}_i$ and $\boldsymbol{\gamma}_{i'}$ are smooth, the transported kernel functions also satisfy the decay property

$$\left|\partial_{\hat{\boldsymbol{x}}}^{\boldsymbol{\alpha}} \partial_{\hat{\boldsymbol{x}}'}^{\boldsymbol{\alpha}'} \hat{K}_{i,i'}(\hat{\boldsymbol{x}}, \hat{\boldsymbol{x}}')\right| \le C_{\boldsymbol{\alpha},\boldsymbol{\alpha}',K,i,i'} \|\boldsymbol{\gamma}_i(\hat{\boldsymbol{x}}) - \boldsymbol{\gamma}_{i'}(\hat{\boldsymbol{x}}')\|^{-(2+2q+|\boldsymbol{\alpha}|+|\boldsymbol{\alpha}'|)} \tag{44}$$

provided that $2 + 2q + |\boldsymbol{\alpha}| + |\boldsymbol{\alpha}'| > 0$. Therefore, in view of (43) and (44), the far-field estimates of Section 4.1 hold true also in case of piecewise smooth Lipschitz manifolds.

7.3.1. *Wavelets supported on the same patch.* Let us look at the easiest situation first. If we consider the interaction between $\psi_{i,\mathbf{j},\mathbf{k}}$ and $\psi_{i,\mathbf{j}',\mathbf{k}'}$, we can use the relations (43) and (44) together with

$$\|\boldsymbol{\gamma}_i(\hat{\boldsymbol{x}}) - \boldsymbol{\gamma}_i(\hat{\boldsymbol{x}}')\| \sim \|\hat{\boldsymbol{x}} - \hat{\boldsymbol{x}}'\|$$

to conclude that the situation on a single patch is equivalent to that of the unit square. Therefore, also the near-field estimates are valid one-to-one and we find at most $\mathcal{O}(4^J)$ nontrivial entries in the matrix block associated with $\Gamma_i \times \Gamma_i$. Especially, the compression error in each matrix block satisfies the same estimates as in Section 5.

7.3.2. *Wavelets supported on patches with a common edge.* Let us now assume that Γ_i and $\Gamma_{i'}$ share a common edge. For the sake of simplicity, we assume that the common edge Σ satisfies

$$\boldsymbol{\gamma}_i\big(\{1\} \times [0,1]\big) = \Sigma = \boldsymbol{\gamma}_{i'}\big(\{0\} \times [0,1]\big),$$

especially that there holds $\boldsymbol{\gamma}_i(1, x_2) = \boldsymbol{\gamma}_{i'}(0, x_2)$ for all $x_2 \in [0,1]$. Otherwise, we can apply suitable rotations such that this assumption holds.

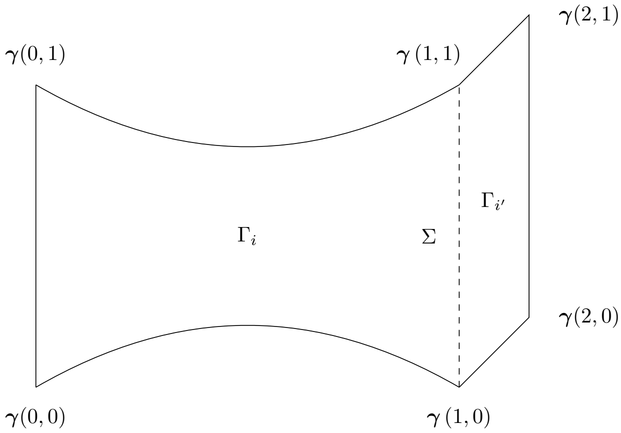

FIGURE 3. Graphical illustration of the parametrization $\boldsymbol{\gamma}$ in case of a common edge.

By gluing the two parametrizations together, we obtain a Lipschitz continuous parametrization $\boldsymbol{\gamma} : [0,2] \times [0,1] \to \Gamma_i \cup \Gamma_{i'}$ such that

$$\boldsymbol{\gamma}(\boldsymbol{x}) = \begin{cases} \boldsymbol{\gamma}_i(\boldsymbol{x}), & \boldsymbol{x} \in [0,1] \times [0,1], \\ \boldsymbol{\gamma}_{i'}(x_1 - 1, x_2), & \boldsymbol{x} \in [1,2] \times [0,1], \end{cases}$$

compare Figure 3 for a graphical illustration. For the near field estimates, we need to interpret the coordinate directions in a meaningful way. This is quite intuitive in Figure 3, since we can simply define the x-direction as the direction across the edge, while the y-direction can be interpreted as the direction parallel to the edge. Especially, we find

$$\|\boldsymbol{\gamma}_i(\hat{\boldsymbol{x}}) - \boldsymbol{\gamma}_{i'}(\hat{\boldsymbol{x}}')\| \sim \left\|\hat{\boldsymbol{x}} - \hat{\boldsymbol{x}}' - \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]\right\|$$

For Theorem 4.3, we followed the arguments in [20], where the smooth part of the bigger wavelet $\psi_{j_1',k_1'}$ is extended to a smooth, compactly supported function. Then, also the operator $\mathcal{A}_1$ is extended to a classical, pseudo-differential operator $\mathcal{A}_1^\sharp$, compare [18]. This is not so easy to do in the current situation, as the kernel $\hat{K}_{i,i'}$ is no longer asymptotically smooth of the order $2q$, since it is only continuous over the edge Σ. Nevertheless, this difficulty can be overcome. For the sake of simplicity, let us define

$$\hat{\Omega}_{j_1,k_1} \times \hat{\Omega}_{j_2,k_2} := \operatorname{supp}\big(\psi_{i,\mathbf{j},\mathbf{k}} \circ \boldsymbol{\gamma}\big), \qquad \hat{\Omega}_{j_1',k_1'} \times \hat{\Omega}_{j_2',k_2'} := \operatorname{supp}\big(\psi_{i',\mathbf{j}',\mathbf{k}'} \circ \boldsymbol{\gamma}\big).$$

Then, there holds $\hat{\Omega}_{j_1,k_1} \subseteq [0,1]$ and $\hat{\Omega}_{j_1',k_1'} \subseteq [1,2]$ in (43). Hence, we either have $\sigma_{x_1} = 0$, in which case we cannot compress in the direction of x_1 at all, or that the smooth extension function $\tilde{f}$ from the proof of Theorem 4.3 is equal to 0, since the two wavelets are located on different patches. Hence, we can ignore everything about the pseudo-differential operator and only consider the complement function.

In the second coordinate direction, the situation looks more complicated, as $\hat{\Omega}_{j_2,k_2}$ and $\hat{\Omega}_{j_2',k_2'}$ may not be disjoint. In this case, we still need to argue with $\hat{\mathcal{A}}_2^\sharp$, which corredponds to the kernel

$$\hat{K}_{i,i'}^{(2)}(\hat{x}_2,\hat{x}_2') := \int_0^1 \int_0^1 \hat{K}_{i,i'}(\hat{\boldsymbol{x}},\hat{\boldsymbol{x}}')\psi_{j_1,k_1}(\hat{x}_1)\psi_{j_1',k_1'}(\hat{x}_1')\,\mathrm{d}\hat{x}_1'\,\mathrm{d}\hat{x}_1.$$

Although $\boldsymbol{\gamma}$ is overall only Lipschitz continuous, it is indeed smooth in the second coordinate direction, as both $\boldsymbol{\gamma}_i$ and $\boldsymbol{\gamma}_{i'}$ are smooth and coincide on the common edge Σ. Thus, the arguments of [20] work in this case as well.

In the proof of Theorem 4.3, we have also made use of vanishing moments hidden in the kernel. This is possible to do here as well. Indeed, if e.g. $j_1' \geq j_1$, then only have to consider a term of the form

$$\operatorname*{ess\,sup}_{\hat{x}_1' \in \hat{\Omega}_{j_1',k_1'}} \left| \int_{\hat{\Omega}_{j_1,k_1}} \psi_{j_1,k_1}(\hat{x}_1)\ \partial_{\hat{x}_1'}^{\tilde{d}} \hat{K}_{i,i'}(\hat{\boldsymbol{x}},\hat{\boldsymbol{x}}')\,\mathrm{d}\hat{x}_1 \right|,$$

where the asymptotic estimate for $\hat{K}_{i,i'}$ holds since, $\boldsymbol{\gamma}_{i'}$ is smooth on $\Omega_{j_1',k_1'} \times [0,1]$ in view of (42). Hence, the same arguments as in the proof of Theorem 4.3 apply.

7.3.3. *Patches with a common vertex.* If the patches Γ_i and $\Gamma_{i'}$ have a common vertex, we can use a similar argument as before. We assume that the common vertex $\boldsymbol{v}$ satisfies – possibly after application of suitable rotations and translations –

$$\boldsymbol{\gamma}_i\big(\big[\begin{smallmatrix}1\\1\end{smallmatrix}\big]\big) = \boldsymbol{v} = \boldsymbol{\gamma}_{i'}\big(\big[\begin{smallmatrix}0\\0\end{smallmatrix}\big]\big).$$

Hence, we may find a Lipschitz continuous parametrization $\boldsymbol{\gamma} : [0,2]^2 \to \Gamma$ such that

$$\boldsymbol{\gamma}\big|_{[0,1]^2} = \boldsymbol{\gamma}_i, \quad \boldsymbol{\gamma}\big|_{[1,2]^2} = \boldsymbol{\gamma}_{i'}.$$

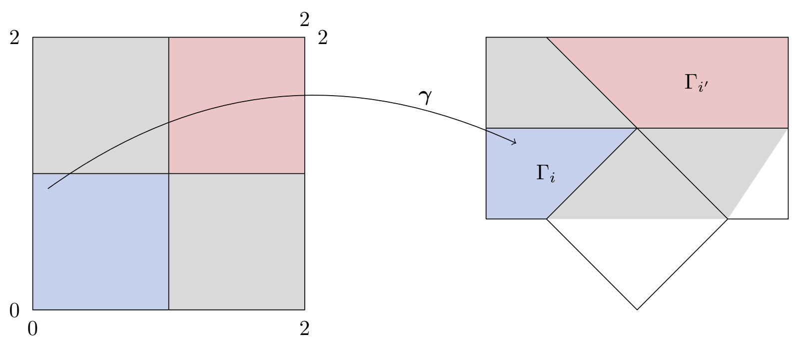

FIGURE 4. A possible Lipschitz continuous extension of the maps $\boldsymbol{\gamma}_i$ and $\boldsymbol{\gamma}_{i'}$ in case of a common vertex.

Concerning the first compression, in view of (43), (44), and

$$\|\boldsymbol{\gamma}_i(\hat{\boldsymbol{x}}) - \boldsymbol{\gamma}_{i'}(\hat{\boldsymbol{x}}')\| \sim \big\|\hat{\boldsymbol{x}} - \hat{\boldsymbol{x}}' - \big[\begin{smallmatrix}1\\1\end{smallmatrix}\big]\big\| \tag{45}$$

we can obviously use the estimates of Section 4.1 here, too.

To proceed with the second compression, we need an estimate like the one in Theorem 4.3, for which we define

$$\hat{K}^{(1)}_{i,i'}(\hat{x}_1,\hat{x}_1') := \int_0^1\int_0^1 \hat{K}_{i,i'}(\hat{\boldsymbol{x}},\hat{\boldsymbol{x}}')\psi_{j_2,k_2}(\hat{x}_2)\psi_{j_2',k_2'}(\hat{x}_2')\,\mathrm{d}x_2'\,\mathrm{d}x_2.$$

In view of (45) and the fact that the restrictions of $\boldsymbol{\gamma}$ to $[0,1]^2$ and $[1,2]^2$ are smooth, we deduce for $\boldsymbol{x} = \boldsymbol{\gamma}_i(\hat{\boldsymbol{x}})$ and $\boldsymbol{x}' = \boldsymbol{\gamma}_i(\hat{\boldsymbol{x}}')$ that

$$\begin{aligned}\left|\partial_x^\alpha\partial_{x'}^{\alpha'}\hat{K}^{(1)}_{i,i'}(\hat{x}_1,\hat{x}_1')\right| &\lesssim 2^{-\frac{1}{2}(j_2+j_2')}2^{-\widetilde{d}\max\{j_2,j_2'\}}\|\boldsymbol{x}-\boldsymbol{x}'\|^{-(2+2q+\widetilde{d}+\alpha+\alpha')}\\ &\lesssim 2^{-\frac{1}{2}(j_2+j_2')}2^{-\widetilde{d}\max\{j_2,j_2'\}}\left\|\hat{\boldsymbol{x}}-\hat{\boldsymbol{x}}'-\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]\right\|^{-(2+2q+\widetilde{d}+\alpha+\alpha')}\\ &\lesssim 2^{-\frac{1}{2}(j_2+j_2')}2^{-\widetilde{d}\max\{j_2,j_2'\}}|\hat{x}_1-\hat{x}_1'-1|^{-(2+2q+\widetilde{d}+\alpha+\alpha')}.\end{aligned}$$

This is exactly the estimate needed for Theorem 4.3 when considered on the interval $[0,2]$ with $\hat{x}_1 \in [0,1]$ and $\widetilde{x}_1 := \hat{x}_1' + 1 \in [1,2]$. Similarly, we can derive such an estimate for the second coordinate direction. As it holds

$$\begin{aligned}\left\langle\psi_{i',\mathbf{j}',\mathbf{k}'},\mathcal{A}\psi_{i,\mathbf{j},\mathbf{k}}\right\rangle_\Gamma &= \int_0^1\int_0^1 \hat{K}^{(1)}_{i,i'}(x_1,x_1')\psi_{j_1,k_1}(x_1)\psi_{j_1',k_1'}(x_1')\,\mathrm{d}x_1'\,\mathrm{d}x_1\\ &= \int_0^1\int_0^1 \hat{K}^{(2)}_{i,i'}(x_2,x_2')\psi_{j_2,k_2}(x_2)\psi_{j_2',k_2'}(x_2')\,\mathrm{d}x_2'\,\mathrm{d}x_2,\end{aligned}$$

it is enough if one realization of the entry can be compressed.

Moreover, on the parameter domain $[0,2]^2$, cf. Figure 4, the preimages of $\Omega_{i,\mathbf{j},\mathbf{k}}$ and $\Omega_{i',\mathbf{j}',\mathbf{k}'}$ either touch each other, in which case we cannot compress, or they are well-seperated in at least one coordinate direction. Hence, in this direction, the smooth extension $\tilde{f}$ of the larger wavelet is 0 as well, so we do only have to consider the complement function. This estimate depends only on the distance between the supports, which is, at least in this coordinate direction, equivalent to the distance in $\mathbb{R}^3$, since

$$\|\boldsymbol{\gamma}_i(\hat{\boldsymbol{x}})-\boldsymbol{\gamma}_{i'}(\hat{\boldsymbol{x}}')\|_2 \sim \left\|\hat{\boldsymbol{x}}-\hat{\boldsymbol{x}}'-\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]\right\|_2 \sim \left\|\hat{\boldsymbol{x}}-\hat{\boldsymbol{x}}'-\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]\right\|_\infty = |\hat{x}_\ell-\hat{x}_\ell'-1|$$

for an $\ell \in \{1,2\}$.

7.3.4. *Well-separated patches.* If Γ_i and $\Gamma_{i'}$ do neither share a common vertex nor a common edge, then, since the domain under consideration has at least a Lipschitz boundary, it holds $\operatorname{dist}(\Gamma_i,\Gamma_{i'}) \gtrsim 1 \geq 2^{-\min\{j_1,j_2,j_1',j_2'\}}$. Therefore, the first compression is possible for all such entries. We also note that in this case, we only need the spatial distance, so we do not have to think about appropriate coordinate directions here.

In the first compression, there are two different possibilities: First, if

$$\mathcal{B}_{\mathbf{j},\mathbf{j}'} \sim 2^{-\min\{j_1,j_2,j_1',j_2'\}},$$

and $h := \min\{1,\operatorname{dist}(\Gamma_i,\Gamma_j)\} > 0$, then we can compress all entries, for which the maximal support size is smaller than h, meaning if

$$\min\{j_1,j_2,j_1',j_2'\} > \lceil\log_2 h\rceil =: C,$$

leaving us at

$$\sum_{\substack{j_1,j_2\leq J\\ j_1',j_2'\leq J}} 2^{|\mathbf{j}|_1+|\mathbf{j}'|_1-\min\{j_1,j_2,j_1',j_2'\}} \lesssim C2^{3C} = \mathcal{O}(1)$$

entries. Note that this is entirely theoretical – it is clear that this constant deteriorates if $h > 0$ is small. If not, then there are only $\mathcal{O}(4^J)$ entries on each patch-patch interaction, which is implied by the proof of Theorem 6.1.

8. Consistency and convergence

In this section, we are going to show that the Galerkin scheme for the compressed operator converges as well as the Galerkin scheme for the uncompressed operator. This means that the wavelet matrix compression under consideration realizes the discretization error accuracy offered by the underlying Galerkin scheme.

Similar to [8], we define the compressed boundary integral operator $\mathcal{A}_J^\varepsilon$ in accordance with

$$\mathcal{A}_J^\varepsilon u := \sum_{\substack{|\mathbf{j}|_\infty \leq J \\ |\mathbf{j}'|_\infty \leq J}} \sum_{\substack{\mathbf{k}\in\nabla_\mathbf{j} \\ \mathbf{k}'\in\nabla_{\mathbf{j}'}}} \big[\mathbf{A}_J^\varepsilon\big]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')} \langle \widetilde{\psi}_{\mathbf{j}',\mathbf{k}'}, u\rangle_\Gamma \widetilde{\psi}_{\mathbf{j},\mathbf{k}},$$

which defines a continuous operator $H^s(\Gamma) \to H^{s-2q}(\Gamma)$ for all $-\widetilde{\gamma} < s < \widetilde{\gamma} + 2q$. Especially, this operator represents the compressed matrix $\mathbf{A}_J^\varepsilon$ in terms of

$$\big\langle \psi_{\mathbf{j}',\mathbf{k}'}, \mathcal{A}_J^\varepsilon \psi_{\mathbf{j},\mathbf{k}}\big\rangle_\Gamma = \big[\mathbf{A}_J^\varepsilon\big]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}, \quad |\mathbf{j}|_\infty, |\mathbf{j}'|_\infty \leq J, \quad \mathbf{k}\in\nabla_\mathbf{j}, \mathbf{k}'\in\nabla_{\mathbf{j}'}.$$

Theorem 8.1 (Consistency)**.** *Let $\mathbf{A}_J^\varepsilon$ denote the compressed matrix of the level J with a parameter a, such that*

$$\max\big\{a^{-(2q+4\widetilde{d})},\, a^{-(2q+2\widetilde{d})},\, a^{-(1+2q+2\widetilde{d})}\big\} \leq \varepsilon.$$

Then, for $q \leq s, t \leq d$, the associated compressed operator $\mathcal{A}_J^\varepsilon$ satisfies the estimate

$$\big|\big\langle (\mathcal{A}-\mathcal{A}_J^\varepsilon)Q_J u,\, Q_J v\big\rangle_\Gamma\big| \lesssim \varepsilon\, 2^{J(2q-s-t)} \|u\|_{H^s(\Gamma)} \|v\|_{H^t(\Gamma)} \tag{46}$$

holds uniformly in J.

Proof. By the definition of the operators $\mathcal{A}$, $\mathcal{A}_J^\varepsilon$, the biorthogonality and the representation formula

$$u = \sum_{\mathbf{j},\mathbf{k}} \langle \widetilde{\psi}_{\mathbf{j},\mathbf{k}}, u\rangle_\Gamma \psi_{\mathbf{j},\mathbf{k}},$$

together with all the block error estimations and the definition of $\sigma_{\mathbf{j},\mathbf{j}'}$ in (21), we obtain that

$$\begin{aligned}
\big\langle (\mathcal{A}-\mathcal{A}_J^\varepsilon)Q_J u,\, Q_J v\big\rangle_\Gamma &= \sum_{\substack{|\mathbf{j}|_\infty \leq J \\ |\mathbf{j}'|_\infty \leq J}} \sum_{\substack{\mathbf{k}\in\nabla_\mathbf{j} \\ \mathbf{k}'\in\nabla_{\mathbf{j}'}}} \langle \widetilde{\psi}_{\mathbf{j},\mathbf{k}}, u\rangle_\Gamma \big[\mathbf{A}-\mathbf{A}_J^\varepsilon\big]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')} \langle \widetilde{\psi}_{\mathbf{j}',\mathbf{k}'}, v\rangle_\Gamma \\
&= \sum_{\substack{|\mathbf{j}|_\infty \leq J \\ |\mathbf{j}'|_\infty \leq J}} \big[u_{\mathbf{j},\mathbf{k}}\big]_\mathbf{k}^\intercal \big[\mathbf{A}-\mathbf{A}_J^\varepsilon\big]_{\mathbf{j},\mathbf{j}'} \big[v_{\mathbf{j}',\mathbf{k}'}\big]_{\mathbf{k}'} \\
&\leq \sum_{\substack{|\mathbf{j}|_\infty \leq J \\ |\mathbf{j}'|_\infty \leq J}} \big\|\big[u_{\mathbf{j},\mathbf{k}}\big]_\mathbf{k}\big\|_2 \, \big\|\big[\mathbf{A}-\mathbf{A}_J^\varepsilon\big]_{\mathbf{j},\mathbf{j}'}\big\|_2 \, \big\|\big[v_{\mathbf{j}',\mathbf{k}'}\big]_{\mathbf{k}'}\big\|_2 \\
&\lesssim \varepsilon\, 2^{2J(q-d')} \sum_{n,n'=0}^{J} 2^{(d'-2\kappa)(n+n')} \left(\sum_{|\mathbf{j}|_\infty = n} 2^{\kappa|\mathbf{j}|_1} \big\|\big[u_{\mathbf{j},\mathbf{k}}\big]_\mathbf{k}\big\|_2\right) \\
&\qquad \cdot \left(\sum_{|\mathbf{j}'|_\infty = n'} 2^{\kappa|\mathbf{j}'|_1} \big\|\big[v_{\mathbf{j}',\mathbf{k}'}\big]_{\mathbf{k}'}\big\|_2\right).
\end{aligned}$$

For each the two sums, we may apply the inequality of Cauchy-Schwarz to obtain that

$$\sum_{|\mathbf{j}|_\infty = n} 2^{\kappa|\mathbf{j}|_1} \big\|\big[u_{\mathbf{j},\mathbf{k}}\big]_\mathbf{k}\big\|_2 \leq \left(\sum_{|\mathbf{j}|_\infty = n} 2^{2\kappa|\mathbf{j}|_1}\right)^{\frac12} \left(\sum_{|\mathbf{j}|_\infty = n} \big\|\big[u_{\mathbf{j},\mathbf{k}}\big]_\mathbf{k}\big\|_2^2\right)^{\frac12}.$$

Herein, the first sum can be estimated by

$$\sum_{|\mathbf{j}|_\infty=n} 2^{2\kappa|\mathbf{j}|_1} = 2^{4\kappa n} + 2\sum_{m=0}^{n-1} 2^{2\kappa(m+n)} \lesssim 2^{4\kappa n}.$$

Moreover, if $s \geq 0$, the second sum can be treated by the approximation property (15) of the spaces V_n, and we obtain that

$$\begin{aligned}\left(\sum_{|\mathbf{j}|_\infty=n} \left\|\left[u_{\mathbf{j},\mathbf{k}}\right]_{\mathbf{k}}\right\|_2^2\right)^{\frac{1}{2}} &\sim \|(Q_n - Q_{n-1})u\|_{L^2(\Gamma)}\\ &\leq \|Q_n u - u\|_{L^2(\Gamma)} + \|u - Q_{n-1}u\|_{L^2(\Gamma)}\\ &\lesssim 2^{-sn}\|u\|_{H^s(\Gamma)}.\end{aligned}$$

Whereas, if $s < 0$, we can use the Bernstein inequality (16) to get

$$\left(\sum_{|\mathbf{j}|_\infty=n} \left\|\left[u_{\mathbf{j},\mathbf{k}}\right]_{\mathbf{k}}\right\|_2^2\right)^{\frac{1}{2}} \sim \|(Q_n - Q_{n-1})u\|_{L^2(\Gamma)} \lesssim 2^{-sn}\|(Q_n - Q_{n-1})u\|_{H^s(\Gamma)} \lesssim 2^{-sn}\|u\|_{H^s(\Gamma)}.$$

After applying the same procedure to v, we finally arrive at

$$\begin{aligned}\left\langle(\mathcal{A} - \mathcal{A}_J^\varepsilon)\mathbb{P}_J u, \mathbb{P}_J v\right\rangle_\Gamma &\lesssim \varepsilon\, 2^{2J(q-d')} \sum_{n,n'=0}^{J} 2^{n(d'-s)} 2^{n'(d'-t)} \|u\|_{H^s(\Gamma)}\|v\|_{H^t(\Gamma)}\\ &\lesssim \varepsilon\, 2^{2J(q-d')}\|u\|_{H^s(\Gamma)}\|v\|_{H^t(\Gamma)} \sum_{n=0}^{J} 2^{n(d'-s)} \sum_{n'=0}^{J} 2^{n'(d'-t)}\\ &\lesssim \varepsilon\, 2^{J(2q-s-t)}\|u\|_{H^s(\Gamma)}\|v\|_{H^t(\Gamma)}\end{aligned}$$

since $q \leq s, t \leq d < d'$. □

Our next goal is to show that the compressed wavelet scheme converges to the original solution. First, similar to [8], we keep in mind that Theorem 8.1 implies that

$$\left|\left\langle(\mathcal{A} - \mathcal{A}_J^\varepsilon)u_J, v_J\right\rangle_\Gamma\right| \lesssim \varepsilon\, \|u_J\|_{H^q(\Gamma)}\|v_J\|_{H^q(\Gamma)}, \quad u_J, v_J \in V_J.$$

In view of the strong ellipticity (4), we can conclude that

$$\left\langle u, \left(\mathcal{A}_J^\varepsilon + (\mathcal{A}_J^\varepsilon)^\star\right)u\right\rangle_\Gamma \geq (c - 2\varepsilon)\|u\|_{H^q(\Gamma)}^2,$$

so the operator $\mathcal{A}_J^\varepsilon$ is strongly elliptic for ε sufficiently small, too. These two properties then imply that the operator $\mathcal{A}_J^\varepsilon$ is stable in the sense that

$$\left\|\mathcal{A}_J^\varepsilon u_J\right\|_{H^{-q}(\Gamma)} \sim \left\|u_J\right\|_{H^q(\Gamma)}.$$

With these two results at hand, we may deduce the following two theorems using the arguments of [8].

Theorem 8.2 (Convergence)**.** *Let ε be sufficiently small such that $\mathcal{A}_J^\varepsilon$ is strongly elliptic. Then, the solution of the compressed matrix equation*

$$u_J = \sum_{|\mathbf{j}|_\infty \leq J} \sum_{\mathbf{k}\in\nabla_\mathbf{j}} u_{\mathbf{j},\mathbf{k}}\psi_{\mathbf{j},\mathbf{k}},$$

where the coefficient vector $\mathbf{u}_J$ satisfies

$$\mathbf{A}_J^\varepsilon \mathbf{u}_J = \mathbf{g}_J, \quad \text{where } \left[\mathbf{g}_J\right]_{\mathbf{j},\mathbf{k}} = \left\langle g, \psi_{\mathbf{j},\mathbf{k}}\right\rangle_\Gamma,$$

converges to the solution u of (5) in $H^q(\Gamma)$ and the estimate

$$\|u - u_J\|_{H^q(\Gamma)} \lesssim 2^{J(q-d)}\|u\|_{H^d(\Gamma)}$$

holds.

Theorem 8.3 (Aubin-Nitsche)**.** *Let all the assumptions of Theorem 8.2 hold and moreover assume that* $\mathcal{A}^\star : H^{t+q}(\Gamma) \to H^{t-q}(\Gamma)$ *is an isomorphism for any* $0 \le t \le d-q$*. Then, we have the error estimate*

$$\|u - u_J\|_{H^{q-t}(\Gamma)} \lesssim 2^{J(q-d-t)} \|u\|_{H^d(\Gamma)}.$$

These two theorems can be shown by using only the consistency, the ellipticity, the stability, and the approximation property, cf. [8].

9. Numerical Computations

In this section, we present numerical experiments to validate the theoretical findings. We use piecewise constant wavelets with three vanishing moments and consider the single layer operator on the unit square. We compute first the full wavelet Galerkin matrix

$$\mathbf{A}_J = \big[\langle \psi_{\mathbf{j}',\mathbf{k}'}, \mathcal{A}\psi_{\mathbf{j},\mathbf{k}}\rangle_\Box\big]_{(\mathbf{j},\mathbf{k}),(\mathbf{j}',\mathbf{k}')}, \qquad |\mathbf{j}|_\infty, |\mathbf{j}'|_\infty \le J,\ \mathbf{k} \in \nabla_{\mathbf{j}},\ \mathbf{k}' \in \nabla_{\mathbf{j}'}.$$

with $\mathcal{A}$ being the single layer operator. Then, the obsolete entries are removed according to the compression scheme from Section 5. The number of nonzero entries are calculated and plotted (blue line) in on the left-hand side of Figure 6. To this end, we have chosen the parameters $\tilde{d} = 3$, $d' = 1.1$, and $\kappa = 10^{-3}$, while for the bandwidth parameter a the different values $a = 0.5, 1.0, 2.0$ have been considered. The system matrix for the anisotropic tensor product wavelet basis and its compressed counterpart with $4^7 = 16\,384$ rows and columns and $a = 1.0$ can be found in Figure 5.

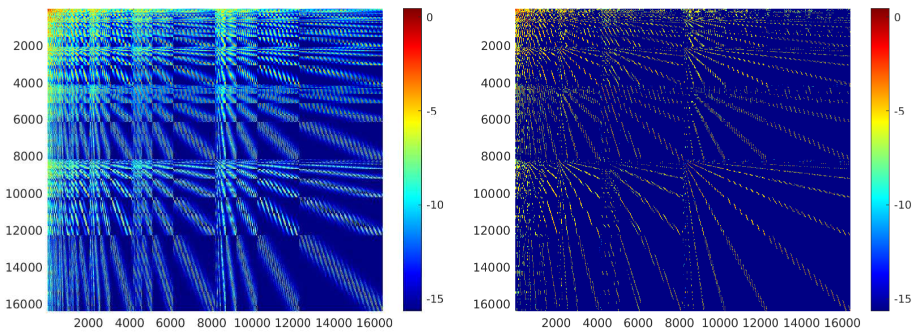

Figure 5. The wavelet Galerkin matrix (left) and its compressed version (right) for $J = 7$. We have used the parameters $a = 1.0$, $\tilde{d} = 3$, $d' = 1.1$, and $\kappa = 10^{-3}$. The colour indicates the absolute values of the matrix entries in a logarithmic scale.

From a theoretical point of view, the number of nontrivial entries in each block can be bounded by

$$N_{\mathbf{j},\mathbf{j}'} \lesssim 2^{|\mathbf{j}|_1 + |\mathbf{j}'|_1} \big(\mathcal{B}^2_{\mathbf{j},\mathbf{j}'} + \min\{2^{-\min\{j_1,j_2,j_1',j_2'\}}, \mathcal{E}_{\mathbf{j},\mathbf{j}'}\} + \mathcal{F}^{x_1}_{\mathbf{j},\mathbf{j}'}\mathcal{F}^{x_2}_{\mathbf{j},\mathbf{j}'}\big). \tag{47}$$

Moreover, in accordance with (40) and (41), an additional summand $2^{\min\{|\mathbf{j}|_1 + \mathbf{j}'|_1\}}$ is added if $\mathbf{j} \le \mathbf{j}'$ or $\mathbf{j}' \le \mathbf{j}$. The number of nontrivial entries in the whole compressed matrix can therefore be also bounded by

$$\mathcal{C} \lesssim \sum_{|\mathbf{j}|_\infty \le J} \sum_{|\mathbf{j}'|_\infty \le J} N_{\mathbf{j},\mathbf{j}'}. \tag{48}$$

This estimate is also found (red line) in the plot on the left-hand side of Figure 6.

On the right-hand side of Figure 6, we have computed the consistency errors arising from the compression scheme. Therein, we have generated random vectors $\mathbf{u}, \mathbf{v}$ and scaled them such that they correspond to coefficient vectors of functions $u, v \in H^1(\Gamma)$ with respect to the anisotropic tensor product wavelet basis. We have used 100 random samples to calculate the quantity $\left|\mathbf{v}^\intercal(\mathbf{A}_J - \mathbf{A}_J^c)\mathbf{u}\right|$ which is the discrete version of (46). We see that the calculations match the expected behaviour.

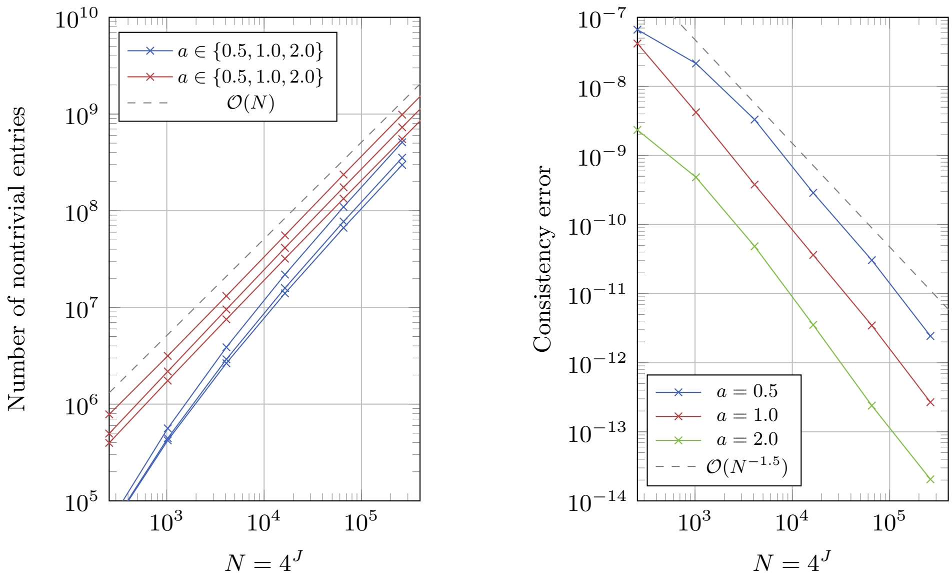

FIGURE 6. Left: Calculation (blue) and estimation (red) of the number of nonzero entries according to (47) and (48). Right: Consistency error according to (46).

10. CONCLUSION

We have developed a matrix compression scheme for the boundary element method using anisotropic tensor product wavelets. In the end, we get a quasi-sparse matrix containing only $\mathcal{O}(N)$ nontrivial entries whilst the approximate solution converges to the exact solution at the rate of the discretisation error. This applies for every integral operator of arbitrary order on the unit square. On a Lipschitz geometry, however, the order of the integral operator is bounded by $2q < 1$ for a conforming method since the underlying wavelet construction yields ansatz functions which are discontinuous across patch boundaries.

Likewise to [20, 21], our compression scheme may be generalized to an arbitrary spatial dimension on the unit cube. One would have to choose the compression parameters according to the location of the two tensor product wavelets with respect to each other. Then, one should combine the first compression in all directions, in which the two wavelets are in the far-field with a second compression for all directions, in which the wavelets are in the near-field. Especially, optimal compression estimates in sparse tensor product spaces are possible.

On the contrary, it is not known yet whether the anisotropic tensor product wavelet basis is $s^\star$-compressible, which was established for the isotropic wavelet basis in [27]. With the $s^\star$-compressibility at hand, it was shown in [9, 11] that adaptive wavelet compression schemes can approximate the solution at the rate of the best N-term approximation at a linear complexity.

Acknowledgement. This research has been supported by the Swiss National Science Foundation (SNSF) through the project "Adaptive Boundary Element Methods Using Anisotropic Wavelets" under grant agreement no. 200021_192041.

References

[1] Alex Barnett and Leslie Greengard. A new integral representation for quasi-periodic fields and its application to two-dimensional band structure calculations. *Journal of Computational Physics*, 229(19):6898–6914, 2010.

[2] Mario Bebendorf and Sergej Rjasanow. Adaptive low-rank approximation of collocation matrices. *Computing*, 70(1):1–24, 2003.

[3] Dietrich Braess. *Finite Elements. Theory, Fast Solvers, and Applications in Solid Mechanics.* Cambridge University Press, Cambridge, second edition, 2001.

[4] Susanne C. Brenner and L. Ridgway Scott. *The Mathematical Theory of Finite Element Methods.* Springer, Berlin, third edition, 2008.

[5] Paul Cazeaux and Olivier Zahm. A fast boundary element method for the solution of periodic many-inclusion problems via hierarchical matrix techniques. *ESAIM: Proceedings and Surveys*, 48:156–168, January 2015.

[6] David Colton and Rainer Kress. *Inverse Acoustic and Electromagnetic Scattering Theory.* Springer, Berlin-Heidelberg, second edition, 1998.

[7] Wolfgang Dahmen. Wavelet and multiscale methods for operator equations. *Acta Numerica*, 6:55–228, 1997.

[8] Wolfgang Dahmen, Helmut Harbrecht, and Reinhold Schneider. Compression techniques for boundary integral equations: Asymptotically optimal complexity estimates. *SIAM Journal on Numerical Analysis*, 43(6):2251–2271, 2006.

[9] Wolfgang Dahmen, Helmut Harbrecht, and Reinhold Schneider. Adaptive methods for boundary integral equations: Complexity and convergence estimates. *Mathematics of Computation*, 76(259):1243–1274, 2007.

[10] Ingrid Daubechies. *Ten Lectures on Wavelets.* CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1992.

[11] Tsogtgerel Gantumur and Rob Stevenson. Computation of singular integral operators in wavelet coordinates. *Computing*, 76:77–107, 2006.

[12] Leslie Greengard and Vladimir Rokhlin. A new version of the fast multipole method for the Laplace equation in three dimensions. *Acta Numerica*, 6:229–269, 1997.

[13] Michael Griebel, Peter Oswald, and Thomas Schiekofer. Sparse grids for boundary integral equations. *Numerische Mathematik*, 83(2):279–312, 1999.

[14] Helmut Harbrecht and Michael Peters. Comparison of fast boundary element methods on parametric surfaces. *Computer Methods in Applied Mechanics and Engineering*, 261–262:39–55, 2013.

[15] Helmut Harbrecht and Reinhold Schneider. Wavelet Galerkin schemes for boundary integral equations. implementation and quadrature. *SIAM Journal on Scientific Computing*, 27(4):1347–1370, 2006.

[16] Helmut Harbrecht and Reinhold Schneider. Rapid solution of boundary integral equations by wavelet galerkin schemes. In Ronald DeVore and Angela Kunoth, editors, *Multiscale, Nonlinear and Adaptive Approximation*, pages 249–294, Berlin-Heidelberg, 2009. Springer.

[17] Helmut Harbrecht and Manuela Utzinger. On adaptive wavelet boundary element methods. *Journal of Computational Mathematics*, 36(1):90–109, 2018.

[18] George Hsiao and Wolfgang L. Wendland. *Boundary Integral Equations*, volume 164 of *Applied Mathematical Sciences.* Springer, Berlin, 2008.

[19] Dalibor Lukáš, Günther Of, Jan Zapletal, and Jiří Bouchala. A boundary element method for homogenization of periodic structures. *Mathematical Methods in the Applied Sciences*, 43(3):1035–1052, 2020.

[20] Nils Reich. *Wavelet Compression of Anisotropic Integrodifferential Operators on Sparse Tensor Product Spaces.* PhD thesis, ETH Zürich, 2008. ETH Diss Nr. 17661.

[21] Nils Reich. Wavelet compression of integral operators on sparse tensor spaces. Construction, consistency and asymptotically optimal complexity. Research Report No. 2008-26, Seminar for Applied Mathematics, ETH Zurich, 2008.

[22] Sergej Rjasanow and Olaf Steinbach. *The Fast Solution of Boundary Integral Equations.* Springer, 2003.

[23] Stefan Sauter and Christoph Schwab. *Boundary Element Methods*, volume 39 of *Springer Series in Computational Mathematics.* Springer, Berlin-Heidelberg, 2011.

[24] Reinhold Schneider. *Multiskalen- und Wavelet-Matrixkompression: analysisbasierte Methoden zur effizienten Lösung großer vollbesetzter Gleichungssysteme.* B. G. Teubner, Stuttgart, 1998.

[25] Elias M. Stein. *Harmonic Analysis.* Princeton University Press, Princeton, NJ, 2002.

[26] Olaf Steinbach. *Numerical Approximation Methods for Elliptic Boundary Value Problems.* Springer, New York, NY, 2008.

[27] Rob Stevenson. On the compressibility of operators in wavelet coordinates. *SIAM Journal on Mathematical Analysis*, 35(5):1110–1132, 2004.

[28] Manuela Utzinger. *An Adaptive Wavelet Method for the Solution of Boundary Integral Equations in Three Dimensions.* PhD thesis, Universität Basel, 2016.

[29] Christoph Winter. Wavelet Galerkin schemes for multidimensional anisotropic integrodifferential operators. *SIAM Journal on Scientific Computing*, 32(3):1545–1566, 2010.

[30] Joseph Wloka. *Partial Differential Equations.* Cambridge University Press, Cambridge, 1987.

Helmut Harbrecht and Remo von Rickenbach, Departement Mathematik und Informatik, Universität Basel, Spiegelgasse 1, 4051 Basel

Email address: {helmut.harbrecht,remo.vonrickenbach}@unibas.ch