Adaptive Wavelet Algorithms
for solving operator equations

Tsogtgerel Gantumur

August 2006

Contents

List of Algorithms

Notations and acronyms

notation	meaning
SPD	symmetric and positive definite
CBS inequality	Cauchy-Bunyakowsky-Schwarz inequality
$\mathbb{N}$, $\mathbb{N}_0$	the natural numbers $1, 2, 3, \ldots$, and $\mathbb{N} \cup \{0\}$, resp.
$\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$	integers, reals, and complex numbers, respectively
$\mathbb{R}_{>0}$, $\mathbb{R}_{\geq 0}$	positive and nonnegative reals, respectively
Ω, $\partial\Omega$	bounded Lipschitz domain in $\mathbb{R}^n$, and its boundary
$L_p(\Omega)$, L_p	the space of functions on Ω for which $\int_\Omega \lvert f\rvert^p$ is finite
$W_p^s(\Omega)$, W_p^s	the Sobolev space with smoothness s measured in L_p
$H^s(\Omega)$, H^s	equal to $W_2^s(\Omega)$
$H_0^s(\Omega)$, H_0^s	the closure of $C_0^\infty(\Omega)$ in $H^s(\Omega)$
$B_q^s(L_p(\Omega))$, $B_q^s(L_p)$	Besov space with smoothness s measured in L_p and secondary index q
$\mathcal{H}$	a separable Hilbert space, typically L_2 or H_0^1
$\mathcal{H}'$	the dual of $\mathcal{H}$
$u, v, w, \ldots$	elements of $\mathcal{H}$
ℓ_2	the space $\ell_2(\nabla)$ with a countable index set ∇
P	the set of finitely supported sequences in ℓ_2, i.e., is equal to $\{\mathbf{v} \in \ell_2 : \#\operatorname{supp}\mathbf{v} < \infty\}$
$\langle \cdot, \cdot \rangle$	the duality product on $\mathcal{H} \times \mathcal{H}'$, or the standard inner product in ℓ_2
$\lVert \cdot \rVert$	the standard norm on ℓ_2, or the induced operator norm on $\ell_2 \to \ell_2$

notation	meaning
$\mathbf{u}, \mathbf{v}, \mathbf{w}$	elements (or vectors, sequences) of $\ell_2(\Lambda)$ with some countable index set $\Lambda \subseteq \nabla$
$\mathbf{v}_\lambda, [\mathbf{v}+\mathbf{w}]_\lambda, \mathbf{w}_\mu$	entries (or coefficients) in elements of $\ell_2(\Lambda)$, thus e.g. $\mathbf{v}_\lambda \in \mathbb{R}$
$\mathbf{v}_1, \mathbf{v}_k, \mathbf{w}_K$	different elements of ℓ_2, thus e.g. $\mathbf{v}_k \in \ell_2$
$\mathcal{B}_N(\mathbf{v})$	a best N-term approximation of $\mathbf{v} \in \ell_2$
$\mathcal{A}^s$	the set of sequences in ℓ_2 that can be approximated by best N-term approximations with the rate s; or in Chapter 6, the set of sequences in ℓ_2 that can be approximated by best *tree* N-term approximations with the rate s
$\tilde{\mathcal{A}}^s$	the set of sequences in ℓ_2 that can be approximated by best *graded tree* N-term approximations with the rate s
$\mathbf{A}, \mathbf{L}, \mathbf{M}$	bounded linear operators of type $\ell_2 \to \ell_2$
$\mathbf{A}$	a symmetric and positive definite matrix
$\kappa(\mathbf{M})$	condition number of $\mathbf{M}$, i.e., $\|\mathbf{M}\|\|\mathbf{M}^{-1}\|$ for an invertible $\mathbf{M}$
$\langle\!\langle \cdot, \cdot \rangle\!\rangle$	inner product defined by $\langle \mathbf{A}\cdot, \cdot \rangle$
$\vert\!\vert\!\vert \cdot \vert\!\vert\!\vert$	the norm defined by $\langle\!\langle \cdot, \cdot \rangle\!\rangle^{\frac{1}{2}}$, called the *energy norm*
$\mathbf{I}_\Lambda$	the trivial inclusion $\ell_2(\Lambda) \to \ell_2(\nabla)$, for $\Lambda \subset \nabla$
$\mathbf{P}_\Lambda$	equal to the adjoint $\mathbf{I}^* : \ell_2(\nabla) \to \ell_2(\Lambda)$, for $\Lambda \subset \nabla$
$f \lesssim g$	$f \leq C \cdot g$ with a constant $C > 0$ that may depend only on *fixed* constants under consideration
$f \gtrsim g$	$g \lesssim f$
$f \eqsim g$	$f \lesssim g$ and $g \lesssim f$
$\oslash$	end of example, definition, or long remark
$\blacksquare$	end of proof

Chapter 1

Introduction

1.1 Background

This thesis treats various aspects of adaptive wavelet algorithms for solving operator equations. For a separable Hilbert space H, a linear functional $f \in H'$, and a boundedly invertible linear operator $A : H \to H'$, we consider the problem of finding $u \in H$ satisfying

$$Au = f.$$

Typically A is given by a variational formulation of a boundary value problem or integral equation, and H is a Sobolev space formulated on some domain or manifold, possibly incorporating essential boundary conditions. Often we will assume that A is self-adjoint and H-elliptic. General operators can be treated, e.g., by forming normal equations, although in particular situations quantitatively more attractive alternatives exist.

In their pioneering works [17, 18], Cohen, Dahmen and DeVore introduced adaptive wavelet paradigms for solving the problem numerically. Utilizing a Riesz basis $\Psi = \{\psi_i \in H : i \in \mathbb{N}\}$ for H, the idea is to transform the original problem into a problem involving the coefficients of u with respect to the basis Ψ. Writing the collection of these coefficients of u as $\mathbf{u} \in \ell_2$, $\mathbf{u}$ has to satisfy

$$\mathbf{A}\mathbf{u} = \mathbf{f},$$

where $\mathbf{A} : \ell_2 \to \ell_2$ is an infinitely sized stiffness matrix with elements $\mathbf{A}_{ik} = [A\psi_k](\psi_i) \in \mathbb{R}$, and $\mathbf{f} \in \ell_2$ is an infinitely sized load vector with elements $\mathbf{f}_i = f(\psi_i) \in \mathbb{R}$. Under certain assumptions concerning the cost of evaluating the entries of the stiffness matrix, the methods from the aforementioned works of Cohen, Dahmen, and DeVore for solving this infinite matrix-vector problem were shown to be of optimal computational complexity. In this thesis, we will verify

those assumptions, extend the scope of problems for which the adaptive wavelet algorithms can be applied directly, and most importantly, develop and analyze modified adaptive algorithms with improved quantitative properties.

In order to solve the infinitely sized problem on a computer (within a given tolerance $\varepsilon > 0$), one should be able to approximate both $\mathbf{f}$ and $\mathbf{Av}$ for finitely supported $\mathbf{v}$. Let $P \subset \ell_2$ denote the set of finite sequences. Then one utilizes some maps $\mathcal{A} : \mathbb{R}_{>0} \times P \to P$ and $\mathcal{F} : \mathbb{R}_{>0} \to P$, both realized by some implementable computational procedures, such that for any $\epsilon > 0$ and for any $\mathbf{v} \in P$,

$$\|\mathcal{A}(\epsilon, \mathbf{v}) - \mathbf{Av}\| \leq \epsilon, \qquad \text{and} \qquad \|\mathcal{F}(\epsilon) - \mathbf{f}\| \leq \epsilon.$$

We know that the sequence $(\mathbf{u}^{(j)})_{j\geq 0}$ given by the Richardson iteration

$$\begin{cases} \mathbf{u}^{(0)} = 0, \\ \mathbf{u}^{(j+1)} = \mathbf{u}^{(j)} + \alpha(\mathbf{f} - \mathbf{Au}^{(j)}), \quad j \in \mathbb{N}, \end{cases}$$

converges to the solution $\mathbf{u}$ for $\alpha \in (0, \frac{2}{\|\mathbf{A}\|})$; however, this iteration is not computable since in general the retrieval of all coefficients of $\mathbf{f}$ and the application of $\mathbf{A}$ requires infinite storage and unlimited computing power. Therefore one has to perform this iteration only approximately, working with finitely supported vectors and matrices only. By using the procedures $\mathcal{A}$ and $\mathcal{F}$ one can design a convergent inexact Richardson iteration. Other Krylov subspace methods can be used as well, where the theory of inexact Krylov subspace methods comes into play.

In [17, 86], it is shown, assuming that the individual entries of the matrix $\mathbf{A}$ can be computed efficiently, how a reasonably fast procedure $\mathcal{A}$ can be realized, essentially by proving that the matrix $\mathbf{A}$ can be approximated well by sparse matrices. The latter is a result of the facts that the wavelets are *locally supported* and have the so-called *cancellation property*, and that the considered operators are *local* (in case of differential operators) or *pseudolocal* (in case of singular integral operators). Based on an inexact Richardson iteration, employing the fast procedure from the above papers, and assuming that on average, an individual entry of the matrix $\mathbf{A}$ can be computed at unit cost, in [18] an iterative adaptive algorithm was developed that has *optimal computational complexity*, meaning that the algorithm approximates the solution using, up to a constant factor, as few degrees of freedom as possible and the computational work stays proportional to the number of degrees of freedom. The average unit cost assumption will be confirmed in Chapters 7 and 8 of this thesis for both differential and singular integral operators.

As an alternative to using the Richardson iteration, in [17] another approach was suggested of using Galerkin approximation in combination with a residual

based *a posteriori* error indicator, leading to an algorithm of optimal computational complexity which is similar in spirit to an adaptive finite element method.

A crucial ingredient for proving the optimal complexity of both algorithms was the *coarsening* step that was applied after every fixed number of iterations. This step consists of removing small coefficients from the current iterand, ensuring that the support of the iterand does not grow too much in comparison to the convergence obtained by the algorithm. As we will show in Chapter 3, it turns out that coarsening is unnecessary for proving optimal computational complexity of algorithms of the type considered in [17]. Since with the new method no information is deleted that has been created by a sequence of computations, we expect that it is more efficient. Numerical experiments from e.g. Chapter 3 and [30] show that removing the coarsening improves the quantitative performance of the algorithm.

For the algorithms we have mentioned, the matrix $\mathbf{A}$ is assumed to be symmetric and positive definite, i.e., the operator A is self-adjoint and H-elliptic. In the general case one may replace the problem by the normal equation $\mathbf{A}^*\mathbf{A}\mathbf{u} = \mathbf{A}^*\mathbf{f}$. From a quantitative point of view, the normal equation is undesirable since the condition number is squared. In some special cases it can be avoided. For example, for saddle point problems one can use the Schur complement system, cf. [25]. For strongly elliptic operators, i.e., the operator A is a compact perturbation of a self-adjoint and H-elliptic operator, we will show in Chapter 5 that the algorithm from Chapter 3 can be applied directly with minor modifications, avoiding the normal equations.

Although the algorithms above described are proven to have asymptotically optimal computational complexity, there are some reasons to expect that the algorithms can be quantitatively improved. Let $\mathbf{w} := \mathcal{A}(\epsilon, \mathbf{v})$ for some $\epsilon > 0$ and finitely supported $\mathbf{v} \in P$. The individual wavelet ψ_i is characterized by its so-called *level* and its location in space. Then, with the commonly used map $\mathcal{A}$, in general, the difference between the highest levels of wavelets that are used in $\mathbf{w}$ and that are used in $\mathbf{v}$ grows as $\epsilon \to 0$, which leads to serious obstacles in practical implementations of the algorithm. If we simply force the level difference not to exceed some fixed number, then the numerical experiments show relatively good performance, see e.g. [5, 54]. In Chapter 6, we will analyze similarly modified algorithms.

1.2 Thesis overview

The thesis is outlined as follows:

Chapter 2 (*Basic principles*) contains a short introduction to the theory of adaptive wavelet algorithms. We start with recalling essential properties of

wavelet bases, and briefly present basic results on best N-term approximation. Then we describe how an optimally convergent algorithm can be constructed using any linearly convergent iteration in the energy space. We include proofs of the most fundamental results, along with references to relevant literature.

In Chapter 3 (*Adaptive Galerkin methods*), an adaptive wavelet method for solving linear operator equations is constructed that is a modification of the method from [17], in the sense that there is no recurrent coarsening of the iterands. In spite of this, it will be shown that the method has optimal computational complexity. Numerical results for a simple model problem indicate that the new method is more efficient than the existing method.

In Chapter 4 (*Using polynomial preconditioners*), we investigate the possibility of using polynomial preconditioners in the context of adaptive wavelet methods. We propose a version of a preconditioned adaptive wavelet algorithm and show that it has optimal computational complexity.

In Chapter 5 (*Adaptive algorithm for nonsymmetric and indefinite elliptic problems*), we modify the adaptive wavelet algorithm from Chapter 3 so that it applies directly, i.e., without forming the normal equation, not only to self-adjoint elliptic operators but also to operators of the form $L = A + B$, where A is self-adjoint elliptic and B is compact, assuming that the resulting operator equation is well-posed. We show that the algorithm has optimal computational complexity.

Aiming at a further improvement of quantitative properties, in Chapter 6 (*Adaptive algorithm with truncated residuals*), a class of adaptive wavelet algorithms for solving elliptic operator equations is introduced, and is proven to have optimal complexity assuming a certain property of the stiffness matrix. This assumption is confirmed for elliptic differential operators.

In Chapter 7 (*Computability of differential operators*), restricting us to differential operators, we develop a numerical integration scheme that computes the entries of the stiffness matrix at the expense of an error that is consistent with the approximation error, whereas in each column the average computational cost per entry is $\mathcal{O}(1)$. As a consequence, we can conclude that the "fully discrete" adaptive wavelet algorithm has optimal computational complexity.

In Chapter 8 (*Computability of singular integral operators*), we prove an analogous result for singular integral operators, by carefully distributing computational costs over the matrix entries in combination with choosing efficient quadrature schemes.

Chapter 9 (*Conclusion*) closes the thesis with a summary and discussion of the presented research topics, as well as with some suggestions for future research.

To help readers who prefer to read the chapters in an order different than linear, Figure 1.1 on the facing page illustrates the logical dependencies between

chapters.

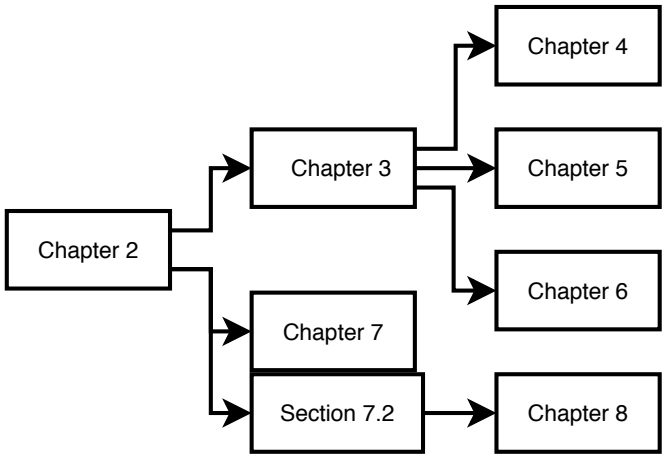

Figure 1.1: *Chapter dependencies*

Chapters 3, 5, 7 and 8 have appeared as separate papers. For this thesis, they have been edited to some extent, varying from small editorial changes to enlargement by extra sections. Some notations have been changed to ensure uniformity. Chapter 3 is based on [46]:

> Ts. Gantumur, H. Harbrecht, and R. P. Stevenson, *An optimal adaptive wavelet method without coarsening of the iterands*, Technical Report 1325, Utrecht University, The Netherlands, March 2005. To appear in Math. Comp.

Chapter 5 is [45]:

> Ts. Gantumur, *An optimal adaptive wavelet method for nonsymmetric and indefinite elliptic problems*, Technical Report 1343, Utrecht University, The Netherlands, January 2006. Submitted.

Chapter 7 is [47]:

> Ts. Gantumur and R. P. Stevenson, *Computation of differential operators in wavelet coordinates*, Math. Comp., 75 (2006), pp. 697–709.

Chapter 8 appeared as [48]:

> Ts. Gantumur and R. P. Stevenson, *Computation of singular integral operators in wavelet coordinates*, Computing, 76 (2006), pp. 77–107.

1.3 Algorithms

Algorithms in this thesis are numbered within sections, and placed between two horizontal lines, preceded by the caption of the algorithm. Some algorithms have a name, which is placed, except for a few instances, at the end of the caption. The *name* of an algorithm ends with the list of *input variables* placed between square brackets, followed by the list of *output variables* separated from the input list by an arrow. For example, $\mathbf{XY}[a, b] \to c$ and $\mathbf{XY}[a, b] \to [c, d]$ are names of different algorithms. In any chapter, each algorithm has a unique name. A few algorithms in different chapters have common names, but it will be clear from the context which algorithm is in focus. At the beginning of an algorithm, the conditions that should be satisfied for the input variables are stated after the keyword **Input**. For algorithms that do not have a name, the input variables are also introduced here. Similarly, conditions that are satisfied for the output variables are stated after the keyword **Output**. After the keyword **Parameter** we declare fixed constants or input parameters that are changed infrequently. In order not to clutter algorithm names too much, these input parameters are not listed within the algorithm name. *Abstract algorithms* are defined only by their key properties, which should be satisfied for any concrete realization of the algorithm. Sometimes we call abstract algorithms *algorithm templates*.

1.4 Notational conventions

While many notations are summarized in the table on page vii, we would like to highlight some specific ones that appear frequently throughout the thesis. In any case, their definitions appear at the first place where they are introduced.

In this thesis, we will encounter function spaces $L_p(\Omega)$, $W_p^s(\Omega)$, etc., with Ω being a bounded Lipschitz domain. Elements of those spaces are indicated by lowercase letters (e.g., u). Capital letters (e.g., S, L) are used to denote subspaces, spaces, or operators.

A large portion of the thesis concerns sequence spaces, such as $\ell_p(\nabla)$ with a countable index set ∇. We use boldface lowercase letters (e.g., $\mathbf{u}$) for elements of a sequence space. To indicate an individual entry in a sequence, Greek subscripts are used, and when a sequence of elements of a sequence space is considered, Roman subscripts are used. For example, if $\mathbf{v} \in \ell_2(\nabla)$ and $\lambda \in \nabla$, then $\mathbf{v}_\lambda \in \mathbb{R}$ is an entry in the sequence $\mathbf{v}$. In contrast, $(\mathbf{v}_k)_{k\in\mathbb{N}}$ can be a sequence of elements of ℓ_2 and so $\mathbf{v}_k \in \ell_2$ for $k \in \mathbb{N}$. Operators on sequence spaces are denoted by boldface capital letters, as in $\mathbf{L} : \ell_2 \to \ell_2$. We use $\|\cdot\|$ to denote both $\|\cdot\|_{\ell_2}$ and $\|\cdot\|_{\ell_2\to\ell_2}$. For an invertible $\mathbf{M} : \ell_2 \to \ell_2$, its condition number is defined by $\kappa(\mathbf{M}) = \|\mathbf{M}\|\|\mathbf{M}^{-1}\|$.

In order to avoid the repeated use of generic but unspecified constants, by $f \lesssim g$ we mean that $f \leq C \cdot g$ with a constant $C > 0$ that may depend only on *fixed* constants under consideration. For example, $|n \sin x| \lesssim 1$ is true uniformly in $x \in \mathbb{R}$ for any *fixed* $n \in \mathbb{N}$. Obviously, $f \gtrsim g$ is defined as $g \lesssim f$, and $f \eqsim g$ as $f \lesssim g$ and $g \lesssim f$.

Chapter 2

Basic principles

2.1 Introduction

In this chapter we will take a short tour of the field of adaptive wavelet algorithms. We introduce and explain various concepts and terms that will be referred to frequently in this thesis.

We begin with recalling essential properties of wavelet bases, and briefly present basic results on best N-term approximation. Using Richardson's iteration as an example, we will describe how an optimally convergent algorithm can be constructed using linearly convergent iterations in the energy space. We then go into the fundamental building blocks of optimally convergent adaptive wavelet algorithms, such as the fast application of operators and the coarsening routine.

We include proofs of the most crucial results, along with references to relevant literature.

2.2 Wavelet bases

A wavelet basis is a basis with certain properties, and one or more of these properties can be emphasized depending on the particular application. In this section, we recall some relevant properties of wavelet bases, for simplicity considering the case of wavelet bases for Sobolev spaces on bounded domains. Although many of the results in this thesis hold in more general and hence abstract settings, we will occasionally return to the setting from this section to discuss how those general ideas could be applied in a concrete setting. On the other hand, we will explicitly state it if we need specific additional properties of wavelet bases. Let $\mathcal{H} := H^t(\Omega)$ be the Sobolev space with some smoothness index $t \in \mathbb{R}$, defined on a bounded Lipschitz domain $\Omega \subset \mathbb{R}^n$, and with ∇ being some countable index

set, let $\Psi = \{\psi_\lambda : \lambda \in \nabla\}$ be a wavelet basis for $\mathcal{H}$.

Riesz basis property

The first important property is that Ψ is a *Riesz basis* of $\mathcal{H}$. Recall that a basis Ψ is Riesz if and only if

$$\|\mathbf{v}\| \eqsim \|\mathbf{v}^T\Psi\|_{\mathcal{H}} \qquad \mathbf{v} \in \ell_2(\nabla), \tag{2.2.1}$$

where we used the shorthand notation $\mathbf{v}^T\Psi := \sum_{\lambda\in\nabla} \mathbf{v}_\lambda\psi_\lambda$. Here $\|\cdot\|$ denotes the standard norm on $\ell_2 := \ell_2(\nabla)$. With $\langle\cdot,\cdot\rangle$ denoting the duality product on $\mathcal{H}\times\mathcal{H}'$, we define the *analysis* and *synthesis* operators by

$$F : \mathcal{H}' \to \ell_2 : g \mapsto \langle g, \Psi\rangle \qquad \text{and} \qquad F' : \ell_2 \to \mathcal{H} : \mathbf{v} \mapsto \mathbf{v}^T\Psi, \tag{2.2.2}$$

respectively, where with $\langle g, \Psi\rangle$ we mean the sequence $(\langle g, \psi_\lambda\rangle)_\lambda$. The Riesz basis property of Ψ ensures that both F and F' are continuous bijections. The collection $\tilde{\Psi} := (F'F)^{-1}\Psi$ is a Riesz basis for $\mathcal{H}'$, called the *dual basis* of Ψ.

Direct and inverse estimates

Another property is that there exists a sequence of subsets $\nabla_0 \subset \nabla_1 \subset \ldots \subset \nabla$ such that with some $d > \gamma > \max\{0, t\}$, the subspaces

$$S_j := \operatorname{span}\{\psi_\lambda : \lambda \in \nabla_j\} \qquad (j \in \mathbb{N}_0),$$

satisfy the *Jackson* (or *direct*) estimate for $r < \gamma$ and $s \in [r, d]$,

$$\inf_{v_j\in S_j} \|v - v_j\|_{H^r} \lesssim 2^{-j(s-r)}\|v\|_{H^s} \qquad (v \in H^s), \tag{2.2.3}$$

as well as the *Bernstein* (or *inverse*) estimate for $r \le s < \gamma$,

$$\|v_j\|_{H^s} \lesssim 2^{j(s-r)}\|v\|_{H^r} \qquad (v_j \in S_j). \tag{2.2.4}$$

Furthermore, the dual sequence $(\tilde{S}_j)_{j\ge0}$ defined via the dual wavelets $\tilde{\Psi}$ and the sequence $(\nabla_j)_{j\ge0}$ also satisfies the analogous estimates with constants $\tilde{d} > \tilde{\gamma} > \max\{0, -t\}$. In particular, the Bernstein estimates give information about the smoothness of the wavelets or their duals, namely, we have $\Psi \subset H^s$ for any $s < \gamma$ and $\tilde{\Psi} \subset H^s$ for any $s < \tilde{\gamma}$.

The Jackson estimate is typically valid when S_j both contains all polynomials of degree less than d, and is spanned by compactly supported functions such that the diameter of the supports is uniformly proportional to 2^{-j}. Likewise the Bernstein estimate is known to hold with $\gamma = r + \frac{3}{2}$ when S_j is spanned by piecewise smooth, globally C^r-functions for some $r \in \{-1, 0, 1, \ldots\}$, where $r = -1$ means that they satisfy no global continuity condition.

Locality

Another important characteristic of wavelets is that they are *local* in the sense that for $\lambda \in \nabla$ and $x \in \Omega$, $j \in \mathbb{N}_0$,

$$\operatorname{diam}(\operatorname{supp}\psi_\lambda) \lesssim 2^{-|\lambda|} \quad \text{and} \quad \#\{|\lambda| = j : B(x, 2^{-j}) \cap \operatorname{supp}\psi_\lambda \neq \varnothing\} \lesssim 1,$$

where the *level number* $|\lambda|$ for $\lambda \in \nabla$ is defined by $|\lambda| = j$ if $\lambda \in \nabla_j \setminus \nabla_{j-1}$ with $\nabla_{-1} := \varnothing$, and $B(x,r) \subset \mathbb{R}^n$ is the ball with radius $r > 0$ centered at $x \in \mathbb{R}^n$. For $j \in \mathbb{N}_0$, the domain Ω can be covered by an order of 2^{jn} balls with radius 2^{-j}, thus the number of wavelets on level j is bounded by some constant multiple of 2^{jn}.

We remark that typically the locality of the dual wavelets is not necessary for wavelet methods for solving operator equations.

Cancellation property

By using that $\langle \tilde{\psi}_\mu, \psi_\lambda \rangle = \delta_{\mu,\lambda}$, with $\delta_{\mu,\lambda}$ the Kronecker delta, we have for $\lambda \in \nabla \setminus \nabla_0$, $g \in H^s(\Omega)$, $g_\lambda \in \tilde{S}_{|\lambda|-1}$,

$$\langle g, \psi_\lambda \rangle = \langle g - g_\lambda, \psi_\lambda \rangle \leq \|g - g_\lambda\|_{H^{-t}} \|\psi_\lambda\|_{H^t},$$

and from the Jackson estimate for the dual sequence $(\tilde{S}_j)_{j\geq 0}$, we infer

$$\langle g, \psi_\lambda \rangle \leq \inf_{g_\lambda \in \tilde{S}_{|\lambda|-1}} \|g - g_\lambda\|_{H^{-t}} \lesssim 2^{-|\lambda|(s+t)} \|g\|_{H^s} \qquad (-t \leq s \leq \tilde{d}).$$

This is an instance of the so-called *cancellation property* of order $\tilde{d}$.

Analogously to the above lines, for $w_j \in W_j := \operatorname{span}\{\psi_\lambda : |\lambda| = j\}$ and $g \in H^{-s}$ with $-r \leq -s \leq \tilde{d}$ for some $-r < \tilde{\gamma}$, we have

$$\langle g, w_j \rangle \leq \inf_{g_j \in \tilde{S}_{j-1}} \|g - g_j\|_{H^{-r}} \|w_j\|_{H^r} \lesssim 2^{j(s-r)} \|g\|_{H^{-s}} \|w_j\|_{H^r},$$

and so, for $r > -\tilde{\gamma}$ and $s \in [-\tilde{d}, r]$,

$$\|w_j\|_{H^s} \lesssim 2^{j(s-r)} \|w_j\|_{H^r} \qquad (w_j \in W_j). \tag{2.2.5}$$

Note that since $W_j \subset S_j$, the above estimate is valid also for $s \in [r, \gamma)$ by the Bernstein estimate (2.2.4) on the preceding page.

Characterization of Besov spaces

Since the next property of wavelets will involve Besov spaces, before stating that property we recall some definitions and facts related to Besov spaces.

For $p \in (0,\infty]$, we introduce the m-th order L_p-modulus of smoothness

$$\omega_m(v,t)_{L_p} := \sup_{|h|\le t} \|\Delta_h^m v\|_{L_p(\Omega_{h,m})},$$

where $\Omega_{h,m} := \{x \in \Omega : x + jh \in \Omega,\ j = 0,\dots m\}$ and Δ_h^m is the m-th order forward difference operator defined recursively by $[\Delta_h^1 v](x) = v(x+h) - v(x)$ and $\Delta_h^m v = \Delta_h^1(\Delta_h^{m-1})v$. Then, for $p, q \in (0,\infty]$ and $s \ge 0$, with $m > s$ being an integer, the Besov space $B_q^s(L_p)$ consists of those $v \in L_p$ for which

$$|v|_{B_q^s(L_p)} := \|(2^{js}\omega_m(v,2^{-j})_{L_p})_{j\ge 0}\|_{\ell_q}$$

is finite. The mapping $\|\cdot\|_{B_q^s(L_p)} := \|\cdot\|_{L_p} + |\cdot|_{B_q^s(L_p)}$ defines a norm when $p, q \ge 1$ and only a quasi-norm otherwise.

We now recall a number of embedding relations between Besov spaces with different indices. Simple embeddings are that $B_{q_1}^s(L_p) \subset B_{q_2}^s(L_p)$ for $q_1 < q_2$, and that $B_q^s(L_{p_1}) \supset B_q^s(L_{p_2})$ for $p_1 < p_2$. We also have $B_{p_1}^s(L_{p_1}) \supset B_{p_2}^s(L_{p_2})$ for $p_1 < p_2$, and $B_{q_1}^{s_1}(L_p) \supset B_{q_2}^{s_2}(L_p)$ for $s_1 < s_2$, regardless of the secondary indices q_1 and q_2. Not so obvious is that

$$B_q^{s_1}(L_{p_1}) \subset B_q^{s_2}(L_{p_2}) \qquad \text{for } s_1 - s_2 = n(\tfrac{1}{p_1} - \tfrac{1}{p_2}) > 0.$$

In particular, combining some of the above relations we have $B_{p_1}^{s_1}(L_{p_1}) \subset B_{p_2}^{s_2}(L_{p_2})$ for $s_1 - s_2 \ge n(\frac{1}{p_1} - \frac{1}{p_2}) > 0$, cf. [16].

It is worth noting that besides the aforementioned definition, there are a number of other natural ways to define Besov spaces, which definitions are all equivalent when $s/n > \max\{1/p - 1, 0\}$, cf. [16]. Besov spaces with negative smoothness index s are defined by duality: for $s < 0$, $B_q^s(L_p) := [B_{q'}^{-s}(L_{p'})]'$ with $1/q + 1/q' = 1$ and $1/p + 1/p' = 1$, so necessarily $p, q \ge 1$.

It is well known that at least when Ω is a bounded Lipschitz domain, one has $B_2^s(L_2) = H^s$ for $s \in \mathbb{R}$ and $B_p^s(L_p) = W_p^s$ for $s > 0$, $s \notin \mathbb{N}$, where $H^s = W_2^s$, and W_p^s denotes the Sobolev space of smoothness s measured in $L_p(\Omega)$.

The norm equivalence (2.2.1) provides a simple criterion to check whether a function is in $\mathcal{H}$ by means of its wavelet coefficients. Similarly, other function spaces also can be characterized by wavelet coefficients of functions. We shall briefly describe such a characterization for Besov spaces. It is known that for any $\mathbf{v} = (\mathbf{v}_\lambda)_{\lambda\in\nabla}$ such that $\mathbf{v}^T\Psi \in B_q^s(L_p)$,

$$\left\| \left(2^{j(s-t+\frac{n}{2}-\frac{n}{p})} \|(\mathbf{v}_\lambda)_{|\lambda|=j}\|_{\ell_p} \right)_{j\ge 0} \right\|_{\ell_q} \eqsim \|\mathbf{v}^T\Psi\|_{B_q^s(L_p)}, \tag{2.2.6}$$

is valid for $p > 0$ and $\max\{0, n(1/p-1)\} < s < \min\{d, \gamma(p)\}$, with

$$\gamma(p) := \sup\{\sigma : \Psi \subset B_{q_0}^\sigma(L_p) \text{ for some } q_0\},$$

at least when $\Psi, \tilde{\Psi} \subset L_\infty$. The equivalence (2.2.6) is also valid for $p \geq 1$ and $-\min\{\tilde{d}, \tilde{\gamma}(p)\} < s < 0$, with $\tilde{\gamma}(p) := \sup\{\sigma : \tilde{\Psi} \subset B^\sigma_{q_0}(L_{1-1/p}) \text{ for some } q_0\}$. It is perhaps most convenient to describe the above conditions as a region in the $(\frac{1}{p}, s)$-plane, see Figure 2.1. Note also that depending on the particular situation, this region may have some more constraints, e.g., when boundary conditions are incorporated into the space. For proofs of (2.2.6) in various circumstances we refer to [16, 29].

An interesting special case of (2.2.6) occurs when $s - t = n(\frac{1}{p} - \frac{1}{2})$ and $p = q$, namely

$$\|\mathbf{v}^T \Psi\|_{B^s_p(L_p)} \eqsim \|\mathbf{v}\|_{\ell_p}. \tag{2.2.7}$$

As noted earlier, the line $s - t = n(\frac{1}{p} - \frac{1}{2})$ is the demarcation line of the embedding $B^s_p(L_p) \subset B^t_2(L_2) \equiv H^t$.

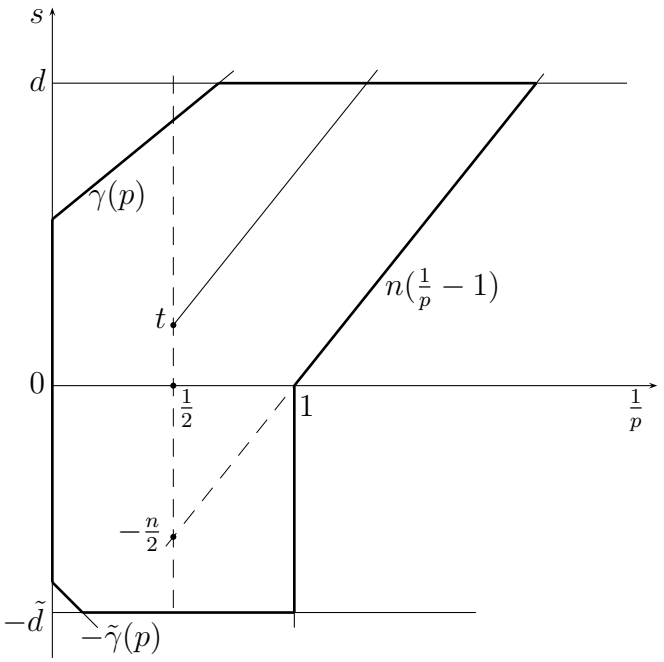

Figure 2.1: *In this so-called DeVore diagram ([39]), the point* $(\frac{1}{p}, s)$ *represents the whole range of Besov spaces* $B^s_q(L_p)$, $0 < q \leq \infty$. *Then the concave polygon bordered by the thick lines is the region for which the norm equivalence* (2.2.6) *is valid. The Besov spaces satisfying the norm equivalence* (2.2.7) *are on the line starting from the point* $(\frac{1}{2}, t)$.

Finally, we would like to note that one side of the estimate (2.2.6) is generally valid for a wider range of parameters p and s. To be specific, for $p > 0$ and

$\max\{0, n(1/p-1)\} < s < d$, we have

$$\sup_{j\geq 0}\left(2^{j(s-t+\frac{n}{2}-\frac{n}{p})}\|(\mathbf{v}_\lambda)_{|\lambda|=j}\|_{\ell_p}\right) \lesssim \|\mathbf{v}^T\Psi\|_{B^s_p(L_p)}, \tag{2.2.8}$$

at least when $\Psi, \tilde{\Psi} \subset L_\infty$, cf. [16].

2.3 Best N-term approximations

In order to assess the quality of approximations generated by adaptive algorithms that we will consider in the sequel, we introduce the following benchmark. With ∇ a countable index set, let $\ell_2 := \ell_2(\nabla)$. For $N \in \mathbb{N}$, we collect all the elements of ℓ_2 whose support size is at most N in

$$X_N := \{\mathbf{v} \in \ell_2 : \#\operatorname{supp}\mathbf{v} \leq N\}, \tag{2.3.1}$$

and define $X_0 := \{0\}$. We will consider approximations to elements of ℓ_2 from the subsets X_N. The subset X_N is not a linear space, meaning that it concerns nonlinear approximation. For $\mathbf{v} \in \ell_2$ and $N \in \mathbb{N}_0$, we define the error of the best approximation of $\mathbf{v}$ from X_N by

$$E_N(\mathbf{v}) := \operatorname{dist}(\mathbf{v}, X_N) = \inf_{\mathbf{v}_N \in X_N}\|\mathbf{v}-\mathbf{v}_N\|. \tag{2.3.2}$$

Any element $\mathbf{v}_N \in X_N$ that realizes this error is called a *best N-term approximation* of $\mathbf{v}$. With $\mathbf{P}_\Lambda : \ell_2 \to \ell_2(\Lambda)$ being the ℓ_2-orthogonal projector onto $\ell_2(\Lambda)$, a best N-term approximation of $\mathbf{v} \in \ell_2$ is equal to $\mathbf{P}_\Lambda\mathbf{v}$ for some set $\Lambda \subset \nabla$ with $\#\Lambda \leq N$, on which $|\mathbf{v}_\lambda|$ takes its largest N values. Note that $\mathbf{P}_\Lambda\mathbf{v}$ is obtained by simply discarding the coefficients $\mathbf{v}_\lambda$ of $\mathbf{v}$ with $\lambda \notin \Lambda$. The set Λ is not necessarily unique. For $N \in \mathbb{N}_0$, we denote an arbitrary best N-term approximation of $\mathbf{v} \in \ell_2$ by $\mathcal{B}_N(\mathbf{v})$ or more briefly, $\mathbf{v}_N$ if there is no risk of confusion. Any result in the thesis shall not depend on the arbitrary choice between best N-term approximations.

For $s \geq 0$, we define the approximation space $\mathcal{A}^s \subset \ell_2$ by

$$\mathcal{A}^s := \{\mathbf{v} \in \ell_2 : |\mathbf{v}|_{\mathcal{A}^s} := \|\mathbf{v}\| + \sup_{N\in\mathbb{N}} N^s E_N(\mathbf{v}) < \infty\}. \tag{2.3.3}$$

Clearly, it is the class of ℓ_2-sequences whose best N-term approximation decays like N^{-s}. It is obvious that $\mathcal{A}^s \subset \mathcal{A}^r$ for $s > r$.

Lemma 2.3.1. *For $s \geq 0$ and for $\mathbf{v}, \mathbf{w} \in \mathcal{A}^s$ a generalized triangle inequality holds,*

$$|\mathbf{v}+\mathbf{w}|_{\mathcal{A}^s} \leq \max\{2^s, 2^{2s-1}\}\big(|\mathbf{v}|_{\mathcal{A}^s} + |\mathbf{w}|_{\mathcal{A}^s}\big),$$

meaning that $|\cdot|_{\mathcal{A}^s}$ is a quasi-norm.

The Aoki-Rolewicz theorem (cf. [4, 68]) states the existence of a quasi-norm $|\cdot|^*_{\mathcal{A}^s} \eqsim |\cdot|^\mu_{\mathcal{A}^s}$ with $\mu = \min\{\frac{1}{s+1}, \frac{1}{2s}\}$, satisfying the standard triangle inequality $|\mathbf{v}+\mathbf{w}|^*_{\mathcal{A}^s} \le |\mathbf{v}|^*_{\mathcal{A}^s} + |\mathbf{w}|^*_{\mathcal{A}^s}$ for $\mathbf{v}, \mathbf{w} \in \mathcal{A}^s$. Moreover, $\mathcal{A}^s$ is complete with respect to the metric defined by $d(\mathbf{v}, \mathbf{w}) = |\mathbf{v}-\mathbf{w}|^*_{\mathcal{A}^s}$, i.e., $\mathcal{A}^s$ is a quasi-Banach space.

Proof. Since $X_N + X_N \subset X_{2N}$ for $N \in \mathbb{N}$, we have

$$E_{2N}(\mathbf{v}+\mathbf{w}) \le \|\mathbf{v}+\mathbf{w}-\mathcal{B}_N(\mathbf{v})-\mathcal{B}_N(\mathbf{w})\| \le E_N(\mathbf{v}) + E_N(\mathbf{w}).$$

Moreover, we have $E_{2N+1}(\cdot) \le E_{2N}(\cdot)$, and $E_1(\cdot) \le \|\cdot\|$, and taking into account that $(2N+1)^s \le \max\{2^s N^s + 1, 2^{2s-1}N^s + 2^{s-1}\}$, we get the generalized triangle inequality.

We remark that the functional $|\cdot|_{\mathcal{A}^s}$ is homogeneous: $|\nu\cdot|_{\mathcal{A}^s} = |\nu||\cdot|_{\mathcal{A}^s}$, $\nu \in \mathbb{R}$, but it is not guaranteed to satisfy the standard triangle inequality, while for $|\cdot|^*_{\mathcal{A}^s}$ the situation is the other way around. Let $(\mathbf{v}_k)_{k\in\mathbb{N}}$ be a Cauchy sequence in $\mathcal{A}^s$. Then obviously it has a limit $\mathbf{v} \in \ell_2$, and with a subsequence $(\mathbf{v}_{k_N})_{N\in\mathbb{N}}$ such that $\|\mathbf{v}-\mathbf{v}_{k_N}\| \le N^{-s}$, we have

$$\begin{aligned} E_N(\mathbf{v}) &\le \|\mathbf{v}-\mathcal{B}_N(\mathbf{v}_{k_N})\| \le \|\mathbf{v}-\mathbf{v}_{k_N}\| + \|\mathbf{v}_{k_N}-\mathcal{B}_N(\mathbf{v}_{k_N})\| \\ &\le N^{-s} + N^{-s}|\mathbf{v}_{k_N}|_{\mathcal{A}^s}. \end{aligned}$$

From the triangle inequality for $|\cdot|^*_{\mathcal{A}^s}$ we have $||\mathbf{w}|^*_{\mathcal{A}^s} - |\mathbf{z}|^*_{\mathcal{A}^s}| \le |\mathbf{w}-\mathbf{z}|^*_{\mathcal{A}^s}$ for $\mathbf{w}, \mathbf{z} \in \mathcal{A}^s$, thus $(|\mathbf{v}_k|^*_{\mathcal{A}^s})_{k\in\mathbb{N}}$ is a Cauchy sequence. This implies the existence of a constant $C > 0$ such that $|\mathbf{v}_k|^\mu_{\mathcal{A}^s} \lesssim |\mathbf{v}_k|^*_{\mathcal{A}^s} \le C$ for $k \in \mathbb{N}$, and so we conclude that $E_N(\mathbf{v}) \lesssim N^{-s}$, or equivalently, $\mathbf{v} \in \mathcal{A}^s$. ∎

Now we consider a relation between $\mathcal{A}^s$ and the classical sequence spaces ℓ_p. To this end, for $p \in (0,2)$, we introduce the *weak ℓ_p* spaces by

$$\ell^*_p := \{\mathbf{v} \in \ell_2 : \|\mathbf{v}\|_{\ell^*_p} := \sup_{j\in\mathbb{N}} j^{1/p}|\gamma_j(\mathbf{v})| < \infty\},$$

where $(\gamma_j(\mathbf{v}))_{j\in\mathbb{N}}$ denotes the *non-increasing rearrangement* of $\mathbf{v}$ in modulus.

Lemma 2.3.2. *Let $s > 0$ and let p be defined by $\frac{1}{p} = s + \frac{1}{2}$. Then we have*

$$\mathcal{A}^s = \ell^*_p, \qquad \text{and} \qquad \|\cdot\|_{\mathcal{A}^s} \eqsim \|\cdot\|_{\ell^*_p},$$

with the equivalency constants depending on s only as $s \to 0$ or $s \to \infty$.

Proof. We include a proof for the reader's convenience. By definition, $\mathbf{v} \in \ell_p^*$ if and only if for some constant $c > 0$, $|\gamma_j(\mathbf{v})| \leq c \cdot j^{-1/p}$, $j \in \mathbb{N}$, and the smallest such c is equal to $\|\mathbf{v}\|_{\ell_p^*}$. For $\mathbf{v} \in \ell_p^*$ and $N \in \mathbb{N}$, we have

$$\begin{aligned}(E_N(\mathbf{v}))^2 = \|\mathbf{v} - \mathcal{B}_N(\mathbf{v})\|^2 = \sum_{j>N} |\gamma_j(\mathbf{v})|^2 \leq \|\mathbf{v}\|_{\ell_p^*}^2 \sum_{j>N} j^{-2/p} \\ \lesssim \tfrac{1}{2/p-1} \|\mathbf{v}\|_{\ell_p^*}^2 N^{1-2/p} = \tfrac{1}{2s} \|\mathbf{v}\|_{\ell_p^*}^2 N^{-2s}.\end{aligned}$$

Conversely, for $\mathbf{v} \in \mathcal{A}^s$ and $N \in \mathbb{N}$ we have

$$|\gamma_{2N}(\mathbf{v})|^2 N \leq \sum_{N<j\leq 2N} |\gamma_j(\mathbf{v})|^2 \leq \|\mathbf{v} - \mathcal{B}_N(\mathbf{v})\|^2 \leq N^{-2s} |\mathbf{v}|_{\mathcal{A}^s}^2,$$

which means that $|\gamma_{2N}(\mathbf{v})| \leq N^{(-s+1/2)} |\mathbf{v}|_{\mathcal{A}^s} = N^{-1/p} |\mathbf{v}|_{\mathcal{A}^s}$. Now we use $\gamma_1(\mathbf{v}) \leq \|\mathbf{v}\|$ and $\gamma_{2N+1}(\mathbf{v}) \leq \gamma_{2N}(\mathbf{v})$ to complete the proof. ∎

Since for $p \leq 1$, ℓ_p^* is not normable, cf. [4], the above result shows that for $s \geq \frac{1}{2}$, $\mathcal{A}^s$ is not normable, meaning that it is only a quasi-Banach space. On the other hand, also from the theory of ℓ_p^*-spaces one infers that for $s < \frac{1}{2}$, there exists a norm equivalent to $|\cdot|_{\mathcal{A}^s}$, so that $\mathcal{A}^s$ is a Banach space with respect to it.

The next observation is that ℓ_p^* is very close to ℓ_p. In fact, for any $p \in (0, 2)$ and $\epsilon > 0$, we have

$$j^{1/p} |\gamma_j(\mathbf{v})| = \big(j |\gamma_j(\mathbf{v})|^p\big)^{1/p} \leq \big(\sum_{k\leq j} |\gamma_k(\mathbf{v})|^p\big)^{1/p} \leq \|\mathbf{v}\|_{\ell_p},$$

and

$$\|\mathbf{v}\|_{\ell_{p+\epsilon}}^{p+\epsilon} = \sum_{j\in\mathbb{N}} |\gamma_j(\mathbf{v})|^{p+\epsilon} \leq \sum_{j\in\mathbb{N}} |\mathbf{v}|_{\ell_p^*}^{p+\epsilon} \cdot j^{-1-\epsilon/p} \leq C \cdot |\mathbf{v}|_{\ell_p^*}^{p+\epsilon},$$

so that

$$\ell_p \hookrightarrow \ell_p^* \hookrightarrow \ell_{p+\epsilon}. \tag{2.3.4}$$

Remark 2.3.3. Let us consider a wavelet basis Ψ for H^t. In view of the above results and the norm equivalence (2.2.7) on page 13, we have that whenever $\mathbf{v}^T\Psi \in B_p^{t+ns}(L_p)$ with $\frac{1}{p} = s + \frac{1}{2}$, $\mathbf{v}$ satisfies $\mathbf{v} \in \mathcal{A}^s$ with $|\mathbf{v}|_{\mathcal{A}^s} \lesssim \|\mathbf{v}^T\Psi\|_{B_p^{t+ns}(L_p)}$. Therefore, the rate of the best N-term approximation of a function in wavelet bases is governed by the Besov regularity of the function.

As we know, the validity of the norm equivalence (2.2.7) imposes certain constraints on the possible values of the parameters. In the present context, those constraints can be rephrased as follows. For $t < -\frac{n}{2}$, the value of s is

restricted by $s \le \frac{1}{2}$, because of the condition $p \ge 1$. For arbitrary t, one needs $t + ns < \min\{d, \gamma(p)\}$ or $s < \min\{\frac{d-t}{n}, \frac{\gamma(p)-t}{n}\}$. If the wavelets are piecewise smooth, globally C^r-functions for some $r \in \{-1, 0, 1, \ldots\}$, where $r = -1$ means that they satisfy no global continuity condition, then it is known that $\gamma(p) = r + 1 + 1/p = r + 1 + s + 1/2 = \gamma(2) + s$, giving the bound $s < \min\{\frac{d-t}{n}, \frac{\gamma(2)-t}{n-1}\}$. So if $r \ge \frac{t-d}{n} + d - \frac{3}{2}$, then the smoothness of the wavelets does not limit the range for which the norm equivalence (2.2.7) is valid. With spline wavelets we have $r = d - 2$, in which case the above requirement reads as $\frac{d-t}{n} \ge \frac{1}{2}$.

On the other hand, we see that only one side of the relation (2.2.7) is sufficient to bound the ℓ_p-norm of a sequence by the Besov norm of the corresponding function. In fact, by using the inequality (2.2.8), for $s \in (0, \frac{d-t}{n})$ and $t > -\frac{n}{2}$, we infer that if $\mathbf{v}^T\Psi \in B_q^{t+ns}(L_p)$ with $\frac{1}{p} < s + \frac{1}{2}$ and $q \in (0, p]$, then $\mathbf{v} \in \mathcal{A}^s$ with $|\mathbf{v}|_{\mathcal{A}^s} \lesssim \|\mathbf{v}^T\Psi\|_{B_p^{t+ns}(L_p)}$. Note that the condition involving $\gamma(p)$ has disappeared.

We sketch here a proof of the aforementioned fact. Let $\mathbf{v}^T\Psi \in B_q^{t+ns}(L_p)$ with $q = p$, and let $C \ge 0$ denote the quantity in the left side of the inequality (2.2.8). Noting that s in (2.2.8) has to be replaced by $t + ns$ here, when $\frac{1}{p} < s + \frac{1}{2}$, we have $\|(\mathbf{v}_\lambda)_{|\lambda|=j}\|_{\ell_p} \le C2^{-jn\delta}$ with $\delta := s + \frac{1}{2} - \frac{1}{p} > 0$. With $(\gamma_j(\mathbf{v}))_{j\ge 0}$ denoting the non-increasing rearrangement of $\mathbf{v}$, we infer

$$2^{jn/p}|\gamma_{2^{jn}}(\mathbf{v})| \le \Big(\sum_{k\le 2^{jn}} |\gamma_k(\mathbf{v})|^p\Big)^{1/p} \lesssim C2^{-jn\delta} = C(2^{jn})^{-\delta}.$$

Now taking into account that $\#\{\lambda : |\lambda| = j\} \lesssim 2^{jn}$, by monotonicity of $(\gamma_k(\mathbf{v}))$, the above estimate implies that $j^{1/p}|\gamma_j(\mathbf{v})| \lesssim j^{-\delta}$ or $\mathbf{v} \in \ell_{\tilde{p}}^*$ with $\frac{1}{\tilde{p}} = \frac{1}{p} + \delta = s + \frac{1}{2}$, so that $\mathbf{v} \in \mathcal{A}^s$. The case $q < p$ follows by embedding. ⊘

Remark 2.3.4. Even though ℓ_p^* is very close to ℓ_p in the sense of (2.3.4), the embedding $\ell_p \hookrightarrow \ell_p^*$ is proper, since for example, a sequence $\mathbf{v}$ with $|\gamma_j(\mathbf{v})| = j^{-1/p}$ is in ℓ_p^* but not in ℓ_p. Hence we see that the space $X^s := \{\mathbf{v}^T\Psi : \mathbf{v} \in \mathcal{A}^s\}$ is slightly bigger than $B_p^{t+ns}(L_p)$, with $\frac{1}{p} = s + \frac{1}{2}$. Actually, given the norm equivalence (2.2.7), the spaces X^α, $\alpha \in (0, s)$, can be characterized by interpolation spaces as $X^\alpha = [H^t, B_p^{t+ns}(L_p)]_{\alpha/s,\infty}$, which, however, is not a Besov space, cf. [16, 39].

On the other hand, defining the "refined" approximation spaces for $s > 0$ and $q \in (0, \infty]$, by

$$\mathcal{A}_q^s := \left\{\mathbf{v} \in \ell_2 : |\mathbf{v}|_{\mathcal{A}_q^s} := \left\|\left(N^{s-1/q}E_N(\mathbf{v})\right)_{N\in\mathbb{N}}\right\|_{\ell_q} < \infty\right\},$$

an extension of Lemma 2.3.2 exists that says that $\mathcal{A}_q^s = \ell_{p,q}$ with $\frac{1}{p} = s + \frac{1}{2}$, where $\ell_{p,q} := \{\mathbf{v} : \|(j^{1/p-1/q}|\gamma(\mathbf{v})|)_{j\in\mathbb{N}}\|_{\ell_q} < \infty\}$ is the Lorentz sequence space. Since $\ell_{p,p} = \ell_p$, in view of the norm equivalence (2.2.7), we have $B_p^{t+ns}(L_p) = \{\mathbf{v}^T\Psi : \mathbf{v} \in \mathcal{A}_p^s\}$ with $\frac{1}{p} = s + \frac{1}{2}$. Note that $\mathcal{A}^s = \mathcal{A}_\infty^s$, and that $\mathcal{A}_{q_1}^s \hookrightarrow \mathcal{A}_{q_2}^s$ for

$0 < q_1 < q_2 \leq \infty$, and $\mathcal{A}^s_{q_1} \hookrightarrow \mathcal{A}^{s-\varepsilon}_{q_2}$ for any $\varepsilon > 0$ and any $q_1, q_2 \in (0, \infty]$. These relations imply (2.3.4) as special cases. For a detailed treatment of related issues in the theory of nonlinear approximation, the reader is referred to [16, 39]. ⊘

Remark 2.3.5. In view of the Jackson estimate (2.2.3) on page 10, membership of a function v in the Sobolev space H^{t+ns} yields an error decay measured in H^t-metric of order $2^{-jns}|v|_{H^{t+ns}}$ for the approximation from the "coarsest level" *linear* subspaces $S_j = \operatorname{span}\{\psi_\lambda : \lambda \in \nabla_j\}$. Since the number of wavelets in ∇_j is of order $N_j \eqsim 2^{jn}$, the error of this linear approximation expressed in terms of the number of degrees of freedom decays like $N_j^{-s}|v|_{H^{t+ns}}$. The condition $v \in B_p^{t+ns}(L_p)$ with $\frac{1}{p} = s + \frac{1}{2}$ involving Besov regularity which is sufficient to guarantee this rate of convergence with nonlinear approximation, is much milder than the condition $v \in H^{t+ns}$ involving Sobolev regularity. Indeed, H^{t+ns} is properly imbedded in $B_p^{t+ns}(L_p)$, and the gap increases when s grows. Assuming a sufficiently smooth right-hand side, for several boundary value problems it was proven that the solution has a much higher Besov regularity than Sobolev regularity [26].

Similar to the previous remark, the Jackson estimate (2.2.3), however, presents only a sufficient condition for the error decay of order N_j^{-s}, and the question arises whether there are functions in H^t outside H^{t+ns} that nevertheless show an error decay of order N_j^{-s} for the linear approximation process. One can show that for $s < \gamma$, such functions do exist, but they are necessarily contained in $H^{t+ns-\varepsilon}$ for arbitrarily small $\varepsilon > 0$.

Note that we have been discussing only a particular type of linear approximation, namely, the approximation from the subspaces S_j. So a natural question is whether there exists a linear approximation process that approximates as good as best N-term approximations. The answer turns out to be negative. By employing the notion of Kolmogorov's N-widths, it has been shown that for *any* sequence of nested linear spaces, the corresponding approximation space $\mathcal{A}^s$ is always properly included in the approximation space $\mathcal{A}^s$ for the best N-term approximation, where the gap between them increases as s grows, cf. [39]. ⊘

We end this section by recalling some facts concerning perturbations of best N-term approximations, which will be often used in the sequel. The following proposition is recalled from [17, 83].

Proposition 2.3.6. *Let $s > 0$ and let $P \subset \ell_2$ denote the set of all finitely supported sequences. Then for any $\mathbf{v} \in \mathcal{A}^s$ and $\mathbf{z} \in P$, we have*

$$|\mathbf{z}|_{\mathcal{A}^s} \lesssim |\mathbf{v}|_{\mathcal{A}^s} + (\#\operatorname{supp}\mathbf{z})^s \|\mathbf{v} - \mathbf{z}\|.$$

Proof. Let $N := \#\operatorname{supp}\mathbf{z}$, then

$$|\mathbf{z}|_{\mathcal{A}^s} \lesssim |\mathbf{z}-\mathcal{B}_N(\mathbf{v})|_{\mathcal{A}^s} + |\mathcal{B}_N(\mathbf{v})|_{\mathcal{A}^s} \lesssim (2N)^s\|\mathbf{z}-\mathcal{B}_N(\mathbf{v})\| + |\mathbf{v}|_{\mathcal{A}^s},$$

where we used $\#\operatorname{supp}(\mathbf{z}-\mathcal{B}_N(\mathbf{v})) \le 2N$ and (2.3.3). The proof is completed by

$$\|\mathbf{z}-\mathcal{B}_N(\mathbf{v})\| \le \|\mathbf{z}-\mathbf{v}\| + \|\mathbf{v}-\mathcal{B}_N(\mathbf{v})\| \le 2\|\mathbf{z}-\mathbf{v}\|. \qquad \blacksquare$$

The following result shows that by removing small coefficients from an approximation $\mathbf{z} \in P$ of $\mathbf{v} \in \mathcal{A}^s$, one can get an approximation nearly as efficient as a best N-term approximation. The proof follows the proof of [28, Proposition 3.4].

Proposition 2.3.7. *Let $\theta > 1$ and $s > 0$. Then for any $\varepsilon > 0$, $\mathbf{v} \in \mathcal{A}^s$, and $\mathbf{z} \in P$ with*

$$\|\mathbf{z}-\mathbf{v}\| \le \varepsilon,$$

for the smallest $N \in \mathbb{N}_0$ such that $\|\mathbf{z}-\mathcal{B}_N(\mathbf{z})\| \le \theta\varepsilon$, it holds that

$$N \lesssim \varepsilon^{-1/s}|\mathbf{v}|_{\mathcal{A}^s}^{1/s},$$

and

$$|\mathcal{B}_N(\mathbf{z})|_{\mathcal{A}^s} \lesssim |\mathbf{v}|_{\mathcal{A}^s}.$$

Proof. When $\|\mathbf{v}\| \le (\theta-1)\varepsilon$, we have $\|\mathbf{z}-0\| \le \theta\varepsilon$, meaning that $N = 0$.

From now on we assume that $\|\mathbf{v}\| > (\theta-1)\varepsilon$. Let $m \in \mathbb{N}_0$ be the largest integer with $E_m(\mathbf{v}) > (\theta-1)\varepsilon$. Such an m exists by our assumption. For $m > 0$, we have

$$(\theta-1)\varepsilon < E_m(\mathbf{v}) \le m^{-s}|\mathbf{v}|_{\mathcal{A}^s},$$

or $m \lesssim \varepsilon^{-1/s}|\mathbf{v}|_{\mathcal{A}^s}^{1/s}$, which is also trivially true for $m = 0$. By the definition of m, we infer $E_{m+1}(\mathbf{v}) \le (\theta-1)\varepsilon$ or $\|\mathbf{z}-\mathcal{B}_{m+1}(\mathbf{v})\| \le \|\mathbf{z}-\mathbf{v}\| + E_{m+1}(\mathbf{v}) \le \theta\varepsilon$, and so $N \le m+1$. The proof of the bound on N is completed by noting that $1 \lesssim (\theta-1)^{1/s} < \varepsilon^{-1/s}\|\mathbf{v}\|^{1/s} \le \varepsilon^{-1/s}|\mathbf{v}|_{\mathcal{A}^s}^{1/s}$. The bound on $|\mathcal{B}_N(\mathbf{z})|_{\mathcal{A}^s}$ follows from an application of Proposition 2.3.6. $\blacksquare$

2.4 Linear operator equations

Let $\mathcal{H}$ and $\mathcal{H}'$ be a separable Hilbert space and its dual respectively. We consider the problem of numerically solving an operator equation, which is formulated as

follows. For a given *boundedly invertible linear operator* $L : \mathcal{H} \to \mathcal{H}'$ and a linear functional $f \in \mathcal{H}'$, find $u \in \mathcal{H}$ such that

$$Lu = f. \tag{2.4.1}$$

We refer to $\mathcal{H}$ as the energy space of the problem. Within this framework we can discuss a quite wide range of problems, including for example weak formulations of partial differential equations, pseudo-differential equations, boundary integral equations, as well as systems of equations of those kinds. Then the corresponding energy space $\mathcal{H}$ is (a closed subspace of) a relevant Sobolev space formulated on a domain or manifold, or a product of relevant Sobolev spaces, cf. [18]. As a well known example, one may think of the weak formulation of an elliptic boundary value problem.

Example 2.4.1 (Elliptic boundary value problems). Let $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain, and with $\Gamma \subseteq \partial\Omega$ being a part of the boundary with nonzero measure, let $\mathcal{H} := H^1_\Gamma(\Omega) \subset H^1(\Omega)$ be the subspace of the Sobolev space $H^1(\Omega)$ of functions with vanishing trace on Γ. Let $L : \mathcal{H} \to \mathcal{H}'$ be defined by

$$\langle Lv, w\rangle = -\sum_{j,k=1}^{n} \langle a_{jk}\partial_k v, \partial_j w\rangle_{L_2} + \sum_{k=1}^{n} \langle b_k \partial_k v, w\rangle_{L_2} + \langle cv, w\rangle_{L_2} \quad v, w \in \mathcal{H},$$

where $\langle\cdot,\cdot\rangle$ is the duality pairing on $\mathcal{H} \times \mathcal{H}'$. If the coefficients satisfy $a_{jk}, b_k, c \in L_\infty$ then $L : \mathcal{H} \to \mathcal{H}'$ is bounded. Moreover, if there exists a constant $\alpha > 0$ such that

$$\textstyle\sum_{j,k=1}^{n} a_{jk}(x)\xi_j\xi_k \geq \alpha \sum_{k=1}^{n} \xi_k^n \qquad \text{for all } \xi \in \mathbb{R}^n \qquad a.e. \text{ in } \Omega,$$

and

$$\textstyle\alpha^2 + \sum_{k=1}^{n} \|b_k\|^2_{L_\infty(\Omega)} \leq 2\alpha \cdot \operatorname{essinf}\{b_0(x) : x \in \Omega\},$$

then the operator L is elliptic on $\mathcal{H}$, meaning that $\langle Lv, v\rangle \gtrsim \|v\|^2_{\mathcal{H}}$ for $v \in \mathcal{H}$. Therefore L is boundedly invertible, cf. [11]. ⊘

Another class of examples comes from a reformulation of boundary value problems on domains as integral equations on the boundary of the domain.

Example 2.4.2 (Single layer operator). Let Γ be a sufficiently smooth closed two dimensional manifold in $\mathbb{R}^3$, and set $\mathcal{H} := H^{\frac{1}{2}}(\Gamma)$. Then the single layer operator $L : \mathcal{H} \to \mathcal{H}'$ defined by

$$\langle Lv, w\rangle = \iint_{\Gamma\times\Gamma} \frac{v(x)w(y)}{4\pi|x-y|} d\Gamma_x d\Gamma_y \qquad v, w \in \mathcal{H},$$

is bounded and $\mathcal{H}$-elliptic, cf. [57]. ⊘

Let $\Psi = \{\psi_\lambda : \lambda \in \nabla\}$ be a *Riesz basis* of $\mathcal{H}$, with $F : \mathcal{H}' \to \ell_2$ and $F' : \ell_2 \to \mathcal{H}$ being the analysis and synthesis operators as defined in (2.2.2), respectively. If we write the solution of (2.4.1) as $u = F'\mathbf{u}$ for some $\mathbf{u} \in \ell_2$, $\mathbf{u}$ must satisfy

$$\mathbf{L}\mathbf{u} = \mathbf{f}, \tag{2.4.2}$$

where the so called *stiffness matrix* $\mathbf{L} := FLF' : \ell_2 \to \ell_2$ is boundedly invertible and the right hand side vector $\mathbf{f} := Ff \in \ell_2$. In the sequel, we also use the notation $\langle \Psi, L\Psi \rangle := FLF'$.

Many of the results in the sequel are formulated specifically for the case that the stiffness matrix $\mathbf{L}$ in (2.4.2) is *symmetric and positive definite* (SPD). For clarity, in the context of those results we will denote the stiffness matrix by $\mathbf{A} := \mathbf{L}$, i.e., we will be considering the equation

$$\mathbf{A}\mathbf{u} = \mathbf{f}, \tag{2.4.3}$$

with $\mathbf{A} : \ell_2 \to \ell_2$ SPD, and $\mathbf{f} \in \ell_2$. For the case that $\mathbf{L}$ is not SPD, in view of transferring the results obtained for (2.4.3) to the general case (2.4.2), one possibility could be to consider the normal equation $\mathbf{L}^T\mathbf{L}\mathbf{u} = \mathbf{L}^T\mathbf{f}$.

For a given subset $\Lambda \subset \nabla$, considering $\ell_2(\Lambda)$ as a linear subspace of ℓ_2, an approximation from $\ell_2(\Lambda)$ to the exact solution of (2.4.3) is given by the *Ritz-Galerkin approximation* that is obtained by requiring that the residual $\mathbf{r} := \mathbf{f} - \mathbf{A}\mathbf{u}_\Lambda$ for the sought approximation $\mathbf{u}_\Lambda \in \ell_2(\Lambda)$ is ℓ_2-orthogonal to the subspace $\ell_2(\Lambda)$, i.e., $\langle \mathbf{f} - \mathbf{A}\mathbf{u}_\Lambda, \mathbf{v}_\Lambda \rangle = 0$ for $\mathbf{v}_\Lambda \in \ell_2(\Lambda)$. Since $\mathbf{A}$ is SPD, $\langle\!\langle \cdot, \cdot \rangle\!\rangle := \langle \mathbf{A}\cdot, \cdot \rangle$ defines an inner product, and $|\!|\!| \cdot |\!|\!| := \langle\!\langle \cdot, \cdot \rangle\!\rangle^{\frac{1}{2}}$ is an equivalent norm in ℓ_2. Then the orthogonality condition $\langle \mathbf{f} - \mathbf{A}\mathbf{u}_\Lambda, \mathbf{v}_\Lambda \rangle = 0$ is equivalent to $\langle\!\langle \mathbf{u} - \mathbf{u}_\Lambda, \mathbf{v}_\Lambda \rangle\!\rangle = 0$, so for any $\mathbf{v}_\Lambda \in \ell_2(\Lambda)$, we have $|\!|\!|\mathbf{u} - \mathbf{v}_\Lambda|\!|\!|^2 = |\!|\!|\mathbf{u} - \mathbf{u}_\Lambda|\!|\!|^2 + |\!|\!|\mathbf{u}_\Lambda - \mathbf{v}_\Lambda|\!|\!|^2$, which is called the *Galerkin orthogonality*. The Galerkin orthogonality immediately implies that the approximation $\mathbf{u}_\Lambda$ is the best approximation to $\mathbf{u}$ from the subspace $\ell_2(\Lambda)$ in the norm $|\!|\!| \cdot |\!|\!|$.

Recalling that $\mathbf{P}_\Lambda : \ell_2 \to \ell_2(\Lambda)$ is the ℓ_2-orthogonal projector onto $\ell_2(\Lambda)$, the Ritz-Galerkin approximation $\mathbf{u}_\Lambda$ can be found by solving the equation $\mathbf{P}_\Lambda\mathbf{A}\mathbf{u}_\Lambda = \mathbf{P}_\Lambda\mathbf{f}$. This equation has a unique solution $\mathbf{u}_\Lambda$ since, as the following lemma implies, the matrix $\mathbf{A}_\Lambda := \mathbf{P}_\Lambda\mathbf{A}\mathbf{I}_\Lambda$ is SPD with $\mathbf{I}_\Lambda := \mathbf{P}_\Lambda^* : \ell_2(\Lambda) \to \ell_2$ being the trivial inclusion of $\ell_2(\Lambda)$ into ℓ_2. Note that $\mathbf{I}_\Lambda\mathbf{v}_\Lambda$ is simply the vector obtained by extending $\mathbf{v}_\Lambda$ by zeros for indices outside Λ. We will return to the Ritz-Galerkin approximation in the next chapter.

Lemma 2.4.3. *Let $\mathbf{A} : \ell_2 \to \ell_2$ be a symmetric and positive definite matrix. Then $|\!|\!| \cdot |\!|\!| := \langle \mathbf{A}\cdot, \cdot \rangle^{\frac{1}{2}}$ is a norm in ℓ_2, satisfying*

$$\|\mathbf{A}^{-1}\|^{-\frac{1}{2}}\|\mathbf{v}\| \le |\!|\!|\mathbf{v}|\!|\!| \le \|\mathbf{A}\|^{\frac{1}{2}}\|\mathbf{v}\|, \tag{2.4.4}$$

and

$$\|\mathbf{A}^{-1}\|^{-\frac{1}{2}}\|\mathbf{I}_\Lambda \mathbf{v}_\Lambda\| \le \|\mathbf{P}_\Lambda \mathbf{A}\mathbf{I}_\Lambda \mathbf{v}_\Lambda\| \le \|\mathbf{A}\|^{\frac{1}{2}}\|\mathbf{I}_\Lambda \mathbf{v}_\Lambda\|, \tag{2.4.5}$$

for any $\mathbf{v} \in \ell_2$, $\Lambda \subseteq \nabla$, *and* $\mathbf{v}_\Lambda \in \ell_2(\Lambda)$.

Proof. Since $\mathbf{A}$ is SPD, so are $\mathbf{A}^{-1}$ and the finite section $\mathbf{A}_\Lambda = \mathbf{P}_\Lambda \mathbf{A}\mathbf{I}_\Lambda$, therefore $\langle \mathbf{A}^{-1}\cdot, \cdot\rangle$ and $\langle \mathbf{A}_\Lambda \cdot, \cdot\rangle$ define inner products in ℓ_2 and $\ell_2(\Lambda)$, respectively. The second inequality in (2.4.4) follows from the CBS (Cauchy-Bunyakowsky-Schwarz) inequality for the standard inner product $\langle \cdot, \cdot\rangle$. The first inequality is derived by using the CBS inequality for $\langle \mathbf{A}^{-1}\cdot, \cdot\rangle$ as

$$\langle \mathbf{A}^{-1}\mathbf{A}\mathbf{v}, \mathbf{v}\rangle \le \langle \mathbf{A}^{-1}\mathbf{A}\mathbf{v}, \mathbf{A}\mathbf{v}\rangle^{\frac{1}{2}} \langle \mathbf{A}^{-1}\mathbf{v}, \mathbf{v}\rangle^{\frac{1}{2}} \le \|\mathbf{v}\|\|\mathbf{A}^{-1}\|^{\frac{1}{2}}\|\mathbf{v}\|.$$

An application of the CBS inequality for $\langle \cdot, \cdot\rangle$ followed by the first inequality in (2.4.4) gives the first inequality in (2.4.5). The second inequality in (2.4.5) is obtained similarly by applying the CBS inequality for $\langle \mathbf{A}_\Lambda \cdot, \cdot\rangle$. ∎

2.5 Convergent iterations in the energy space

Let us consider the following iteration in the sequence space ℓ_2 to solve our discrete problem (2.4.2)

$$\mathbf{u}_l = \mathcal{K}\mathbf{u}_{l-1}, \qquad l = 1, 2, \ldots \tag{2.5.1}$$

where $\mathbf{u}_0 \in \ell_2$ is an initial guess and $\mathcal{K} : \ell_2 \to \ell_2$ is continuous. The map $\mathcal{K}$ depends on the operator $\mathbf{L}$ and the right hand side $\mathbf{f}$. We assume that for some $\rho < 1$,

$$\|\mathbf{u}_l - \mathbf{u}\|_\star \le \rho^l \|\mathbf{u}_0 - \mathbf{u}\|_\star \qquad \text{for all } \mathbf{u}^0 \in \ell_2, \tag{2.5.2}$$

where the norm $\|\cdot\|_\star$ satisfies

$$\alpha_\star \|\mathbf{v}\| \le \|\mathbf{v}\|_\star \le \beta_\star \|\mathbf{v}\| \qquad \mathbf{v} \in \ell_2, \tag{2.5.3}$$

with constants $\alpha_\star, \beta_\star > 0$. We will call the map $\mathcal{K}$ the *iterator* and the result vectors $\mathbf{u}_l$ the *iterands*.

For symmetric and positive definite (SPD) systems, typical examples are the steepest descent, and the Richardson iteration. In addition, general problems can be transferred to SPD problems using the formulation of normal equations, although in special cases more efficient formulations can be achieved, for example Uzawa type algorithms for saddle point problems. Therefore, for the moment ignoring the question of quantitative performance, there is no loss of generality when we focus on SPD matrices $\mathbf{L} = \mathbf{A}$.

Example 2.5.1 (The Richardson iteration). Let $\mathbf{A} : \ell_2 \to \ell_2$ be an SPD matrix. We consider here the Richardson iteration for the linear equation (2.4.3),

$$\mathcal{K}\mathbf{v} := \mathbf{v} + \omega(\mathbf{f} - \mathbf{A}\mathbf{v}). \tag{2.5.4}$$

Using the positive definiteness and the boundedness of the matrix $\mathbf{A}$, for any $\mathbf{v} \in \ell_2$ the following estimate is obtained.

$$\begin{aligned} \|\mathbf{u} - \mathcal{K}\mathbf{v}\| &= \|(\mathbf{I} - \omega\mathbf{A})(\mathbf{u} - \mathbf{v})\| \\ &\leq \max\{|1 - \omega\lambda_{\max}|, |1 - \omega\lambda_{\min}|\} \cdot \|\mathbf{u} - \mathbf{v}\|, \end{aligned}$$

with $\lambda_{\max} := \|\mathbf{A}\|$ and $\lambda_{\min} := \|\mathbf{A}^{-1}\|^{-1}$. Therefore, if $\rho := \max\{|1 - \omega\lambda_{\min}|, |1 - \omega\lambda_{\max}|\} < 1$ or equivalently, $\omega \in (0, 2/\lambda_{\max})$ then Richardson's iteration converges:

$$\|\mathbf{u} - \mathcal{K}\mathbf{v}\| \leq \rho\|\mathbf{u} - \mathbf{v}\|.$$

Furthermore, with $\kappa(\mathbf{A}) := \|\mathbf{A}\|\|\mathbf{A}^{-1}\|$, the minimum value of the error reduction factor ρ and the corresponding damping parameter ω are:

$$\rho_{\text{opt}} = \tfrac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}} = \tfrac{\kappa(\mathbf{A}) - 1}{\kappa(\mathbf{A}) + 1} \qquad \text{when} \quad \omega_{\text{opt}} = \tfrac{2}{\lambda_{\max} + \lambda_{\min}}. \qquad \oslash$$

Example 2.5.2 (Steepest descent method). Let $\mathbf{A} : \ell_2 \to \ell_2$ be a SPD matrix. We consider the steepest descent iteration for the linear equation (2.4.3),

$$\mathcal{K}\mathbf{v} := \mathbf{v} + \frac{\langle \mathbf{r}, \mathbf{r} \rangle}{\langle \mathbf{A}\mathbf{r}, \mathbf{r} \rangle}\mathbf{r}, \tag{2.5.5}$$

where $\mathbf{r} := \mathbf{f} - \mathbf{A}\mathbf{v} \neq 0$ is the residual for $\mathbf{v}$. With the equivalent norm $|\!|\!| \cdot |\!|\!| := \langle \mathbf{A}\cdot, \cdot \rangle^{\frac{1}{2}}$, this iteration satisfies, cf. e.g. [66],

$$|\!|\!|\mathbf{u} - \mathcal{K}\mathbf{v}|\!|\!| \leq \frac{\kappa(\mathbf{A}) - 1}{\kappa(\mathbf{A}) + 1}|\!|\!|\mathbf{u} - \mathbf{v}|\!|\!|. \qquad \oslash$$

Now let us turn our attention to general iterations (2.5.1). In view of the above examples, the exact iteration cannot be expected to be implementable since in general the iterands are infinite dimensional vectors. However, since we can approximate any ℓ_2-sequence by finite ones within any finite accuracy, we shall consider the approximate application of the iterator within finite accuracies. Postponing the question of how to do so, first we will discuss how a perturbation affects the exact iteration (2.5.1). Let $P \subset \ell_2$ be the set of all finitely supported sequences and let $\tilde{\mathcal{K}} : \mathbb{R}_{>0} \times P \to P$ be a mapping such that

$$\|\tilde{\mathcal{K}}(\epsilon, \mathbf{v}) - \mathcal{K}\mathbf{v}\| \leq \epsilon \qquad \text{for all } \epsilon > 0,\ \mathbf{v} \in P. \tag{2.5.6}$$

Then we consider the following approximate iteration:

$$\tilde{\mathbf{u}}_l = \tilde{\mathcal{K}}(\epsilon_l, \tilde{\mathbf{u}}_{l-1}), \qquad l = 1, 2, \dots \tag{2.5.7}$$

with the initial guess $\tilde{\mathbf{u}}_0 \in P$ and control parameters $(\epsilon_l)_l$.

Lemma 2.5.3. *Let the initial guesses of the iterations* (2.5.1) *and* (2.5.7) *satisfy* $\mathbf{u}_0 = \tilde{\mathbf{u}}_0$. *Then the error of the approximate iteration* (2.5.7) *is, with* $\epsilon_0 := \|\mathbf{u}_0 - \mathbf{u}\|$,

$$\|\tilde{\mathbf{u}}_l - \mathbf{u}\| \le \tfrac{\beta_\star}{\alpha_\star} \textstyle\sum_{k=0}^{l} \rho^k \epsilon_{l-k},$$

with the constants $\alpha_\star$ *and* $\beta_\star$ *from* (2.5.3). *In particular, by taking* $\epsilon_i := \gamma\epsilon_0\rho^i/l$, $i = 1, \dots, l$, *with some* $\gamma > 0$, *we can ensure* $\|\tilde{\mathbf{u}}_l - \mathbf{u}\| \le (1+\gamma)\epsilon_0\rho^l\beta_\star/\alpha_\star$.

Proof. By using (2.5.3), (2.5.6), and (2.5.2), the distance between the two iterations can be estimated as

$$\begin{aligned} e_l &:= \|\tilde{\mathbf{u}}_l - \mathbf{u}_l\|_\star = \|\tilde{\mathcal{K}}(\epsilon_l, \tilde{\mathbf{u}}_{l-1}) - \mathcal{K}\mathbf{u}_{l-1}\|_\star \\ &\le \beta_\star \|\tilde{\mathcal{K}}(\epsilon_l, \tilde{\mathbf{u}}_{l-1}) - \mathcal{K}\tilde{\mathbf{u}}_{l-1}\| + \|\mathcal{K}\tilde{\mathbf{u}}_{l-1} - \mathcal{K}\mathbf{u}_{l-1}\|_\star \\ &\le \beta_\star\epsilon_l + \rho e_{l-1} \le \beta_\star \textstyle\sum_{k=0}^{l-1} \rho^k \epsilon_{l-k}. \end{aligned}$$

Hence the error of the approximate iteration is

$$\|\tilde{\mathbf{u}}_l - \mathbf{u}\| \le \tfrac{1}{\alpha_\star}\left(\|\tilde{\mathbf{u}}_l - \mathbf{u}_l\|_\star + \|\mathbf{u}_l - \mathbf{u}\|_\star\right) \le \tfrac{\beta_\star}{\alpha_\star} \textstyle\sum_{k=0}^{l} \rho^k \epsilon_{l-k}. \qquad \blacksquare$$

2.6 Optimal complexity with coarsening of the iterands

Lemma 2.5.3 shows that the approximate iteration (2.5.7) can be organized such that for any given target tolerance $\varepsilon > 0$, it produces an approximation $\mathbf{u}_\varepsilon \in P$ with $\|\mathbf{u} - \mathbf{u}_\varepsilon\| \le \varepsilon$. We are interested in *adaptive* solution methods, where $\operatorname{supp} \mathbf{u}_\varepsilon$ depends on both the exact solution $\mathbf{u}$ and the target tolerance ε. The method may use low level wavelets where the solution is smooth, and higher level wavelets only where the solution has singularities. This is an analogy to non-uniform meshes arising from local refinements in adaptive finite element methods. For *non-adaptive* methods a sequence $\Lambda_0 \subset \Lambda_1 \subset \ldots \subset \nabla$ is fixed *a priori*, and the goal is to find the smallest i such that there is an approximation $\mathbf{u}_\varepsilon \in \ell_2(\Lambda_i)$ with $\|\mathbf{u} - \mathbf{u}_\varepsilon\| \le \varepsilon$.

In any case, it is obvious that with $N := \#\operatorname{supp} \mathbf{u}_\varepsilon$, $\|\mathbf{u} - \mathbf{u}_\varepsilon\| \ge E_N(\mathbf{u})$. In this regard, the rate of convergence of best N-term approximations delivers a yardstick against which the convergence rate of a solution method can be measured. Recall

that whenever $\mathbf{u} \in \mathcal{A}^s$, the smallest N such that $E_N(\mathbf{u}) \leq \varepsilon$ satisfies $N \lesssim \varepsilon^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}$. Let a method define a map $(\mathbf{u}, \varepsilon) \mapsto \mathbf{u}_\varepsilon$, where of course, the solution $\mathbf{u}$ is given only implicitly. Then, for $s > 0$, we say that the method converges at the *optimal rate* s, when $\mathbf{u} \in \mathcal{A}^s$ implies $\#\operatorname{supp} \mathbf{u}_\varepsilon \lesssim \varepsilon^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}$. Our goal is to construct methods which converge at the optimal rate for a reasonably wide range of s, with the additional property that the method takes a number of arithmetic operations bounded by an absolute multiple of $\varepsilon^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}$. This additional property is called the property of *optimal computational complexity*.

Since for non-adaptive methods the approximations take place in the linear spaces $\ell_2(\Lambda_i)$, these methods converge at most with the same rate as that of the corresponding linear approximation process. In view of Remark 2.3.5 on page 18, we see that adaptive methods have potentially large advantages over their non-adaptive counterparts.

We now return to the discussion of constructing optimally convergent methods. To this end, a central idea is the idea of *coarsening*, which was introduced in the pioneering work [17]. Given *some* approximation $\mathbf{z} \in P$ with $\|\mathbf{u} - \mathbf{z}\| \leq \varepsilon$, Proposition 2.3.7 on page 19 states that with a constant $\theta > 1$, and the *smallest* $N \in \mathbb{N}_0$ such that $\|\mathbf{z} - \mathcal{B}_N(\mathbf{z})\| \leq \theta\varepsilon$, obviously $\|\mathbf{u} - \mathcal{B}_N(\mathbf{z})\| \leq (1+\theta)\varepsilon$ and $N \lesssim \varepsilon^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}$ whenever $\mathbf{u} \in \mathcal{A}^s$ for some $s > 0$. The name coarsening comes from the fact that removing small coefficients from $\mathbf{z}$ most likely results in removing unnecessarily fine level wavelets from regions where the solution is smooth, hence leaving only coarser level wavelets. This idea reduces the issue of optimal convergence rate to that of convergence: any linearly convergent method can be made optimally convergent with the help of an appropriate coarsening procedure. As it turns out, the remaining issue of optimal computational complexity can be dealt with by coarsening the iterands at least once in every fixed number of iterations. Of course, appropriate (but mild) requirements have to be made on the computational cost of the underlying convergent method.

In view of implementing the coarsening routine, for $\mathbf{z} \in P$, determining $\mathcal{B}_N(\mathbf{z})$ generally requires sorting of the coefficients in $\mathbf{z}$, which takes at least the order of $m \log m$ operations, with $m = \#\operatorname{supp} \mathbf{z}$. Although it is not likely that in practice this log-factor harms the efficiency of the algorithm, for a full proof of optimality we need to get rid of this log-factor. The observation is that instead of determining $\mathcal{B}_N(\mathbf{z})$, it suffices to find some index set $\Lambda \subset \operatorname{supp} \mathbf{z}$ such that $\|\mathbf{z} - \mathbf{P}_\Lambda \mathbf{z}\| \leq \theta\varepsilon$ and $\#\Lambda$ is at most a constant multiple of N, after which one can use $\mathbf{P}_\Lambda \mathbf{z}$ as a "coarsened" $\mathbf{z}$. To this end, we introduce a quasi-sorting algorithm which uses the so-called *bins* or *buckets* to store entries with roughly equal values. In the context of adaptive wavelet algorithms, this sorting algorithm was first used in [3, 83], see also [63].

Algorithm 2.6.1 Quasi-sorting algorithm $\mathbf{BSORT}[\mathbf{z}, \varepsilon] \to \{\mathbf{b}_i\}_{0 \le i \le q}$

Parameter: Let $\beta \in (0,1)$ be a constant.
Input: $\mathbf{z} \in P$ and $\varepsilon > 0$.
Output: $\mathbf{b}_i \in P$, $\mathbf{z}|_{\operatorname{supp} \mathbf{b}_i} = \mathbf{b}_i$ for all i, and $\mathbf{z} = \sum_i \mathbf{b}_i$, and $\|\mathbf{b}_q\| \le \varepsilon$.
1: $N := \#\operatorname{supp} \mathbf{z}$, $M := \|\mathbf{z}\|_{\ell_\infty}$;
2: Let $q \in \mathbb{N}_0$ be the smallest integer with $\beta^q M \le \varepsilon/\sqrt{N}$;
3: From the elements of $\mathbf{z}$, construct the vectors $\mathbf{b}_0, \dots, \mathbf{b}_q$ as follows:
4: $\mathbf{b}_0 := 0, \dots, \mathbf{b}_q := 0$;
5: For $\lambda \in \operatorname{supp} \mathbf{z}$ and $0 \le i < q$, set $[\mathbf{b}_i]_\lambda := \mathbf{z}_\lambda$ when $|\mathbf{z}_\lambda| \in (\beta^{i+1} M, \beta^i M]$; set $[\mathbf{b}_q]_\lambda := \mathbf{z}_\lambda$ when $|\mathbf{z}_\lambda| \le \beta^q M$.

For future reference, we state the following straightforward result, cf. [46, 83].

Lemma 2.6.2. *The number of arithmetic operations and storage locations needed for* $\{\mathbf{b}_i\} := \mathbf{BSORT}[\mathbf{z}, \varepsilon]$ *can be bounded by an absolute multiple of*

$$\#\operatorname{supp} \mathbf{z} + q + 1 \lesssim \#\operatorname{supp} \mathbf{z} + \log\left(\varepsilon^{-1}\|\mathbf{z}\|\right) + 1. \tag{2.6.1}$$

Moreover, $\|\mathbf{b}_q\| \le \varepsilon$, *and for* $0 \le i < q$, *any two nonzero entries from the vector* $\mathbf{b}_i$ *differ at most a factor* $1/\beta$ *in modulus.*

Proof. The only thing that might need a proof is (2.6.1). We have

$$\begin{aligned} q + 1 &\lesssim 1 + \log\left(\varepsilon^{-1}\|\mathbf{z}\|_{\ell_\infty}(\#\operatorname{supp} \mathbf{z})^{\frac{1}{2}}\right) \\ &\le 1 + \log\left(\varepsilon^{-1}\|\mathbf{z}\|_{\ell_\infty}\right) + \tfrac{1}{2}\log(\#\operatorname{supp} \mathbf{z}) \\ &\lesssim 1 + \log\left(\varepsilon^{-1}\|\mathbf{z}\|\right) + \#\operatorname{supp} \mathbf{z}. \end{aligned}$$

∎

Now we are ready to define the *coarsening routine* that for a given $\mathbf{z} \in P$, finds a $\mathbf{P}_\Lambda \mathbf{z}$ such that $\|\mathbf{z} - \mathbf{P}_\Lambda \mathbf{z}\| \le \varepsilon$, and where $\#\Lambda$ is minimal modulo some constant factor.

Algorithm 2.6.3 Clean-up step $\mathbf{COARSE}[\mathbf{z}, \varepsilon] \to \tilde{\mathbf{z}}$

Input: Let $\mathbf{z} \in P$ and $\varepsilon > 0$.
Output: $\tilde{\mathbf{z}} \in P$ and $\|\tilde{\mathbf{z}} - \mathbf{z}\| \le \varepsilon$.
1: $\{\mathbf{b}_i\}_{0 \le i \le q} := \mathbf{BSORT}[\mathbf{z}, \varepsilon]$;
2: Create $\tilde{\mathbf{z}}$ by collecting nonzero entries first from $\mathbf{b}_0$ and when it is exhausted from $\mathbf{b}_1$ and so on, until $\|\tilde{\mathbf{z}} - \mathbf{z}\| \le \varepsilon$ is satisfied.

Lemma 2.6.4. *For* $\mathbf{z} \in P$ *and* $\varepsilon > 0$, $\tilde{\mathbf{z}} := \mathbf{COARSE}[\mathbf{z}, \varepsilon]$ *terminates with* $\|\tilde{\mathbf{z}} - \mathbf{z}\| \le \varepsilon$ *and* $\tilde{\mathbf{z}} = \mathbf{P}_{[\operatorname{supp} \tilde{\mathbf{z}}]}\mathbf{z}$. *Moreover, the output satisfies*

$$\#\operatorname{supp} \tilde{\mathbf{z}} \lesssim \min\{N : E_N(\mathbf{z}) \le \varepsilon\} = \min\{\#\Lambda : \|\mathbf{z} - \mathbf{P}_\Lambda \mathbf{z}\| \le \varepsilon\}, \tag{2.6.2}$$

with $E_N(\cdot)$ from (2.3.2) on page 14. The number of arithmetic operations and storage locations needed for this routine can be bounded by an absolute multiple of $\#\operatorname{supp}\mathbf{z} + \log(\varepsilon^{-1}\|\mathbf{z}\|) + 1$. Note that for any fixed $s > 0$, $\log(\varepsilon^{-1}\|\mathbf{z}\|) \lesssim \varepsilon^{1/s}\|\mathbf{z}\|^{1/s} \leq \varepsilon^{1/s}|\mathbf{z}|_{\mathcal{A}^s}^{1/s}$.

Proof. We will prove only (2.6.2). Assume that $\tilde{\mathbf{z}} \neq 0$, and let β be the constant inside **BSORT**. Since $\|\mathbf{b}_q\| \leq \varepsilon$, the last entry added to $\tilde{\mathbf{z}}$ originates from $\mathbf{b}_i$ with $i < q$. Then a minimal set Λ that satisfies $\|\mathbf{P}_\Lambda\mathbf{z} - \mathbf{z}\| \leq \varepsilon$ contains all the entries from the vectors $\mathbf{b}_0, \ldots, \mathbf{b}_{i-1}$, as any entry in any of these vectors is greater in magnitude than any entry in $\mathbf{b}_i$. Since any two nonzero entries from $\mathbf{b}_i$ differ less than a factor $1/\beta$ in modulus, the cardinality of the contribution from $\mathbf{b}_i$ to $\operatorname{supp}\tilde{\mathbf{z}}$ is at most a factor $1/\beta^2$ larger than that to Λ, so that $\#\operatorname{supp}\tilde{\mathbf{z}} \leq \beta^{-2}\#\Lambda$.∎

The following is a key ingredient in proving optimal complexity of adaptive algorithms with coarsening of the iterands. Given Proposition 2.3.7 on page 19 and Lemma 2.6.4, the proof is straightforward.

Corollary 2.6.5. *Let $\theta > 1$ and $s > 0$. Then for any $\varepsilon > 0$, $\mathbf{v} \in \mathcal{A}^s$, and $\mathbf{z} \in P$ with*

$$\|\mathbf{z} - \mathbf{v}\| \leq \varepsilon,$$

for $\tilde{\mathbf{z}} := \mathbf{COARSE}[\mathbf{z}, \theta\varepsilon]$ it holds that

$$\#\operatorname{supp}\tilde{\mathbf{z}} \lesssim \varepsilon^{-1/s}|\mathbf{v}|_{\mathcal{A}^s}^{1/s},$$

obviously $\|\tilde{\mathbf{z}} - \mathbf{v}\| \leq (1+\theta)\varepsilon$, and

$$|\tilde{\mathbf{z}}|_{\mathcal{A}^s} \lesssim |\mathbf{v}|_{\mathcal{A}^s}.$$

In view of the discussion at the beginning of this section, the above result shows that this coarsening routine can be used in adaptive algorithms. Before presenting an optimal adaptive algorithm with coarsening of the iterands, we assume to have the following routine available, which can be thought of as some convergent method, not necessarily being optimal. In the subsequent sections we will consider a number of realizations of this routine, including the approximate Richardson and steepest descent iterations.

Algorithm 2.6.6 Algorithm template $\mathbf{ITERATE}[\mathbf{v}, \nu, \eta] \to \mathbf{w}$

Parameters: Let $\|\cdot\|_\star$ and $\alpha_\star, \beta_\star > 0$ be such that $\alpha_\star\|\mathbf{z}\| \leq \|\mathbf{z}\|_\star \leq \beta_\star\|\mathbf{z}\|$ for $\mathbf{z} \in \ell_2$.
Input: Let $\eta > 0$, $\mathbf{v} \in P$ and $\nu \geq \|\mathbf{u} - \mathbf{v}\|_\star$.
Output: $\mathbf{w} \in P$ with $\|\mathbf{u} - \mathbf{w}\|_\star \leq \eta$

Now we are ready to present our adaptive wavelet algorithm. Note that inside this algorithm we will only call **ITERATE** for $\nu/\eta \lesssim 1$.

Algorithm 2.6.7 Method **SOLVE**$[\varepsilon] \to \mathbf{u}_j$ with coarsening

---

Parameters: Let $\chi > 0$ and $\theta > 1$ be constants with $\chi(1+\theta)(\beta_\star/\alpha_\star) < 1$.
Input: $\varepsilon > 0$.
Output: $\mathbf{u}_j \in P$ with $\|\mathbf{u} - \mathbf{u}_j\|_\star \le \varepsilon$.

1: $\mathbf{u}_0 := 0$, $\nu_0 := \beta_\star \|\mathbf{L}^{-1}\| \|\mathbf{f}\|$, $j := 0$;
2: **while** $\nu_j > \varepsilon$ **do**
3: $\quad j := j + 1$;
4: $\quad \mathbf{v}_j := \mathbf{ITERATE}[\mathbf{u}_{j-1}, \nu_{j-1}, \chi\nu_{j-1}]$;
5: $\quad \mathbf{u}_j := \mathbf{COARSE}[\mathbf{v}_j, \theta\chi\nu_{j-1}/\alpha_\star]$;
6: $\quad \nu_j := \chi\nu_{j-1}(1+\theta)(\beta_\star/\alpha_\star)$;
7: **end while**

Theorem 2.6.8. *For any $\varepsilon > 0$, $\mathbf{u}_\varepsilon := \mathbf{SOLVE}[\varepsilon]$ terminates with $\|\mathbf{u} - \mathbf{u}_\varepsilon\|_\star \le \varepsilon$. Moreover, if $\mathbf{u} \in \mathcal{A}^s$ for some $s > 0$, then $\#\operatorname{supp} \mathbf{u}_\varepsilon \lesssim \varepsilon^{-1/s} |\mathbf{u}|_{\mathcal{A}^s}^{1/s}$. In addition, let $\varepsilon \lesssim \|\mathbf{f}\|$, and assume that for any $\mathbf{v} \in P$ and $\eta \gtrsim \nu \ge \|\mathbf{u} - \mathbf{v}\|_\star$, $\mathbf{w} := \mathbf{ITERATE}[\mathbf{v}, \nu, \eta]$ satisfies*

$$\#\operatorname{supp} \mathbf{w} \lesssim \#\operatorname{supp} \mathbf{v} + \eta^{-1/s} |\mathbf{u}|_{\mathcal{A}^s}^{1/s} \qquad \text{and} \qquad |\mathbf{w}|_{\mathcal{A}^s} \lesssim (\#\operatorname{supp} \mathbf{v})^s \eta + |\mathbf{u}|_{\mathcal{A}^s},$$

where the number of arithmetic operations and storage locations required by this call of **ITERATE** *can be bounded by an absolute multiple of*

$$\eta^{-1/s} |\mathbf{u}|_{\mathcal{A}^s}^{1/s} + \#\operatorname{supp} \mathbf{v} + 1.$$

Then, the number of arithmetic operations and storage locations required by the call is bounded by some absolute multiple of $\varepsilon^{-1/s} |\mathbf{u}|_{\mathcal{A}^s}^{1/s}$.

Proof. We first indicate the need for the condition $\varepsilon \lesssim \|\mathbf{f}\|$. If $\varepsilon \not\lesssim \|\mathbf{f}\|$, then $\varepsilon^{-1/s} |\mathbf{u}|_{\mathcal{A}^s}^{1/s}$ might be arbitrarily small, whereas **SOLVE** takes in any case some arithmetic operations. Without this condition, the total work can be bounded by an absolute multiple of $\varepsilon^{-1/s} |\mathbf{u}|_{\mathcal{A}^s}^{1/s} + 1$.

We have $\nu_0 \ge \|\mathbf{u}\|_\star$. Now suppose that in the j-th iteration, **ITERATE** was called with a valid parameter ν_{j-1}. Then from the properties of the subroutine **ITERATE**, we have

$$\begin{aligned} \|\mathbf{u} - \mathbf{u}_j\|_\star &\le \beta_\star \|\mathbf{u} - \mathbf{u}_j\| \le \beta_\star (\|\mathbf{u} - \mathbf{v}_j\| + \theta\chi\nu_{j-1}/\alpha_\star) \\ &\le (\beta_\star/\alpha_\star)(1+\theta)\chi\nu_{j-1} = \nu_j, \end{aligned}$$

from which the first statement of the theorem follows.

Since $\|\mathbf{u} - \mathbf{v}_j\| \leq \nu_j/\alpha_\star$, Corollary 2.6.5 on page 27 implies $\#\operatorname{supp}\mathbf{u}_j \lesssim \nu_{j-1}^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}$. So if **SOLVE** terminates directly after the K-th iteration with $K > 0$, meaning that $\nu_K \leq \varepsilon$ and $\nu_{K-1} > \varepsilon$, then we have the second statement of the theorem. The case $K = 0$ is trivial.

Now we will confirm the bound on the cost of the algorithm. By the third assumption on **ITERATE**, the cost of the j-th call of **ITERATE** is of order $\nu_{j-1}^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + 1$. Taking into account the cost of **COARSE** and the first assumption on **ITERATE**, the total cost of the j-th iteration can be bounded by an absolute multiple of

$$\nu_{j-1}^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + \log\left(\nu_{j-1}^{-1}\|\mathbf{v}_j\|\right) + 1 \lesssim \nu_{j-1}^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + \nu_{j-1}^{-1/s}|\mathbf{v}_j|_{\mathcal{A}^s}^{1/s} + 1.$$

By the second assumption on **ITERATE**, we have

$$|\mathbf{v}_j|_{\mathcal{A}^s} \lesssim (\#\operatorname{supp}\mathbf{u}_{j-1})^s\nu_{j-1} + |\mathbf{u}|_{\mathcal{A}^s} \lesssim |\mathbf{u}|_{\mathcal{A}^s}.$$

From $\nu_j \leq \nu_0 \lesssim \|\mathbf{u}\|$, we have $\nu_{j-1}^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} \gtrsim \|\mathbf{u}\|^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} \gtrsim 1$. The proof is completed by the geometric decrease of ν_j. ∎

2.7 Adaptive application of operators. Computability

When implementing an approximate Richardson iteration, for a given approximation $\mathbf{w} \in P$, we need to compute the residual $\mathbf{f} - \mathbf{L}\mathbf{w}$ approximately. We will accomplish this by computing the two terms separately, by assuming that the succeeding two subroutines are available.

Algorithm 2.7.1 Algorithm template $\mathbf{APPLY}[\mathbf{M}, \mathbf{v}, \varepsilon] \to \mathbf{w}$

Input: Let $\mathbf{M} : \ell_2 \to \ell_2$ be bounded, $\mathbf{v} \in P$ and $\varepsilon > 0$.
Output: $\mathbf{w} \in P$ and $\|\mathbf{w} - \mathbf{M}\mathbf{v}\| \leq \varepsilon$.

Algorithm 2.7.2 Algorithm template $\mathbf{RHS}[\mathbf{g}, \varepsilon] \to \mathbf{g}_\varepsilon$

Input: Let $\mathbf{g} \in \ell_2$ and $\varepsilon > 0$.
Output: $\mathbf{g}_\varepsilon \in P$ and $\|\mathbf{g}_\varepsilon - \mathbf{g}\| \leq \varepsilon$.

Prior to considering how to implement such subroutines, we need to state some more requirements in the form of definitions.

Definition 2.7.3 (Admissibility of the stiffness matrix). Let $s^* > 0$. A bounded linear $\mathbf{M} : \ell_2 \to \ell_2$ is called *s^*-admissible*, when for a suitable routine **APPLY**, for each $s \in (0, s^*)$, for all $\mathbf{v} \in P$ and $\varepsilon > 0$, with $\mathbf{w}_\varepsilon := \mathbf{APPLY}[\mathbf{M}, \mathbf{v}, \varepsilon]$ the following is valid:

(i) $\#\operatorname{supp} \mathbf{w}_\varepsilon \lesssim \varepsilon^{-1/s} |\mathbf{v}|_{\mathcal{A}^s}^{1/s}$;

(ii) the number of arithmetic operations and storage locations required by the call is bounded by some absolute multiple of $\varepsilon^{-1/s} |\mathbf{v}|_{\mathcal{A}^s}^{1/s} + \#\operatorname{supp} \mathbf{v} + 1$. ⊘

Definition 2.7.4 (Admissibility of the right hand side). Let $s^* > 0$. A vector $\mathbf{g} \in \ell_2$ is called *s^*-admissible*, when for a suitable routine **RHS**, for each $s \in (0, s^*)$, for all $\varepsilon > 0$, with $\mathbf{g}_\varepsilon := \mathbf{RHS}[\mathbf{g}, \varepsilon]$ the following is valid:

(i) $\#\operatorname{supp} \mathbf{g}_\varepsilon \lesssim \varepsilon^{-1/s} |\mathbf{g}|_{\mathcal{A}^s}^{1/s}$;

(ii) the number of arithmetic operations and storage locations required by the call is bounded by some absolute multiple of $\varepsilon^{-1/s} |\mathbf{g}|_{\mathcal{A}^s}^{1/s} + 1$. ⊘

We recall the following result from [18, 28].

Proposition 2.7.5. *Let $\mathbf{M} : \ell_2 \to \ell_2$ be s^*-admissible for some $s^* > 0$. Then, for any $s \in (0, s^*)$, $\mathbf{M} : \mathcal{A}^s \to \mathcal{A}^s$ is bounded, and for $\mathbf{w}_\varepsilon := \mathbf{APPLY}[\mathbf{M}, \mathbf{v}, \varepsilon]$, we have $|\mathbf{w}_\varepsilon|_{\mathcal{A}^s} \lesssim |\mathbf{v}|_{\mathcal{A}^s}$ uniformly in $\varepsilon > 0$ and $\mathbf{v} \in P$.*

Similarly, if $\mathbf{g} \in \ell_2$ is s^*-admissible for some $s^* > 0$, then for any $s \in (0, s^*)$, $\mathbf{g} \in \mathcal{A}^s$, and for $\mathbf{g}_\varepsilon := \mathbf{RHS}[\mathbf{g}, \varepsilon]$, we have $|\mathbf{g}_\varepsilon|_{\mathcal{A}^s} \lesssim |\mathbf{g}|_{\mathcal{A}^s}$ uniformly in $\varepsilon > 0$.

Proof. It is immediately clear that $\mathbf{g} \in \mathcal{A}^s$. Next we will show that for any $s \in (0, s^*)$, $\mathbf{M} : \mathcal{A}^s \to \mathcal{A}^s$ is bounded. Let $C > 0$ be a constant such that for $\mathbf{w}_\varepsilon := \mathbf{APPLY}[\mathbf{M}, \mathbf{v}, \varepsilon]$, $\#\operatorname{supp} \mathbf{w}_\varepsilon \le C\varepsilon^{-1/s} |\mathbf{v}|_{\mathcal{A}^s}^{1/s}$. Let $\mathbf{v} \in \mathcal{A}^s$ and $N \in \mathbb{N}$ be given. For $\bar{\varepsilon} := C^s |\mathcal{B}_N(\mathbf{v})|_{\mathcal{A}^s} N^{-s}$, let $\mathbf{w}_{\bar{\varepsilon}} := \mathbf{APPLY}[\mathbf{M}, \mathcal{B}_N(\mathbf{v}), \bar{\varepsilon}]$. Then, by (2.3.3), we have

$$\begin{aligned} \|\mathbf{M}\mathbf{v} - \mathbf{w}_{\bar{\varepsilon}}\| &\le \|\mathbf{M}\mathcal{B}_N(\mathbf{v}) - \mathbf{w}_{\bar{\varepsilon}}\| + \|\mathbf{M}\| \|\mathbf{v} - \mathcal{B}_N(\mathbf{v})\| \\ &\le C^s |\mathcal{B}_N(\mathbf{v})|_{\mathcal{A}^s} N^{-s} + \|\mathbf{M}\| N^{-s} |\mathbf{v}|_{\mathcal{A}^s} \lesssim N^{-s} |\mathbf{v}|_{\mathcal{A}^s}. \end{aligned}$$

Since $\#\operatorname{supp} \mathbf{w}_\varepsilon \le N$, from (2.3.3) we infer that $|\mathbf{M}\mathbf{v}|_{\mathcal{A}^s} \lesssim |\mathbf{v}|_{\mathcal{A}^s}$.

With $\mathbf{w}_\varepsilon$ as above, by using Proposition 2.3.6 on page 18 we have $|\mathbf{w}_\varepsilon|_{\mathcal{A}^s} \lesssim |\mathbf{M}\mathbf{v}|_{\mathcal{A}^s} + (\#\operatorname{supp} \mathbf{w}_\varepsilon)^s \varepsilon \le |\mathbf{M}\mathbf{v}|_{\mathcal{A}^s} + C^s |\mathbf{v}|_{\mathcal{A}^s} \lesssim |\mathbf{v}|_{\mathcal{A}^s}$. Similarly, for $\mathbf{g}_\varepsilon := \mathbf{RHS}[\mathbf{g}, \varepsilon]$, we have $|\mathbf{g}_\varepsilon|_{\mathcal{A}^s} \lesssim |\mathbf{g}|_{\mathcal{A}^s} + (\#\operatorname{supp} \mathbf{g}_\varepsilon)^s \varepsilon \lesssim |\mathbf{g}|_{\mathcal{A}^s}$. ∎

With the subroutines **APPLY** and **RHS** at hand, we can define an approximate Richardson iteration that defines a valid procedure **ITERATE** in Algorithm 2.6.7 on page 28, and so provides an optimal adaptive algorithm. This algorithm was first introduced in the pioneering work [18].

Algorithm 2.7.6 The Richardson method $\mathbf{RICHARDSON}[\mathbf{v}, \nu, \eta] \to \mathbf{w}$

Parameters: Let ω be the damping parameter of Richardson's iteration (Example 2.5.1 on page 23), let $\rho < 1$ be the corresponding error reduction factor, and let $l \in \mathbb{N}$ be the smallest number such that $2\nu\rho^l \leq \eta$.

Input: Let $\mathbf{v} \in P$, $\nu \geq \|\mathbf{u} - \mathbf{v}\|$, and $\eta > 0$.

Output: $\mathbf{w} \in P$ with $\|\mathbf{u} - \mathbf{w}\| \leq \eta$.

1: $\mathbf{v}_0 := \mathbf{v}$;
2: **for** $i = 1$ to l **do**
3: $\quad \epsilon_i := \nu\rho^i/l$;
4: $\quad \mathbf{v}_i := \mathbf{RHS}[\omega\mathbf{f}, \epsilon_i/2] + \mathbf{APPLY}[\mathbf{I} - \omega\mathbf{A}, \mathbf{v}_{i-1}, \epsilon_i/2]$;
5: **end for**
6: $\mathbf{w} := \mathbf{v}_l$.

Theorem 2.7.7. *Let* $\mathbf{A}$ *be symmetric and positive definite, and let both* $\mathbf{A}$ *and* $\mathbf{f}$ *be* s^**-admissible for some* $s^* > 0$*. Then, for* $\mathbf{v} \in P$, $\nu \geq \|\mathbf{f} - \mathbf{A}\mathbf{v}\|$*, and* $\eta > 0$, $\mathbf{w} := \mathbf{RICHARDSON}[\mathbf{v}, \nu, \eta]$ *terminates with* $\|\mathbf{u} - \mathbf{w}\| \leq \eta$*. Moreover, the procedure* $\mathbf{ITERATE} := \mathbf{RICHARDSON}$ *with* $\|\cdot\|_\star := \|\cdot\|$ *satisfies the conditions of Theorem 2.6.8 on page 28 for any* $s \in (0, s^*)$*, meaning that* **RICHARDSON** *defines an optimal adaptive algorithm for* $s \in (0, s^*)$*.*

Proof. An application of Lemma 2.5.3 on page 24 guarantees that $\|\mathbf{u} - \mathbf{w}\| \leq 2\nu\rho^l \leq \eta$, latter inequality by construction.

As for the conditions of Theorem 2.6.8 on page 28, recall that we need to prove that for any $s \in (0, s^*)$ and for $\eta \gtrsim \nu \geq \|\mathbf{u} - \mathbf{v}\|$,

$$\#\operatorname{supp}\mathbf{w} \lesssim \#\operatorname{supp}\mathbf{v} + \eta^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}, \qquad |\mathbf{w}|_{\mathcal{A}^s} \lesssim (\#\operatorname{supp}\mathbf{v})^s\eta + |\mathbf{u}|_{\mathcal{A}^s},$$

and that the number of arithmetic operations and storage locations required by this call of **RICHARDSON** can be bounded by an absolute multiple of

$$\eta^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + \#\operatorname{supp}\mathbf{v} + 1.$$

For $1 \leq i \leq l$, from $\|\mathbf{u} - \mathbf{v}_i\| \leq \nu$ and Proposition 2.3.6 on page 18 we have

$$|\mathbf{v}_i|_{\mathcal{A}^s}^{1/s} \lesssim |\mathbf{u}|_{\mathcal{A}^s}^{1/s} + (\#\operatorname{supp}\mathbf{v}_i)\nu^{1/s}.$$

From $\nu\rho^{l-1} \gtrsim \eta$ we get $(1/\rho)^{l-1} \lesssim \nu/\eta \lesssim 1$ or $l \lesssim 1$, and so $\epsilon_i \gtrsim \nu\rho^{l-1}/l \gtrsim \eta/l \gtrsim \eta$. By using this and the s^*-admissibility of $\mathbf{f}$ and $\mathbf{A}$, we infer

$$\#\operatorname{supp}\mathbf{v}_i \lesssim \eta^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + \eta^{-1/s}|\mathbf{v}_{i-1}|_{\mathcal{A}^s}^{1/s}.$$

Taking into account the condition $\nu \lesssim \eta$, and repeatedly using the above two estimates, we get for $1 \le i \le l$,

$$|\mathbf{v}_i|_{\mathcal{A}^s}^{1/s} \lesssim |\mathbf{u}|_{\mathcal{A}^s}^{1/s} + |\mathbf{v}_0|_{\mathcal{A}^s}^{1/s},$$

and

$$\#\operatorname{supp}\mathbf{v}_i \lesssim \eta^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + |\mathbf{v}_0|_{\mathcal{A}^s}^{1/s}.$$

From Proposition 2.3.6 on page 18, we have $|\mathbf{v}_0|_{\mathcal{A}^s} \lesssim |\mathbf{u}|_{\mathcal{A}^s} + (\#\operatorname{supp}\mathbf{v}_0)^s\nu \lesssim |\mathbf{u}|_{\mathcal{A}^s} + (\#\operatorname{supp}\mathbf{v}_0)^s\eta$. By using the above estimates, and for bounding the cost of the algorithm, the s^*-admissibility of $\mathbf{f}$ and $\mathbf{A}$, we complete the proof. ∎

Now we address the question of how to implement the subroutine **APPLY**. We need the the notion of matrix computability.

Definition 2.7.8 (Computability). $\mathbf{M}$ is called *s^*-computable*, when for each $j \in \mathbb{N}_0$, we can *construct* an infinite matrix $\mathbf{M}_j$ having in each column and in each row at most $\alpha_j 2^j$ non-zero entries, *whose computation takes $\mathcal{O}(\alpha_j 2^j)$ arithmetic operations*, such that $\|\mathbf{M} - \mathbf{M}_j\| \le C_j$, where $(\alpha_j)_{j\in\mathbb{N}_0}$ is summable and for any $s < s^*$, $(C_j 2^{js})_{j\in\mathbb{N}_0}$ is summable. We call the matrices $\mathbf{M}_j$ the *compressed matrices*. ⊘

For a discussion on why s^*-computability can be expected for the stiffness matrices $\mathbf{M} = \mathbf{L}$, e.g., corresponding to Example 2.4.1 on page 20, we refer to the forthcoming Remark 2.7.13 on page 34.

Theorem 2.7.10 (cf. Proposition 3.8 of [83]). *If a matrix $\mathbf{M} : \ell_2 \to \ell_2$ is s^*-computable for some $s^* > 0$, then it is s^*-admissible.*

Proof. We employ the routine **APPLY** as presented in Algorithm 2.7.9 on the facing page. From (2.7.3), (2.7.1) and (2.7.2), we have

$$\begin{aligned}\|\mathbf{M}\mathbf{v} - \mathbf{w}\| &\le \sum_{k=0}^{\ell} \|\mathbf{M} - \mathbf{M}_{j-k}\|\|\mathbf{z}_j\| + \|\mathbf{M}\|\|\mathbf{v} - \sum_{k=0}^{\ell}\mathbf{z}_j\| \\ &\le \sum_{k=0}^{\ell} C_{j-k}\|\mathbf{z}_k\| + \varepsilon/2 \le \varepsilon.\end{aligned}$$

Algorithm 2.7.9 Realization of $\mathbf{APPLY}[\mathbf{M}, \mathbf{v}, \varepsilon] \to \mathbf{w}$

Parameters: For $j \in \mathbb{N}_0$, let C_j be such that $\|\mathbf{M} - \mathbf{M}_j\| \le C_j$.
Input: Let $\mathbf{M} : \ell_2 \to \ell_2$ be bounded linear, $\mathbf{v} \in P$ and $\varepsilon > 0$.
Output: $\mathbf{w} \in P$ and $\|\mathbf{w} - \mathbf{M}\mathbf{v}\| \le \varepsilon$.

1: $\{\mathbf{b}_i\}_{0 \le i \le q} := \mathbf{BSORT}[\mathbf{v}, \varepsilon/(2\|\mathbf{M}\|)]$;

2: For $k = 0, 1, \ldots$, generate vectors $\mathbf{z}_k$ by subsequently collecting $2^k - \lfloor 2^{k-1} \rfloor$ nonzero entries from $\cup_i \mathbf{b}_i$, starting from $\mathbf{b}_0$ and when it is exhausted from $\mathbf{b}_1$ and so on, until for some $k = \ell$ either $\cup_i \mathbf{b}_i$ becomes empty or

$$\|\mathbf{M}\| \|\mathbf{v} - \sum_{k=0}^{\ell} \mathbf{z}_k\| \le \varepsilon/2; \tag{2.7.1}$$

3: Compute the smallest $j \ge \ell$ such that

$$\sum_{k=0}^{\ell} C_{j-k} \|\mathbf{z}_k\| \le \varepsilon/2; \tag{2.7.2}$$

4: Compute

$$\mathbf{w} := \mathbf{M}_j \mathbf{z}_0 + \mathbf{M}_{j-1} \mathbf{z}_1 + \ldots + \mathbf{M}_{j-\ell} \mathbf{z}_\ell. \tag{2.7.3}$$

Let $s \in (0, s^*)$ be given. The number of operations needed for generating the vectors $\mathbf{z}_k$ is of order $\#\operatorname{supp} \mathbf{v} + \log(\varepsilon^{-1}\|\mathbf{v}\|) + 1 \lesssim \#\operatorname{supp} \mathbf{v} + \varepsilon^{-1/s} |\mathbf{v}|_{\mathcal{A}^s}^{1/s} + 1$. Both the number of operations needed for the evaluation of (2.7.3) and $\#\operatorname{supp} \mathbf{w}$ can be bounded by a multiple of $\sum_{k=0}^{\ell} \alpha_{j-k} 2^{j-k} 2^k \lesssim 2^j$.

Now we will bound 2^j. With $\mathbf{v}_k := \sum_{m=0}^{k} \mathbf{z}_m$, we have $\#\operatorname{supp} \mathbf{v}_k = 2^k$. Let $\tilde{\mathbf{v}}_k$ be constructed as follows: Create $\tilde{\mathbf{v}}_k$ by extracting nonzero entries first from $\mathbf{b}_0$ and when it is zero from $\mathbf{b}_1$ and so on, until $\|\mathbf{v} - \tilde{\mathbf{v}}_k\| \le \min\{\|\mathbf{v} - \mathbf{v}_k\|, \varepsilon/(2\|\mathbf{M}\|)\}$ is satisfied. Note that $\|\mathbf{v} - \mathbf{v}_k\| < \varepsilon/(2\|\mathbf{M}\|)$ for $k < \ell$. Then by construction, we have $2^{k-1} < \#\operatorname{supp} \tilde{\mathbf{v}}_k$. Since $\|\mathbf{b}_q\| \le \varepsilon/(2\|\mathbf{M}\|)$ by Lemma 2.6.2 on page 26, for $k \le \ell$, the last entry added to $\tilde{\mathbf{v}}_k$ originates from $\mathbf{b}_{i_k}$ with some $i_k < q$. Moreover, for $k \le \ell$, a minimal set Λ_k that satisfies $\|\mathbf{v} - \mathbf{P}_{\Lambda_k} \mathbf{v}\| \le \min\{\|\mathbf{v} - \mathbf{v}_k\|, \varepsilon/(2\|\mathbf{M}\|)\}$ contains all the entries from the vectors $\mathbf{b}_0, \ldots, \mathbf{b}_{i_k - 1}$. Since any two nonzero entries from $\mathbf{b}_{i_k}$ differ less than a factor $1/\beta$ in modulus, with β the constant inside **BSORT**, the cardinality of the contribution from $\mathbf{b}_{i_k}$ to $\operatorname{supp} \tilde{\mathbf{v}}_k$ is at most a factor $1/\beta^2$ larger than that to Λ_k, so that $\#\operatorname{supp} \tilde{\mathbf{v}}_k \le \beta^{-2} \#\Lambda_k$. By the same reasoning as in the proof of Proposition 2.3.7 on page 19, we conclude that $\#\Lambda_k \lesssim \|\mathbf{v} - \mathbf{v}_k\|^{-1/s} |\mathbf{v}|_{\mathcal{A}^s}^{1/s}$ or

$$\|\mathbf{v} - \mathbf{v}_k\| \lesssim (\#\Lambda_k)^{-s} |\mathbf{v}|_{\mathcal{A}^s} \lesssim 2^{-ks} |\mathbf{v}|_{\mathcal{A}^s}, \qquad \text{for } k < \ell,$$

and $2^{\ell-1} \lesssim \#\Lambda_\ell \lesssim \varepsilon^{-1/s}|\mathbf{v}|_{\mathcal{A}^s}$. The latter estimate gives a suitable bound on 2^j for $j = \ell$.

For $j > \ell$, from the definition of j we have

$$\varepsilon/2 < \sum_{k=0}^{\ell} C_{j-1-k}\|\mathbf{z}_k\| \leq \sum_{k=0}^{\ell} C_{j-1-k}\|\mathbf{v} - \mathbf{v}_{k-1}\| \lesssim \sum_{k=0}^{\ell} C_{j-1-k}2^{-(k-1)s}|\mathbf{v}|_{\mathcal{A}^s}$$
$$\lesssim 2^{-js}|\mathbf{v}|_{\mathcal{A}^s} \sum_{k=0}^{\ell} C_{j-1-k}2^{(j-k-1)s} \lesssim 2^{-js}|\mathbf{v}|_{\mathcal{A}^s},$$

which completes the proof. ∎

The notion of matrix compressibility further simplifies the notion of computability, by isolating the costs for computing matrix entries.

Definition 2.7.11 (Compressibility). $\mathbf{M}$ is called s^*-*compressible*, when for each $j \in \mathbb{N}_0$, there exists an infinite matrix $\mathbf{M}_j$, *constructed by dropping entries from* $\mathbf{M}$, such that in each column and in each row it has $\mathcal{O}(2^j)$ non-zero entries, and such that for any $s < s^*$, $\|\mathbf{M} - \mathbf{M}_j\| \lesssim 2^{-js}$. ⊘

Lemma 2.7.12 (cf. Remark 2.4 of [86]). *Let* $\mathbf{M}$ *be* s^**-compressible, and let the matrices* $\mathbf{M}_j$ *be as in Definition 2.7.11. In addition, for* $j \in \mathbb{N}_0$, *assume that each column and each row of* $\mathbf{M}_j$ *can be computed at the expense of* $\mathcal{O}(2^j)$ *arithmetic operations. Then* $\mathbf{M}$ *is* s^**-computable.*

Proof. For $j \in \mathbb{N}_0$, let $\tilde{\mathbf{M}}_j := \mathbf{M}_{\lceil j+\log_2 \alpha_j \rceil}$ with $\alpha_j = j^{-\epsilon}$ for some $\epsilon > 1$. Then the number of nonzero entries, as well as the cost of computing these entries in each column and each row of $\tilde{\mathbf{M}}_j$ is of order $2^j\alpha_j \lesssim 2^j j^{-\epsilon}$. Since for any $s < s' < s^*$ we have

$$2^{-js}\|\mathbf{M} - \tilde{\mathbf{M}}_j\| \lesssim 2^{-js}2^{-(j+\log \alpha_j)s'} = 2^{-j(s'-s)}\alpha_j^{-s'}$$

and $\sum_j 2^{-j(s'-s)}\alpha_j^{-s'} < \infty$, the proof is established. ∎

Remark 2.7.13. In this remark, we comment on why s^*-compressibility of a matrix $\mathbf{M}$ can be expected when $\mathbf{M}$ is the stiffness matrix corresponding to a differential operator in a wavelet basis. For simplicity, with $\Omega \subset \mathbb{R}^n$ a bounded Lipschitz domain and $H := H_0^1(\Omega)$, we will consider the Laplace operator $-\Delta : H \to H'$ and a wavelet basis Ψ for H. An element of the stiffness matrix is given by $\mathbf{M}_{\lambda\mu} = \langle \psi_\lambda, -\Delta\psi_\mu \rangle := \int_\Omega \nabla\psi_\lambda \nabla\psi_\mu$. First note that the matrix $\mathbf{M}$ is not sparse, since any wavelet will necessarily intersect with infinitely many higher level wavelets. Let us look more closely into the interactions between wavelets on different levels. Let $\mathbf{M}_{[j,k]} := (\mathbf{M}_{\lambda\mu})_{|\lambda|=j,|\mu|=k}$ be the block of $\mathbf{M}$ corresponding

to the interaction between the j-th and k-th levels. Then the number of rows or columns of $\mathbf{M}_{[j,k]}$ is of order 2^{jn} or 2^{kn}, respectively. For a given λ with $|\lambda| = j$, by the locality of the wavelets, the number of indices μ with $|\mu| = k$ for which $\operatorname{supp}\psi_\lambda \cap \operatorname{supp}\psi_\mu \neq \varnothing$ is of order $\max\{1, 2^{(k-j)n}\}$. We see that the block $\mathbf{M}_{[j,k]}$ is sparse (or nearly sparse) when the difference $|j-k|$ is small, and that the sparseness diminishes as the difference increases. Our strategy to compress the matrix $\mathbf{M}$ will be to discard blocks $\mathbf{M}_{[j,k]}$ for which $|j-k|$ is larger than a certain threshold. For $J \in \mathbb{N}$, let $\mathbf{M}_J$ be the matrix obtained from $\mathbf{M}$ by keeping only the blocks $\mathbf{M}_{[j,k]}$ with $|j-k| \leq J$. Then, the number of nonzero entries in each row and column of $\mathbf{M}_J$ is of order

$$\sum_{|j-k|\leq J} \max\{1, 2^{(k-j)n}\} \lesssim J + 2^{Jn} \lesssim 2^{Jn}. \tag{2.7.4}$$

Now we will estimate the error $\|\mathbf{M} - \mathbf{M}_J\|$. For any $r > 0$, $-\Delta : H^{1+r} \to H^{-1+r}$ is bounded. Using this, and the estimate (2.2.5) on page 11, for $w_j \in W_j$, $w_k \in W_k$, and $r \in (0, \tilde{d}+1] \cap (0, \gamma - 1)$, we have

$$\begin{aligned}\langle w_j, -\Delta w_k\rangle &\leq \|w_j\|_{H^{1-r}} \|\Delta w_k\|_{H^{-1+r}} \lesssim \|w_j\|_{H^{1-r}} \|w_k\|_{H^{1+r}} \\ &\lesssim 2^{r(k-j)} \|w_j\|_{H^1} \|w_k\|_{H^1}.\end{aligned}$$

and analogously by the self-adjointness of the Laplacian,

$$\langle w_j, -\Delta w_k\rangle = \langle -\Delta w_j, w_k\rangle \lesssim 2^{r(j-k)} \|w_j\|_{H^1} \|w_k\|_{H^1}.$$

So for r in the above range, and for arbitrary $\mathbf{v} \in \ell_2(\nabla_j \setminus \nabla_{j-1})$ and $\mathbf{w} \in \ell_2(\nabla_k \setminus \nabla_{k-1})$, we have

$$\begin{aligned}\langle \mathbf{v}, \mathbf{M}_{[j,k]}\mathbf{w}\rangle = \left\langle \mathbf{v}^T\Psi, -\Delta(\mathbf{w}^T\Psi)\right\rangle &\lesssim 2^{-r|j-k|} \left\|\mathbf{v}^T\Psi\right\|_{H^1} \left\|\mathbf{w}^T\Psi\right\|_{H^1} \\ &\lesssim 2^{-r|j-k|} \|\mathbf{v}\| \|\mathbf{w}\|,\end{aligned}$$

or $\|\mathbf{M}_{[j,k]}\| \lesssim 2^{-r|j-k|}$. Furthermore, with $\mathbf{P}_{[i]} := \mathbf{P}_{\nabla_i \setminus \nabla_{i-1}}$, for arbitrary $\mathbf{v}, \mathbf{w} \in \ell_2$ we have

$$\begin{aligned}\langle \mathbf{v}, (\mathbf{M} - \mathbf{M}_J)\mathbf{w}\rangle &= \sum_{|j-k|>J} \langle \mathbf{P}_{[j]}\mathbf{v}, \mathbf{M}_{[j,k]}\mathbf{P}_{[k]}\mathbf{w}\rangle \\ &\lesssim \sum_{|j-k|>J} 2^{-r|j-k|} \|\mathbf{P}_{[j]}\mathbf{v}\| \|\mathbf{P}_{[k]}\mathbf{w}\| \\ &\lesssim 2^{-rJ} \sqrt{\sum_{j=0}^{\infty} \|\mathbf{P}_{[j]}\mathbf{v}\|^2} \sqrt{\sum_{k=0}^{\infty} \|\mathbf{P}_{[k]}\mathbf{w}\|^2} \\ &= 2^{-rJ} \|\mathbf{v}\| \|\mathbf{w}\|,\end{aligned}$$

where in the third line we used $\|(2^{-r|j-k|})_{j,k}\|_{\ell_2\to\ell_2} < \infty$. We conclude that $\|\mathbf{M}-\mathbf{M}_J\| \lesssim 2^{-rJ}$ for $r \in (0, \tilde{d}+1] \cap (0, \gamma-1)$, and this, together with (2.7.4), implies that $\mathbf{M}$ is s^*-compressible with $s^* = \max\{\frac{\tilde{d}+1}{n}, \frac{\gamma-1}{n}\}$. ⊘

Remark 2.7.14. In view of Remark 2.3.3 on page 16, since, by imposing whatever smoothness conditions on the solution u generally the convergence rate of best N-term approximations cannot be higher than $\frac{d-t}{n}$, it is fully satisfactory if an adaptive wavelet algorithm is optimal for $s \in (0, \frac{d-t}{n}]$. To this end, considering Theorem 2.7.7 on page 31, it is necessary to show that the stiffness matrix $\mathbf{L}$ is s^*-computable for some $s^* > \frac{d-t}{n}$, since otherwise for a solution u that has sufficient Besov regularity, the computability will be the limiting factor. So in particular, since $\gamma < d$, the value of s^* from the previous remark is not satisfactory.

For both differential and singular integral operators, and piecewise polynomial wavelets that are sufficiently smooth and have sufficiently many vanishing moments, s^*-compressiblity for some $s^* > \frac{d-t}{n}$ has been demonstrated in [86]. These results are quoted in Chapter 7 and Chapter 8. For simplicity thinking of the Laplacian as in the previous remark, the key to obtaining these improved results on compressibility can be understood as follows. For piecewise polynomial wavelets, for a given ψ_λ, most of the wavelets ψ_μ, especially when $|\mu| \gg |\lambda|$, will have their support inside some patch on which ψ_λ is infinitely smooth, hence by the cancellation property giving an improved bound on the corresponding matrix entry, and only for ψ_μ with a support that intersects with the $(n-1)$-dimensional singular supports of ψ_λ, the estimate of the corresponding entry has to rely on the global smoothness parameter γ.

Yet, only in a few special cases, e.g., in the case of a differential operator with constant coefficients, entries of $\mathbf{L}$ can be computed exactly, in $\mathcal{O}(1)$ operations, so that s^*-compressibility immediately implies s^*-computability. In general, numerical quadrature is required to approximate the entries. In Chapter 7 and Chapter 8, considering both differential and singular integral operators, we will show that $\mathbf{L}$ is s^*-computable for the same value of s^* as for which it was shown to be s^*-compressible. ⊘

Remark 2.7.15. In view of Definition 2.7.4 on page 30, s^*-admissibility of $\mathbf{f}$ requires the availability of a sequence of approximations for $\mathbf{f}$ that converges with the rate s for any $s < s^*$. By Proposition 2.7.5, if $\mathbf{u} \in \mathcal{A}^s$ and $\mathbf{L}$ is s^*-admissible for some $s^* > s$, then $\mathbf{f} = \mathbf{L}\mathbf{u} \in \mathcal{A}^s$ with $|\mathbf{f}|_{\mathcal{A}^s} \lesssim |\mathbf{u}|_{\mathcal{A}^s}$, and so $\sup_N N^s\|\mathbf{f} - \mathcal{B}_N(\mathbf{f})\| \lesssim |\mathbf{u}|_{\mathcal{A}^s}$, which, however does not tell how to *construct* an approximation $\mathbf{g}$ which is qualitatively as good as $\mathcal{B}_N(\mathbf{f})$ with a comparable support size. In general, a realization of such an approximation depends on the right-hand side at hand, and it can be practically achieved by exploiting the local smoothness of f and the cancellation properties of the wavelets. See §3.4 for an example.

2.8 Approximate steepest descent iterations

In Algorithm 2.7.6 on page 31, being an approximate Richardson iteration, the user needs to provide estimates of the error reduction factor ρ and the optimal value of the damping parameter ω. For doing so, one has to estimate the extremal eigenvalues of $\mathbf{A}$. Since, in view of Example 2.5.2 on page 23, without requiring any user defined parameters, the steepest descent method automatically achieves the best error reduction factor of Richardson's iteration, a suitable implementation of the approximate steepest descent method would release the user from the task of accurately estimating the extremal eigenvalues. In the context of adaptive wavelet algorithms, the steepest descent method was first studied in [15]. In [28], the analysis was extended to the case where wavelet frames are used instead of a basis.

The following perturbation result on the steepest descent iteration is a quotation of [28, Proposition 3.2].

Proposition 2.8.1. *In the setting of Example 2.5.2 on page 23, for any* $\rho \in \left(\frac{\kappa(\mathbf{A})-1}{\kappa(\mathbf{A})+1}, 1\right)$*, there exists a* $\delta = \delta(\rho)$ *small enough, such that if* $\|\mathbf{r}-\tilde{\mathbf{r}}\| \leq \delta\|\tilde{\mathbf{r}}\|$ *and* $\|\mathbf{A}\tilde{\mathbf{r}} - \mathbf{z}\| \leq \delta\|\tilde{\mathbf{r}}\|$*, then with*

$$\mathbf{w} = \mathbf{v} + \frac{\langle \tilde{\mathbf{r}}, \tilde{\mathbf{r}} \rangle}{\langle \mathbf{z}, \tilde{\mathbf{r}} \rangle} \tilde{\mathbf{r}},$$

we have

$$|||\mathbf{u} - \mathbf{w}||| \leq \rho |||\mathbf{u} - \mathbf{v}|||, \qquad \textit{and} \qquad \frac{\langle \tilde{\mathbf{r}}, \tilde{\mathbf{r}} \rangle}{\langle \mathbf{z}, \tilde{\mathbf{r}} \rangle} \lesssim 1.$$

In view of this proposition, we introduce an algorithm that computes the residual with some prescribed relative error δ, unless the residual itself is less than the prescribed final tolerance $\varepsilon > 0$. Moreover, the residual is computed within an absolute error $\xi > 0$.

Algorithm 2.8.2 Residual computation $\mathbf{RES}[\mathbf{v}, \xi, \delta, \varepsilon] \to [\tilde{\mathbf{r}}, \nu]$

Input: $\mathbf{v} \in P$, $\delta \in (0,1)$, and $\xi, \varepsilon > 0$.

Output: $\tilde{\mathbf{r}} \in P$ and $\nu > 0$, such that with $\mathbf{r} := \mathbf{f} - \mathbf{L}\mathbf{v}$, $\nu \geq \|\mathbf{r}\|$, $\|\mathbf{r} - \tilde{\mathbf{r}}\| \leq \xi$, and either $\nu \leq \varepsilon$ or $\|\mathbf{r} - \tilde{\mathbf{r}}\| \leq \delta\|\tilde{\mathbf{r}}\|$.

1: $\zeta := 2\xi$;
2: **repeat**
3: $\quad \zeta := \zeta/2$;
4: $\quad \tilde{\mathbf{r}} := \mathbf{RHS}[\mathbf{f}, \zeta/2] - \mathbf{APPLY}[\mathbf{L}, \mathbf{v}, \zeta/2]$;
5: **until** $\nu := \|\tilde{\mathbf{r}}\| + \zeta \leq \varepsilon$ or $\zeta \leq \delta\|\tilde{\mathbf{r}}\|$.

Remark 2.8.3. If **RES** is called with a parameter ξ that it is outside $[\frac{\delta}{1+\delta}\|\mathbf{f}-\mathbf{L}\mathbf{v}\|, \frac{\delta}{1-\delta}\|\mathbf{f}-\mathbf{L}\mathbf{v}\|]$, then so is ζ at the first evaluation of $\tilde{\mathbf{r}}$, and from $\|\tilde{\mathbf{r}}\|-\zeta \leq \|\mathbf{f}-\mathbf{L}\mathbf{v}\| \leq \|\tilde{\mathbf{r}}\|+\zeta$, one infers that in this case either the second test in the **until**-clause will fail anyway, meaning that the first iteration of the **repeat**-loop is not of any use, or that the second test in the **until**-clause is always passed, but possibly with a tolerance that is unnecessarily small. We conclude that there is not much sense in calling **RES** with a value of ξ that is far outside $[\frac{\delta}{1+\delta}\|\mathbf{f}-\mathbf{L}\mathbf{v}\|, \frac{\delta}{1-\delta}\|\mathbf{f}-\mathbf{L}\mathbf{v}\|]$.

The following result can be extracted from [46, Theorem 2.4] or [28, Theorem 3.7].

Proposition 2.8.4. $[\tilde{\mathbf{r}},\nu] := \mathbf{RES}[\mathbf{v},\xi,\delta,\varepsilon]$ *terminates with* $\nu \geq \|\mathbf{r}\|$, *and either* $\nu \leq \varepsilon$ *or* $\|\mathbf{r}-\tilde{\mathbf{r}}\| \leq \delta\|\tilde{\mathbf{r}}\|$, *where* $\mathbf{r} := \mathbf{f}-\mathbf{L}\mathbf{v}$. *In addition, we have* $\nu \gtrsim \min\{\xi,\varepsilon\}$, $\|\mathbf{r}-\tilde{\mathbf{r}}\| \leq \xi$, *and in case* $\nu > \varepsilon$, $\nu \leq (1+\delta)\|\tilde{\mathbf{r}}\|$. *Furthermore, if, for some* $s>0$, $\mathbf{u} \in \mathcal{A}^s$, *and* $\mathbf{L}$ *and* $\mathbf{f}$ *are* s^**-admissible with* $s^* > s$, *then*

$$\#\operatorname{supp}\tilde{\mathbf{r}} \lesssim \min\{\xi,\nu\}^{-1/s}\left(|\mathbf{v}|_{\mathcal{A}^s}^{1/s}+|\mathbf{u}|_{\mathcal{A}^s}^{1/s}\right), \qquad |\tilde{\mathbf{r}}|_{\mathcal{A}^s} \lesssim |\mathbf{v}|_{\mathcal{A}^s}+|\mathbf{u}|_{\mathcal{A}^s},$$

and the number of arithmetic operations and storage locations required by the call is bounded by some absolute multiple of

$$\min\{\xi,\nu\}^{-1/s}\left(|\mathbf{v}|_{\mathcal{A}^s}^{1/s}+|\mathbf{u}|_{\mathcal{A}^s}^{1/s}+\xi^{1/s}(\#\operatorname{supp}\mathbf{v}+1)\right).$$

Proof. If at evaluation of the **until**-clause, $\zeta > \delta\|\tilde{\mathbf{r}}\|$, then $\|\tilde{\mathbf{r}}\|+\zeta < (\delta^{-1}+1)\zeta$. Since ζ is halved in each iteration, we infer that, if not by $\zeta \leq \delta\|\tilde{\mathbf{r}}\|$, **RES** will terminate by $\|\tilde{\mathbf{r}}\|+\zeta \leq \varepsilon$.

Since after any evaluation of $\tilde{\mathbf{r}}$ inside the algorithm, $\|\tilde{\mathbf{r}}-\mathbf{r}\| \leq \zeta$, any value of ν determined inside the algorithm is an upper bound on $\|\mathbf{r}\|$.

If the loop terminates in the first iteration, or the algorithm terminates with $\nu > \varepsilon$, then $\nu \gtrsim \min\{\xi,\varepsilon\}$. In the other case, let $\tilde{\mathbf{r}}^{\mathrm{old}} := \mathbf{RHS}[\zeta]-\mathbf{APPLY}[\mathbf{w},\zeta]$. We have $\|\tilde{\mathbf{r}}^{\mathrm{old}}\|+2\zeta > \varepsilon$ and $2\zeta > \delta\|\tilde{\mathbf{r}}^{\mathrm{old}}\|$, so that $\nu \geq \zeta > (2\delta^{-1}+2)^{-1}(\|\mathbf{r}^{\mathrm{old}}\|+2\zeta) > \frac{\delta\varepsilon}{2+2\delta}$.

The bound on $\#\operatorname{supp}\tilde{\mathbf{r}}$ and $|\tilde{\mathbf{r}}|_{\mathcal{A}^s}$ easily follows from the s^*-admissibility of $\mathbf{L}$ and $\mathbf{f}$, once we have shown that $\zeta \gtrsim \min\{\xi,\nu\}$. When the loop terminates in the first iteration, we have $\zeta=\xi$, and when the algorithm terminates with $\zeta \geq \delta\|\tilde{\mathbf{r}}\|$, we have $\zeta \gtrsim \nu$. In the other case, we have $\delta\|\tilde{\mathbf{r}}^{\mathrm{old}}\| < 2\zeta$ with $\tilde{\mathbf{r}}^{\mathrm{old}}$ as above, and so from $\|\tilde{\mathbf{r}}-\tilde{\mathbf{r}}^{\mathrm{old}}\| \leq \zeta+2\zeta$, we infer $\|\tilde{\mathbf{r}}\| \leq \|\tilde{\mathbf{r}}^{\mathrm{old}}\|+3\zeta < (2\delta^{-1}+3)\zeta$, so that $\nu < (2\delta^{-1}+4)\zeta$.

By the geometrical decrease of ζ inside the algorithm, and the s^*-admissi-bility of $\mathbf{L}$ and $\mathbf{f}$, the total cost of the call of **RES** can be bounded by some multiple

of $\zeta^{-1/s}(|\mathbf{v}|_{\mathcal{A}^s}^{1/s} + |\mathbf{u}|_{\mathcal{A}^s}^{1/s}) + K(\#\operatorname{supp}\mathbf{v} + 1)$, with ζ, $\tilde{\mathbf{r}}$ and ν having their values at termination and K being the number of calls of **APPLY** that were made. Taking into account the initial value of ζ, and again its geometrical decrease inside the algorithm, we have $K(\#\operatorname{supp}\mathbf{v} + 1) = K\xi^{-1/s}\xi^{1/s}(\#\operatorname{supp}\mathbf{v} + 1) \lesssim \zeta^{-1/s}\xi^{1/s}(\#\operatorname{supp}\mathbf{v} + 1)$. ∎

We are now ready to present the approximate steepest descent method.

Algorithm 2.8.5 Method of steepest descent $\mathbf{SD}[\mathbf{v}, \nu_0, \varepsilon] \to \mathbf{v}_i$

Parameters: Let $\delta = \delta(\rho)$ be the constant as in Proposition 2.8.1 on page 37 with some $\rho < 1$. Let $\theta > 0$ be a fixed constant.
Input: $\mathbf{v} \in P$, $\nu_0 \geq \|\mathbf{f} - \mathbf{A}\mathbf{v}\|$, and $\varepsilon > 0$.
Output: $\mathbf{v}_i \in P$ with $\|\mathbf{f} - \mathbf{A}\mathbf{v}_i\| \leq \varepsilon$.

$i := 0$, $\mathbf{v}_1 := \mathbf{v}$;
loop
  $i := i + 1$;
  $[\tilde{\mathbf{r}}_i, \nu_i] := \mathbf{RES}[\mathbf{v}_i, \theta\nu_{i-1}, \delta, \varepsilon]$, with $\mathbf{L} := \mathbf{A}$ inside **RES**;
  if $\nu_i \leq \varepsilon$ **then**
    Terminate the subroutine.
  end if
  $\mathbf{z}_i := \mathbf{APPLY}[\tilde{\mathbf{r}}_i, \delta\|\tilde{\mathbf{r}}_i\|]$;
  $\mathbf{v}_{i+1} := \mathbf{v}_i + \frac{\langle\tilde{\mathbf{r}}_i, \tilde{\mathbf{r}}_i\rangle}{\langle\mathbf{z}_i, \tilde{\mathbf{r}}_i\rangle}\tilde{\mathbf{r}}_i$;
end loop

Remark 2.8.6. We will see that at the call of $\mathbf{RES}[\mathbf{v}_i, \theta\nu_{i-1}, \delta, \varepsilon]$, it holds that $\|\mathbf{f} - \mathbf{A}\mathbf{v}_i\| \lesssim \nu_{i-1}$. Although for any fixed $\theta > 0$ the following theorem is valid, in view of Remark 2.8.3 on page 38 a suitable tuning of θ will result in quantitatively better results. Ideally, θ has the largest value for which the **repeat**-loop inside **RES** always terminates in one iteration.

Theorem 2.8.7. *Let* $\mathbf{A}$ *be symmetric and positive definite, and let both* $\mathbf{A}$ *and* $\mathbf{f}$ *be* s^**-admissible for some* $s^* > 0$. *Then* $\mathbf{w} := \mathbf{SD}[\mathbf{v}, \nu_0, \varepsilon]$ *terminates with* $\|\mathbf{f} - \mathbf{A}\mathbf{w}\| \leq \varepsilon$. *Moreover, the procedure* **ITERATE** := **SD** *with* $\|\cdot\|_\star := \|\mathbf{A}\cdot\|$ *satisfies the conditions of Theorem 2.6.8 on page 28 for any* $s \in (0, s^*)$, *meaning that, incorporated in the method* **SOLVE** *from Algorithm 2.6.7 on page 28,* **SD** *defines an optimal adaptive algorithm for* $s \in (0, s^*)$.

Proof. From the properties of **RES**, for any $\mathbf{v}_i$ determined inside the loop, we have $\nu_i \geq \|\mathbf{r}_i\|$, and with $\mathbf{r}_i := \mathbf{f} - \mathbf{A}\mathbf{v}_i$, either $\nu_i \leq \varepsilon$ or $\|\mathbf{r}_i - \tilde{\mathbf{r}}_i\| \leq \delta\|\tilde{\mathbf{r}}_i\|$. As long as $\nu_i > \varepsilon$, from $(1-\delta)\|\tilde{\mathbf{r}}_i\| \leq \|\mathbf{r}_i\| \leq (1+\delta)\|\tilde{\mathbf{r}}_i\|$ and $\nu_i \leq (1+\delta)\|\tilde{\mathbf{r}}_i\|$, we have $\nu_i \eqsim \|\tilde{\mathbf{r}}_i\| \eqsim \|\mathbf{r}_i\|$, and Proposition 2.8.1 on page 37 shows that $|||\mathbf{u} - \mathbf{v}_{i+1}||| \leq$

$\rho|\!|\!|\mathbf{u}-\mathbf{v}_i|\!|\!|$, or $\nu_i \lesssim \rho^i \nu_0$. This proves that the loop terminates after a finite number of iterations, say directly after the K-th call of **RES**.

As for the conditions of Theorem 2.6.8 on page 28, recall that we need to prove that for any $s \in (0, s^*)$ and for $\varepsilon \gtrsim \nu_0 \geq \|\mathbf{f} - \mathbf{A}\mathbf{v}\|$,

$$\#\operatorname{supp}\mathbf{w} \lesssim \#\operatorname{supp}\mathbf{v} + \varepsilon^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}, \qquad |\mathbf{w}|_{\mathcal{A}^s} \lesssim (\#\operatorname{supp}\mathbf{v})^s\varepsilon + |\mathbf{u}|_{\mathcal{A}^s},$$

and that the number of arithmetic operations and storage locations required by this call of **SD** can be bounded by an absolute multiple of

$$\varepsilon^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + \#\operatorname{supp}\mathbf{v} + 1.$$

Since, by $\nu_0 \lesssim \varepsilon$, K is uniformly bounded and $\|\mathbf{u} - \mathbf{v}_1\| \leq \nu_0 \lesssim \varepsilon$, for $1 \leq i < K$ it follows from Proposition 2.8.4 on page 38 that

$$|\mathbf{v}_{i+1}|_{\mathcal{A}^s} \lesssim |\mathbf{v}_i|_{\mathcal{A}^s} + |\mathbf{u}|_{\mathcal{A}^s} \lesssim |\mathbf{v}_1|_{\mathcal{A}^s} + |\mathbf{u}|_{\mathcal{A}^s} \lesssim |\mathbf{u}|_{\mathcal{A}^s} + (\#\operatorname{supp}\mathbf{v}_1)^s\varepsilon,$$

and therefore

$$\begin{aligned}\#\operatorname{supp}\mathbf{v}_{i+1} &\lesssim \#\operatorname{supp}\mathbf{v}_i + \varepsilon^{-1/s}\left(|\mathbf{v}_i|_{\mathcal{A}^s}^{1/s} + |\mathbf{u}|_{\mathcal{A}^s}^{1/s}\right)\\ &\lesssim \#\operatorname{supp}\mathbf{v}_1 + \varepsilon^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}.\end{aligned}$$

Note that the above two estimates are trivially true for $i = 0$.

For $1 \leq i < K$, Proposition 2.8.4 on page 38 shows that $|\tilde{\mathbf{r}}_i|_{\mathcal{A}^s} \lesssim |\mathbf{v}_i|_{\mathcal{A}^s} + |\mathbf{u}|_{\mathcal{A}^s} \lesssim (\#\operatorname{supp}\mathbf{v}_1)^s\varepsilon + |\mathbf{u}|_{\mathcal{A}^s}$, and using this we infer that the cost of the i-th iteration is bounded by an absolute multiple of

$$\begin{aligned}&\varepsilon^{-1/s}\left(|\mathbf{v}_i|_{\mathcal{A}^s}^{1/s} + |\mathbf{u}|_{\mathcal{A}^s}^{1/s} + \varepsilon^{1/s}(\#\operatorname{supp}\mathbf{v}_i + 1)\right)\\ &+ \varepsilon^{-1/s}\left((\#\operatorname{supp}\mathbf{v}_1)\varepsilon^{1/s} + |\mathbf{u}|_{\mathcal{A}^s}^{1/s}\right) \lesssim \varepsilon^{-1/s}|\mathbf{u}|_{\mathcal{A}^s} + \#\operatorname{supp}\mathbf{v}_1 + 1.\end{aligned}$$

The cost of the K-th call of **RES** can be bounded by some multiple of the same expression, and the proof is completed by the uniform boundedness of K. ∎

Remark 2.8.8. In Algorithm 2.8.5 on the preceding page, if we remove Line 8 and replace the statement in Line 9 by $\mathbf{v}_{i+1} := \mathbf{v}_i + \omega\tilde{\mathbf{r}}_i$, with ω having a value for which Richardson's iteration converges (cf. Example 2.5.1 on page 23), then we get another implementation of Richardson's iteration. The results of Theorem 2.8.7 on the previous page carries over to this case in a straightforward manner. The point is now we use *a posteriori* tolerances, whereas in Algorithm 2.7.6 on page 31 we used *a priori* tolerances.

Remark 2.8.9. The Chebyshev iteration can be used to accelerate the convergence of the aforementioned methods. Then a convergence proof is obtained by following the analysis in [49], with the help of the spectral theory for bounded self-adjoint operators.

Chapter 3

Adaptive Galerkin methods

3.1 Introduction

We consider the equation (2.4.3) on page 21, which is repeated here for convenience:

$$\mathbf{A}\mathbf{u} = \mathbf{f},$$

where $\mathbf{A} : \ell_2 \to \ell_2$ is an SPD matrix, and $\mathbf{f} \in \ell_2$. For $\Lambda \subset \nabla$, we call the solution $\mathbf{u}_\Lambda \in \ell_2(\Lambda)$ of the system

$$\mathbf{P}_\Lambda \mathbf{A} \mathbf{I}_\Lambda \mathbf{u}_\Lambda = \mathbf{P}_\Lambda \mathbf{f},$$

the *Galerkin solution* on Λ. We are going to exploit the fact that it is the best approximation in energy norm from $\ell_2(\Lambda)$, i.e.,

$$|||\mathbf{u} - \mathbf{u}_\Lambda||| = \inf_{\mathbf{v}_\Lambda \in \ell_2(\Lambda)} |||\mathbf{u} - \mathbf{v}_\Lambda|||,$$

and furthermore that $\mathbf{u}_\Lambda$ can be accurately approximated at relatively low cost. To this end, obviously we need some way to generate the index set Λ, or a sequence of increasingly larger index sets, that gives rise to an accurate approximation to the exact solution $\mathbf{u}$. One could use e.g. an approximate steepest descent iteration to create a sequence of index sets as follows: For a given approximation $\mathbf{v} \in P$, compute the next approximate steepest descent iterand $\mathbf{w} \in P$ as in Proposition 2.8.1 on page 37. Then take $\Lambda := \operatorname{supp} \mathbf{w}$ and compute the Galerkin solution $\mathbf{u}_\Lambda$ on Λ, to update $\mathbf{w}$. Now Proposition 2.8.1 guarantees convergence:

The work in this chapter is a joint work with Helmut Harbrecht and Rob Stevenson, see Section 1.2

$\|\mathbf{u}-\mathbf{u}_\Lambda\| \le \|\mathbf{u}-\mathbf{w}\| \le \rho\|\mathbf{u}-\mathbf{v}\|$ with $\rho < 1$. In fact, there is no need to compute $\mathbf{w}$; it suffices to compute the approximate residual $\tilde{\mathbf{r}}$ for $\mathbf{v}$, and then set $\Lambda := \operatorname{supp}\mathbf{v} \cup \operatorname{supp}\tilde{\mathbf{r}}$.

Since $\mathbf{u}_\Lambda$ is the best approximation to $\mathbf{u}$ in the energy norm, an analysis based on Proposition 2.8.1 is likely not sharp, however. An improved analysis can be made by employing the Galerkin orthogonality: $\|\mathbf{u}-\mathbf{u}_\Lambda\|^2 + \|\mathbf{u}_\Lambda-\mathbf{v}\|^2 = \|\mathbf{u}-\mathbf{v}\|^2$. This orthogonality shows the equivalence between the error reduction $\|\mathbf{u}-\mathbf{u}_\Lambda\| \le \xi\|\mathbf{u}-\mathbf{v}\|$ for some $\xi \in (0,1)$, and the so-called *saturation property* $\|\mathbf{u}_\Lambda-\mathbf{v}\| \ge (1-\xi^2)^{\frac{1}{2}}\|\mathbf{u}-\mathbf{v}\|$. It is well known, and recalled below in Lemma 3.2.1, that for a given initial approximation $\mathbf{v}$, any set $\Lambda \supset \operatorname{supp}\mathbf{v}$ satisfying $\|\mathbf{P}_\Lambda(\mathbf{f}-\mathbf{A}\mathbf{v})\| \ge \mu\|\mathbf{f}-\mathbf{A}\mathbf{v}\|$ for some constant $\mu \in (0,1)$, realizes the saturation property: $\|\mathbf{u}_\Lambda-\mathbf{v}\| \ge \kappa(\mathbf{A})^{-\frac{1}{2}}\mu\|\mathbf{u}-\mathbf{v}\|$. In [17], this property, combined with coarsening of the iterands, was used to obtain the first optimal adaptive wavelet algorithm.

The main point of this chapter is that we will show that if μ is less than $\kappa(\mathbf{A})^{-\frac{1}{2}}$, and Λ is the *smallest* set containing $\operatorname{supp}\mathbf{v}$ that satisfies the condition $\|\mathbf{P}_\Lambda(\mathbf{f}-\mathbf{A}\mathbf{v})\| \ge \mu\|\mathbf{f}-\mathbf{A}\mathbf{v}\|$, then, *without coarsening* of the iterands, these approximations converge with a rate that is guaranteed for best N-term approximations. Both conditions on the selection of Λ can be qualitatively understood as follows: The idea to realize the saturation property is the use of the coefficients of the residual vector as local error indicators. In case $\kappa(\mathbf{A}) = 1$, the residual is just a multiple of the error, but when $\kappa(\mathbf{A}) \gg 1$, only the largest coefficients can be used as reliable indicators about where the error is large. Of course, applying a larger set of indicators cannot reduce the convergence rate, but it may hamper optimal computational complexity. Notice the similarity with adaptive finite element methods where the largest local error indicators are used for marking elements for further refinement.

As we will see, the above result holds also true when the residuals and the Galerkin solutions are determined only inexactly, assuming a proper decay of the tolerances as the iteration proceeds, and when the cardinality of Λ is only minimal up to some constant factor. Using both generalizations, again a method of optimal computational complexity is obtained.

One might argue that picking the largest coefficients of the (approximate) residual vector is another instance of coarsening, but on a different place in the algorithm. The principle behind it, however, is very different from that behind coarsening of the iterands. What is more, since with the new method no information is deleted that has been created by a sequence of computations, we expect that it is more efficient.

Another modification to the method from [17] we will make is that for each call of **APPLY** or **RHS**, we will use as a tolerance some fixed multiple of the norm of

the current approximate residual, instead of using an a priori prescribed tolerance. Since it seems hard to avoid that a priori tolerances get either unnecessarily smaller, making the calls costly, or larger so that the perturbed iteration due to the inexact evaluations converges significantly slower than the unperturbed one, also here we expect to obtain a quantitative improvement.

This chapter is organized as follows. Before introducing our adaptive algorithm without coarsening of the iterands, in the next section, we will formulate an adaptive Galerkin algorithm as a valid instance of the subroutine **ITERATE** that is intended to be combined with coarsening of the iterands as in Algorithm 2.6.7 on page 28. Then in Section 3.3, we will introduce the adaptive algorithm without coarsening of the iterands and prove its optimality. We tested our adaptive wavelet solver for the Poisson equation on the interval. The results reported in the last section show that in this simple example the new method is indeed much more efficient than the inexact Richardson method with coarsening of the iterands. We would like to mention that in [30], numerical results based on tree approximations are given for singular integral equations on the boundary of three dimensional domains.

3.2 Adaptive Galerkin iterations

The next lemma is well known:

Lemma 3.2.1. *Let $\mu \in (0,1]$ be a constant. Let $\mathbf{v} \in \ell_2$ and let $\nabla \supseteq \Lambda \supset \operatorname{supp} \mathbf{v}$ be such that*

$$\|\mathbf{P}_\Lambda(\mathbf{f} - \mathbf{A}\mathbf{v})\| \geq \mu \|\mathbf{f} - \mathbf{A}\mathbf{v}\|. \tag{3.2.1}$$

Then, for $\mathbf{u}_\Lambda \in \ell_2(\Lambda)$ being the solution of the Galerkin system $\mathbf{P}_\Lambda \mathbf{A} \mathbf{u}_\Lambda = \mathbf{P}_\Lambda \mathbf{f}$, and with $\kappa(\mathbf{A}) := \|\mathbf{A}\|\|\mathbf{A}^{-1}\|$, we have

$$|||\mathbf{u} - \mathbf{u}_\Lambda||| \leq \left[1 - \kappa(\mathbf{A})^{-1}\mu^2\right]^{\frac{1}{2}} |||\mathbf{u} - \mathbf{v}|||.$$

Proof. We have

$$\begin{aligned} |||\mathbf{u}_\Lambda - \mathbf{v}||| &\geq \|\mathbf{A}\|^{-\frac{1}{2}} \|\mathbf{A}(\mathbf{u}_\Lambda - \mathbf{v})\| \geq \|\mathbf{A}\|^{-\frac{1}{2}} \|\mathbf{P}_\Lambda(\mathbf{f} - \mathbf{A}\mathbf{v})\| \\ &\geq \|\mathbf{A}\|^{-\frac{1}{2}} \mu \|\mathbf{f} - \mathbf{A}\mathbf{v}\| \geq \kappa(\mathbf{A})^{-\frac{1}{2}} \mu |||\mathbf{u} - \mathbf{v}|||, \end{aligned}$$

which, with $\kappa(\mathbf{A})^{-\frac{1}{2}}\mu$ reading as some arbitrary positive constant, is known as the saturation property of the space $\ell_2(\Lambda)$ containing $\mathbf{v}$. The proof is completed by using the Galerkin orthogonality $|||\mathbf{u} - \mathbf{v}|||^2 = |||\mathbf{u} - \mathbf{u}_\Lambda|||^2 + |||\mathbf{u}_\Lambda - \mathbf{v}|||^2$. ∎

In this lemma it was assumed to have full knowledge about the exact residual, and furthermore that the arising Galerkin system is solved exactly. As the following result shows, however, linear convergence is retained with an inexact evaluation of the residuals and an inexact solution of the Galerkin systems, in case the relative errors are sufficiently small.

Proposition 3.2.2. *Let* $0 < \delta < \alpha \leq 1$ *and* $0 < \gamma < \frac{1}{3}\kappa(\mathbf{A})^{-\frac{1}{2}}(\alpha - \delta)$. *Let* $\mathbf{v}, \tilde{\mathbf{r}} \in \ell_2$, $\nabla \supseteq \Lambda \supset \operatorname{supp} \mathbf{v}$, $\mathbf{w} \in \ell_2(\Lambda)$ *be such that, with* $\mathbf{r} := \mathbf{f} - \mathbf{A}\mathbf{v}$, $\|\mathbf{r} - \tilde{\mathbf{r}}\| \leq \delta\|\tilde{\mathbf{r}}\|$, $\|\mathbf{P}_\Lambda \tilde{\mathbf{r}}\| \geq \alpha\|\tilde{\mathbf{r}}\|$, *and* $\|\mathbf{P}_\Lambda(\mathbf{f} - \mathbf{A}\mathbf{w})\| \leq \gamma\|\tilde{\mathbf{r}}\|$. *Then, with* $\beta := \gamma\kappa(\mathbf{A})^{\frac{1}{2}}/(\alpha - \delta)$, *we have*

$$|||\mathbf{u} - \mathbf{w}||| \leq \left(1 - (1-\beta)(1-3\beta)\kappa(\mathbf{A})^{-1}\left(\tfrac{\alpha-\delta}{1+\delta}\right)^2\right)^{\frac{1}{2}} |||\mathbf{u} - \mathbf{v}|||.$$

Proof. From $\|\mathbf{r}\| \leq (1+\delta)\|\tilde{\mathbf{r}}\|$ and $\|\mathbf{P}_\Lambda \tilde{\mathbf{r}}\| \leq \|\mathbf{P}_\Lambda \mathbf{r}\| + \delta\|\tilde{\mathbf{r}}\|$ we have

$$\|\mathbf{P}_\Lambda \mathbf{r}\| \geq (\alpha - \delta)\|\tilde{\mathbf{r}}\| \geq \tfrac{\alpha-\delta}{1+\delta}\|\mathbf{r}\|,$$

so that Lemma 3.2.1 shows that

$$|||\mathbf{u} - \mathbf{u}_\Lambda||| \leq \left[1 - \kappa(\mathbf{A})^{-1}(\tfrac{\alpha-\delta}{1+\delta})^2\right]^{\frac{1}{2}} |||\mathbf{u} - \mathbf{v}|||.$$

One can simply estimate $|||\mathbf{u} - \mathbf{w}||| \leq |||\mathbf{u} - \mathbf{u}_\Lambda||| + |||\mathbf{u}_\Lambda - \mathbf{v}|||$, but a sharper result can be derived by using that $\mathbf{u} - \mathbf{w}$ is nearly $\langle\!\langle \cdot, \cdot \rangle\!\rangle$-orthogonal to $\ell_2(\Lambda)$, with $\langle\!\langle \cdot, \cdot \rangle\!\rangle := \langle \mathbf{A}\cdot, \cdot \rangle$. We have

$$\begin{aligned} |||\mathbf{u}_\Lambda - \mathbf{w}||| &\leq \|\mathbf{A}^{-1}\|^{\frac{1}{2}} \|\mathbf{P}_\Lambda \mathbf{A}(\mathbf{u}_\Lambda - \mathbf{w})\| = \|\mathbf{A}^{-1}\|^{\frac{1}{2}} \|\mathbf{P}_\Lambda(\mathbf{f} - \mathbf{A}\mathbf{w})\| \\ &\leq \|\mathbf{A}^{-1}\|^{\frac{1}{2}} \gamma\|\tilde{\mathbf{r}}\| \leq \|\mathbf{A}^{-1}\|^{\frac{1}{2}} \tfrac{\gamma}{\alpha-\delta}\|\mathbf{P}_\Lambda \mathbf{r}\| \leq \beta |||\mathbf{u}_\Lambda - \mathbf{v}|||. \end{aligned}$$

Using the Galerkin orthogonality $\mathbf{u} - \mathbf{u}_\Lambda \perp_{\langle\!\langle \cdot, \cdot \rangle\!\rangle} \ell_2(\Lambda)$, we have

$$\begin{aligned} \langle\!\langle \mathbf{u} - \mathbf{w}, \mathbf{w} - \mathbf{v} \rangle\!\rangle &= \langle\!\langle \mathbf{u}_\Lambda - \mathbf{w}, \mathbf{w} - \mathbf{v} \rangle\!\rangle \\ &\leq |||\mathbf{u}_\Lambda - \mathbf{w}||| |||\mathbf{w} - \mathbf{v}||| \leq \beta |||\mathbf{u}_\Lambda - \mathbf{v}||| |||\mathbf{w} - \mathbf{v}|||. \end{aligned}$$

Now by writing

$$|||\mathbf{u} - \mathbf{v}|||^2 = |||\mathbf{u} - \mathbf{w}|||^2 + |||\mathbf{w} - \mathbf{v}|||^2 + 2\langle\!\langle \mathbf{u} - \mathbf{w}, \mathbf{w} - \mathbf{v} \rangle\!\rangle,$$

and, for obtaining the second line in the following multi-line formula, twice applying

$$|||\mathbf{w} - \mathbf{v}||| \geq |||\mathbf{u}_\Lambda - \mathbf{v}||| - |||\mathbf{w} - \mathbf{u}_\Lambda||| \geq (1 - \beta)|||\mathbf{u}_\Lambda - \mathbf{v}|||,$$

and for the third line, using $\|\mathbf{u}_\Lambda - \mathbf{v}\| \geq \frac{\alpha-\delta}{1+\delta}\|\mathbf{u} - \mathbf{v}\|$, we find that

$$\begin{aligned}\|\mathbf{u} - \mathbf{v}\|^2 &\geq \|\mathbf{u} - \mathbf{w}\|^2 + \|\mathbf{w} - \mathbf{v}\|\big(\|\mathbf{w} - \mathbf{v}\| - 2\beta\|\mathbf{u}_\Lambda - \mathbf{v}\|\big)\\ &\geq \|\mathbf{u} - \mathbf{w}\|^2 + (1-\beta)(1-3\beta)\|\mathbf{u}_\Lambda - \mathbf{v}\|^2\\ &\geq \|\mathbf{u} - \mathbf{w}\|^2 + (1-\beta)(1-3\beta)\kappa(\mathbf{A})^{-1}(\tfrac{\alpha-\delta}{1+\delta})^2\|\mathbf{u} - \mathbf{v}\|^2,\end{aligned}$$

which completes the proof. ∎

An important ingredient of the adaptive method is the approximate solution of the Galerkin system on $\ell_2(\Lambda)$ for $\Lambda \subset \nabla$. Given an approximation $\mathbf{g}_\Lambda$ for $\mathbf{P}_\Lambda\mathbf{f}$, there are various possibilities to iteratively solving the system $\mathbf{P}_\Lambda\mathbf{A}\mathbf{I}_\Lambda\mathbf{u}_\Lambda = \mathbf{g}_\Lambda$ starting with some initial approximation $\mathbf{v}_\Lambda$ for $\mathbf{u}_\Lambda$. Thinking of Λ being an extension of $\operatorname{supp}\mathbf{v}$ as created in Proposition 3.2.2 on page 46, obviously we will take $\mathbf{v}_\Lambda = \mathbf{v}$. Note that even when the underlying operator is a differential operator, due to the fact that Λ can be in principle an arbitrary subset of ∇, it cannot be expected that the exact application of $\mathbf{A}_\Lambda := \mathbf{P}_\Lambda\mathbf{A}\mathbf{I}_\Lambda$ to a vector takes $\mathcal{O}(\#\Lambda)$ operations. So in order to end up with a method of optimal complexity we have to approximate this matrix-vector product. Instead of relying on the adaptive routine **APPLY** throughout the iteration, after approximately computing the initial residual using the **APPLY** routine, the following routine **GALSOLVE** iterates using some fixed, non-adaptive approximation for $\mathbf{A}_\Lambda$. The accuracy of this approximation depends only on the *factor* with which one wants to reduce the norm of the residual. This approach can be expected to be particularly efficient when the approximate computation of the entries of $\mathbf{A}$ is relatively expensive, as with singular integral operators. As can be deduced from [41], it is even possible in the course of the iteration to gradually diminish the accuracy of the approximation for $\mathbf{A}_\Lambda$.

Algorithm 3.2.3 Galerkin system solver $\mathbf{GALSOLVE}[\Lambda, \mathbf{g}_\Lambda, \mathbf{v}_\Lambda, \nu, \varepsilon] \to \mathbf{w}_\Lambda$

Parameters: Let $\mathbf{A} : \ell_2 \to \ell_2$ be SPD and s^*-computable for some $s^* > 0$. With $\mathbf{A}_j$ the compressed matrices from Definition 2.7.8 on page 32, let j be such that

$$\sigma := \|\mathbf{A} - \mathbf{A}_j\|\|\mathbf{A}^{-1}\| \leq \tfrac{\varepsilon}{3\varepsilon+3\nu}.$$

Input: Let $\Lambda \subset \nabla$, $\#\Lambda < \infty$, $\mathbf{g}_\Lambda, \mathbf{v}_\Lambda \in \ell_2(\Lambda)$, $\nu \geq \|\mathbf{g}_\Lambda - \mathbf{A}_\Lambda\mathbf{v}_\Lambda\|$, $\varepsilon > 0$.
Output: $\|\mathbf{g}_\Lambda - \mathbf{A}_\Lambda\mathbf{w}_\Lambda\| \leq \varepsilon$.

1: $\mathbf{B} := \mathbf{P}_\Lambda\frac{1}{2}(\mathbf{A}_j + \mathbf{A}_j^T)\mathbf{I}_\Lambda$;
2: $\mathbf{r}_0 := \mathbf{g}_\Lambda - \mathbf{P}_\Lambda\big(\mathbf{APPLY}[\mathbf{A}, \mathbf{v}_\Lambda, \frac{\varepsilon}{3}]\big)$;
3: To find an $\mathbf{x}$ with $\|\mathbf{r}_0 - \mathbf{B}\mathbf{x}\| \leq \frac{\varepsilon}{3}$, apply a suitable iterative method for solving $\mathbf{B}\mathbf{x} = \mathbf{r}_0$, e.g., Conjugate Gradients or Conjugate Residuals;
4: $\mathbf{w}_\Lambda := \mathbf{v}_\Lambda + \mathbf{x}$.

Proposition 3.2.4. *Let* $\mathbf{A}$ *be* s^**-computable for some* $s^* > 0$*. Then* $\mathbf{w}_\Lambda := \mathbf{GALSOLVE}[\Lambda, \mathbf{g}_\Lambda, \mathbf{v}_\Lambda, \nu, \varepsilon]$ *terminates with* $\|\mathbf{g}_\Lambda - \mathbf{A}_\Lambda \mathbf{v}_\Lambda\| \leq \varepsilon$*, and for any* $s < s^*$*, the number of arithmetic operations and storage locations required by the call is bounded by some absolute multiple of* $\varepsilon^{-1/s}|\mathbf{v}_\Lambda|_{\mathcal{A}^s}^{1/s} + c(\nu/\varepsilon)\#\Lambda + 1$*, where* $c : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ *is some non-decreasing function.*

Proof. Using that $\langle \mathbf{A}_\Lambda \mathbf{z}_\Lambda, \mathbf{z}_\Lambda \rangle \geq \|\mathbf{A}^{-1}\|^{-1}\|\mathbf{v}_\Lambda\|^2$ for $\mathbf{z} \in \ell_2(\Lambda)$, and $\|\mathbf{A}_\Lambda - \mathbf{B}\| \leq \|\mathbf{A} - \mathbf{A}_j\| = \sigma\|\mathbf{A}^{-1}\|^{-1} < \frac{1}{3}\|\mathbf{A}^{-1}\|^{-1}$, we infer that $\mathbf{B}$ is SPD with respect to the canonical scalar product on $\ell_2(\Lambda)$, and that $\kappa(\mathbf{B}) \lesssim 1$ uniformly in ε and ν. Writing $\mathbf{B}^{-1} = (\mathbf{I} - \mathbf{A}_\Lambda^{-1}(\mathbf{A}_\Lambda - \mathbf{B}))^{-1}\mathbf{A}_\Lambda^{-1}$, we find that $\|\mathbf{B}^{-1}\| \leq \frac{\|\mathbf{A}_\Lambda^{-1}\|}{1-\|\mathbf{A}_\Lambda^{-1}\|\|\mathbf{A}_\Lambda - \mathbf{B}\|}$ and so $\|\mathbf{A}_\Lambda - \mathbf{B}\|\|\mathbf{B}^{-1}\| \leq \frac{\sigma}{1-\sigma}$.

We have $\|\mathbf{r}_0\| \leq \nu + \frac{\varepsilon}{3}$. Writing

$$\mathbf{g}_\Lambda - \mathbf{A}_\Lambda \mathbf{w}_\Lambda = (\mathbf{g}_\Lambda - \mathbf{A}_\Lambda \mathbf{v}_\Lambda - \mathbf{r}_0) + (\mathbf{r}_0 - \mathbf{B}\mathbf{x}) + (\mathbf{B} - \mathbf{A}_\Lambda)\mathbf{B}^{-1}(\mathbf{r}_0 + \mathbf{B}\mathbf{x} - \mathbf{r}_0),$$

we find

$$\|\mathbf{A}_\Lambda \mathbf{w}_\Lambda - \mathbf{g}_\Lambda\| \leq \tfrac{\varepsilon}{3} + \tfrac{\varepsilon}{3} + \tfrac{\sigma}{1-\sigma}(\nu + \tfrac{\varepsilon}{3} + \tfrac{\varepsilon}{3}) \leq \varepsilon.$$

The s^*-computability of $\mathbf{A}$ show that the cost of the computation of $\mathbf{r}_0$ is bounded by some multiple of $\varepsilon^{-1/s}|\mathbf{v}_\Lambda|_{\mathcal{A}^s}^{1/s} + \#\Lambda$. Since $\mathbf{B}$ is sparse and can be constructed in $\mathcal{O}(\#\Lambda)$ operations, and the required number of iterations of the iterative method is bounded, everything only dependent on an upper bound for ν/ε, the proof is complete. ∎

As announced in the introduction of this chapter, before introducing our adaptive algorithm without coarsening of the iterands, we present an adaptive Galerkin algorithm which, combined with coarsening of the iterands as in Algorithm 2.6.7 on page 28, provides an optimal adaptive algorithm. We use the subroutine **RES** given in Algorithm 2.8.2 on page 37 for the computation of the approximate residuals with sufficiently small relative errors.

Algorithm 3.2.5 Adaptive Galerkin method $\mathbf{GALERKIN}[\mathbf{v}, \nu_0, \varepsilon] \to \mathbf{v}_i$

Parameters: Let $0 < \delta < 1$ and $0 < \gamma < \frac{1}{6}\kappa(\mathbf{A})^{-\frac{1}{2}}(1-\delta)$. Let $\theta > 0$ be a fixed constant.
Input: Let $\mathbf{v} \in P$, $\nu_0 \geq \|\mathbf{f} - \mathbf{A}\mathbf{v}\|$, and $\varepsilon > 0$.
Output: $\mathbf{v}_i \in P$ with $\|\mathbf{f} - \mathbf{A}\mathbf{v}_i\| \leq \varepsilon$.

$i := 0$, $\mathbf{v}_1 := \mathbf{v}$;
loop
  $i := i + 1$;
  $[\tilde{\mathbf{r}}_i, \nu_i] := \mathbf{RES}[\mathbf{v}_i, \theta\nu_{i-1}, \delta, \varepsilon]$, with $\mathbf{L} := \mathbf{A}$ inside **RES**;
  if $\nu_i \leq \varepsilon$ **then**
    Terminate the subroutine.
  end if
  $\Lambda_{i+1} := \operatorname{supp} \mathbf{v}_i \cup \operatorname{supp} \tilde{\mathbf{r}}_i$;
  $\mathbf{g}_{i+1} := \mathbf{P}_{\Lambda_{i+1}} (\mathbf{RHS}[\mathbf{f}, \gamma\|\tilde{\mathbf{r}}_i\|])$;
  $\mathbf{v}_{i+1} := \mathbf{GALSOLVE}[\Lambda_{i+1}, \mathbf{g}_{i+1}, \mathbf{v}_i, \gamma\|\tilde{\mathbf{r}}_i\| + \nu_i, \gamma\|\tilde{\mathbf{r}}_i\|]$;
end loop

Remark 3.2.6. Given $\mathbf{v}_i$, the index set Λ_{i+1} is the same as the support of the next iterand in an approximate steepest descent iteration. Although one could apply Proposition 2.8.1 on page 37 to analyze its convergence, we use Proposition 3.2.2 on page 46 to get a sharper result. It is clear that the above algorithm corresponds to the case $\alpha = 1$ in Proposition 3.2.2. In the next section, we will explore the possibility $\alpha < 1$.

Theorem 3.2.7. *Let* $\mathbf{A}$ *be* s^**-computable, and let* $\mathbf{f}$ *be* s^**-admissible for some* $s^* > 0$. *Then* $\mathbf{w} := \mathbf{GALERKIN}[\mathbf{v}, \nu_0, \varepsilon]$ *terminates with* $\|\mathbf{f} - \mathbf{A}\mathbf{w}\| \leq \varepsilon$. *Moreover, the procedure* **ITERATE** := **GALERKIN** *with* $\|\cdot\|_\star := \|\mathbf{A}\cdot\|$ *satisfies the conditions of Theorem 2.6.8 on page 28 for any* $s \in (0, s^*)$, *meaning that* **SOLVE** *presented in Algorithm 2.6.7 on page 28 using this* **ITERATE** *defines an optimal adaptive algorithm for* $s \in (0, s^*)$.

Proof. From the properties of **RES**, for any $\mathbf{v}_i$ determined inside the loop, with $\mathbf{r}_i := \mathbf{f} - \mathbf{A}\mathbf{v}_i$, we have $\nu_i \geq \|\mathbf{r}_i\|$, and either $\nu_i \leq \varepsilon$ or $\|\mathbf{r}_i - \tilde{\mathbf{r}}_i\| \leq \delta\|\tilde{\mathbf{r}}_i\|$. We have $\|\mathbf{P}_{\Lambda_{i+1}}\tilde{\mathbf{r}}_i\| \geq \alpha\|\tilde{\mathbf{r}}_i\|$ with $\alpha = 1$, $\Lambda_{i+1} \supseteq \operatorname{supp} \mathbf{v}_i$, and

$$\|\mathbf{g}_{i+1} - \mathbf{P}_{\Lambda_{i+1}}\mathbf{A}\mathbf{v}_i\| \leq \|\mathbf{P}_{\Lambda_{i+1}}(\mathbf{g}_{i+1} - \mathbf{f})\| + \|\mathbf{P}_{\Lambda_{i+1}}(\mathbf{f} - \mathbf{A}\mathbf{v}_i)\| \leq \gamma\|\tilde{\mathbf{r}}_i\| + \nu_i.$$

As long as $\nu_i > \varepsilon$, from $(1-\delta)\|\tilde{\mathbf{r}}_i\| \leq \|\mathbf{r}_i\| \leq (1+\delta)\|\tilde{\mathbf{r}}_i\|$ and $\nu_i \leq (1+\delta)\|\tilde{\mathbf{r}}_i\|$, we have $\nu_i \eqsim \|\tilde{\mathbf{r}}_i\| \eqsim \|\mathbf{r}_i\|$, and Proposition 3.2.2 shows that $|||\mathbf{u} - \mathbf{v}_{i+1}||| \leq \rho|||\mathbf{u} - \mathbf{v}_i|||$ with some $\rho \in [0, 1)$, or $\nu_{i+1} \lesssim \rho^{i-k}\nu_k$ for $0 \leq k \leq i+1$. This proves that the

loop terminates after a finite number of iterations, say directly after the K-th call of **RES**.

As for the conditions of Theorem 2.6.8, recall that we need to prove that for any $s \in (0, s^*)$ and for $\varepsilon \gtrsim \nu_0 \geq \|\mathbf{f} - \mathbf{A}\mathbf{v}\|$,

$$\#\operatorname{supp}\mathbf{w} \lesssim \#\operatorname{supp}\mathbf{v} + \varepsilon^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}, \qquad |\mathbf{w}|_{\mathcal{A}^s} \lesssim (\#\operatorname{supp}\mathbf{v})^s\varepsilon + |\mathbf{u}|_{\mathcal{A}^s},$$

and that the number of arithmetic operations and storage locations required by this call of **GALERKIN** can be bounded by an absolute multiple of

$$\varepsilon^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + \#\operatorname{supp}\mathbf{v} + 1.$$

These conditions are trivially true for $K = 1$, and from now on we will assume that $K > 1$. For $1 \leq i < K$, from Proposition 2.3.6 on page 18 we have, with $\Lambda_1 := \operatorname{supp}\mathbf{v}_1$,

$$|\mathbf{v}_i|_{\mathcal{A}^s} \lesssim |\mathbf{u}|_{\mathcal{A}^s} + (\#\Lambda_i)^s\nu_i, \tag{3.2.2}$$

and since $\#\Lambda_{i+1} - \#\Lambda_i \leq \#\operatorname{supp}\tilde{\mathbf{r}}_i$ for $1 \leq i < K$, by applying Proposition 2.8.4 on page 38 we have, for $1 \leq k < K$,

$$\begin{aligned}\#\Lambda_{k+1} &\leq \#\Lambda_1 + \textstyle\sum_{i=1}^{k}(\#\Lambda_{i+1} - \#\Lambda_i)\\ &\lesssim \#\Lambda_1 + \textstyle\sum_{i=1}^{k}\nu_i^{-1/s}\left(|\mathbf{v}_i|_{\mathcal{A}^s}^{1/s} + |\mathbf{u}|_{\mathcal{A}^s}^{1/s}\right)\\ &\lesssim \#\Lambda_1 + \nu_k^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + \textstyle\sum_{i=1}^{k}\#\Lambda_i,\end{aligned} \tag{3.2.3}$$

where in the last line we have used (3.2.2) and the fact that ν_i is geometrically decreasing.

We have $\nu_i \gtrsim \min\{\theta\nu_{i-1}, \varepsilon\} \gtrsim \varepsilon$ for $1 < i \leq K$. We claim that $\#\Lambda_{k+1} \lesssim \#\Lambda_1 + \nu_k^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}$ for $1 \leq k < K$, and prove it by induction. Since ν_i is geometrically decreasing, using (3.2.3) we infer

$$\begin{aligned}\#\Lambda_{k+1} &\lesssim \#\Lambda_1 + \nu_k^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + \textstyle\sum_{i=1}^{k}\left(\#\Lambda_1 + \nu_{i-1}^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}\right)\\ &\lesssim k\#\Lambda_1 + \nu_k^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s},\end{aligned} \tag{3.2.4}$$

which proves the claim since we have $K \lesssim 1$ by the condition $\varepsilon \gtrsim \nu_0$. This claim, together with (3.2.2) and $\nu_i \eqsim \varepsilon$, proves the bounds on $\#\operatorname{supp}\mathbf{w}$ and $|\mathbf{w}|_{\mathcal{A}^s}$.

Now it remains to bound the cost of the algorithm. For $1 \leq i \leq K$, we have

$$|\mathbf{v}_i|_{\mathcal{A}^s}^{1/s} \lesssim |\mathbf{u}|_{\mathcal{A}^s}^{1/s} + (\#\Lambda_1)\nu_i^{1/s}, \tag{3.2.5}$$

from (3.2.2) with the help of (3.2.4) when $K > 1$. By Proposition 2.8.4 on page 38, the cost of the i-th call of **RES** for $1 \le i \le K$ is bounded by

$$\begin{aligned}
&\nu_i^{-1/s}\left(|\mathbf{v}_i|_{\mathcal{A}^s}^{1/s} + |\mathbf{u}|_{\mathcal{A}^s}^{1/s} + \nu_{i-1}^{1/s}(\#\operatorname{supp}\mathbf{v}_i + 1)\right)\\
&\lesssim \nu_i^{-1/s}\left(|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + (\#\Lambda_i)\nu_i^{1/s} + \nu_{i-1}^{1/s}(\#\Lambda_1 + 1)\right)\\
&\lesssim \nu_i^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + (\tfrac{\nu_{i-1}}{\nu_i})^{1/s}(\#\Lambda_1 + 1),
\end{aligned}$$

where we used (3.2.2), $\#\operatorname{supp}\mathbf{v}_i \le \#\Lambda_i \lesssim \#\Lambda_1 + \nu_{i-1}^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}$, and $\nu_i \lesssim \nu_{i-1}$. Now taking into account that $\nu_i \gtrsim \varepsilon \gtrsim \nu_0 \gtrsim \nu_{i-1}$, and summing over $1 \le i \le K$, we conclude that the total cost of the $K \lesssim 1$ calls of **RES** is bounded by an absolute multiple of $\varepsilon^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + \#\Lambda_1 + 1$.

By Proposition 2.8.4 and (3.2.5), we have $\operatorname{supp}\tilde{\mathbf{r}}_i \lesssim \nu_i^{1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + \#\Lambda_1$. The cost of the i-th iteration with the cost of **RES** removed, for $1 \le i < K$, is bounded by an absolute multiple of

$$\begin{aligned}
&1 + \#\Lambda_1 + \nu_i^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + \nu_i^{-1/s}|\mathbf{v}_i|_{\mathcal{A}^s}^{1/s} + \#\Lambda_{i+1}c(1 + \tfrac{2\nu_i}{\gamma\|\tilde{\mathbf{r}}_i\|})\\
&\lesssim 1 + \#\Lambda_1 + \nu_i^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s},
\end{aligned}$$

where we have used $\nu_i \eqsim \|\tilde{\mathbf{r}}_i\|$, (3.2.5), (3.2.4), and $c : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ is the non-decreasing function from Proposition 3.2.4 on page 48. The proof is completed by summing the above cost over $1 \le i < K$ and using that $K \lesssim 1$. ∎

Remark 3.2.8. Inside the call of $[\tilde{\mathbf{r}}_i, \nu_i] := \mathbf{RES}[\mathbf{v}_i, \theta\nu_{i-1}, \delta, \varepsilon]$ that is made in **GALERKIN**, we search an approximation

$$\tilde{\mathbf{r}}_{i,\zeta} := \mathbf{RHS}[\mathbf{f}, \zeta/2] - \mathbf{APPLY}[\mathbf{A}, \mathbf{v}_i, \zeta/2]$$

for $\mathbf{r}_i := \mathbf{f} - \mathbf{A}\mathbf{v}_i$ with a $\zeta \le \delta\|\tilde{\mathbf{r}}_{i,\zeta}\|$ that is as large as possible in order to minimize the support of $\tilde{\mathbf{r}}_{i,\zeta}$ outside $\operatorname{supp}\mathbf{v}_i$. When $i > 0$, because of the preceding calls of **RHS** and **GALSOLVE**, we have a set $\Lambda_i \supset \operatorname{supp}\mathbf{v}_i$ and a $\tilde{\mathbf{r}}_{i-1}$ with $\|\mathbf{P}_{\Lambda_i}\mathbf{r}_i\| \le \epsilon_i := \gamma\|\tilde{\mathbf{r}}_{i-1}\|$. In this remark, we investigate whether it is possible to benefit from this information to obtain an approximation for the residual with relative error not exceeding δ whose support extends less outside $\operatorname{supp}\mathbf{v}_i$.

Let $\tilde{\mathbf{r}}_{i,\zeta}^I := \mathbf{P}_{\Lambda_i}\tilde{\mathbf{r}}_{i,\zeta}$ and $\tilde{\mathbf{r}}_{i,\zeta}^E := \mathbf{P}_{\nabla\setminus\Lambda_i}\tilde{\mathbf{r}}_{i,\zeta}$, and similarly $\mathbf{r}_i^I$ and $\mathbf{r}_i^E$. From

$$\begin{aligned}
\zeta^2 &\ge \|\mathbf{r}_i - \tilde{\mathbf{r}}_{i,\zeta}\|^2 = \|\mathbf{r}_i^I - \tilde{\mathbf{r}}_{i,\zeta}^I\|^2 + \|\mathbf{r}_i^E - \tilde{\mathbf{r}}_{i,\zeta}^E\|^2\\
&\ge (\|\tilde{\mathbf{r}}_{i,\zeta}^I\| - \epsilon_i)^2 + \|\mathbf{r}_i^E - \tilde{\mathbf{r}}_{i,\zeta}^E\|^2,
\end{aligned}$$

we have

$$\|\mathbf{r}_i - \tilde{\mathbf{r}}_{i,\zeta}^E\| = (\|\mathbf{r}_i^E - \tilde{\mathbf{r}}_{i,\zeta}^E\|^2 + \|\mathbf{r}_i^I\|^2)^{\frac{1}{2}} \le (\zeta^2 - (\|\tilde{\mathbf{r}}_{i,\zeta}^I\| - \epsilon_i)^2 + \epsilon_i^2)^{\frac{1}{2}} =: \breve{\zeta}.$$

So, alternatively, instead of $\tilde{\mathbf{r}}_{i,\zeta}$, we may use $\tilde{\mathbf{r}}^E_{i,\zeta}$ as an approximation for $\mathbf{r}_i$, and thus stop the routine **RES** as soon as $\nu_i := \|\tilde{\mathbf{r}}^E_{i,\zeta}\| + \breve{\zeta} \le \varepsilon$ or $\breve{\zeta} \le \delta\|\tilde{\mathbf{r}}^E_{i,\zeta}\|$, and use $\tilde{\mathbf{r}}^E_{i,\zeta}$ also for the determination of Λ_{i+1}. Since for any ζ and $\tilde{\mathbf{r}}_{i,\zeta}$ with $\tilde{\mathbf{r}}^I_{i,\zeta} \neq 0$ and $\zeta < \|\tilde{\mathbf{r}}_{i,\zeta}\|$ it holds that $\breve{\zeta}\|\mathbf{r}_{i,\zeta}\| < \zeta\|\mathbf{r}^E_{i,\zeta}\|$ if ϵ_i is small enough, under this condition the alternative test is passed more easily. This may even be a reason to decrease the parameter γ.

The approach discussed in this remark has been applied in the experiments reported in [30]. ⊘

3.3 Optimal complexity *without* coarsening of the iterands

Now we come to the main part of this chapter. So far we relied on coarsening of the iterands to control their support sizes. Below we will show that, after a small change, **GALERKIN** produces approximate solutions with optimal convergence rate without such coarsening.

In the following key lemma, it is shown that for sufficiently small μ and $\mathbf{u} \in \mathcal{A}^s$, for a set Λ as in Lemma 3.2.1 on page 45 that has *minimal* cardinality, $\#(\Lambda \setminus \operatorname{supp} \mathbf{v})$ can be bounded in terms of $\|\mathbf{f} - \mathbf{A}\mathbf{v}\|$ and $|\mathbf{u}|_{\mathcal{A}^s}$ only, i.e., *independently* of $|\mathbf{v}|_{\mathcal{A}^s}$ and the value of s^* (cf. (3.2.3) on page 50 and [17, §4.2-4.3]).

Lemma 3.3.1. *Let $\mu \in (0, \kappa(\mathbf{A})^{-\frac{1}{2}})$ be a constant, $\mathbf{v} \in P$, and for some $s > 0$, $\mathbf{u} \in \mathcal{A}^s$. Then the* smallest *set $\Lambda \supset \operatorname{supp} \mathbf{v}$ with*

$$\|\mathbf{P}_\Lambda(\mathbf{f} - \mathbf{A}\mathbf{v})\| \ge \mu\|\mathbf{f} - \mathbf{A}\mathbf{v}\|$$

satisfies

$$\#(\Lambda \setminus \operatorname{supp} \mathbf{v}) \lesssim \|\mathbf{f} - \mathbf{A}\mathbf{v}\|^{-1/s}|\mathbf{u}|^{1/s}_{\mathcal{A}^s}. \tag{3.3.1}$$

Proof. Let $\lambda > 0$ be a constant with $\mu \le \kappa(\mathbf{A})^{-\frac{1}{2}}(1 - \|\mathbf{A}\|\lambda^2)^{\frac{1}{2}}$. Let N be the smallest integer such that a best N-term approximation $\mathbf{u}_N$ for $\mathbf{u}$ satisfies $\|\mathbf{u} - \mathbf{u}_N\| \le \lambda|||\mathbf{u} - \mathbf{v}|||$. Since $|||\mathbf{u} - \mathbf{v}||| \ge \|\mathbf{A}\|^{-\frac{1}{2}}\|\mathbf{f} - \mathbf{A}\mathbf{v}\|$, we have

$$N \lesssim \|\mathbf{f} - \mathbf{A}\mathbf{v}\|^{-1/s}|\mathbf{u}|^{1/s}_{\mathcal{A}^s}.$$

With $\breve{\Lambda} := \operatorname{supp} \mathbf{v} \cup \operatorname{supp} \mathbf{u}_N$, the solution of $\mathbf{P}_{\breve{\Lambda}}\mathbf{A}\mathbf{u}_{\breve{\Lambda}} = \mathbf{P}_{\breve{\Lambda}}\mathbf{f}$ satisfies

$$|||\mathbf{u} - \mathbf{u}_{\breve{\Lambda}}||| \le |||\mathbf{u} - \mathbf{u}_N||| \le \|\mathbf{A}\|^{\frac{1}{2}}\|\mathbf{u} - \mathbf{u}_N\| \le \|\mathbf{A}\|^{\frac{1}{2}}\lambda|||\mathbf{u} - \mathbf{v}|||,$$

and so by Galerkin orthogonality, $\|\mathbf{u}_{\breve{\Lambda}} - \mathbf{v}\| \geq (1 - \|\mathbf{A}\|\lambda^2)^{\frac{1}{2}}\|\mathbf{u} - \mathbf{v}\|$, giving

$$\begin{aligned}\|\mathbf{P}_{\breve{\Lambda}}(\mathbf{f} - \mathbf{A}\mathbf{v})\| &= \|\mathbf{P}_{\breve{\Lambda}}(\mathbf{A}\mathbf{u}_{\breve{\Lambda}} - \mathbf{A}\mathbf{v})\| \geq \|\mathbf{A}^{-1}\|^{-\frac{1}{2}}\|\mathbf{u}_{\breve{\Lambda}} - \mathbf{v}\| \\ &\geq \|\mathbf{A}^{-1}\|^{-\frac{1}{2}}(1 - \|\mathbf{A}\|\lambda^2)^{\frac{1}{2}}\|\mathbf{u} - \mathbf{v}\| \\ &\geq \kappa(\mathbf{A})^{-\frac{1}{2}}(1 - \|\mathbf{A}\|\lambda^2)^{\frac{1}{2}}\|\mathbf{f} - \mathbf{A}\mathbf{v}\| \\ &\geq \mu\|\mathbf{f} - \mathbf{A}\mathbf{v}\|.\end{aligned}$$

Since $\breve{\Lambda} \supset \operatorname{supp}\mathbf{v}$, by definition of Λ we conclude that

$$\#(\Lambda \setminus \operatorname{supp}\mathbf{v}) \leq \#(\breve{\Lambda} \setminus \operatorname{supp}\mathbf{v}) \leq N \lesssim \|\mathbf{f} - \mathbf{A}\mathbf{v}\|^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}. \qquad \blacksquare$$

Before proceeding further, let us briefly describe how the above lemma can be used to prove the optimal convergence rate of the adaptive algorithm in an ideal setting. For some constant $\mu \in (0, \kappa(\mathbf{A})^{-\frac{1}{2}})$ and $i \in \mathbb{N}_0$, we define Λ_{i+1} to be the *smallest* set with $\|\mathbf{P}_\Lambda(\mathbf{f} - \mathbf{A}\mathbf{u}_{\Lambda_i})\| \geq \mu\|\mathbf{f} - \mathbf{A}\mathbf{u}_{\Lambda_i}\|$, where $\Lambda_0 := \varnothing$ and $\mathbf{u}_{\Lambda_i}$ is the Galerkin solution in the subspace $\ell_2(\Lambda_i)$. By Lemma 3.2.1 on page 45, we have a fixed error reduction: $\|\mathbf{u} - \mathbf{u}_{\Lambda_{i+1}}\| \leq \rho\|\mathbf{u} - \mathbf{u}_{\Lambda_i}\|$ for $i \in \mathbb{N}_0$, with a constant $\rho < 1$. Now assuming that $\mathbf{u} \in \mathcal{A}^s$ with some $s > 0$, by the preceding lemma and the geometric decrease of $\|\mathbf{f} - \mathbf{A}\mathbf{u}_{\Lambda_i}\| \eqsim \|\mathbf{u} - \mathbf{u}_{\Lambda_i}\|$, for $i \in \mathbb{N}_0$ we have

$$\begin{aligned}\#\Lambda_k &= \textstyle\sum_{i=0}^{k-1} \#(\Lambda_{i+1} \setminus \Lambda_i) \lesssim \sum_{i=0}^{k-1} \|\mathbf{f} - \mathbf{A}\mathbf{u}_{\Lambda_i}\|^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} \\ &\lesssim \|\mathbf{f} - \mathbf{A}\mathbf{u}_{\Lambda_{k-1}}\|^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s},\end{aligned}$$

or, $\|\mathbf{u} - \mathbf{u}_{\Lambda_k}\| \lesssim (\#\Lambda_k)^{-s}|\mathbf{u}|_{\mathcal{A}^s}$, which, in view of the assumption $\mathbf{u} \in \mathcal{A}^s$, is modulo some constant factor the best possible bound on the error.

In view of realizing the above discussed idea for an algorithm with an inexact evaluation of the residuals and an inexact solution of the Galerkin systems, we will modify Algorithm 3.2.5 on page 49 so that in Line 8, the set $\Lambda_{i+1} \supset \operatorname{supp}\mathbf{v}_i$ is chosen to be such that $\|\mathbf{P}_{\Lambda_{i+1}}\tilde{\mathbf{r}}_i\| \geq \alpha\|\tilde{\mathbf{r}}_i\|$ with $\#(\Lambda_{i+1} \setminus \operatorname{supp}\mathbf{v}_i)$ minimal modulo some constant factor. We define the following routine to perform the latter task.

Algorithm 3.3.2 Index set expansion $\mathbf{RESTRICT}[\Lambda, \mathbf{r}, \alpha] \to \tilde{\Lambda}$

Input: $\Lambda \subset \nabla$, $\#\Lambda < \infty$, $\mathbf{r} \in P$, $\alpha \in (0, 1)$.
Output: $\tilde{\Lambda} \supseteq \Lambda$ and $\|\mathbf{P}_{\tilde{\Lambda}}\mathbf{r}\| \geq \alpha\|\mathbf{r}\|$.
1: $\tilde{\mathbf{r}} := \mathbf{COARSE}[\mathbf{r}|_{\nabla\setminus\Lambda}, \sqrt{1-\alpha^2}\|\mathbf{r}\|]$;
2: $\tilde{\Lambda} := \Lambda \cup \operatorname{supp}\tilde{\mathbf{r}}$.

Lemma 3.3.3. *The output of* $\tilde{\Lambda} := \mathbf{RESTRICT}[\Lambda, \mathbf{r}, \alpha]$ *satisfies* $\tilde{\Lambda} \supseteq \Lambda$ *and* $\|\mathbf{P}_{\tilde{\Lambda}}\mathbf{r}\| \geq \alpha\|\mathbf{r}\|$. *Moreover, the output satisfies*

$$\#\tilde{\Lambda} - \#\Lambda \lesssim \min\{\#\breve{\Lambda} - \#\Lambda : \|\mathbf{P}_{\breve{\Lambda}}\mathbf{r}\| \geq \alpha\|\mathbf{r}\| \text{ and } \nabla \supset \breve{\Lambda} \supseteq \Lambda\}, \qquad (3.3.2)$$

and the number of arithmetic operations and storage locations needed for this routine can be bounded by an absolute multiple of $\#\Lambda + \#\operatorname{supp}\mathbf{r} + 1$.

Proof. We have $\|\mathbf{r} - \mathbf{P}_{\tilde{\Lambda}}\mathbf{r}\| = \|\mathbf{r}|_{\nabla\setminus\Lambda} - \tilde{\mathbf{r}}\| \leq \sqrt{1-\alpha^2}\|\mathbf{r}\|$, which is equivalent to $\|\mathbf{P}_{\tilde{\Lambda}}\mathbf{r}\| \geq \alpha\|\mathbf{r}\|$. The work bound immediately follows from Lemma 2.6.4 on page 26. Since $\Lambda \cap \operatorname{supp}\tilde{\mathbf{r}} = \varnothing$, by applying Lemma 2.6.4 we have

$$\begin{aligned}\#\tilde{\Lambda} - \#\Lambda = \#\operatorname{supp}\tilde{\mathbf{r}} &\lesssim \min\{\#\breve{\Lambda} : \|\mathbf{r}|_{\nabla\setminus\Lambda} - \mathbf{P}_{\breve{\Lambda}}\mathbf{r}|_{\nabla\setminus\Lambda}\| \leq \sqrt{1-\alpha^2}\|\mathbf{r}\|\} \\ &= \min\{\#\breve{\Lambda} : \|\mathbf{r} - \mathbf{P}_{\breve{\Lambda}\cup\Lambda}\mathbf{r}\| \leq \sqrt{1-\alpha^2}\|\mathbf{r}\|\},\end{aligned}$$

and observing that the minimum is obtained when $\breve{\Lambda} \cap \Lambda = \varnothing$, the proof is completed. ∎

Now we present our modification of Algorithm 3.2.5. Recall that Algorithm 3.2.5 was intended for a reduction of the error with a fixed factor, to be used inside the algorithm **SOLVE** from Chapter 2, i.e., Algorithm 2.6.7 on page 28. The modification given below will be an optimal solver in its own as the forthcoming Theorem 3.3.5 shows.

Algorithm 3.3.4 Method **SOLVE**$[\varepsilon] \to \mathbf{v}_i$ without coarsening of the iterands

---

Parameters: Let $0 < \delta < \alpha < 1$ and $0 < \gamma < \frac{1}{6}\kappa(\mathbf{A})^{-\frac{1}{2}}(\alpha - \delta)$. Let $\theta > 0$ be a fixed constant, and let $\nu_0 > 0$ and $\mathbf{v} = 0$.

Input: $\varepsilon > 0$.

Output: $\mathbf{v}_i \in P$ with $\|\mathbf{f} - \mathbf{A}\mathbf{v}_i\| \leq \varepsilon$.

Description: The body of this algorithm is identical to Algorithm 3.2.5 on page 49, except that we replace the statement in Line 8 by

$\Lambda_{i+1} := \mathbf{RESTRICT}[\Lambda_i, \tilde{\mathbf{r}}_i, \alpha]$;

Using perturbation arguments, we will prove that **SOLVE** has optimal computational complexity.

Theorem 3.3.5. *Let* $\mathbf{A}$ *be* s^**-computable, and let* $\mathbf{f}$ *be* s^**-admissible for some* $s^* > 0$. *Then* $\mathbf{u}_\varepsilon := \mathbf{SOLVE}[\varepsilon]$ *terminates with* $\|\mathbf{f} - \mathbf{A}\mathbf{u}_\varepsilon\| \leq \varepsilon$. *In addition, let the parameters inside* **SOLVE** *satisfy* $\frac{\alpha+\delta}{1-\delta} < \kappa(\mathbf{A})^{-\frac{1}{2}}$. *If* $\nu_0 \eqsim \|\mathbf{f}\| \gtrsim \varepsilon$, *and for some* $s < s^*$, $\mathbf{u} \in \mathcal{A}^s$, *then* $\operatorname{supp}\mathbf{u}_\varepsilon \lesssim \varepsilon^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}$ *and the number of arithmetic operations and storage locations required by the call is bounded by some absolute multiple of the same expression.*

Proof. By the same reasoning in the proof of Theorem 3.2.7 on page 49, **SOLVE**$[\varepsilon]$ terminates say, after K iterations, and with some $\rho \in (0,1)$, we have

$$\nu_i \lesssim \rho^{i-k}\nu_k \qquad \text{for } 0 \leq k \leq i \leq K. \tag{3.3.3}$$

Here we will use the notations from the proof of Theorem 3.2.7. With $\mu = \frac{\alpha+\delta}{1-\delta}$, for $1 \le i < K$ let $\breve{\Lambda}_{i+1} \supset \operatorname{supp} \mathbf{v}_i$ be the *smallest* set with $\|\mathbf{P}_{\breve{\Lambda}_{i+1}} \mathbf{r}_i\| \ge \mu \|\mathbf{r}_i\|$. Then

$$\mu\|\tilde{\mathbf{r}}_i\| \le \mu\|\mathbf{r}_i\| + \mu\delta\|\tilde{\mathbf{r}}_i\| \le \|\mathbf{P}_{\breve{\Lambda}_{i+1}} \mathbf{r}_i\| + \mu\delta\|\tilde{\mathbf{r}}_i\| \le \|\mathbf{P}_{\breve{\Lambda}_{i+1}} \tilde{\mathbf{r}}_i\| + (1+\mu)\delta\|\tilde{\mathbf{r}}_i\|$$

or $\|\mathbf{P}_{\breve{\Lambda}_{i+1}} \tilde{\mathbf{r}}_i\| \ge \alpha \|\tilde{\mathbf{r}}_i\|$. By the property (3.3.2) of **RESTRICT** we have $\#(\Lambda_{i+1} \setminus \operatorname{supp} \mathbf{v}_i) \lesssim \#(\breve{\Lambda}_{i+1} \setminus \operatorname{supp} \mathbf{v}_i)$. Since $\mu < \kappa(\mathbf{A})^{-\frac{1}{2}}$ by the condition on α and δ, and $\|\mathbf{f} - \mathbf{A}\mathbf{v}_i\| \eqsim \nu_i$, an application of Lemma 3.3.1 on page 52 shows that $\#(\breve{\Lambda}_{i+1} \setminus \operatorname{supp} \mathbf{v}_i) \lesssim \nu_i^{-1/s} |\mathbf{u}|_{\mathcal{A}^s}^{1/s}$.

Since with $\Lambda_1 := \varnothing$, $\operatorname{supp} \mathbf{v}_i \subseteq \Lambda_i$ and $\Lambda_i \subset \Lambda_{i+1}$, for $1 \le k \le K$ by (3.3.1) we have

$$\#\operatorname{supp} \mathbf{v}_k \le \#\Lambda_k = \sum_{i=1}^{k-1} \#(\Lambda_{i+1} \setminus \Lambda_i) \lesssim (\sum_{i=1}^{k-1} \nu_i^{-1/s}) |\mathbf{u}|_{\mathcal{A}^s}^{1/s} \lesssim \nu_{k-1}^{-1/s} |\mathbf{u}|_{\mathcal{A}^s}^{1/s}. \quad (3.3.4)$$

From $|\mathbf{v}_k|_{\mathcal{A}^s} \lesssim |\mathbf{u}|_{\mathcal{A}^s} + (\#\operatorname{supp} \mathbf{v}_k)^s \|\mathbf{v}_k - \mathbf{u}\|$ (Proposition 2.3.6 on page 18), we infer that $|\mathbf{v}_k|_{\mathcal{A}^s} \lesssim |\mathbf{u}|_{\mathcal{A}^s}$.

By Proposition 2.8.4 on page 38, the cost of the i-th call of **RES** for $1 \le i \le K$ is bounded by an absolute multiple of

$$\nu_i^{-1/s} \left(|\mathbf{v}_i|_{\mathcal{A}^s}^{1/s} + |\mathbf{u}|_{\mathcal{A}^s}^{1/s} + \nu_{i-1}^{1/s} (\#\operatorname{supp} \mathbf{v}_i + 1) \right) \lesssim \nu_i^{-1/s} |\mathbf{u}|_{\mathcal{A}^s}^{1/s},$$

where we used (3.3.4), and $1 \lesssim \nu_{k-1}^{-1/s} |\mathbf{u}|_{\mathcal{A}^s}^{1/s}$ by $\nu_{k-1} \lesssim \nu_0 \lesssim \|\mathbf{f}\| \lesssim |\mathbf{u}|_{\mathcal{A}^s}$.

The cost of the k-th call for $k < K$ of the subroutines **RESTRICT**, **RHS** or **GALSOLVE** is bounded by an absolute multiple of $\#\Lambda_{k+1} + \#\operatorname{supp} \tilde{\mathbf{r}}_i \lesssim \nu_k^{-1/s} |\mathbf{u}|_{\mathcal{A}^s}^{1/s}$, $\nu_k^{-1/s} |\mathbf{u}|_{\mathcal{A}^s}^{1/s}$, or $\nu_k^{-1/s} (|\mathbf{v}_k|_{\mathcal{A}^s}^{1/s} + |\mathbf{u}|_{\mathcal{A}^s}^{1/s}) + \#\Lambda_{k+1} \lesssim \nu_k^{-1/s} |\mathbf{u}|_{\mathcal{A}^s}^{1/s}$, respectively. From (3.3.3) and $\nu_K \gtrsim \min\{\nu_{K-1}, \varepsilon\} \gtrsim \varepsilon$ by Proposition 2.8.4, where the second inequality follows from $\nu_{K-1} > \varepsilon$ when $K > 0$, and by assumption when $K = 0$, the proof is completed. ∎

3.4 Numerical experiment

We consider the variational formulation of the following problem of order $2t = 2$ on the interval $[0, 1]$, i.e., $n = 1$, with periodic boundary conditions

$$-\Delta u + u = f \qquad \text{on } \mathbb{R}/\mathbb{Z}. \quad (3.4.1)$$

We define the right-hand side f by $f(v) = 4v(\frac{1}{2}) + \int_0^1 g(x)v(x)dx$, with

$$g(x) = (16\pi^2 + 1)\cos(4\pi x) - 4 + \begin{cases} 2x^2, & \text{if } x \in [0, 1/2), \\ 2(1-x)^2, & \text{if } x \in [1/2, 1], \end{cases} \quad (3.4.2)$$

so that the solution u is given by

$$u(x) = \cos(4\pi x) + \begin{cases} 2x^2, & \text{if } x \in [0, 1/2), \\ 2(1-x)^2, & \text{if } x \in [1/2, 1], \end{cases} \tag{3.4.3}$$

see Figure 3.1.

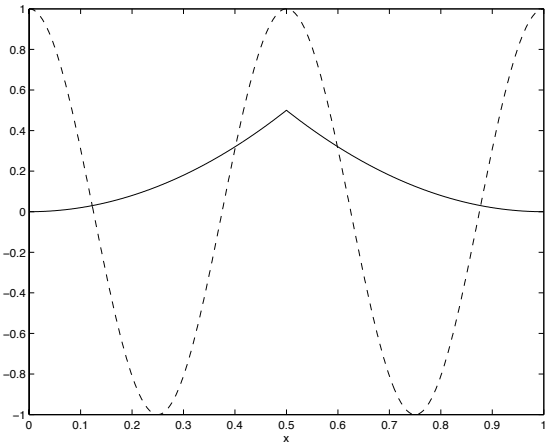

Figure 3.1: *The solution u is the sum of both functions illustrated.*

We use the periodized B-spline wavelets of order $d = 3$ with $\tilde{d} = 3$ vanishing moments from [21] normalized in the $H^1(0,1)$-norm. The solution u is in $H^{s+1}(\mathbb{R}/\mathbb{Z})$ only for $s < \frac{1}{2}$. On the other hand, since u can be shown to be in $B_p^{s+1}(L_p(\mathbb{R}/\mathbb{Z}))$ for any $s > 0$ with $\frac{1}{p} = \frac{1}{2} + s$, we deduce that the corresponding discrete solution $\mathbf{u}$ is in $\mathcal{A}^s$ for any $s < \frac{d-t}{n} = 2$.

Each entry of the infinite stiffness matrix $\mathbf{A}$ can be computed in $\mathcal{O}(1)$ operations. By applying the compression rules from [86], which are quoted in Theorem 7.3.3 on page 127, we see that $\mathbf{A}$ is s^*-compressible with $s^* = t + \tilde{d} = 4$.

For developing a routine **RHS**, we split $f = \langle f_1, \cdot \rangle_{L_2} + f_2$, where $f_1(x) = (16\pi^2 + 1)\cos(4\pi x) - 4$. Correspondingly, we split $\mathbf{f} = \mathbf{f}_1 + \mathbf{f}_2$, and given a tolerance ε, we approximate both infinite vectors within tolerance $\varepsilon/2$ by, for suitable $\ell_1(\varepsilon)$, $\ell_2(\varepsilon)$, dropping all coefficients with indices λ with $|\lambda| > \ell_1(\varepsilon)$ or $|\lambda| > \ell_2(\varepsilon)$, respectively.

From

$$\begin{aligned} |\langle \psi_\lambda, f_1 \rangle_{L_2}| &\le \|\psi_\lambda\|_{L_1(0,1)} \inf_{p \in P_2} \|f_1 - p\|_{L_\infty(\operatorname{supp}\psi_\lambda)}, \\ \|\psi_\lambda\|_{L_1(0,1)} &\le \operatorname{diam}(\operatorname{supp}\psi_\lambda))^{\frac{1}{2}} \|\psi_\lambda\|_{L_2(0,1)}, \\ \operatorname{diam}(\operatorname{supp}\psi_\lambda) &= 5 \cdot 2^{-|\lambda|}, \\ \|\psi_\lambda\|_{L_2(0,1)} &= [4^{|\lambda|}\|\psi'\|^2_{L_2(\mathbb{R})}/\|\psi\|^2_{L_2(\mathbb{R})} + 1]^{-1}, \end{aligned}$$

where ψ is the "mother wavelet", $\#\{\lambda : |\lambda| = k\} = 2^k$, and $\inf_{p\in P_2} \|f_1 - p\|_{L_\infty(\operatorname{supp}\psi_\lambda)} \leq (\frac{\pi}{4}\operatorname{diam}(\operatorname{supp}\psi_\lambda))^3\|f_1'''\|_{L_\infty(0,1)}/3!$ (Jackson estimate), we find an upper bound for the error $\sqrt{\sum_{|\lambda|>\ell_1(\varepsilon)} |\langle\psi_\lambda, f_1\rangle_{L_2}|^2}$ which is $\eqsim 2^{-4\ell_1(\varepsilon)}$. Setting this upper bound equal to $\varepsilon/2$, and solving for $\ell_1(\varepsilon)$ gives an approximation for $\mathbf{f}_1$ of length $\eqsim 2^{\ell_1(\varepsilon)} \eqsim \varepsilon^{-1/4}$. Note that in view of the admissibility assumption we made on $\mathbf{f}$, a vector length $\eqsim \varepsilon^{-1/2}$ would have been sufficient. Such a length would have been found with wavelets that have 1 vanishing moment.

From

$$\begin{aligned}
|\langle\psi_\lambda, f_2\rangle| &\leq (4 + \|g - f_1\|_{L_1(\operatorname{supp}\psi_\lambda)})\|\psi_\lambda\|_{L_\infty(0,1)},\\
\|g - f_1\|_{L_1(\operatorname{supp}\psi_\lambda)} &\leq 5\cdot 2^{-|\lambda|},\\
\|\psi_\lambda\|_{L_\infty(0,1)} &= [2^{|\lambda|/2}\|\psi\|_{L_\infty(\mathbb{R})}/\|\psi\|_{L_2(\mathbb{R})}]\|\psi_\lambda\|_{L_2(0,1)},\\
\#\{|\lambda| = k : &\tfrac{1}{2} \text{ is an interior point of } \operatorname{supp}\psi_\lambda\} = 9,
\end{aligned}$$

and the fact that $\langle\psi_\lambda, f_2\rangle$ vanishes when λ is not in any of these sets, we find an upper bound for the error $\sqrt{\sum_{|\lambda|>\ell_2(\varepsilon)} |\langle\psi_\lambda, f_2\rangle|^2}$ which is $\eqsim 2^{-\ell_2(\varepsilon)/2}$. Setting this upper bound equal to $\varepsilon/2$ and solving for $\ell_2(\varepsilon)$ gives an approximation for $\mathbf{f}_2$ of length $\leq 9(\ell_2(\varepsilon) + 1) = \mathcal{O}(|\log(\varepsilon)| + 1)$, which is asymptotically even much smaller than the bound we found in the $\mathbf{f}_1$ case. Since each entry of $\mathbf{f}$ can be computed in $\mathcal{O}(1)$ operations, in view of Definition 2.7.4 on page 30, we conclude that $\mathbf{f}$ is s^*-admissible with $s^* = 4$.

We will compare the results of Algorithm 3.3.4 with those of Algorithm 2.6.7 on page 28 that uses the subroutine **RICHARDSON** from Algorithm 2.7.6 on page 31 as **ITERATE**, which we refer to as being the CDD2 method.

We tested our Algorithm 3.3.4 or CDD2 with parameters $\alpha = 0.4$, $\delta = 0.012618$, and $\gamma = 0.009581$, or $K = 5$ and $\theta = 2.5$, respectively. Inside the ranges where the methods are proven to be of optimal computational complexity, these parameters are close to the values that give the best quantitative results. Actually, since these ranges result from a succession of worst case analyses, we may expect that outside them, i.e., concerning Algorithm 3.3.4 for larger α, δ and γ, more efficient algorithms are obtained. The numerical results, given in Figure 3.2 on the next page, illustrate the optimal computational complexity of both Algorithm 3.3.4 and CDD2. Note that the time measurements do not start at zero, but after $10^0 = 1$ second. The results show that in this example the new method needs less than a factor $\frac{1}{10}$ in computing time to achieve the same accuracy.

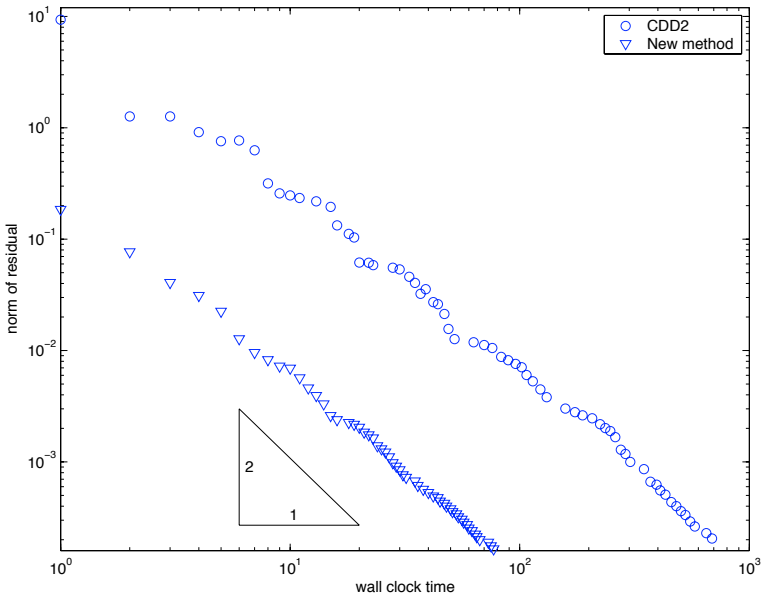

Figure 3.2: *Convergence histories*

Chapter 4

Using polynomial preconditioners

4.1 Introduction

In this chapter, we carry on with considering the linear equation

$$\mathbf{A}\mathbf{u} = \mathbf{f}, \tag{4.1.1}$$

where $\mathbf{A} : \ell_2 \to \ell_2$ is an SPD matrix and $\mathbf{f} \in \ell_2$. As we saw in the foregoing chapters, the quantitative properties of the adaptive algorithms for solving this system depend on the condition number $\kappa(\mathbf{A}) := \|\mathbf{A}\|\|\mathbf{A}^{-1}\|$, which in turn depends on the underlying wavelet basis. While constructing a wavelet basis with favourable quantitative properties is a rather delicate task, preconditioning the equation (4.1.1) without any reference to the original continuous problem seems an attractive possibility to improve the conditioning of the system. In this chapter, with a *preconditioner* $\mathbf{S} : \ell_2 \to \ell_2$ such that $\kappa(\mathbf{SA}) < \kappa(\mathbf{A})$, instead of (4.1.1) we will consider the following linear equation,

$$\mathbf{SAu} = \mathbf{Sf}. \tag{4.1.2}$$

Apart from the diagonal one, perhaps the simplest preconditioner is the inverse of a finite section of the stiffness matrix $\mathbf{A}$. One could use the LU-decomposition to preserve the symmetry. For instance, when the coarsest level functions of the wavelet basis adversely affect the condition of the system harmfully, one can invert the stiffness matrix restricted to the coarsest level and use the inverse as a preconditioner. Since we can approximate the action of the stiffness matrix, the next idea would be to use polynomial preconditioners. In this chapter, we investigate the use of polynomial preconditioners.

This chapter is organized as follows. In the next section, we recall some results on polynomial preconditioners and put forward ways to use them in our

setting. Then in Section 4.3, we show that the adaptive wavelet algorithms with polynomial preconditioners are again of optimal computational complexity.

4.2 Polynomial preconditioners

In the context of linear algebraic equations, polynomial preconditioners have been studied extensively, see e.g. [1, 58, 70]. In this section, we recall some of the results regarding common polynomial preconditioners and analyze them in our setting.

Now the preconditioning matrix $\mathbf{S}$ is a *polynomial* in $\mathbf{A}$, namely, we assume that $\mathbf{S} = p(\mathbf{A})$ for some polynomial p of degree $k \geq 0$. $p(\mathbf{A})$ commutes $\mathbf{A}$, thus $p(\mathbf{A})\mathbf{A}$ is symmetric. Moreover, if p is positive on the spectrum of $\mathbf{A}$, $p(\mathbf{A})\mathbf{A}$ is positive definite. To be more practical, if

$$\|\mathbf{I} - p(\mathbf{A})\mathbf{A}\| \leq \rho \qquad \text{for some } \rho < 1, \tag{4.2.1}$$

then $p(\mathbf{A})\mathbf{A}$ is positive definite, and we have the bound

$$\kappa[p(\mathbf{A})\mathbf{A}] \leq \frac{1+\rho}{1-\rho}. \tag{4.2.2}$$

Neumann series polynomials

A simple choice for p is a polynomial based on Neumann series. With some $\omega \in (0, \frac{2}{\|\mathbf{A}\|})$ and $\mathbf{N} := \mathbf{I} - \omega\mathbf{A}$, we have

$$(\omega\mathbf{A})^{-1} = \mathbf{I} + \mathbf{N} + \mathbf{N}^2 + \ldots,$$

and truncating this series we define a polynomial preconditioner

$$p_k(\mathbf{A}) := \omega(\mathbf{I} + \mathbf{N} + \ldots + \mathbf{N}^k). \tag{4.2.3}$$

One easily identifies the application of $p_k(\mathbf{A})$ with k iterations of a damped Richardson method. As for (4.2.1), we have

$$\|\mathbf{I} - p_k(\mathbf{A})\mathbf{A}\| = \|\omega(\mathbf{N}^{k+1} + \mathbf{N}^{k+2} + \ldots)\mathbf{A}\| = \|\mathbf{N}^{k+1}\| \leq \|\mathbf{N}\|^{k+1}.$$

Min-max polynomials

If the coefficients of the preconditioning polynomial p are given, in general the action of $p(\mathbf{A})$ is computed using k applications of $\mathbf{A}$. Therefore we shall try to

minimize the condition number (4.2.2) keeping k as small as possible. By first complexifying ℓ_2 and then applying the spectral theorem, cf. [69], we have

$$\|p(\mathbf{A})\mathbf{A}\mathbf{x}\|^2 = \int_{\sigma(\mathbf{A})} [p(\lambda)\lambda]^2 dE_{\mathbf{x},\mathbf{x}}(\lambda) \leq \|\mathbf{x}\|^2 \cdot \max_{\lambda\in\sigma(\mathbf{A})} [p(\lambda)\lambda]^2, \tag{4.2.4}$$

where $\sigma(\mathbf{A})$ is the spectrum and E is the spectral decomposition of $\mathbf{A}$. This immediately implies $\|p(\mathbf{A})\mathbf{A}\| \leq \max_{\lambda\in\sigma(\mathbf{A})} |p(\lambda)\lambda|$. Similarly, we can estimate

$$\|[p(\mathbf{A})\mathbf{A}]^{-1}\| \leq \max_{\lambda\in\sigma(\mathbf{A})} \left|[p(\lambda)\lambda]^{-1}\right| = \left(\min_{\lambda\in\sigma(\mathbf{A})} |p(\lambda)\lambda|\right)^{-1},$$

where we assumed that p is nonzero on the spectrum of $\mathbf{A}$. Now let $\sigma(\mathbf{A}) \subset [c,d]$ with $c, d > 0$. Then we have

$$\kappa[p(\mathbf{A})\mathbf{A}] \leq \frac{\max_{\lambda\in[c,d]} |p(\lambda)\lambda|}{\min_{\lambda\in[c,d]} |p(\lambda)\lambda|} \qquad \text{for } p \text{ nonzero on } [c,d]. \tag{4.2.5}$$

We consider the problem of minimizing the above upper bound over the polynomials of degree k. Recall the following result from [1, 58].

Lemma 4.2.1. *Let p_k be the polynomial defined by*

$$p_k(\lambda)\lambda = 1 - \frac{C_k(\frac{d+c-2\lambda}{d-c})}{C_k(\frac{d+c}{d-c})}, \tag{4.2.6}$$

where C_k is the k-th Chebyshev polynomial of the first kind. Then

$$\frac{\max_{\lambda\in[c,d]} |p_k(\lambda)\lambda|}{\min_{\lambda\in[c,d]} |p_k(\lambda)\lambda|} = \frac{|C_k(\frac{d+c}{d-c})| + 1}{|C_k(\frac{d+c}{d-c})| - 1} \leq \frac{\max_{\lambda\in[c,d]} |q(\lambda)\lambda|}{\min_{\lambda\in[c,d]} |q(\lambda)\lambda|}$$

for all $q \in P_k[c,d]$, with equality if and only if q is a scalar multiple of p_k.

With $[c,d]$ as above, estimating the norm in (4.2.1) as in (4.2.4), we get

$$\|\mathbf{I} - p(\mathbf{A})\mathbf{A}\| \leq \max_{\lambda\in[c,d]} |1 - p(\lambda)\lambda|. \tag{4.2.7}$$

It turns out that the same polynomial p_k from (4.2.6) minimizes the upper bound (4.2.7) over $P_k[c,d]$, and these polynomials are called the *min-max* polynomials. Furthermore, with the min-max polynomial p_k we have $\|\mathbf{I} - p_k(\mathbf{A})\mathbf{A}\| \leq |C_k(\frac{d+c}{d-c})|^{-1} < 1$ for $k \geq 0$, ensuring the positive definiteness of $p_k(\mathbf{A})\mathbf{A}$.

For min-max polynomials, the application of $p_k(\mathbf{A})$ is equivalent to k iterations of the Chebyshev method. In light of this fact, the coefficients of the polynomial p_k can be computed, e.g. using three term recurrences.

Least-squares polynomials

An alternative family of preconditioning polynomials can be obtained by minimizing some quadratic norm of $1-p(\lambda)\lambda$ instead of the uniform norm, cf. (4.2.7). With a positive weight function $w : [c, d] \to \mathbb{R}_{>0}$, we consider the following minimization problem

$$\int_c^d |1 - p(\lambda)\lambda|^2 w(\lambda) d\lambda \to \min \quad \text{over } p \in P_k[c, d]. \tag{4.2.8}$$

We call the solution of this problem a *least-squares* polynomial. Unlike the min-max polynomial, the least-square polynomial is biased in its suppression of the eigenvalues of $\mathbf{A}$. For example, when $w \equiv 1$, the larger eigenvalues are mapped closer to 1 than the small ones. If the eigenvalue distribution of $\mathbf{A}$ were known, one could choose w to emphasize the dense parts of the spectrum.

We recall the following result from [58].

Lemma 4.2.2. *Let $s_i \in P_{k+1}[c, d]$, $0 \le i \le k+1$, be orthonormal with respect to the weight function w, normalized so that $s_i(0) > 0$ for $0 \le i \le k+1$. Assume that each s_i, $0 \le i \le k+1$, obtains its maximum on $[c, d]$ at c. Then with $J_{k+1}(\mu, \lambda) := \sum_{j=0}^{k+1} s_j(\mu) s_j(\lambda)$, the solution p_k to the problem (4.2.8) is given by*

$$p_k(\lambda)\lambda = 1 - \frac{J_{k+1}(0, \lambda)}{J_{k+1}(0, 0)}, \tag{4.2.9}$$

and $p_k(\lambda) > 0$ for $\lambda \in [c, d]$.

This lemma is applicable to a wide range of weights, including the Jacobi weights

$$w_{\alpha,\beta}(\lambda) = (d - \lambda)^\alpha (\lambda - c)^\beta,$$

when $\alpha, \beta \ge -\frac{1}{2}$. In this case, the function $J_{k+1}(0, \cdot)$ is a shifted and scaled Jacobi polynomial, cf. [1]. Since the polynomial is known explicitly, the norm (4.2.1) can be estimated by using e.g. the estimate (4.2.7).

Approximate preconditioning

Since it is not possible to compute the application of $\mathbf{A}$ exactly, the action of the polynomial preconditioner $p(\mathbf{A})$ must be approximated. For the approximate application of an s^*-computable matrix $\mathbf{A}$ with $s^* > 0$, we distinguish between two possibilities: (a) we can use the subroutine **APPLY**; or (b) we can apply some approximation $\mathbf{A}_j$ as in Definition 2.7.8. Let p_k be the polynomial given

by $p_k(\lambda) = a_0 + a_1\lambda + \ldots + a_k\lambda^k$ and let $\mathbf{S} = p_k(\mathbf{A})$. We consider the following algorithm which implements possibility (a).

Algorithm 4.2.3 Polynomial preconditioner $\mathbf{PREC}_a[\mathbf{r}, \xi] \to \mathbf{d}$

Parameters: Let ε_i, $i = 1, \ldots, k$, be such that $\sum_{i=1}^{k} \varepsilon_i \|\mathbf{A}\|^{i-1} \leq \xi\|\mathbf{r}\|$ and $\varepsilon_i \gtrsim \xi\|\mathbf{r}\|$.
Input: Let $\mathbf{r} \in P$ and $\xi > 0$.
Output: $\mathbf{d} \in P$ and $\|\mathbf{Sr} - \mathbf{d}\| \leq \xi\|\mathbf{r}\|$.
1: $\tilde{\mathbf{b}}_k := a_k\mathbf{r}$;
2: **for** $i = k, \ldots, 1$ **do**
3: $\quad \tilde{\mathbf{b}}_{k-1} := a_{i-1}\mathbf{r} + \mathbf{APPLY}[\mathbf{A}, \tilde{\mathbf{b}}_i, \varepsilon_i]$;
4: **end for**
5: $\mathbf{d} := \tilde{\mathbf{b}}_0$.

Correspondingly, possibility (b) suggests the following algorithm.

Algorithm 4.2.4 Polynomial preconditioner $\mathbf{PREC}_b[\mathbf{r}, \xi] \to \mathbf{d}$

Parameters: Let J satisfy $\|\mathbf{A} - \mathbf{A}_J\| \sum_{i=1}^{k} i|a_i|\|\mathbf{A}\|^{i-1} \leq \xi$ and $2^J \lesssim \xi^{-1/s}$, for some $s > 0$, with $\mathbf{A}_J$ a compressed matrix as in Definition 2.7.8.
Input: Let $\mathbf{r} \in P$ and $\xi > 0$.
Output: $\mathbf{d} \in P$ and $\|\mathbf{Sr} - \mathbf{d}\| \leq \xi\|\mathbf{r}\|$.
1: $\tilde{\mathbf{b}}_k := a_k\mathbf{r}$;
2: **for** $i = k, \ldots, 1$ **do**
3: $\quad \tilde{\mathbf{b}}_{k-1} := a_{i-1}\mathbf{r} + \mathbf{A}_J\tilde{\mathbf{b}}_i$;
4: **end for**
5: $\mathbf{d} := \tilde{\mathbf{b}}_0$.

Definition 4.2.5. A subroutine $\mathbf{PREC}[\mathbf{r}, \xi] \to \mathbf{d}$ is said to have *linear complexity* when for any $\xi > 0$ and a finite dimensional vector $\mathbf{r} \in \ell_2$, $\mathbf{d} := \mathbf{PREC}[\mathbf{r}, \xi]$ terminates with $\|\mathbf{Sr} - \mathbf{d}\| \leq \xi\|\mathbf{r}\|$, and for a non-increasing function $c : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$, $\#\operatorname{supp}\mathbf{d} \lesssim c(\xi)\#\operatorname{supp}\mathbf{r}$ and the number of arithmetic operations and storage locations required by the call being bounded by an absolute multiple of $c(\xi)\#\operatorname{supp}\mathbf{r} + 1$. ⊘

Proposition 4.2.6. *Let* $\mathbf{A}$ *be* s^**-computable with some* $s^* > 0$*. Then the subroutines* $\mathbf{PREC}_a[\mathbf{r}, \xi]$ *and* $\mathbf{PREC}_b[\mathbf{r}, \xi]$ *are both of linear complexity.*

Proof. For the subroutine $\mathbf{PREC}_a$, the choice $\varepsilon_i = \frac{\xi\|\mathbf{r}\|}{k\|\mathbf{A}\|^{i-1}}$, $1 \leq i \leq k$, satisfies the assumptions on the parameters ε_i. For the subroutine $\mathbf{PREC}_b$, we can choose J satisfying the first condition and $\|\mathbf{A} - \mathbf{A}_J\| \gtrsim \xi$. Choosing s such that $s < s^*$,

we have $2^{-Js} \gtrsim \xi$, thus J satisfies the second condition. With $\mathbf{b}_k := a_k\mathbf{r}$, define $\mathbf{b}_{i-1} := a_{i-1}\mathbf{r} + \mathbf{A}\mathbf{b}_i$ recursively for $i = k, \ldots, 1$. Note that $\mathbf{S}\mathbf{r} = \mathbf{b}_0$.

We consider $\mathbf{PREC}_a$ first. From

$$\begin{aligned}\tilde{\mathbf{b}}_{i-1} - \mathbf{b}_{i-1} &= \mathbf{APPLY}[\mathbf{A}, \tilde{\mathbf{b}}_i, \varepsilon_i] - \mathbf{A}\mathbf{b}_i \\ &= \mathbf{APPLY}[\mathbf{A}, \tilde{\mathbf{b}}_i, \varepsilon_i] - \mathbf{A}\tilde{\mathbf{b}}_i + \mathbf{A}(\tilde{\mathbf{b}}_i - \mathbf{b}_i),\end{aligned}$$

we have

$$\|\tilde{\mathbf{b}}_{i-1} - \mathbf{b}_{i-1}\| \le \varepsilon_i + \|\mathbf{A}\|\|\tilde{\mathbf{b}}_i - \mathbf{b}_i\| \qquad \text{for } i = 1, \ldots, k,$$

giving $\|\mathbf{d} - \mathbf{S}\mathbf{r}\| = \|\tilde{\mathbf{b}}_0 - \mathbf{b}_0\| \le \sum_{i=1}^k \varepsilon_i \|\mathbf{A}\|^{i-1} \le \xi\|\mathbf{r}\|$.

Taking into account that $\|\mathbf{APPLY}[\mathbf{A}, \cdot, \varepsilon]\| \le \|\mathbf{A}\|\|\cdot\| + \varepsilon$ for any $\varepsilon > 0$, and $\varepsilon_i \lesssim \xi\|\mathbf{r}\|$, we can derive that $\|\tilde{\mathbf{b}}_i\| \lesssim \|\mathbf{r}\|$ where the constant absorbed by "$\lesssim$" possibly depends on ξ. For any $\mathbf{x} \in \mathcal{A}^s$, from (2.3.3) we have

$$|\mathbf{x}|_{\mathcal{A}^s} = \sup_N N^s \|\mathbf{x} - \mathcal{B}_N(\mathbf{x})\| \le (\#\operatorname{supp}\mathbf{x})^s \|\mathbf{x}\|.$$

Accordingly, we infer

$$\#\operatorname{supp}\tilde{\mathbf{b}}_{i-1} \lesssim \#\operatorname{supp}\mathbf{r} + \varepsilon_i^{-1/s}|\tilde{\mathbf{b}}_i|_{\mathcal{A}^s}^{1/s} \lesssim \#\operatorname{supp}\mathbf{r} + \varepsilon_i^{-1/s}\|\tilde{\mathbf{b}}_i\|^{1/s}(\#\operatorname{supp}\mathbf{b}_i),$$

for $i = 1, \ldots, k$. Employing this bound for $i = k, \ldots, 1$, we obtain

$$\begin{aligned}\#\operatorname{supp}\mathbf{d} &= \#\operatorname{supp}\tilde{\mathbf{b}}_0 \\ &\lesssim \#\operatorname{supp}\mathbf{r}\left(1 + \varepsilon_1^{-1/s}\|\mathbf{r}\|^{1/s} + \ldots + \varepsilon_1^{-1/s}\ldots\varepsilon_k^{-1/s}\|\mathbf{r}\|^{k/s}\right).\end{aligned}$$

We employ the assumption $\varepsilon_i \gtrsim \xi\|\mathbf{r}\|$ or $\varepsilon_i^{-1}\|\mathbf{r}\| \lesssim \xi^{-1}$ to conclude the first part of the proof.

Now we turn to the subroutine $\mathbf{PREC}_b$. We have for $i = 0, \ldots, k-1$

$$\|\tilde{\mathbf{b}}_i\| \le |a_i|\|\mathbf{r}\| + \|\mathbf{A}_J\|\|\tilde{\mathbf{b}}_{i+1}\| \le |a_i|\|\mathbf{r}\| + \|\mathbf{A}\|\|\tilde{\mathbf{b}}_{i+1}\|,$$

or $\|\tilde{\mathbf{b}}_i\| \le \|\mathbf{r}\| \sum_{j=i}^k |a_j|\|\mathbf{A}\|^{j-i}$. On the other hand, we have for $i = 1, \ldots, k$

$$\begin{aligned}\|\tilde{\mathbf{b}}_{i-1} - \mathbf{b}_{i-1}\| = \|\mathbf{A}_J\tilde{\mathbf{b}}_i - \mathbf{A}\mathbf{b}_i\| &\le \|\mathbf{A}_J - \mathbf{A}\|\|\tilde{\mathbf{b}}_i\| + \|\mathbf{A}\|\|\tilde{\mathbf{b}}_i - \mathbf{b}_i\| \\ &\le \|\mathbf{A}_J - \mathbf{A}\|\|\mathbf{r}\| \textstyle\sum_{j=i}^k |a_j|\|\mathbf{A}\|^{j-i} + \|\mathbf{A}\|\|\tilde{\mathbf{b}}_i - \mathbf{b}_i\|.\end{aligned}$$

This yields

$$\begin{aligned}\|\mathbf{d} - \mathbf{S}\mathbf{r}\| = \|\tilde{\mathbf{b}}_0 - \mathbf{b}_0\| &\le \|\mathbf{A}_J - \mathbf{A}\|\|\mathbf{r}\| \sum_{i=1}^k \|\mathbf{A}\|^{i-1} \sum_{j=i}^k |a_j|\|\mathbf{A}\|^{j-i} \\ &\le \|\mathbf{A}_J - \mathbf{A}\|\|\mathbf{r}\| \sum_{j=1}^k \sum_{i=1}^j |a_j|\|\mathbf{A}\|^{j-1} \le \|\mathbf{A}_J - \mathbf{A}\|\|\mathbf{r}\| \sum_{j=1}^k j|a_j|\|\mathbf{A}\|^{j-1},\end{aligned}$$

where by assumption the last expression is bounded by $\xi\|\mathbf{r}\|$.

For the support size, we have $\#\operatorname{supp}\mathbf{b}_{i-1} \lesssim \#\operatorname{supp}\mathbf{r} + 2^J \#\operatorname{supp}\mathbf{b}_i$, giving that

$$\#\operatorname{supp}\mathbf{d} = \#\operatorname{supp}\mathbf{b}_0 \lesssim (1 + 2^J + \ldots + 2^{Jk})\#\operatorname{supp}\mathbf{r} \lesssim 2^{Jk}\#\operatorname{supp}\mathbf{r}.$$

Finally, we use the assumption $2^J \lesssim \xi^{-1/s}$ to complete the proof. ∎

4.3 Preconditioned adaptive algorithm

Throughout this section, we assume that p_k is a polynomial of degree k such that $\mathbf{S} = p_k(\mathbf{A})$ is positive definite and $\mathbf{PREC}[\mathbf{r}, \xi] \to \mathbf{d}$ is an algorithm of linear complexity to approximate the action of $\mathbf{S}$. We analyze here the preconditioning of the algorithm from the preceding chapter. First we define the routine for approximately solving the preconditioned Galerkin system $\mathbf{P}_\Lambda\mathbf{S}\mathbf{A}\mathbf{u}_\Lambda = \mathbf{P}_\Lambda\mathbf{S}\mathbf{f}$, with $\Lambda \subset \nabla$.

Algorithm 4.3.1 Galerkin system solver $\mathbf{GALSOLVE}[\Lambda, \mathbf{v}_\Lambda, \nu, \eta, \varepsilon] \to \mathbf{w}_\Lambda$

Parameters: Let $\mathbf{A}$ be s^*-computable for some $s^* > 0$. With $\mathbf{A}_j$ the compressed matrices from Definition 2.7.8, let J be such that

$$\varrho := \|\mathbf{S}\mathbf{A} - p_k(\mathbf{A}_J)\mathbf{A}_J\|\|(\mathbf{S}\mathbf{A})^{-1}\| \leq \tfrac{\alpha_\varrho\varepsilon}{\eta+(1-\alpha_\varrho)\varepsilon}.$$

Let $\alpha_d, \alpha_r, \alpha_g, \alpha_\varrho > 0$ be constants such that $\alpha_d+\alpha_r+\alpha_g+\alpha_\varrho = 1$ and $\alpha_\varrho \leq \frac{1}{2}$.

Input: Let $\Lambda \subset \nabla$, $\#\Lambda < \infty$, $\mathbf{v}_\Lambda \in \ell_2(\Lambda)$, $\varepsilon > 0$, $\nu \geq \|\mathbf{f} - \mathbf{A}\mathbf{v}_\Lambda\|$ and $\eta \geq \|\mathbf{P}_\Lambda\mathbf{S}(\mathbf{f} - \mathbf{A}\mathbf{v}_\Lambda)\|$.

Output: $\mathbf{w}_\Lambda \in \ell_2(\Lambda)$ and $\|\mathbf{P}_\Lambda\mathbf{S}(\mathbf{f} - \mathbf{A}\mathbf{w}_\Lambda)\| \leq \varepsilon$.

1: $\mathbf{B} := \mathbf{P}_\Lambda \frac{1}{2}\left[p_k(\mathbf{A}_J)\mathbf{A}_J + p_k(\mathbf{A}_J^*)\mathbf{A}_J^*\right]\mathbf{I}_\Lambda$;

2: $\tilde{\mathbf{r}} := \mathbf{RHS}[\mathbf{f}, \frac{\varepsilon_r}{2}] - \mathbf{APPLY}[\mathbf{A}, \mathbf{v}_\Lambda, \frac{\varepsilon_r}{2}]$ with $\varepsilon_r := \frac{\alpha_r\nu\varepsilon}{\alpha_d\varepsilon+\nu\|\mathbf{S}\|}$;

3: $\tilde{\mathbf{d}} := \mathbf{P}_\Lambda\left(\mathbf{PREC}[\tilde{\mathbf{r}}, \frac{\alpha_d\varepsilon}{\nu}]\right)$;

4: To find an $\tilde{\mathbf{x}}$ with $\|\tilde{\mathbf{d}} - \mathbf{B}\tilde{\mathbf{x}}\| \leq \alpha_g\varepsilon$, apply a suitable iterative method for solving $\mathbf{B}\mathbf{x} = \tilde{\mathbf{d}}$, e.g., Conjugate Gradients or Conjugate Residuals;

5: $\mathbf{w}_\Lambda := \mathbf{v}_\Lambda + \tilde{\mathbf{x}}$.

Proposition 4.3.2. *Let $\mathbf{A}$ be s^*-computable, and let $\mathbf{f}$ be s^*-admissible for some $s^* > 0$. Then $\mathbf{w}_\Lambda := \mathbf{GALSOLVE}[\Lambda, \mathbf{v}_\Lambda, \nu, \eta, \varepsilon]$ terminates with $\|\mathbf{P}_\Lambda\mathbf{S}(\mathbf{f} - \mathbf{A}\mathbf{w}_\Lambda)\| \leq \varepsilon$, and for any $s < s^*$, the number of arithmetic operations and storage locations required by the call is bounded by an absolute multiple of $\varepsilon^{-1/s}(|\mathbf{v}_\Lambda|_{\mathcal{A}^s}^{1/s} + |\mathbf{u}|_{\mathcal{A}^s}^{1/s}) + c(\eta/\varepsilon)\#\Lambda$, with $c : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ being some non-decreasing function.*

Proof. Since the proof of Proposition 3.2.4 works for this proposition with slight adjustments, we comment here only on some points. From $\|\mathbf{SA} - p_k(\mathbf{A}_J)\mathbf{A}_J\| \leq \varrho\|(\mathbf{SA})^{-1}\|^{-1}$, we imply that $\mathbf{B}$ is SPD, and that $\kappa(\mathbf{B}) \lesssim 1$ uniformly in η and ε. To prove the first claim of the theorem, one can use

$$\|\mathbf{P}_\Lambda\mathbf{S}(\mathbf{f} - \mathbf{Aw}) - \mathbf{d}\| = \|\mathbf{P}_\Lambda\mathbf{S}(\mathbf{f} - \mathbf{Aw} - \tilde{\mathbf{r}}) + \mathbf{P}_\Lambda\mathbf{S}\tilde{\mathbf{r}} - \mathbf{d}\| \leq \|\mathbf{S}\|\varepsilon_r + \xi\|\tilde{\mathbf{r}}\|,$$

and $\|\tilde{\mathbf{r}}\| \leq \nu + \varepsilon_r$. $\mathbf{B}$ is sparse, and thus the work bound follows. ∎

Algorithm 4.3.3 Preconditioned adaptive method $\mathbf{SOLVE}[\nu_0, \varepsilon] \to \mathbf{v}_i$

Parameters: Let α, δ, and ξ be some positive constants such that with $\tilde{\delta} := \frac{\delta\|\mathbf{S}\|+\xi}{\|\mathbf{S}^{-1}\|^{-1}-\xi}$, $0 < \tilde{\delta} < \alpha < 1$ and $\frac{\alpha+\tilde{\delta}}{1-\tilde{\delta}} < \kappa(\mathbf{SA})^{-\frac{1}{2}}$. Let $0 < \gamma < \frac{1}{3}\kappa(\mathbf{SA})^{-\frac{1}{2}}(\alpha - \tilde{\delta})$ and $\theta > 0$ be constants.

Input: Let $\nu_0 \gtrsim \varepsilon > 0$.

Output: $\mathbf{v}_i \in P$ and $\|\mathbf{f} - \mathbf{Av}_i\| \leq \nu_i \leq \varepsilon$.

1: $i := 0$, $\mathbf{v}_1 := 0$;
2: **loop**
3: $i := i + 1$;
4: $[\tilde{\mathbf{r}}_i, \nu_i] := \mathbf{RES}[\mathbf{v}_i, \theta\nu_{i-1}, \delta, \varepsilon]$;
5: **if** $\nu_i \leq \varepsilon$ **then**
6: Terminate the subroutine.
7: **end if**
8: $\tilde{\mathbf{d}}_i := \mathbf{PREC}[\tilde{\mathbf{r}}_i, \xi]$;
9: $\Lambda_{i+1} := \mathbf{RESTRICT}[\operatorname{supp}\mathbf{v}_i, \tilde{\mathbf{d}}_i, \alpha]$;
10: $\eta_i := \|\tilde{\mathbf{d}}_i\| + (\xi + \delta\|\mathbf{S}\|)\|\tilde{\mathbf{r}}_i\|$;
11: $\mathbf{v}_{i+1} := \mathbf{GALSOLVE}[\Lambda_{i+1}, \mathbf{v}_i, \nu_i, \eta_i, \gamma\|\tilde{\mathbf{d}}_i\|]$;
12: **end loop**

We now define the preconditioned adaptive wavelet solver.

Theorem 4.3.4. *Let* $\mathbf{A}$ *be* s^**-computable, and let* $\mathbf{f}$ *be* s^**-admissible for some* $s^* > 0$. *Then* $\mathbf{u}_\varepsilon := \mathbf{SOLVE}[\nu_0, \varepsilon]$ *terminates with* $\|\mathbf{f} - \mathbf{Au}_\varepsilon\| \leq \varepsilon$. *Moreover, if* $\nu_0 \eqsim \|\mathbf{f}\| \gtrsim \varepsilon$, *and for some* $s < s^*$, $\mathbf{u} \in \mathcal{A}^s$, *then* $\#\operatorname{supp}\mathbf{u}_\varepsilon \lesssim \varepsilon^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}$ *and the number of arithmetic operations and storage locations required by the call is bounded by some absolute multiple of the same expression.*

Proof. From the properties of **RES**, for any $\mathbf{v}_i$ determined inside the loop, with $\mathbf{r}_i := \mathbf{f} - \mathbf{Av}_i$, we have $\nu_i \geq \|\mathbf{r}_i\|$, and either $\nu_i \leq \varepsilon$ or $\|\mathbf{r}_i - \tilde{\mathbf{r}}_i\| \leq \delta\|\tilde{\mathbf{r}}_i\|$. Moreover, we have $\nu_i \gtrsim \min\{\theta\nu_{i-1}, \varepsilon\} \gtrsim \varepsilon$ for $i \geq 1$. Now suppose that for an $i > 0$, **RES** terminates with $\nu_i > \varepsilon$ and thus with $\|\mathbf{r}_i - \tilde{\mathbf{r}}_i\| \leq \delta\|\tilde{\mathbf{r}}_i\|$. Then from

$(1-\delta)\|\tilde{\mathbf{r}}_i\| \leq \|\mathbf{r}_i\| \leq (1+\delta)\|\tilde{\mathbf{r}}_i\|$ and $\nu_i \leq (1+\delta)\|\tilde{\mathbf{r}}_i\|$, we have $\nu_i \eqsim \|\tilde{\mathbf{r}}_i\| \eqsim \|\mathbf{r}_i\|$, and

$$\|\mathbf{S}\mathbf{r}_i - \tilde{\mathbf{d}}_i\| = \|\mathbf{S}(\mathbf{r}_i - \tilde{\mathbf{r}}_i) + \mathbf{S}\tilde{\mathbf{r}}_i - \tilde{\mathbf{d}}_i\| \leq \|\mathbf{S}\|\delta\|\tilde{\mathbf{r}}_i\| + \xi\|\tilde{\mathbf{r}}_i\|,$$

so η_i is an upper bound on $\|\mathbf{S}\mathbf{r}_i\|$. Furthermore, we have

$$\|\tilde{\mathbf{r}}_i\| \leq \|\mathbf{S}^{-1}\|\|\mathbf{S}\tilde{\mathbf{r}}_i\| \leq \|\mathbf{S}^{-1}\| \left(\|\tilde{\mathbf{d}}_i\| + \xi\|\tilde{\mathbf{r}}_i\|\right),$$

implying that $\|\mathbf{S}\mathbf{r}_i - \tilde{\mathbf{d}}_i\| \leq \tilde{\delta}\|\tilde{\mathbf{d}}_i\|$ and $\eta_i \leq (1+\tilde{\delta})\|\tilde{\mathbf{d}}_i\|$. Similarly to the above case with ν_i, we now infer $\eta_i \eqsim \|\tilde{\mathbf{d}}_i\| \eqsim \|\mathbf{S}\mathbf{r}_i\|$, and since $\|\mathbf{S}\mathbf{r}_i\| \eqsim \|\mathbf{r}_i\|$, we have $\eta_i \eqsim \nu_i$. With the norm $|\!|\!|\cdot|\!|\!| := \langle \mathbf{S}\mathbf{A}\cdot, \cdot\rangle^{\frac{1}{2}}$ which is equivalent to the standard norm $\|\cdot\|$ on ℓ_2, Proposition 3.2.2 shows that $|\!|\!|\mathbf{u}-\mathbf{v}_{i+1}|\!|\!| \leq \rho|\!|\!|\mathbf{u}-\mathbf{v}_i|\!|\!|$ for some $\rho \in [0,1)$, or $\nu_{i+1} \lesssim \rho^{i-k}\nu_k$ for $0 \leq k \leq i+1$. This proves that the loop terminates after a finite number of iterations, say directly after the K-the call of **RES**.

Since **PREC** is of linear complexity we have $\#\operatorname{supp}\tilde{\mathbf{d}}_i \lesssim \#\operatorname{supp}\tilde{\mathbf{r}}_i$, and the cost of the i-th call of **PREC** is of order $\#\operatorname{supp}\tilde{\mathbf{r}}_i + 1$. The rest of the proof is completely analogous to the analysis in the proof of Theorem 3.3.5. ∎

Chapter 5

Adaptive algorithm for nonsymmetric and indefinite elliptic problems

5.1 Introduction

Let $\mathcal{H}$ be a real Hilbert space and let $\mathcal{H}'$ denote its dual. Given a boundedly invertible linear operator $L : \mathcal{H} \to \mathcal{H}'$ and a linear functional $f \in \mathcal{H}'$, we consider the problem of finding $u \in \mathcal{H}$ such that

$$Lu = f.$$

As an example of $\mathcal{H}$ one can think of the Sobolev space H^t on a domain or manifold, possibly incorporating essential boundary conditions. Then the weak formulation of (scalar) linear differential or integral equations of order $2t$ leads to the above type of equations.

Let $\Psi = \{\psi_\lambda \in \mathcal{H} : \lambda \in \nabla\}$ be a *Riesz basis* of wavelet type for $\mathcal{H}$ with a countable index set ∇. We consider Ψ formally as a column vector whose entries are elements of $\mathcal{H}$. Let $u = \mathbf{u}^T\Psi$ with $\mathbf{u}$ a column vector in $\ell_2 := \ell_2(\nabla)$. Then, as we already have seen, the above problem is *equivalent* to finding $\mathbf{u} \in \ell_2$ satisfying the infinite matrix-vector system

$$\mathbf{Lu} = \mathbf{f}, \tag{5.1.1}$$

where $\mathbf{L} := \langle \psi_\lambda, L\psi_\mu\rangle_{\lambda,\mu\in\nabla} : \ell_2 \to \ell_2$ is boundedly invertible and $\mathbf{f} := \langle f, \psi_\lambda\rangle_{\lambda\in\nabla} \in \ell_2$. Here $\langle\cdot,\cdot\rangle$ denotes the duality product on $\mathcal{H} \times \mathcal{H}'$.

In the foregoing chapters, we have encountered a number of adaptive methods for solving the above type of equations. The methods apply under the condition that $\mathbf{L}$ is *symmetric, positive definite* (SPD), which is equivalent to $\langle Lv, w\rangle = \langle v, Lw\rangle$, $v, w \in \mathcal{H}$, and $\langle Lv, v\rangle \gtrsim \|v\|^2_{\mathcal{H}}$, $v \in \mathcal{H}$, i.e., that L is self-adjoint and

$\mathcal{H}$-elliptic. For the case that L does not have both properties, as was suggested in [18], one can reformulate $\mathbf{Lu} = \mathbf{f}$ as an equivalent well-posed infinite matrix-vector problem with a symmetric, positive definite system matrix, as via the normal equations, or, in case the equation represents a saddle point problem, by using the reformulation as a positive definite system introduced in [10].

Throughout this chapter, we will consider the operators of type $L = A + B$ where A is self-adjoint and $\mathcal{H}$-elliptic, and B is compact. Now in general $\mathbf{L}$ is no longer SPD, hence the above mentioned adaptive wavelet methods cannot be applied directly. One can consider the normal equation $\mathbf{L}^T\mathbf{Lu} = \mathbf{L}^T\mathbf{f}$; however, the main disadvantage of this approach is that the condition number of the system is squared, while the quantitative properties of the methods depend sensitively on the conditioning of the system. In this chapter, we will modify the adaptive wavelet algorithm from Chapter 3 so that it applies directly to the system $\mathbf{Lu} = \mathbf{f}$, avoiding the normal equations. The analysis in Chapter 3 extensively uses the Galerkin orthogonality, which in the present case has to be replaced by only a *quasi-orthogonality* property. It should be mentioned that this quasi-orthogonality property has been used in [62] in a convergence proof of an adaptive finite element method. By proving the quasi-orthogonality property for the present general setting and extending the complexity analysis in Chapter 3, we will show that our algorithm has optimal computational complexity.

This chapter is organized as follows. In the following section, we derive results on Ritz-Galerkin approximations to the exact solution, and in the last section, the adaptive wavelet algorithm is constructed and analyzed.

5.2 Ritz-Galerkin approximations

Let $\mathcal{H} \hookrightarrow \mathcal{Y}$ be separable real Hilbert spaces with compact embedding, and let $a : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ and $b : \mathcal{Y} \times \mathcal{H} \to \mathbb{R}$ be bounded bilinear forms. We assume that the bilinear form a is symmetric and elliptic, which implies that $|||\cdot||| := a(\cdot,\cdot)^{\frac{1}{2}}$ is an equivalent norm on $\mathcal{H}$, i.e.,

$$|||v||| \eqsim \|v\|_{\mathcal{H}} \qquad v \in \mathcal{H}. \tag{5.2.1}$$

In particular, the operator $A : \mathcal{H} \to \mathcal{H}'$ defined by $\langle Av, w\rangle = a(v, w)$ for $v, w \in \mathcal{H}$, is boundedly invertible. Moreover, since $B : \mathcal{H} \to \mathcal{H}'$ defined by $\langle Bv, w\rangle = b(v, w)$ for $v, w \in \mathcal{H}$, is compact, the linear operator $L := A + B$ is a Fredholm operator of index zero. Therefore, assuming that L is injective, $L : \mathcal{H} \to \mathcal{H}'$ is boundedly invertible, in particular meaning that the linear operator equation

$$Lu = f, \tag{5.2.2}$$

has a unique solution for $f \in \mathcal{H}'$.

For our analysis we will need the following mild *regularity assumption* on the adjoint L' of L: There is a Hilbert space $\mathcal{X} \hookrightarrow \mathcal{H}$ with compact embedding, such that $(L')^{-1} : \mathcal{Y}' \to \mathcal{X}$ is bounded. The following lemma gives a means to check this assumption.

Lemma 5.2.1. *Let either $A^{-1} : \mathcal{Y}' \to \mathcal{X}$ or $L^{-1} : \mathcal{Y}' \to \mathcal{X}$ be bounded. Then $(L')^{-1} : \mathcal{Y}' \to \mathcal{X}$ is bounded.*

Proof. We treat the first case only. The other case is analogous. The operator B extends to a bounded mapping from $\mathcal{Y}$ to $\mathcal{H}'$. So $L' = A + B' : \mathcal{X} \to \mathcal{Y}'$ is bounded. Now consider the equation $L'u = f$. We know that there exists a unique solution $u \in \mathcal{H}$ with $\|u\|_{\mathcal{H}} \lesssim \|f\|_{\mathcal{H}'}$ and thus $\|B'u\|_{\mathcal{Y}'} \lesssim \|u\|_{\mathcal{H}} \lesssim \|f\|_{\mathcal{H}'} \lesssim \|f\|_{\mathcal{Y}'}$. From $Au = f - B'u$, we now infer that $\|u\|_{\mathcal{X}} \lesssim \|f\|_{\mathcal{Y}'}$. ∎

Example 5.2.2. For some Lipschitz domain $\Omega \subset \mathbb{R}^n$, with $\mathcal{H} := H_0^1(\Omega)$ let $L : \mathcal{H} \to \mathcal{H}'$ be defined by

$$\langle Lv, w\rangle = -\sum_{j,k=1}^{n} \langle a_{jk}\partial_k v, \partial_j w\rangle_{L_2} + \sum_{k=1}^{n} \langle b_k \partial_k v, w\rangle_{L_2} + \langle cv, w\rangle_{L_2} \quad v, w \in \mathcal{H}.$$

If the coefficients satisfy $a_{jk}, b_k, c \in L_\infty$ then $L : \mathcal{H} \to \mathcal{H}'$ is bounded. Moreover, if the matrix $[a_{jk}]$ is symmetric and uniformly positive definite a.e. in Ω, then the bilinear form $a(\cdot,\cdot) := -\sum_{j,k=1}^{n} \langle a_{jk}\partial_k \cdot, \partial_j \cdot\rangle_{L_2}$ is symmetric and satisfies (5.2.1). If either $b_k = 0$, $1 \le k \le n$ and $c \ge 0$ a.e. or $c \ge \beta > 0$ a.e., then the generalized maximum principle implies that L is injective, cf. [81]. Also if $L = A - \eta^2$ for a constant $\eta \in \mathbb{R}$, then the injectivity is guaranteed as long as η^2 is not an eigenvalue of A. With $\mathcal{Y}_\sigma := (L_2(\Omega), H_0^1(\Omega))_{1-\sigma,2}$ for some $\sigma \in (0,1]$, where $(X,Y)_{\theta,p}$ denotes the intermediate space between X and Y obtained by the real interpolation method, the bilinear form $b(\cdot,\cdot) := \sum_{k=1}^{n} \langle b_k \partial_k \cdot, \cdot\rangle_{L_2} + \langle c\cdot, \cdot\rangle_{L_2} : \mathcal{Y}_\sigma \times \mathcal{H} \to \mathbb{R}$ is bounded for any $\sigma \in (0,1]$. If the coefficients a_{jk}, $1 \le j,k \le n$, are Lipschitz continuous, then with $\mathcal{X}_\sigma := (H_0^1(\Omega), H^2(\Omega) \cap H_0^1(\Omega))_{\sigma,2}$ it is known that $A^{-1} : \mathcal{Y}_\sigma' \to \mathcal{X}_\sigma$ is bounded for any $\sigma \in (0, \frac{1}{2})$, cf. [75]. Furthermore, the embeddings $\mathcal{X}_\sigma \hookrightarrow \mathcal{H} \hookrightarrow \mathcal{Y}_\sigma$ are compact. From Lemma 5.2.1 we conclude that all aforementioned conditions are satisfied. ⊘

Example 5.2.3. Let L be the operator considered in the above example. We assume that the domain Ω is Lipschitz, the coefficients a_{jk}, b_k, c are constant and that the matrix $[a_{jk}]$ is symmetric and positive definite. Then the single layer and hypersingular boundary integral operators corresponding to the differential operator L can be written as the sum of a bounded $\mathcal{H}$-elliptic operator $A : \mathcal{H} \to \mathcal{H}'$ and a compact operator $B : \mathcal{H} \to \mathcal{H}'$, see [23]. With Γ being the boundary of the underlying domain Ω, here the energy space is $\mathcal{H} = H^t(\Gamma)$ with $t = -\frac{1}{2}$ for the

single layer operator and $t = \frac{1}{2}$ for the hypersingular integral operator. A close inspection of the proofs of [24, Theorem 3.9] and [23, Theorem 2] reveals that in both cases, the operator A is self-adjoint and that with $\mathcal{Y}_\sigma := H^{t-\sigma}(\Gamma)$ where t has the appropriate value depending on the case, the operator B can be extended to a bounded operator $\mathcal{Y}_\sigma \to \mathcal{H}'$ for any $\sigma \in (0, \frac{1}{2}]$. Assuming the injectivity of $L : \mathcal{H} \to \mathcal{H}'$, in [23] it is shown that with $\mathcal{X}_\sigma := H^{t+\sigma}(\Gamma)$, $L^{-1} : \mathcal{Y}'_\sigma \to \mathcal{X}_\sigma$ is bounded for any $\sigma \in [0, \frac{1}{2}]$. The injectivity depends on the particular case at hand, see [61] for some important cases. ⊘

We consider a sequence of finite dimensional closed subspaces $V_0 \subset V_1 \subset \ldots \subset \mathcal{H}$ satisfying

$$\inf_{v_j \in V_j} \|v - v_j\|_{\mathcal{H}} \leq \alpha_j \|v\|_{\mathcal{X}} \qquad v \in \mathcal{X}, \tag{5.2.3}$$

with $\lim_{j\to\infty} \alpha_j = 0$.

Remark 5.2.4. Such a sequence exists since the embedding $\mathcal{X} \hookrightarrow \mathcal{H}$ is compact, cf. [77].

Example 5.2.5. Let $\mathcal{H} = H^t$ and $\mathcal{X} = H^{t+\sigma}$. Then for standard finite element or spline spaces V_j subordinate to dyadic subdivisions of an initial mesh, the approximation property (5.2.3) is satisfied with $\alpha_j \eqsim 2^{-j\sigma}$, for any $t < \gamma$ and $\sigma \leq d - t$, where d is the polynomial order of the spaces V_j, and $\gamma = \sup_j\{s : V_j \subset H^s\}$, see e.g. [64]. ⊘

For a finite dimensional closed subspace $S \subset \mathcal{H}$ such that $V_j \subseteq S$ for some j, we consider the Ritz-Galerkin problem

$$\langle Lu_S, v_S\rangle = \langle f, v_S\rangle \qquad \text{for all } v_S \in S. \tag{5.2.4}$$

It is well known that for j being sufficiently large, a unique solution u_S to the above problem exists, and that u_S is a near best approximation to u in the energy norm $|\!|\!| \cdot |\!|\!|$. In the weaker norm $\|\cdot\|_{\mathcal{Y}}$, convergence of higher order than (5.2.3) can be obtained via an Aubin-Nitsche duality argument, cf. [76]. These results are recalled in the following lemma, where for convenience we also include a proof.

Lemma 5.2.6. *There is an absolute constant $j_0 \in \mathbb{N}_0$ (not depending on S) such that for all $j \geq j_0$,* (5.2.4) *has a unique solution with*

$$|\!|\!| u - u_S |\!|\!| \leq [1 + O(\alpha_j)] \inf_{v\in S} |\!|\!| u - v |\!|\!|. \tag{5.2.5}$$

Moreover, for $j \geq j_0$ we have

$$\|u - u_S\|_{\mathcal{Y}} \leq O(\alpha_j) |\!|\!| u - u_S |\!|\!|. \tag{5.2.6}$$

Proof. Suppose that a solution u_S to (5.2.4) exists. Then we trivially have

$$\langle L(u - u_S), v_S \rangle = 0 \qquad \forall v_S \in S. \tag{5.2.7}$$

Using this and the boundedness of $b : \mathcal{Y} \times \mathcal{H} \to \mathbb{R}$, for arbitrary $v_S \in S$ we get

$$\begin{aligned} |\!|\!| u - u_S |\!|\!|^2 &= \langle L(u - u_S), u - u_S \rangle - b(u - u_S, u - u_S) \\ &= \langle L(u - u_S), u - v_S \rangle - b(u - u_S, u - u_S) \\ &= a(u - u_S, u - v_S) + b(u - u_S, u_S - v_S) \\ &\leq |\!|\!| u - u_S |\!|\!| |\!|\!| u - v_S |\!|\!| + O(1) \|u - u_S\|_{\mathcal{Y}} \|u_S - v_S\|_{\mathcal{H}}. \end{aligned} \tag{5.2.8}$$

We estimate $\|u - u_S\|_{\mathcal{Y}}$ by an Aubin-Nitsche duality argument. For $w \in \mathcal{Y}'$ we infer that

$$\begin{aligned} \langle u - u_S, w \rangle &= \langle L(u - u_S), (L')^{-1} w - w_S \rangle \\ &\leq \|L\|_{\mathcal{H} \to \mathcal{H}'} \|u - u_S\|_{\mathcal{H}} \|(L')^{-1} w - w_S\|_{\mathcal{H}} \\ &\leq \|L\|_{\mathcal{H} \to \mathcal{H}'} \|u - u_S\|_{\mathcal{H}} \alpha_j \|(L')^{-1} w\|_{\mathcal{X}} \\ &\leq \|L\|_{\mathcal{H} \to \mathcal{H}'} \|u - u_S\|_{\mathcal{H}} \alpha_j \|(L')^{-1}\|_{\mathcal{Y}' \to \mathcal{X}} \|w\|_{\mathcal{Y}'}, \end{aligned}$$

where we used (5.2.7), (5.2.3) and the boundedness of $(L')^{-1} : \mathcal{Y}' \to \mathcal{X}$. We have

$$\|u - u_S\|_{\mathcal{Y}} = \sup_{w \in \mathcal{Y}'} \frac{\langle u - u_S, w \rangle}{\|w\|_{\mathcal{Y}'}},$$

and subsequently using (5.2.1) we arrive at (5.2.6). Substituting (5.2.6) into (5.2.8), we get

$$|\!|\!| u - u_S |\!|\!| \leq |\!|\!| u - v_S |\!|\!| + O(\alpha_j) \|u_S - v_S\|_{\mathcal{H}}.$$

For the last term, from the triangle inequality and (5.2.1), we have

$$\|v - u_S\|_{\mathcal{H}} \lesssim |\!|\!| u - u_S |\!|\!| + |\!|\!| u - v_S |\!|\!|.$$

Now choosing j_0 sufficiently large, we finally obtain (5.2.5).

Since (5.2.4) is a finite dimensional system, existence and uniqueness are equivalent. To see the uniqueness, it is sufficient to prove that $f = 0$ implies $u_S = 0$. By linearity and invertibility of L, we have $u = 0$ if $f = 0$, and so (5.2.5) implies that $u_S = 0$. The proof is completed. ∎

The following observation concerning quasi-orthogonality is an easy generalization of [62, Lemma 2.1].

Lemma 5.2.7. *For some $j \geq j_0$ with j_0 being the absolute constant from Lemma 5.2.6, let $S_0 \subset S_1 \subset \mathcal{H}$ be finite dimensional subspaces satisfying $V_j \subseteq S_0$. Let $u_0 \in S_0$ and $u_1 \in S_1$ be the solutions to the Galerkin problems $\langle Lu_0, v\rangle = \langle f, v\rangle$ $\forall v \in S_0$ and $\langle Lu_1, v\rangle = \langle f, v\rangle$ $\forall v \in S_1$, respectively. Then we have*

$$\left|\|\!|u-u_0|\!\|^2 - \|\!|u-u_1|\!\|^2 - \|\!|u_1-u_0|\!\|^2\right| \leq O(\alpha_j)\left(\|\!|u-u_0|\!\|^2 + \|\!|u-u_1|\!\|^2\right). \tag{5.2.9}$$

Proof. We have $\|\!|u-u_0|\!\|^2 = \|\!|u-u_1|\!\|^2 + \|\!|u_1-u_0|\!\|^2 + 2a(u-u_1, u_1-u_0)$. Using (5.2.7), boundedness of $b : \mathcal{Y} \times \mathcal{H} \to \mathbb{R}$, and the triangle inequality, we estimate the absolute value of the last term as

$$\begin{aligned} |2a(u-u_1, u_1-u_0)| &= |2b(u-u_1, u_1-u_0)| \\ &\lesssim \|u-u_1\|_{\mathcal{Y}} \|u_1-u_0\|_{\mathcal{H}} \\ &\leq \|u-u_1\|_{\mathcal{Y}} \left(\|u-u_1\|_{\mathcal{H}} + \|u-u_0\|_{\mathcal{H}}\right) \end{aligned}$$

Now using (5.2.6), and applying the inequality $2ab \leq a^2+b^2$, $a, b \in \mathbb{R}$, we conclude the proof by

$$\begin{aligned} |2a(u-u_1, u_1-u_0)| &\leq O(\alpha_j)\left(\|\!|u-u_1|\!\|^2 + \|\!|u-u_1|\!\|\|\!|u-u_0|\!\|\right) \\ &\leq O(\alpha_j)\left(\|\!|u-u_1|\!\|^2 + \|\!|u-u_0|\!\|^2\right). \end{aligned}$$

■

Using a *Riesz basis* for $\mathcal{H}$, we will now transform (5.2.2) into an equivalent infinite matrix-vector system in ℓ_2. Let $\Psi = \{\psi_\lambda : \lambda \in \nabla\}$ be a Riesz basis for $\mathcal{H}$ of wavelet type. We assume that for some $\nabla_0 \subset \nabla_1 \subset \ldots \subset \nabla$, the subspaces defined by $V_j = \operatorname{span}\{\psi_\lambda : \lambda \in \nabla_j\}$, $j \in \mathbb{N}_0$, satisfies (5.2.3) with $\lim_{j\to\infty} \alpha_j = 0$.

Example 5.2.8. With the spaces V_j described in Example 5.2.5, wavelet bases satisfying the above condition have been constructed e.g. in [14, 22, 33, 34, 56, 85]. ⊘

Writing $u = \mathbf{u}^T\Psi$ for some $\mathbf{u} \in \ell_2$, $\mathbf{u}$ satisfies

$$\mathbf{L}\mathbf{u} = \mathbf{f}, \tag{5.2.10}$$

where $\mathbf{L} := \langle \psi_\lambda, L\psi_\mu\rangle_{\lambda,\mu\in\nabla} : \ell_2 \to \ell_2$ is boundedly invertible and $\mathbf{f} := \langle f, \psi_\lambda\rangle_{\lambda\in\nabla} \in \ell_2$. Similarly to $\mathbf{L}$, we define also the matrices

$$\begin{aligned} \mathbf{A} &:= \langle \psi_\lambda, A\psi_\mu\rangle_{\lambda,\mu\in\nabla} = a(\psi_\mu, \psi_\lambda)_{\lambda,\mu\in\nabla} \qquad \text{and} \\ \mathbf{B} &:= \langle \psi_\lambda, B\psi_\mu\rangle_{\lambda,\mu\in\nabla} = b(\psi_\mu, \psi_\lambda)_{\lambda,\mu\in\nabla}, \end{aligned}$$

so that $\mathbf{L} = \mathbf{A} + \mathbf{B}$. The matrix $\mathbf{A}$ is symmetric positive definite, so $\langle \mathbf{A}\cdot, \cdot\rangle$ is an inner product on ℓ_2, and the induced norm $\|\!|\cdot|\!\|$ satisfies

$$\|\!|\mathbf{v}|\!\|^2 := \langle \mathbf{A}\mathbf{v}, \mathbf{v}\rangle = a(\mathbf{v}^T\Psi, \mathbf{v}^T\Psi) = \|\!|\mathbf{v}^T\Psi|\!\|^2 \qquad \mathbf{v} \in \ell_2.$$

Furthermore, one can verify that for any $\mathbf{v} \in \ell_2$, $\Lambda \subseteq \nabla$, $\mathbf{v}_\Lambda \in \ell_2(\Lambda)$,

$$\|\mathbf{A}\mathbf{v}\| \leq \|\mathbf{A}\|^{\frac{1}{2}} \|\!|\mathbf{v}|\!\| \leq \|\mathbf{A}\| \|\mathbf{v}\|, \qquad \|\!|\mathbf{v}_\Lambda|\!\| \leq \|\mathbf{A}^{-1}\|^{\frac{1}{2}} \|\mathbf{P}_\Lambda \mathbf{A} \mathbf{I}_\Lambda \mathbf{v}_\Lambda\|, \tag{5.2.11}$$

where $\mathbf{P}_\Lambda : \ell_2 \to \ell_2(\Lambda)$ is the orthogonal projector onto $\ell_2(\Lambda)$, and $\mathbf{I}_\Lambda$ denotes the trivial inclusion $\ell_2(\Lambda) \to \ell_2$. For any $\mathbf{v}, \mathbf{w} \in \ell_2$, we have $\langle \mathbf{B}\mathbf{v}, \mathbf{w} \rangle = b(\mathbf{v}^T\Psi, \mathbf{w}^T\Psi) \lesssim \|\mathbf{v}^T\Psi\|_{\mathcal{Y}} \|\mathbf{w}^T\Psi\|_{\mathcal{H}} \lesssim \|\mathbf{v}^T\Psi\|_{\mathcal{Y}} \|\mathbf{w}\|$, implying the following estimate which will be used often in the rest of this section.

$$\|\mathbf{B}\mathbf{v}\| = \sup_{0 \neq \mathbf{w} \in \ell_2} \frac{\langle \mathbf{B}\mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{w}\|} \lesssim \|\mathbf{v}^T\Psi\|_{\mathcal{Y}} \qquad \mathbf{v} \in \ell_2. \tag{5.2.12}$$

For some $\Lambda \subset \nabla$, let $S = \operatorname{span}\{\psi_\lambda : \lambda \in \Lambda\} \subset \mathcal{H}$. Then $u_S = (\mathbf{I}_\Lambda \mathbf{u}_\Lambda)^T \Psi \in S$ is the solution to the Galerkin problem (5.2.4) if and only if $\mathbf{u}_\Lambda \in \ell_2(\Lambda)$ satisfies

$$\mathbf{P}_\Lambda \mathbf{L} \mathbf{I}_\Lambda \mathbf{u}_\Lambda = \mathbf{P}_\Lambda \mathbf{f}. \tag{5.2.13}$$

In the following, we will refer to $\mathbf{u}_\Lambda$ as the *Galerkin solution* with respect to the index set Λ. From Lemma 5.2.6 we know that this solution exists and is unique when $\nabla_j \subseteq \Lambda$ for some $j \geq j_0$.

Lemma 5.2.9. *Let $\mathbf{P}_\Lambda$ and $\mathbf{I}_\Lambda$ be as above. Then for any $\Lambda \supseteq \nabla_j$ for some $j \geq j_0$ we have*

$$\|(\mathbf{P}_\Lambda \mathbf{L} \mathbf{I}_\Lambda)^{-1}\| \leq \|\mathbf{A}^{-1}\| \left[1 + \|\mathbf{B}\mathbf{L}^{-1}\| + O(\alpha_j)\right].$$

Proof. Recalling that $\mathbf{L}(\mathbf{u} - \mathbf{u}_\Lambda) \perp \ell_2(\Lambda)$ and that $\mathbf{A} = \mathbf{L} - \mathbf{B}$, we have

$$\begin{aligned} \|\mathbf{u}_\Lambda\|^2 \leq \|\mathbf{A}^{-1}\| \|\!|\mathbf{u}_\Lambda|\!\|^2 &= \|\mathbf{A}^{-1}\| \left[\langle \mathbf{L}\mathbf{u}_\Lambda, \mathbf{u}_\Lambda \rangle - \langle \mathbf{B}\mathbf{u}_\Lambda, \mathbf{u}_\Lambda \rangle\right] \\ &= \|\mathbf{A}^{-1}\| \left[\langle \mathbf{L}\mathbf{u}, \mathbf{u}_\Lambda \rangle - \langle \mathbf{B}\mathbf{u}, \mathbf{u}_\Lambda \rangle + \langle \mathbf{B}(\mathbf{u} - \mathbf{u}_\Lambda), \mathbf{u}_\Lambda \rangle\right]. \end{aligned}$$

Here and in the following, we write $\mathbf{u}_\Lambda$ to mean $\mathbf{I}_\Lambda \mathbf{u}_\Lambda$ as well, i.e., $\mathbf{u}_\Lambda$ is extended by zeros outside the index set Λ. Now applying the Cauchy-Bunyakowsky-Schwarz inequality gives

$$\|\mathbf{u}_\Lambda\| \leq \|\mathbf{A}^{-1}\| \left[\|\mathbf{L}\mathbf{u}\| + \|\mathbf{B}\mathbf{u}\| + \|\mathbf{B}(\mathbf{u} - \mathbf{u}_\Lambda)\|\right]. \tag{5.2.14}$$

For the last term in the brackets, using the estimates (5.2.12), (5.2.6) and (5.2.5), we have

$$\begin{aligned} \|\mathbf{B}(\mathbf{u} - \mathbf{u}_\Lambda)\| \lesssim \|u - \mathbf{u}_\Lambda^T \Psi\|_{\mathcal{Y}} &\leq O(\alpha_j) \|\!|u - \mathbf{u}_\Lambda^T \Psi|\!\| \\ &\leq O(\alpha_j) \inf_{\mathbf{v} \in \ell_2(\Lambda)} \|\!|\mathbf{u} - \mathbf{v}|\!\| \leq O(\alpha_j) \|\mathbf{u}\|. \end{aligned}$$

We substitute it into (5.2.14) to get

$$\begin{aligned}\|\mathbf{u}_\Lambda\| &\le \|\mathbf{A}^{-1}\| \left[\|\mathbf{L}\mathbf{u}\| + \|\mathbf{B}\mathbf{u}\| + O(\alpha_j)\|\mathbf{u}\|\right] \\ &\le \|\mathbf{A}^{-1}\| \left[1 + \|\mathbf{B}\mathbf{L}^{-1}\| + O(\alpha_j)\|\mathbf{L}^{-1}\|\right] \|\mathbf{f}\|.\end{aligned}$$

Since this estimate holds in particular for arbitrary $\mathbf{f} = \mathbf{P}_\Lambda \mathbf{f}$, taking into account that $\mathbf{u}_\Lambda = (\mathbf{P}_\Lambda \mathbf{L} \mathbf{I}_\Lambda)^{-1} \mathbf{P}_\Lambda \mathbf{f}$ the proof is completed. ∎

The following lemma generalizes Lemma 3.2.1 to the present case of nonsymmetric and indefinite operators, and provides a way to extend a given set $\Lambda_0 \subset \nabla$ such that the error of the Galerkin solution with respect to the extended set is reduced by a constant factor.

Lemma 5.2.10. *Suppose that $\mathbf{u}_0 \in \ell_2(\Lambda_0)$ is the solution to $\mathbf{P}_{\Lambda_0}\mathbf{L}\mathbf{I}_{\Lambda_0}\mathbf{u}_0 = \mathbf{P}_{\Lambda_0}\mathbf{f}$ with $\Lambda_0 \supseteq \nabla_j$ for j sufficiently large. For a constant $\mu \in (0,1)$, let $\nabla \supset \Lambda_1 \supset \Lambda_0$ be such that*

$$\|\mathbf{P}_{\Lambda_1}(\mathbf{f} - \mathbf{L}\mathbf{u}_0)\| \ge \mu \|\mathbf{f} - \mathbf{L}\mathbf{u}_0\|. \tag{5.2.15}$$

Then, for $\mathbf{u}_1 \in \ell_2(\Lambda_1)$ being the solution to $\mathbf{P}_{\Lambda_1}\mathbf{L}\mathbf{I}_{\Lambda_1}\mathbf{u}_1 = \mathbf{P}_{\Lambda_1}\mathbf{f}$, it holds that

$$|||\mathbf{u} - \mathbf{u}_1||| \le \left[1 - \kappa(\mathbf{A})^{-1}\mu^2 + O(\alpha_j)\right]^{\frac{1}{2}} |||\mathbf{u} - \mathbf{u}_0|||.$$

Proof. In this proof, we use the notations $u_0 = \mathbf{u}_0^T \Psi$ and $u_1 = \mathbf{u}_1^T \Psi$. We have

$$\|\mathbf{L}(\mathbf{u}_1 - \mathbf{u}_0)\|^2 = \|\mathbf{A}(\mathbf{u}_1 - \mathbf{u}_0)\|^2 + 2\langle \mathbf{A}(\mathbf{u}_1 - \mathbf{u}_0), \mathbf{B}(\mathbf{u}_1 - \mathbf{u}_0)\rangle + \|\mathbf{B}(\mathbf{u}_1 - \mathbf{u}_0)\|^2.$$

The first term on the right hand side is bounded from above by using the first inequality from (5.2.11). We estimate the second term by using (5.2.12) as

$$\begin{aligned}|2\langle \mathbf{A}(\mathbf{u}_1 - \mathbf{u}_0), \mathbf{B}(\mathbf{u}_1 - \mathbf{u}_0)\rangle| &\le 2\|\mathbf{A}(\mathbf{u}_1 - \mathbf{u}_0)\| \|\mathbf{B}(\mathbf{u}_1 - \mathbf{u}_0)\| \\ &\lesssim |||u_1 - u_0||| \, \|u_1 - u_0\|_{\mathcal{Y}}.\end{aligned}$$

For the third term we have $\|\mathbf{B}(\mathbf{u}_1 - \mathbf{u}_0)\|^2 \lesssim \|u_1 - u_0\|_{\mathcal{Y}}^2$. Combining these estimates, and taking into account (5.2.6), we conclude that

$$\begin{aligned}\|\mathbf{L}(\mathbf{u}_1 - \mathbf{u}_0)\|^2 &\le \|\mathbf{A}\| |||u_1 - u_0|||^2 + O(1) |||u_1 - u_0||| \, \|u_1 - u_0\|_{\mathcal{Y}} \\ &\le \|\mathbf{A}\| |||u_1 - u_0|||^2 + O(\alpha_j) \left(|||u - u_0|||^2 + |||u - u_1|||^2\right).\end{aligned} \tag{5.2.16}$$

On the other hand, we have

$$\|\mathbf{L}(\mathbf{u} - \mathbf{u}_0)\|^2 = \|\mathbf{A}(\mathbf{u} - \mathbf{u}_0)\|^2 + 2\langle \mathbf{A}(\mathbf{u} - \mathbf{u}_0), \mathbf{B}(\mathbf{u} - \mathbf{u}_0)\rangle + \|\mathbf{B}(\mathbf{u} - \mathbf{u}_0)\|^2.$$

The first term can be bounded from below by using the last inequality in (5.2.11) with $\Lambda = \nabla$. By using (5.2.12) and (5.2.6), we bound the second term as

$$|2\langle \mathbf{A}(\mathbf{u}-\mathbf{u}_0), \mathbf{B}(\mathbf{u}-\mathbf{u}_0)\rangle| \lesssim |\!|\!| u-u_0 |\!|\!| \, \|u-u_0\|_{\mathcal{Y}} \leq O(\alpha_j) |\!|\!| u-u_0 |\!|\!|^2. \quad (5.2.17)$$

Estimating the third term by zero, we infer

$$\|\mathbf{L}(\mathbf{u}-\mathbf{u}_0)\|^2 \geq \|\mathbf{A}^{-1}\|^{-1} |\!|\!| u-u_0 |\!|\!|^2 - O(\alpha_j) |\!|\!| u-u_0 |\!|\!|^2. \quad (5.2.18)$$

By hypothesis we have $\|\mathbf{L}(\mathbf{u}_1-\mathbf{u}_0)\| \geq \|\mathbf{P}_{\Lambda_1}\mathbf{L}(\mathbf{u}_1-\mathbf{u}_0)\| = \|\mathbf{P}_{\Lambda_1}(\mathbf{f}-\mathbf{L}\mathbf{u}_0)\| \geq \mu\|\mathbf{L}(\mathbf{u}-\mathbf{u}_0)\|$. Combining this with (5.2.16) and (5.2.18), we get

$$\begin{aligned} \|\mathbf{A}\| |\!|\!| u_1-u_0 |\!|\!|^2 + O(\alpha_j) |\!|\!| u-u_1 |\!|\!|^2 & \\ \geq \mu^2 \|\mathbf{A}^{-1}\|^{-1} |\!|\!| u-u_0 |\!|\!|^2 - O(\alpha_j) |\!|\!| u-u_0 |\!|\!|^2. & \end{aligned}$$

Now by using that $|\!|\!| u_1-u_0 |\!|\!| \leq |\!|\!| u-u_0 |\!|\!|^2 - |\!|\!| u-u_1 |\!|\!|^2 + O(\alpha_j)(|\!|\!| u-u_0 |\!|\!|^2 + |\!|\!| u-u_1 |\!|\!|^2)$ by (5.2.9), and choosing j sufficiently large we finish the proof. ∎

In the following lemma it is shown that for sufficiently small μ and $\mathbf{u} \in \mathcal{A}^s$, for a set Λ_1 as in Lemma 5.2.10 that has minimal cardinality, $\#(\Lambda_1 \backslash \Lambda_0)$ can be bounded in terms of $\|\mathbf{f}-\mathbf{L}\mathbf{u}_0\|$ and $|\mathbf{u}|_{\mathcal{A}^s}$ only, cf. Lemma 3.3.1.

Lemma 5.2.11. *For some $s > 0$ let $\mathbf{u} \in \mathcal{A}^s$, and let $\mu \in (0, \kappa(\mathbf{A})^{-\frac{1}{2}})$. Assume that $\mathbf{u}_0 \in \ell_2(\Lambda_0)$ is the solution to $\mathbf{P}_{\Lambda_0}\mathbf{L}\mathbf{I}_{\Lambda_0}\mathbf{u}_0 = \mathbf{P}_{\Lambda_0}\mathbf{f}$ with $\Lambda_0 \supseteq \nabla_j$ for a sufficiently large j. Then, the* smallest *set $\Lambda_1 \supset \Lambda_0$ with*

$$\|\mathbf{P}_{\Lambda_1}(\mathbf{f}-\mathbf{L}\mathbf{u}_0)\| \geq \mu\|\mathbf{f}-\mathbf{L}\mathbf{u}_0\|$$

satisfies

$$\#(\Lambda_1 \backslash \Lambda_0) \lesssim \|\mathbf{f}-\mathbf{L}\mathbf{u}_0\|^{-1/s} |\mathbf{u}|_{\mathcal{A}^s}^{1/s}.$$

Proof. With a constant $\lambda > 0$ to be chosen later, let N be such that a best N-term approximation $\mathbf{u}_N$ for $\mathbf{u}$ satisfies $\|\mathbf{u}-\mathbf{u}_N\| \leq \lambda |\!|\!| \mathbf{u}-\mathbf{u}_0 |\!|\!|$. Since $\mathbf{L}$ is boundedly invertible we have $|\!|\!| \mathbf{u}-\mathbf{u}_0 |\!|\!| \gtrsim \|\mathbf{f}-\mathbf{L}\mathbf{u}_0\|$ and thus, in view of (2.3.3), $N \lesssim \|\mathbf{f}-\mathbf{L}\mathbf{u}_0\|^{-1/s} |\mathbf{u}|_{\mathcal{A}^s}^{1/s}$. Let $\Lambda := \Lambda_0 \cup \operatorname{supp} \mathbf{u}_N \supset \Lambda_0$. We are going to show that for a suitable λ, and j sufficiently large, $\|\mathbf{P}_\Lambda(\mathbf{f}-\mathbf{L}\mathbf{u}_0)\| \geq \mu\|\mathbf{f}-\mathbf{L}\mathbf{u}_0\|$. Then by definition of Λ_1 we may conclude that

$$\#(\Lambda_1 \backslash \Lambda_0) \lesssim \#(\Lambda \backslash \Lambda_0) \leq N \lesssim \|\mathbf{f}-\mathbf{L}\mathbf{u}_0\|^{-1/s} |\mathbf{u}|_{\mathcal{A}^s}^{1/s}.$$

Now we will show that the above claim is valid. The solution to the equation $\mathbf{P}_\Lambda \mathbf{L}\mathbf{I}_\Lambda \mathbf{u}_\Lambda = \mathbf{P}_\Lambda \mathbf{f}$ satisfies

$$\begin{aligned} |\!|\!| \mathbf{u}-\mathbf{u}_\Lambda |\!|\!| &\leq [1+O(\alpha_j)] |\!|\!| \mathbf{u}-\mathbf{u}_N |\!|\!| \leq [1+O(\alpha_j)] \|\mathbf{A}\|^{\frac{1}{2}} \|\mathbf{u}-\mathbf{u}_N\| \\ &\leq \lambda [1+O(\alpha_j)] \|\mathbf{A}\|^{\frac{1}{2}} |\!|\!| \mathbf{u}-\mathbf{u}_0 |\!|\!|, \end{aligned} \quad (5.2.19)$$

where we have used (5.2.5) and the second inequality from (5.2.11). We have

$$\|\mathbf{P}_\Lambda \mathbf{L}(\mathbf{u}_\Lambda - \mathbf{u}_0)\|^2 \geq \|\mathbf{P}_\Lambda \mathbf{A}(\mathbf{u}_\Lambda - \mathbf{u}_0)\|^2 + 2\langle \mathbf{P}_\Lambda \mathbf{A}(\mathbf{u}_\Lambda - \mathbf{u}_0), \mathbf{B}(\mathbf{u}_\Lambda - \mathbf{u}_0)\rangle.$$

The first term in the right hand side can be bounded from below by using the last inequality from (5.2.11). Estimating the second term as

$$\begin{aligned}\langle \mathbf{P}_\Lambda \mathbf{A}(\mathbf{u}_\Lambda - \mathbf{u}_0), \mathbf{B}(\mathbf{u}_\Lambda - \mathbf{u}_0)\rangle &\lesssim |\!|\!|\mathbf{u}_\Lambda - \mathbf{u}_0|\!|\!| \|\mathbf{B}(\mathbf{u}_\Lambda - \mathbf{u}_0)\| \\ &\leq O(\alpha_j)\left(|\!|\!|\mathbf{u} - \mathbf{u}_\Lambda|\!|\!|^2 + |\!|\!|\mathbf{u} - \mathbf{u}_0|\!|\!|^2\right),\end{aligned}$$

we get

$$\|\mathbf{P}_\Lambda \mathbf{L}(\mathbf{u}_\Lambda - \mathbf{u}_0)\|^2 \geq \|\mathbf{A}^{-1}\|^{-1} |\!|\!|\mathbf{u}_\Lambda - \mathbf{u}_0|\!|\!|^2 - O(\alpha_j)\left(|\!|\!|\mathbf{u} - \mathbf{u}_\Lambda|\!|\!|^2 + |\!|\!|\mathbf{u} - \mathbf{u}_0|\!|\!|^2\right).$$

Now by using that $|\!|\!|\mathbf{u}_\Lambda - \mathbf{u}_0|\!|\!| \geq |\!|\!|\mathbf{u} - \mathbf{u}_0|\!|\!|^2 - |\!|\!|\mathbf{u} - \mathbf{u}_\Lambda|\!|\!|^2 - O(\alpha_j)(|\!|\!|\mathbf{u} - \mathbf{u}_0|\!|\!|^2 + |\!|\!|\mathbf{u} - \mathbf{u}_\Lambda|\!|\!|^2)$ by (5.2.9), and applying (5.2.19), we have

$$\begin{aligned}\|\mathbf{P}_\Lambda \mathbf{L}(\mathbf{u}_\Lambda - \mathbf{u}_0)\|^2 &\geq [1 - O(\alpha_j)]\|\mathbf{A}^{-1}\|^{-1} |\!|\!|\mathbf{u} - \mathbf{u}_0|\!|\!|^2 \\ &\quad - [1 + O(\alpha_j)]\|\mathbf{A}^{-1}\|^{-1} |\!|\!|\mathbf{u} - \mathbf{u}_\Lambda|\!|\!|^2 - O(\alpha_j)\left[|\!|\!|\mathbf{u} - \mathbf{u}_\Lambda|\!|\!|^2 + |\!|\!|\mathbf{u} - \mathbf{u}_0|\!|\!|^2\right] \\ &\geq [1 - O(\alpha_j)]\,\|\mathbf{A}^{-1}\|^{-1} |\!|\!|\mathbf{u} - \mathbf{u}_0|\!|\!|^2 - [1 + O(\alpha_j)]\,\|\mathbf{A}^{-1}\|^{-1} |\!|\!|\mathbf{u} - \mathbf{u}_\Lambda|\!|\!|^2 \\ &\geq \left[1 - \lambda^2\|\mathbf{A}\| - O(\alpha_j)\right]\|\mathbf{A}^{-1}\|^{-1} |\!|\!|\mathbf{u} - \mathbf{u}_0|\!|\!|^2.\end{aligned}$$

On the other hand, we have

$$\|\mathbf{L}(\mathbf{u} - \mathbf{u}_0)\|^2 \leq [1 + O(\alpha_j)]\,\|\mathbf{A}\| |\!|\!|\mathbf{u} - \mathbf{u}_0|\!|\!|^2.$$

Combining the last two estimates we infer

$$\|\mathbf{P}_\Lambda(\mathbf{f} - \mathbf{L}\mathbf{u}_0)\|^2 \geq \kappa(\mathbf{A})^{-1}\left[1 - \lambda^2\|\mathbf{A}\| - O(\alpha_j)\right]\|\mathbf{f} - \mathbf{L}\mathbf{u}_0\|^2.$$

Choose a value of the constant $\lambda > 0$ such that $\kappa(\mathbf{A})^{-\frac{1}{2}}(1 - \lambda^2\|\mathbf{A}\|)^{\frac{1}{2}} > \mu$. Then for j sufficiently large, we have $\|\mathbf{P}_\Lambda(\mathbf{f} - \mathbf{L}\mathbf{u}_0)\| \geq \mu\|\mathbf{f} - \mathbf{L}\mathbf{u}_0\|$, thus completing the proof. ∎

5.3 Adaptive algorithm for nonsymmetric and indefinite elliptic problems

In this section, we will formulate an adaptive wavelet algorithm for solving (5.1.1) and analyse its convergence behaviour. To give a rough idea before going through the rigorous treatment, the algorithm starts with an initial index set Λ and computes an approximate residual of the exact Galerkin solution with respect to the

index set Λ. Having computed the approximate residual, we use Lemma 5.2.10 and Lemma 5.2.11 to extend the set Λ such that the error in the new Galerkin solution is a constant factor smaller where the cardinality of the extention is up to a constant factor minimal, and this process is repeated until the computed residual is satisfactorily small.

We need to choose a way to compute the Galerkin solution $\mathbf{u}_\Lambda$ on a given finite set Λ. Computing the Galerkin solution requires inverting the system (5.2.13). In view of obtaining a method of optimal complexity, we will solve the system approximately using an iterative method. Here we formulate a subroutine to solve the Galerkin system (5.2.13) approximately.

Algorithm 5.3.1 Galerkin system solver $\mathbf{GALSOLVE}[\Lambda, \mathbf{v}_\Lambda, \nu, \varepsilon] \to \mathbf{w}_\Lambda$

Parameters: Let $\mathbf{L}$ be s^*-computable and let $\mathbf{f}$ be s^*-admissible for some $s^* > 0$.

With $\mathbf{L}_j$ the compressed matrices from Definition 2.7.8, let J be such that

$$\varrho := \|\mathbf{L} - \mathbf{L}_J\|\|\mathbf{A}^{-1}\| \left[2 + \|\mathbf{B}\mathbf{L}^{-1}\|\right] \leq \tfrac{\varepsilon}{4\varepsilon + 4\nu}.$$

Input: $\Lambda \subset \nabla$, $\#\Lambda < \infty$, $\mathbf{v}_\Lambda \in \ell_2(\Lambda)$, $\varepsilon > 0$, and $\nu \geq \|\mathbf{P}_\Lambda(\mathbf{f} - \mathbf{L}\mathbf{v}_\Lambda)\|$.
Output: $\mathbf{w}_\Lambda \in \ell_2(\Lambda)$ and $\|\mathbf{P}_\Lambda(\mathbf{f} - \mathbf{L}\mathbf{w}_\Lambda)\| \leq \varepsilon$.

1: $\tilde{\mathbf{L}}_\Lambda := \mathbf{P}_\Lambda \mathbf{L}_J \mathbf{I}_\Lambda$;
2: $\tilde{\mathbf{r}}_\Lambda := \mathbf{P}_\Lambda(\mathbf{RHS}[\mathbf{f}, \frac{\varepsilon}{4}] - \mathbf{APPLY}[\mathbf{L}, \mathbf{v}_\Lambda, \frac{\varepsilon}{4}])$;
3: To find an $\tilde{\mathbf{x}}$ with $\|\tilde{\mathbf{r}}_\Lambda - \tilde{\mathbf{L}}_\Lambda \tilde{\mathbf{x}}\| \leq \frac{\varepsilon}{4}$, apply a suitable iterative method for solving $\tilde{\mathbf{L}}_\Lambda \mathbf{x} = \tilde{\mathbf{r}}_\Lambda$, e.g., Conjugate Gradients to the Normal Equations;
4: $\mathbf{w}_\Lambda := \mathbf{v}_\Lambda + \tilde{\mathbf{x}}$.

Proposition 5.3.2. *Let $\mathbf{L}$ be s^*-computable and let $\mathbf{f}$ be s^*-admissible for some $s^* > 0$. Then, if $\Lambda \supseteq \nabla_j$ with j sufficiently large, $\mathbf{w}_\Lambda := \mathbf{GALSOLVE}[\Lambda, \mathbf{v}_\Lambda, \delta, \varepsilon]$ satisfies $\|\mathbf{P}_\Lambda(\mathbf{f} - \mathbf{L}\mathbf{w}_\Lambda)\| \leq \varepsilon$, and for any $s < s^*$, the number of arithmetic operations and storage locations required by the call is bounded by some absolute multiple of $\varepsilon^{-1/s}(|\mathbf{v}_\Lambda|_{\mathcal{A}^s}^{1/s} + |\mathbf{u}|_{\mathcal{A}^s}^{1/s}) + c(\nu/\varepsilon)\#\Lambda$, with $c : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ being some non-decreasing function.*

Proof. In this proof, j is assumed to be sufficiently large whenever needed. With the shorthand notation $\mathbf{L}_\Lambda = \mathbf{P}_\Lambda \mathbf{L} \mathbf{I}_\Lambda$, using Lemma 5.2.9 and estimating $1 + O(\alpha_j) \leq 2$, we have

$$\begin{aligned}\|\mathbf{L}_\Lambda^{-1}(\tilde{\mathbf{L}}_\Lambda - \mathbf{L}_\Lambda)\| &\leq \|\mathbf{L}_\Lambda^{-1}\|\|\mathbf{L}_J - \mathbf{L}\| \\ &\leq \|\mathbf{A}^{-1}\| \left[1 + \|\mathbf{B}\mathbf{L}^{-1}\| + O(\alpha_j)\right] \|\mathbf{L}_J - \mathbf{L}\| \leq \varrho < 1.\end{aligned}$$

This implies that $\mathbf{I} + \mathbf{L}_\Lambda^{-1}(\tilde{\mathbf{L}}_\Lambda - \mathbf{L}_\Lambda)$ is invertible with $\|(\mathbf{I} + \mathbf{L}_\Lambda^{-1}(\tilde{\mathbf{L}}_\Lambda - \mathbf{L}_\Lambda))^{-1}\| \leq \frac{1}{1-\varrho}$. Now writing $\tilde{\mathbf{L}}_\Lambda = \mathbf{L}_\Lambda(\mathbf{I} + \mathbf{L}_\Lambda^{-1}(\tilde{\mathbf{L}}_\Lambda - \mathbf{L}_\Lambda))$ and using Lemma 5.2.9 again, we

conclude that $\tilde{\mathbf{L}}_\Lambda$ is invertible with

$$\|\tilde{\mathbf{L}}_\Lambda^{-1}\| \leq \frac{1}{1-\varrho}\|\mathbf{L}_\Lambda^{-1}\| \leq \frac{1}{1-\varrho}\|\mathbf{A}^{-1}\|\left[2+\|\mathbf{B}\mathbf{L}^{-1}\|\right]. \tag{5.3.1}$$

We have

$$\|\tilde{\mathbf{L}}_\Lambda - \mathbf{L}_\Lambda\|\|\tilde{\mathbf{L}}_\Lambda^{-1}\| \leq \|\mathbf{L}_J - \mathbf{L}\|\frac{1}{1-\varrho}\|\mathbf{A}^{-1}\|\left[2+\|\mathbf{B}\mathbf{L}^{-1}\|\right] \leq \frac{\varrho}{1-\varrho},$$

and $\|\tilde{\mathbf{r}}_\Lambda\| \leq \|\mathbf{P}_\Lambda(\mathbf{f}-\mathbf{L}\mathbf{v}_\Lambda)\| + \|\mathbf{P}_\Lambda(\mathbf{f}-\mathbf{L}\mathbf{v}_\Lambda) - \tilde{\mathbf{r}}_\Lambda\| \leq \nu + \frac{\varepsilon}{2}$. Setting $\mathbf{r}_\Lambda := \mathbf{P}_\Lambda(\mathbf{f}-\mathbf{L}\mathbf{v}_\Lambda)$ and writing

$$\begin{aligned}\mathbf{P}_\Lambda(\mathbf{f}-\mathbf{L}\mathbf{w}_\Lambda) &= \mathbf{r}_\Lambda - \mathbf{P}_\Lambda\mathbf{L}\tilde{\mathbf{x}}\\ &= (\mathbf{r}_\Lambda - \tilde{\mathbf{r}}_\Lambda) + (\tilde{\mathbf{r}}_\Lambda - \tilde{\mathbf{L}}_\Lambda\tilde{\mathbf{x}}) + (\tilde{\mathbf{L}}_\Lambda - \mathbf{P}_\Lambda\mathbf{L})\tilde{\mathbf{L}}_\Lambda^{-1}(\tilde{\mathbf{r}}_\Lambda + \tilde{\mathbf{L}}_\Lambda\tilde{\mathbf{x}} - \tilde{\mathbf{r}}_\Lambda),\end{aligned}$$

we find

$$\|\mathbf{P}_\Lambda(\mathbf{f}-\mathbf{L}\mathbf{w}_\Lambda)\| \leq \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{4} + \tfrac{\varrho}{1-\varrho}(\nu + \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{4}) \leq \varepsilon.$$

The properties of **APPLY** and **RHS** show that the cost of the computation of $\tilde{\mathbf{r}}_\Lambda$ is bounded by some multiple of $\varepsilon^{-1/s}(|\mathbf{v}_\Lambda|_{\mathcal{A}^s}^{1/s} + |\mathbf{u}|_{\mathcal{A}^s}^{1/s})$. We know that $\|\tilde{\mathbf{L}}_\Lambda\| \lesssim 1$ uniformly in ε and ν. So taking into account (5.3.1) we have $\kappa(\tilde{\mathbf{L}}_\Lambda) \lesssim 1$ uniformly in ε and δ. Since $\tilde{\mathbf{L}}_\Lambda$ is sparse and can be constructed in $\mathcal{O}(\#\Lambda)$ operations, where the proportionality coefficient is only dependent on an upper bound for ν/ε, and the required number of iterations of the iterative method is bounded, the proof is completed. ∎

Remark 5.3.3. If the symmetric part of $\mathbf{L}$ is positive definite, then the spectrum of $\tilde{\mathbf{L}}_\Lambda$ lies in the open right half of the complex plane, and so one can use the GMRES method for the solution of the linear system in **GALSOLVE**, cf. [40, 71]. In this case, the proof of the preceding theorem works verbatim.

Next, we combine the above subroutines into an algorithm which approximately computes the residual $\mathbf{f} - \mathbf{L}\mathbf{u}_\Lambda$ for a given set $\Lambda \subset \nabla$. We get an approximate Galerkin solution as a byproduct because we use **GALSOLVE** to approximate the Galerkin solution $\mathbf{u}_\Lambda$.

Algorithm 5.3.4 Galerkin residual $\mathbf{GALRES}[\Lambda, \mathbf{w}_0, \rho_0, \varepsilon] \to [\mathbf{r}_k, \mathbf{w}_k, \rho_k]$

Parameters: Let $\delta, \gamma \in (0,1)$ and $\theta > 0$ be constants.
Input: $\rho_0 \geq \|\mathbf{f} - \mathbf{L}\mathbf{w}_0\|$.
Output: $\|\mathbf{f} - \mathbf{L}\mathbf{w}_k\| \leq \rho$, and either $\rho \leq \varepsilon$ or $\|\mathbf{f} - \mathbf{L}\mathbf{u}_\Lambda - \mathbf{r}_k\| \leq \delta\|\mathbf{r}_k\|$.

1: $k := 0$, $\zeta_0 := \theta\rho_0$, $\nu_0 := \rho_0$;
2: **repeat**
3: $k := k+1$, $\zeta_k := \zeta_{k-1}/2$;
4: $\nu_k := \gamma\zeta_k \left(\|\mathbf{L}\|\|\mathbf{A}^{-1}\|\left[2 + \|\mathbf{B}\mathbf{L}^{-1}\|\right]\right)^{-1}$;
5: $\mathbf{w}_k := \mathbf{GALSOLVE}[\Lambda, \mathbf{w}_{k-1}, \nu_{k-1}, \nu_k]$;
6: $\mathbf{r}_k := \mathbf{RHS}[(1-\gamma)\zeta_k/2] - \mathbf{APPLY}[\mathbf{w}_k, (1-\gamma)\zeta_k/2]$;
7: $\nu_k := \min\{\nu_{k-1}, \nu_k\}$;
8: **until** $\rho_k := \|\mathbf{r}_k\| + (1-\gamma)\zeta_k \leq \varepsilon$ or $\zeta_k \leq \delta\|\mathbf{r}_k\|$.

Remark 5.3.5. In the above algorithm, as opposed to Algorithm 3.2.5, we are forced to place the Galerkin solver inside the loop that computes the current residual with a sufficient accuracy. The reason is that in Lemma 5.2.10 and Lemma 5.2.11 the vector $\mathbf{u}_0$ must be the Galerkin solution on its support, whereas in the corresponding Lemma 3.2.1 and Lemma 3.3.1 this vector could be arbitrary.

Remark 5.3.6. In view of Remark 2.8.3, taking into account that ρ_0 is an upper bound on the residual of $\mathbf{w}_0$, a reasonable choice for the value of θ is $\theta \approx \frac{2\omega}{(1+\omega)(1-\gamma)}$.

Proposition 5.3.7. *Let* $\mathbf{L}$ *be* s^**-computable and let* $\mathbf{f}$ *be* s^**-admissible for some* $s^* > 0$. *Then, if* $\Lambda \supseteq \nabla_j$ *for some sufficiently large* j, *then the outputs of* $[\mathbf{r}, \mathbf{w}, \rho] := \mathbf{GALRES}[\Lambda, \mathbf{w}_0, \rho_0, \varepsilon]$ *satisfy* $\|\mathbf{f} - \mathbf{L}\mathbf{w}\| \leq \rho$, *and either* $\rho \leq \varepsilon$ *or* $\|\mathbf{f} - \mathbf{L}\mathbf{u}_\Lambda - \mathbf{r}\| \leq \delta\|\mathbf{r}\|$. *Furthermore, under the same condition we have* $\rho \gtrsim \min\{\rho_0, \varepsilon\}$. *In addition, if for some* $s < s^*$, $\mathbf{u} \in \mathcal{A}^s$, *then* $\#\operatorname{supp}\mathbf{r} \lesssim \rho^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + (\rho_0/\rho)^{1/s}\#\Lambda$ *and the number of arithmetic operations and storage locations required by the call is bounded by some absolute multiple of* $\rho^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + (\rho_0/\rho)^{1/s}(\#\Lambda + 1)$.

Proof. If at evaluation of the until-clause for the k-th iteration, $\zeta_k > \delta\|\mathbf{r}_k\|$, then $\rho_k = \|\mathbf{r}_k\| + (1-\gamma)\zeta_k < (\delta^{-1} + 1 - \gamma)\zeta_k$. Since ζ_k is halved in each iteration, we infer that, if not by $\zeta_k \leq \delta\|\mathbf{r}_k\|$, the inner loop will terminate by $\rho_k \leq \varepsilon$.

Let K be the value of k at the termination of the loop. First we will show $\rho \gtrsim \min\{\rho_0, \varepsilon\}$. When the loop terminates in the first iteration, i.e., when $K = 1$, we have $\rho_1 = \|\mathbf{r}_1\| + (1-\gamma)\zeta_1 \gtrsim \rho_0$. In the case the loop terminates with $\rho_K \leq \varepsilon$ we have $\|\mathbf{r}_{K-1}\| + 2(1-\gamma)\zeta_K > \varepsilon$ and $2\zeta_K > \delta\|\mathbf{r}_{K-1}\|$, so we conclude

$$\rho_K \geq (1-\gamma)\zeta_K > \frac{(1-\gamma)\delta(\|\mathbf{r}_{K-1}\| + 2(1-\gamma)\zeta_K)}{2 + 2\delta(1-\gamma)} > \frac{(1-\gamma)\delta\varepsilon}{2 + 2\delta(1-\gamma)}.$$

Since after any evaluation of $\mathbf{r}_k$ inside the algorithm, $\|\mathbf{r}_k - (\mathbf{f} - \mathbf{L}\mathbf{w}_k)\| \leq (1-\gamma)\zeta_k$, for any $1 \leq k \leq K$, ρ_k is an upper bound on $\|\mathbf{f} - \mathbf{L}\mathbf{w}_k\|$. Together with the condition on ρ_0 this guarantees that the subroutine **GALSOLVE** is called with a valid parameter ν_{k-1}. By applying Lemma 5.2.9 for sufficiently large j, we have

$$\begin{aligned}\|\mathbf{r}_k - (\mathbf{f} - \mathbf{L}\mathbf{u}_\Lambda)\| &\leq \|\mathbf{r}_k - (\mathbf{f} - \mathbf{L}\mathbf{w}_k)\| + \|\mathbf{L}(\mathbf{u}_\Lambda - \mathbf{w}_k)\| \\ &\leq (1-\gamma)\zeta_k + \|\mathbf{L}\|\|(\mathbf{P}_\Lambda \mathbf{L}\mathbf{I}_\Lambda)^{-1}\|\|\mathbf{P}_\Lambda(\mathbf{f} - \mathbf{L}\mathbf{w}_k)\| \\ &\leq (1-\gamma)\zeta_k + \|\mathbf{L}\|\|\mathbf{A}^{-1}\|[1 + \|\mathbf{B}\mathbf{L}^{-1}\| + O(\alpha_j)]\nu_k \leq \zeta_k,\end{aligned}$$

and therefore the condition $\zeta_k \leq \delta\|\mathbf{r}_k\|$ implies $\|\mathbf{r}_k - (\mathbf{f} - \mathbf{L}\mathbf{u}_\Lambda)\| \leq \delta\|\mathbf{r}_k\|$. This proves the first part of the theorem.

The properties of **RHS**, **APPLY** and **GALSOLVE** imply that the cost of k-th iteration can be bounded by some multiple of $\zeta_k^{-1/s}(|\mathbf{w}_{k-1}|_{\mathcal{A}^s}^{1/s} + |\mathbf{u}|_{\mathcal{A}^s}^{1/s} + |\mathbf{w}_k|_{\mathcal{A}^s}^{1/s}) + c(\frac{\nu_{k-1}}{\nu_k})\#\Lambda + \#\Lambda + 1$, where $c(\cdot)$ is the non-decreasing function from Proposition 5.3.2. Since any vector $\mathbf{w}_k$ determined inside the algorithm satisfies $\|\mathbf{u} - \mathbf{w}_k\| \lesssim \rho_0$, from $|\mathbf{w}_k|_{\mathcal{A}^s} \lesssim |\mathbf{u}|_{\mathcal{A}^s} + (\#\operatorname{supp}\mathbf{w}_k)^s\|\mathbf{w}_k - \mathbf{u}\|$ (Proposition 2.3.6), we infer that $|\mathbf{w}_k|_{\mathcal{A}^s} \lesssim |\mathbf{u}|_{\mathcal{A}^s} + (\#\Lambda)^s\rho_0$. At any iteration the ratio $\frac{\nu_{k-1}}{\nu_k}$ can be bounded by a multiple of $\max\{\frac{\nu_0}{\nu_1}, 2\} \lesssim \frac{\rho_0}{\zeta_1} + 1 \lesssim 1$. By the geometric decrease of ζ_k inside the loop, the above considerations imply that the total cost of the algorithm can be bounded by some multiple of $\zeta_K^{-1/s}(|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + \rho_0^{1/s}\#\Lambda) + K(\#\Lambda + 1)$. Taking into account the value of ζ_0, and the geometric decrease of ζ_i inside the loop, we have $K(\#\Lambda + 1) = K\rho_0^{-1/s}\rho_0^{1/s}(\#\Lambda + 1) \lesssim \zeta_K^{-1/s}\rho_0^{1/s}(\#\Lambda + 1)$. The number of nonzero coefficients in $\mathbf{r}_K$ is bounded by an absolute multiple of $\zeta_K^{-1/s}(|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + \rho_0^{1/s}\#\Lambda)$ so the theorem is proven upon showing that $\zeta_K \gtrsim \rho_K$. When the loop terminates in the first iteration, i.e., when $K = 1$, we have $\rho_1 = \|\mathbf{r}_1\| + (1-\gamma)\zeta_1 \leq \|\mathbf{f} - \mathbf{L}\mathbf{w}_0\| + 2(1-\gamma)\zeta_1 \lesssim \rho_0 + \zeta_1 \lesssim \zeta_1$, and when the loop terminates with $\zeta_K \geq \delta\|\mathbf{r}_K\|$, we have $\rho_K = \|\mathbf{r}_K\| + (1-\gamma)\zeta_K \leq (\frac{1}{\delta} + 1 - \gamma)\zeta_K$. In the other case, we have $\delta\|\mathbf{r}_{K-1}\| \leq 2\zeta_K$, and so from $\|\mathbf{r}_K - \mathbf{r}_{K-1}\| \leq \zeta_K + 2\zeta_K$, we infer $\|\mathbf{r}_K\| \leq \|\mathbf{r}_{K-1}\| + 3\zeta_K \leq (\frac{2}{\delta} + 3)\zeta_K$, so that $\rho_K \leq (\frac{2}{\delta} + 4 - \gamma)\zeta_K$. ∎

We now define our adaptive wavelet solver.

Algorithm 5.3.8 Adaptive Galerkin method **SOLVE**$[\varepsilon] \to \mathbf{w}_k$

Parameters: Let j be a sufficiently large fixed integer, let $\rho_0 \geq \|\mathbf{f}\|$, and $\alpha \in (0,1)$ be constants.
Input: $\varepsilon > 0$.
Output: $\mathbf{w}_k \in P$ and $\|\mathbf{f} - \mathbf{L}\mathbf{w}_k\| \leq \varepsilon$.

1: $k := 0$, $\mathbf{w}_0 := 0$, $\Lambda_1 := \nabla_j$;
2: **loop**
3: $k := k + 1$;
4: $[\mathbf{r}_k, \mathbf{w}_k, \rho_k] :=$**GALRES**$[\Lambda_k, \mathbf{w}_{k-1}, \rho_{k-1}, \varepsilon]$;
5: **if** $\rho_k \leq \varepsilon$ **then**
6: Terminate the subroutine.
7: **end if**
8: $\Lambda_{k+1} :=$ **RESTRICT**$[\Lambda_k, \mathbf{r}_k, \alpha]$;
9: **end loop**

Theorem 5.3.9. *Let* $\mathbf{L}$ *be* s^**-computable and let* $\mathbf{f}$ *be* s^**-admissible with some* $s^* > 0$. *Then* $\mathbf{w} :=$ **SOLVE**$[\varepsilon]$ *terminates with* $\|\mathbf{f} - \mathbf{L}\mathbf{w}\| \leq \varepsilon$. *In addition, let the parameters* α *and* ρ_0 *in* **SOLVE**, *and* δ *in* **GALRES**, *be selected such that* $\frac{\alpha+\delta}{1-\delta} < \kappa(\mathbf{A})^{-\frac{1}{2}}$, $\alpha < \delta$, *and* $\rho_0 \lesssim \|\mathbf{f}\|$, *and let* $\varepsilon \lesssim \|\mathbf{f}\|$. *Then, if for some* $s < s^*$, $\mathbf{u} \in \mathcal{A}^s$, *we have* $\#\operatorname{supp}\mathbf{w} \lesssim \varepsilon^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}$ *and the number of arithmetic operations and storage locations required by the call is bounded by some absolute multiple of the same expression.*

Proof. Before we come to the actual proof, first we indicate the need for the conditions involving ρ_0, $\|\mathbf{f}\|$ and ε. If $\rho_0 \not\lesssim \|\mathbf{f}\|$ we cannot bound the cost of the first call of **GALRES**. If $\varepsilon \not\lesssim \|\mathbf{f}\|$, then $\varepsilon^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}$ might be arbitrarily small, whereas **SOLVE** takes in any case some arithmetic operations.

Abbreviating $\mathbf{P}_{\Lambda_k}$ as $\mathbf{P}_k$, let $\mathbf{u}_k \in \ell_2(\Lambda_k)$ be the solution of the Galerkin system $\mathbf{P}_k\mathbf{L}\mathbf{u}_k = \mathbf{P}_k\mathbf{f}$. Assume that the k-th call of **GALRES** terminates with $\rho_k > \varepsilon$ and thus with $\|\mathbf{f} - \mathbf{L}\mathbf{u}_k - \mathbf{r}_k\| \leq \delta\|\mathbf{r}_k\|$. Then we have

$$\begin{aligned}\alpha\|\mathbf{r}_k\| \leq \|\mathbf{P}_{k+1}\mathbf{r}_k\| &= \|\mathbf{P}_{k+1}[\mathbf{r}_k - (\mathbf{f} - \mathbf{L}\mathbf{u}_k) + (\mathbf{f} - \mathbf{L}\mathbf{u}_k)]\| \\ &\leq \delta\|\mathbf{r}_k\| + \|\mathbf{P}_{k+1}(\mathbf{f} - \mathbf{L}\mathbf{u}_k)\|,\end{aligned}$$

giving $\|\mathbf{P}_{k+1}(\mathbf{f} - \mathbf{L}\mathbf{u}_k)\| \geq (\alpha - \delta)\|\mathbf{r}_k\|$. Defining $\nu_k := \|\mathbf{r}_k\| + \|\mathbf{f} - \mathbf{L}\mathbf{u}_k - \mathbf{r}_k\|$ we have $\|\mathbf{f} - \mathbf{L}\mathbf{u}_k\| \leq \nu_k \leq (1+\delta)\|\mathbf{r}_k\|$, and using this we obtain

$$\|\mathbf{P}_{k+1}(\mathbf{f} - \mathbf{L}\mathbf{u}_k)\| \geq \tfrac{\alpha-\delta}{1+\delta}\nu_k \geq \tfrac{\alpha-\delta}{1+\delta}\|\mathbf{f} - \mathbf{L}\mathbf{u}_k\|,$$

so that Lemma 5.2.10 shows that

$$|||\mathbf{u}-\mathbf{u}_{k+1}||| \leq [1-\kappa(\mathbf{A})^{-1}(\tfrac{\alpha-\delta}{1+\delta})^2+O(\alpha_j)]^{\frac{1}{2}}|||\mathbf{u}-\mathbf{u}_k|||.$$

Taking into account that $\nu_k \leq (1+\delta)\|\mathbf{r}_k\| < (1+\delta)\rho_k$ and that $\|\mathbf{f}-\mathbf{L}\mathbf{u}_k\| \geq \|\mathbf{P}_{k+1}(\mathbf{f}-\mathbf{L}\mathbf{u}_k)\| \gtrsim \nu_k$, we have $\rho_k \eqsim \nu_k \eqsim \|\mathbf{f}-\mathbf{L}\mathbf{u}_k\| \eqsim |||\mathbf{u}-\mathbf{u}_k|||$ as long as $\rho_k > \varepsilon$. By the conditions that $\alpha > \delta$ and that j is sufficiently large, it holds that $\rho_k \lesssim \xi^{k-1}\rho_1$ for certain $\xi < 1$, so that **SOLVE** terminates, say directly after the K-th iteration. This proves the first part of the theorem.

With $\mu = \frac{\alpha+\delta}{1-\delta}$, for $1 \leq k < K$ let $\nabla \supset \Lambda \supset \Lambda_k$ be the *smallest* set with

$$\|\mathbf{P}_\Lambda(\mathbf{f}-\mathbf{L}\mathbf{u}_k)\| \geq \mu\|\mathbf{f}-\mathbf{L}\mathbf{u}_k\|.$$

Since $\mu < \kappa(\mathbf{A})^{\frac{1}{2}}$ by the condition on δ and α, and $\|\mathbf{f}-\mathbf{L}\mathbf{u}_k\| \leq \nu_k$, an application of Lemma 5.2.11 shows that $\#(\Lambda\backslash\Lambda_k) \lesssim \nu_k^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}$. On the other hand, using Proposition 5.3.7 twice we have $\mu\|\mathbf{r}_k\| \leq \mu\|\mathbf{f}-\mathbf{L}\mathbf{u}_k\|+\mu\delta\|\mathbf{r}_k\| \leq \|\mathbf{P}_\Lambda(\mathbf{f}-\mathbf{L}\mathbf{u}_k)\|+\mu\delta\|\mathbf{r}_k\| \leq \|\mathbf{P}_\Lambda\mathbf{r}_k\|+(1+\mu)\delta\|\mathbf{r}_k\|$ or $\|\mathbf{P}_\Lambda\mathbf{r}_k\| \geq \alpha\|\mathbf{r}_k\|$. Thus by construction of Λ_{k+1} we conclude that

$$\#(\Lambda_{k+1}\backslash\Lambda_k) \lesssim \#(\Lambda\backslash\Lambda_k) \lesssim \nu_k^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} \lesssim \rho_k^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} \quad \text{for } 1 \leq k < K.$$

Since $\Lambda_1 \lesssim 1 \lesssim \rho_0^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}$ by $\rho_0 \lesssim \|\mathbf{f}\| \lesssim |\mathbf{u}|_{\mathcal{A}^s}$, with $\Lambda_0 := \varnothing$ we have for $1 \leq k \leq K$,

$$\#\Lambda_k = \sum_{i=0}^{k-1}\#(\Lambda_{i+1}\backslash\Lambda_i) \lesssim (\sum_{i=0}^{k-1}\rho_i^{-1/s})|\mathbf{u}|_{\mathcal{A}^s}^{1/s} \lesssim \rho_{k-1}^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}. \tag{5.3.2}$$

In view of Lemma 3.3.3, we infer that the cost of determining the set Λ_{k+1} is of order $\#\Lambda_k + \#\operatorname{supp}\mathbf{r}_k + 1$. From Proposition 5.3.7, we have $\#\operatorname{supp}\mathbf{r}_k \lesssim \rho_k^{-1/s}|\mathbf{u}|_{\ell_\tau^w}^{1/s} + (\rho_{k-1}/\rho_k)^{1/s}\#\Lambda_k$ and that the cost of the k-th call of **GALRES** is of order $\rho_k^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + (\rho_{k-1}/\rho_k)^{1/s}(\#\Lambda_k+1)$, implying that the cost of the k-th iteration of **SOLVE** can be bounded by an absolute multiple of $\rho_k^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + (\rho_{k-1}/\rho_k)^{1/s}(\#\Lambda_k+1)+\#\Lambda_k+1$. Now by using (5.3.2) and $1 \lesssim \rho_0^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}$, and taking into account the geometric decrease of ρ_k we conclude that the total cost of the algorithm can be bounded by an absolute multiple of $\rho_K^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}$. From Proposition 5.3.7 we have $\rho_K \gtrsim \min\{\rho_{K-1},\varepsilon\} \gtrsim \varepsilon$, where the second inequality follows from $\rho_{K-1} > \varepsilon$ when $K > 1$ and by assumption when $K = 1$. This completes the proof. ∎

Chapter 6

Adaptive algorithm with truncated residuals

6.1 Introduction

In this chapter, we return to the equation (2.4.3) on page 21, which is recalled here for convenience:

$$\mathbf{A}\mathbf{u} = \mathbf{f}, \tag{6.1.1}$$

where $\mathbf{A} : \ell_2 \to \ell_2$ is an SPD matrix, and $\mathbf{f} \in \ell_2$.

In Chapter 3, we presented Algorithm 3.3.4 on page 54 for solving (6.1.1). The algorithm consists of a loop over the following steps: For a given iterand $\mathbf{v} \in P$, compute the residual $\mathbf{r} := \mathbf{f} - \mathbf{A}\mathbf{v}$ approximately, and then with a constant α from a suitable range, choose an index set $\Lambda \supset \operatorname{supp} \mathbf{v}$ with (nearly) minimal cardinality such that $\|\mathbf{P}_\Lambda \mathbf{r}\| \geq \alpha \|\mathbf{r}\|$ with $\mathbf{r}$ replaced by the approximately computed residual. The next iterand of the iteration is determined by an inexact solution of the Galerkin system on $\ell_2(\Lambda)$. Optimality of the adaptive algorithm was proven in Theorem 3.3.5 on page 54.

In the approximate computation of the residual $\mathbf{r} := \mathbf{f} - \mathbf{A}\mathbf{v}$, among other things, one uses the subroutine **APPLY** as in Algorithm 2.7.9 on page 33. The subroutine **APPLY** employs columns of a compressed matrix $\mathbf{A}_j$ with increasing accuracy (thus with increasing j) as the corresponding entry of $\mathbf{v}$ gets large in absolute value. As indicated in Remark 2.7.13 on page 34 (see also Theorems 7.3.3 and 8.2.4), when j increases, the compressed matrix $\mathbf{A}_j$ involves blocks of $\mathbf{A}$ corresponding to the interactions between wavelets with level difference proportional to j. As a result, it becomes possible that the difference between the highest levels of wavelets that are used in the approximate residual and that are used in the iterand (i.e. $\mathbf{v}$) grows when the iteration proceeds. This makes

very deep refinements feasible but also leads to serious obstacles in practical implementations of the algorithm. Moreover, numerical experiments show that in terms of cardinality, only a tiny part of the support of the approximately computed residual constitutes the index set Λ for the next iterand.

An alternative approach would be to simply compute the "truncated" residual $\mathbf{r}^\star := \mathbf{P}_{\Lambda^\star}\mathbf{r}$ for some index set $\Lambda^\star$, and choose an index set $\Lambda \supset \operatorname{supp}\mathbf{v}$ with (nearly) minimal cardinality such that $\|\mathbf{P}_\Lambda \mathbf{r}^\star\| \geq \alpha\|\mathbf{r}^\star\|$. Of course, the point is that one has to choose the "activable" set $\Lambda^\star$ appropriately. To our knowledge, this approach was first suggested in [54] in the context of adaptive wavelet Galerkin BEM. In [5], the same idea was applied for designing adaptive wavelet algorithms for solving elliptic boundary value problems. Numerical experiments in both papers show relatively good performances. In this chapter, we analyze adaptive wavelet algorithms of the above discussed type, i.e., adaptive algorithms with truncated residuals.

Throughout foregoing chapters, we have been considering approximations for $\mathbf{u}$ from $\ell_2(\Lambda)$, where Λ can be *any* finite subset of ∇. In this chapter, a slightly restricted type of wavelet approximation is employed, in the sense that only sets Λ are considered that are *trees*, roughly meaning that if $\lambda \in \Lambda$, then for any $\lambda' \in \nabla$ with $\operatorname{supp}\psi_\lambda \subset \operatorname{supp}\psi_{\lambda'}$, also $\lambda' \in \Lambda$. From the purpose of constructing the activable sets efficiently, tree approximation arises almost naturally. In the context of adaptive wavelet algorithms, tree approximation is often used, cf. [19, 20, 30]. It is claimed that working with trees has advantages in view of obtaining an efficient implementation, whereas, on the other hand, best tree N-term approximations converge towards $\mathbf{u}$ with a rate N^{-s} under regularity conditions that are only slightly stronger than that for unrestricted best N-term approximations.

This chapter is organized as follows. In the next section, we recall some relevant facts on best N-term approximations with tree constraints. An adaptive algorithm with truncated residuals is proposed and proven to be optimal under some assumptions in Section 6.3. Then Section 6.4 provides a way to verify these assumptions for second order elliptic boundary value problems. In the last section we extend a certain result on completion of trees to graded trees, which is often used in Section 6.4. Since this result concerns general trees and it can be used not only in a wavelet context, we presented it such that it stands on its own independently of other sections in this chapter.

6.2 Tree approximations

We assume that a parent-child relation is defined on the index set ∇. We assume that every element $\lambda \in \nabla$ has a uniformly bounded number of children, and has at most one parent. We say that $\lambda \in \nabla$ is a *descendant* of $\mu \in \nabla$ and write $\lambda \succ \mu$

if λ is a child of a descendant of μ *or* is a child of μ. The relations $\prec$ (ascendant of), $\succeq$ (descendant of or equal to), and $\preceq$ (ascendant of or equal to) are defined accordingly. The *level* or *generation* of an element $\lambda \in \nabla$, denoted by $|\lambda| \in \mathbb{N}_0$, is the number of its ascendants. Obviously, $\lambda \succ \mu$ implies $|\lambda| > |\mu|$. We call the set $\nabla_0 := \{\lambda \in \nabla : |\lambda| = 0\}$ the set of *root*, and assume that $\#\nabla_0 < \infty$.

A subset $\Lambda \subseteq \nabla$ is said to be a *tree* if with every member $\lambda \in \Lambda$ all its ascendants are included in Λ. For a tree Λ, those $\lambda \in \Lambda$ whose children are not contained in Λ are called *leaves* of Λ, and the set of all leaves of Λ is denoted by $\partial\Lambda$. Similarly, those $\lambda \notin \Lambda$ whose parent belongs to Λ are called *outer leaves* of Λ and the set of all outer leaves of Λ is denoted by $\mathcal{L}(\Lambda)$.

For $N \in \mathbb{N}_0$, we collect all trees with at most N elements in the set

$$\mathcal{T}_N := \{\Lambda \subset \nabla : \#\Lambda \leq N, \, \Lambda \text{ is a tree}\},$$

and collect all the elements of ℓ_2 whose support is a tree with cardinality N in

$$X_N := \{\mathbf{v} \in \ell_2 : \mathbf{v} \in \ell_2(\Lambda) \text{ for some } \Lambda \in \mathcal{T}_N\}. \tag{6.2.1}$$

Obviously, we have $\mathcal{T}_0 = \varnothing$, $\mathcal{T}_N \subset \mathcal{T}_{N+1}$ and $X_N \subset X_{N+1}$. The set of all *finite* trees is denoted by $\mathcal{T} := \cup_{N\in\mathbb{N}_0}\mathcal{T}_N$. We will consider here approximations of elements of ℓ_2 from the subsets X_N. The subset X_N is not a linear space, meaning that we deal with a nonlinear approximation. For $\mathbf{v} \in \ell_2$ and $N \in \mathbb{N}_0$, we define the best approximation error when approximating $\mathbf{v}$ from X_N by

$$E_N(\mathbf{v}) := \operatorname{dist}(\mathbf{v}, X_N) = \inf_{\mathbf{v}_N \in X_N} \|\mathbf{v} - \mathbf{v}_N\|. \tag{6.2.2}$$

Any element $\mathbf{v}_N \in X_N$ that achieves this error is called a *best tree N-term approximation* of $\mathbf{v}$. For any $N \in \mathbb{N}_0$ a best tree N-term approximation exists since X_N is a finite union of linear spaces. In particular, with $\mathbf{P}_\Lambda : \ell_2 \to \ell_2(\Lambda)$ being the ℓ_2-orthogonal projector onto $\ell_2(\Lambda)$, a best tree N-term approximation of $\mathbf{v} \in \ell_2$ is equal to $\mathbf{P}_\Lambda\mathbf{v}$ for some $\Lambda \in \mathcal{T}_N$.

The following functional can be shown to be a quasi-norm for $s \in \mathbb{R}$

$$|\mathbf{v}|_{\mathcal{A}^s} := \|\mathbf{v}\| + \sup_{N\in\mathbb{N}} N^s E_N(\mathbf{v}), \tag{6.2.3}$$

where "quasi-" refers to the fact that it only satisfies a generalized triangle inequality, cf. Lemma 2.3.1 on page 14. For $s > 0$, we define the approximation space $\mathcal{A}^s \subset \ell_2$ by collecting all the vectors for which the above quasi-norm is finite. Clearly, it is precisely the set of elements whose best tree N-term approximation error decays like N^{-s}. The space $\mathcal{A}^s$ can be shown to be a quasi-Banach space with the quasi-norm (6.2.3).

Remark 6.2.1. Let Ψ be a suitable wavelet basis with the approximation order d for the Sobolev space H^t defined on a domain $\Omega \subseteq \mathbb{R}^n$, possibly incorporating essential boundary conditions. Then, if $0 < s < \frac{d-t}{n}$ and $v \in B_p^{t+ns}(L_p)$ for $p > (\frac{1}{2} + s)^{-1}$, the vector of expansion coefficients $\mathbf{v}$ of v in the basis Ψ satisfies $\mathbf{v} \in \mathcal{A}^s$, cf. [20].

Apart from tree approximations, in the following we will also consider a seemingly general class of approximations. For $N \geq N_0$ with a constant $N_0 \in \mathbb{N}$, let $\tilde{\mathcal{T}}_N$ be a set of subsets of the index set ∇ satisfying

$$\tilde{\mathcal{T}}_N \subseteq \mathcal{T}_N \subset \tilde{\mathcal{T}}_{cN}, \tag{6.2.4}$$

where $c \in \mathbb{N}$ is a constant. We call an index set $\Lambda \subset \nabla$ a *graded tree* if $\Lambda \in \tilde{\mathcal{T}}_N$ for some N. Using this terminology, the condition (6.2.4) can be read as follows: Any graded tree is a tree, and any tree with cardinality N can be extended to a graded tree with cardinality at most cN. The set of all graded trees is denoted by $\tilde{\mathcal{T}} := \cup_{N \geq N_0} \tilde{\mathcal{T}}_N$.

Analogously to the above lines, by using $\tilde{\mathcal{T}}_N$ we define the spaces $\tilde{X}_N$, the best approximation error $\tilde{E}_N(\cdot)$ from $\tilde{X}_N$, and the approximation spaces $\tilde{\mathcal{A}}^s$. Now we call a best approximation from $\tilde{X}_N$ a *best graded tree N-term approximation*. The condition (6.2.4) implies $\tilde{X}_N \subset X_N \subset \tilde{X}_{cN}$, and from this we have $\tilde{E}_N(\mathbf{v}) \geq E_N(\mathbf{v}) \geq \tilde{E}_{cN}(\mathbf{v})$ for $N \geq N_0$. Finally, since $N^s E_N(\mathbf{v}) \lesssim \|\mathbf{v}\|$ for $N < N_0$, we conclude that $\tilde{\mathcal{A}}^s = \mathcal{A}^s$ with $|\cdot|_{\tilde{\mathcal{A}}^s} \eqsim |\cdot|_{\mathcal{A}^s}$.

The following result is a trivial adaptation of Proposition 2.3.6 to tree approximations, which will be often used in the sequel.

Remark 6.2.2. Let $s > 0$. Then for any $\mathbf{v} \in \mathcal{A}^s$ and $\mathbf{z} \in \ell_2(\Lambda)$ with Λ being a finite tree, we have $|\mathbf{z}|_{\mathcal{A}^s} \lesssim |\mathbf{v}|_{\mathcal{A}^s} + (\#\Lambda)^s \|\mathbf{v} - \mathbf{z}\|$.

6.3 Adaptive algorithm with truncated residuals

6.3.1 The basic scheme

For a given index set $\Lambda \subseteq \nabla$, the Galerkin approximation $\mathbf{u}_\Lambda$ from $\ell_2(\Lambda)$ to the solution of (6.1.1) is the solution of the Galerkin system

$$\mathbf{A}_\Lambda \mathbf{u}_\Lambda = \mathbf{f}_\Lambda, \tag{6.3.1}$$

where, recalling that $\mathbf{P}_\Lambda : \ell_2 \to \ell_2(\Lambda)$ is the ℓ_2-orthogonal projector onto $\ell_2(\Lambda)$, $\mathbf{f}_\Lambda := \mathbf{P}_\Lambda \mathbf{f}$, and $\mathbf{A}_\Lambda := \mathbf{P}_\Lambda \mathbf{A} \mathbf{I}_\Lambda : \ell_2(\Lambda) \to \ell_2(\Lambda)$ with $\mathbf{I}_\Lambda := \mathbf{P}_\Lambda^* : \ell_2(\Lambda) \to \ell_2$ being the trivial inclusion. The Galerkin approximation $\mathbf{u}_\Lambda$ is the best approximation from $\ell_2(\Lambda)$ to $\mathbf{u}$ in the energy norm $|\!|\!| \cdot |\!|\!| := \langle \mathbf{A} \cdot, \cdot \rangle^{\frac{1}{2}}$. Here and in the following,

for any $\Sigma_1 \subset \Sigma_2 \subseteq \nabla$, we consider $\ell_2(\Sigma_1)$ as a subspace of $\ell_2(\Sigma_2)$, implicitly identifying $\mathbf{v} \in \ell_2(\Sigma_1)$ with $\mathbf{P}_{\Sigma_2}\mathbf{I}_{\Sigma_1}\mathbf{v}$.

Let $\mathbf{v} \in \ell_2(\Lambda)$ be some specified approximation (possibly $\mathbf{v} = \mathbf{u}_\Lambda$) to $\mathbf{u}$, and let $\breve{\Lambda} \supset \Lambda$. Then Lemma 3.2.1 on page 45 provides a way to guarantee that $\mathbf{u}_{\breve{\Lambda}}$ has an error that is a constant factor smaller than the error in $\mathbf{v}$. We recall this lemma in the following, adjusting to the case when the index sets are trees.

Lemma 6.3.1. *Let $\alpha \in (0, 1]$, $\mathbf{v} \in \ell_2(\Lambda)$ and $\breve{\Lambda} \supset \Lambda$, where Λ and $\breve{\Lambda}$ are trees, such that*

$$\|\mathbf{P}_{\breve{\Lambda}}(\mathbf{f} - \mathbf{A}\mathbf{v})\| \geq \alpha\|\mathbf{f} - \mathbf{A}\mathbf{v}\|.$$

Then, for $\mathbf{u}_{\breve{\Lambda}} \in \ell_2(\breve{\Lambda})$ being the Galerkin approximation to $\mathbf{u}$ from $\ell_2(\breve{\Lambda})$, and with $\kappa(\mathbf{A}) := \|\mathbf{A}\|\|\mathbf{A}^{-1}\|$, we have

$$|||\mathbf{u} - \mathbf{u}_{\breve{\Lambda}}||| \leq [1 - \kappa(\mathbf{A})^{-1}\alpha^2]^{\frac{1}{2}}|||\mathbf{u} - \mathbf{v}|||.$$

In Chapter 3, the above result was used to construct a convergent algorithm consisting of a loop over the following two steps: Compute the residual $\mathbf{r} := \mathbf{f} - \mathbf{A}\mathbf{v}$ approximately, and then choose $\breve{\Lambda}$ such that $\|\mathbf{P}_{\breve{\Lambda}}\mathbf{r}\| \geq \alpha\|\mathbf{r}\|$ with $\mathbf{r}$ replaced by the approximately computed residual. For the sake of efficiency one evidently has to choose the set $\breve{\Lambda}$ with minimal or nearly minimal cardinality. An optimal convergence rate was proven in Theorem 3.2.7 on page 49 when a coarsening step is applied after each fixed number of iterations, which removes small entries from the current approximation. Later in Theorem 3.3.5 on page 54, by using Lemma 3.3.1 on page 52, it was shown that this coarsening step is unnecessary to get an optimal convergence rate.

Although the latter algorithm was proven to have an optimal convergence rate, there are reasons to expect the algorithm can be quantitatively improved. As we discussed in the introduction, at least with the current approaches of approximating the residual, it is possible that the difference between the highest levels of wavelets that are used in the approximate residual and that are used in the iterand (i.e. $\mathbf{v}$) grows when the iteration proceeds. This leads to serious obstacles in practical implementations of the algorithm. Moreover, the above lemma requires the parameter α to be small, meaning that a small fraction of the residual is actually captured by the set $\breve{\Lambda}$. Numerical experiments show that in terms of cardinality, only a tiny part of the support of the approximately computed residual is used to expand the current index set. Taking into account that finding the smallest set $\breve{\Lambda}$ involves finding the biggest entries in $\mathbf{r}$, it appears that one might be able to save a considerable amount of resources if one knows where to look for the biggest entries in $\mathbf{r}$. This is the basic motivation behind the development in this chapter, which is more explicitly expressed in the following.

Suppose that for any finite tree $\Lambda \subset \nabla$ and any $\mathbf{v} \in \ell_2(\Lambda)$, prior to computing the residual $\mathbf{r} = \mathbf{f} - \mathbf{A}\mathbf{v}$, we know how to find a tree $\Lambda^\star \supset \Lambda$ such that $\|\mathbf{P}_{\Lambda^\star}\mathbf{r}\| \geq \eta\|\mathbf{r}\|$ with an absolute constant $\eta > 0$. Then, by only computing the part $\mathbf{r}_{\Lambda^\star} := \mathbf{P}_{\Lambda^\star}\mathbf{r}$ of the residual, and choosing a tree $\breve{\Lambda}$, with the smallest possible support, such that $\|\mathbf{P}_{\breve{\Lambda}}\mathbf{r}_{\Lambda^\star}\| \geq \alpha\|\mathbf{r}_{\Lambda^\star}\|$ for some $\alpha \in (0,1]$, we can guarantee that $\|\mathbf{P}_{\breve{\Lambda}}\mathbf{r}\| \geq \alpha\eta\|\mathbf{r}\|$. Therefore, employing Lemma 6.3.1 we obtain convergence.

As for the convergence rate, a straightforward adaptation of Lemma 3.3.1 does not give a bound on $\#(\breve{\Lambda} \setminus \Lambda)$ that is independent of $|\mathbf{v}|_{\mathcal{A}^s}$. Yet, the following modification offers such a bound, which result can be thought of as being an analogy to [88, Lemma 5.1] in the adaptive finite element setting.

Lemma 6.3.2. *Let be given a map $\mathcal{V}$ that sends trees $\Lambda \subset \bar{\Lambda}$ to a tree $\Lambda^\star = \mathcal{V}(\Lambda, \bar{\Lambda})$ such that*

$$\|\!|\mathbf{u}_{\Lambda^\star} - \mathbf{u}_\Lambda|\!\| \geq \eta\|\!|\mathbf{u}_{\bar{\Lambda}} - \mathbf{u}_\Lambda|\!\|, \tag{6.3.2}$$

where $\eta > 0$ is a constant, and $\mathbf{u}_{\Lambda^\star}$, $\mathbf{u}_{\bar{\Lambda}}$, and $\mathbf{u}_\Lambda$ are the Galerkin approximations to $\mathbf{u}$ from the corresponding subspaces. Assume that $\mathcal{V}$ is such that for any trees $\Lambda \subset \bar{\Lambda}$,

$$\Lambda \subset \mathcal{V}(\Lambda, \bar{\Lambda}) \subseteq \mathcal{V}(\Lambda, \nabla), \qquad \#\mathcal{V}(\Lambda, \nabla) \lesssim \#\Lambda,$$

and

$$\#\left(\mathcal{V}(\Lambda, \bar{\Lambda}) \setminus \Lambda\right) \lesssim \#(\bar{\Lambda} \setminus \Lambda),$$

for the latter assuming $\bar{\Lambda}$ is finite.

Let $\alpha \in (0, \eta\kappa(\mathbf{A})^{-\frac{1}{2}})$ be a constant, Λ be a finite tree , and for some $s > 0$, let $\mathbf{u} \in \mathcal{A}^s$. Then, with $\Lambda^\star := \mathcal{V}(\Lambda, \nabla)$ and $\mathbf{r}_{\Lambda^\star} := \mathbf{P}_{\Lambda^\star}(\mathbf{f} - \mathbf{A}\mathbf{u}_\Lambda)$, the smallest *tree $\breve{\Lambda} \supset \Lambda$ with*

$$\|\mathbf{P}_{\breve{\Lambda}}\mathbf{r}_{\Lambda^\star}\| \geq \alpha\|\mathbf{r}_{\Lambda^\star}\|$$

satisfies

$$\#\left(\breve{\Lambda} \setminus \Lambda\right) \lesssim \|\mathbf{u} - \mathbf{u}_\Lambda\|^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}.$$

Proof. Let $\lambda > 0$ be a constant with $\alpha = \eta\kappa(\mathbf{A})^{-\frac{1}{2}}(1 - \|\mathbf{A}\|\lambda^2)^{\frac{1}{2}}$. Let Λ' be a smallest tree such that $\|\mathbf{u} - \mathbf{P}_{\Lambda'}\mathbf{u}\| \leq \lambda\|\!|\mathbf{u} - \mathbf{u}_\Lambda|\!\|$. Since $\|\!|\mathbf{u} - \mathbf{u}_\Lambda|\!\| \geq \|\mathbf{A}^{-1}\|^{-\frac{1}{2}}\|\mathbf{u} - \mathbf{u}_\Lambda\|$, we have

$$\#\Lambda' \lesssim \|\mathbf{u} - \mathbf{u}_\Lambda\|^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}.$$

With $\bar{\Lambda} := \Lambda \cup \Lambda'$, we have

$$\|\!|\mathbf{u} - \mathbf{u}_{\bar{\Lambda}}|\!\| \leq \|\!|\mathbf{u} - \mathbf{P}_{\Lambda'}\mathbf{u}|\!\| \leq \|\mathbf{A}\|^{\frac{1}{2}}\|\mathbf{u} - \mathbf{P}_{\Lambda'}\mathbf{u}\| \leq \|\mathbf{A}\|^{\frac{1}{2}}\lambda\|\!|\mathbf{u} - \mathbf{u}_\Lambda|\!\|,$$

and so by Galerkin orthogonality, $|||\mathbf{u}_{\bar\Lambda}-\mathbf{u}_\Lambda||| \geq (1-\|\mathbf{A}\|\lambda^2)^{\frac12}|||\mathbf{u}-\mathbf{u}_\Lambda|||$. Now with $\bar\Lambda^\star := \mathcal{V}(\Lambda,\bar\Lambda)$, we infer

$$\begin{aligned}\|\mathbf{P}_{\bar\Lambda^\star}\mathbf{r}_{\Lambda^\star}\| &= \|\mathbf{P}_{\bar\Lambda^\star}\mathbf{A}(\mathbf{u}_{\bar\Lambda^\star}-\mathbf{u}_\Lambda)\| \geq \|\mathbf{A}^{-1}\|^{-\frac12}|||\mathbf{u}_{\bar\Lambda^\star}-\mathbf{u}_\Lambda||| \\ &\geq \|\mathbf{A}^{-1}\|^{-\frac12}\eta|||\mathbf{u}_{\bar\Lambda}-\mathbf{u}_\Lambda||| \geq \|\mathbf{A}^{-1}\|^{-\frac12}\eta(1-\|\mathbf{A}\|\lambda^2)^{\frac12}|||\mathbf{u}-\mathbf{u}_\Lambda||| \\ &\geq \kappa(\mathbf{A})^{-\frac12}\eta(1-\|\mathbf{A}\|\lambda^2)^{\frac12}\|\mathbf{f}-\mathbf{A}\mathbf{u}_\Lambda\| \geq \alpha\|\mathbf{r}\| \geq \alpha\|\mathbf{r}_{\Lambda^\star}\|.\end{aligned}$$

Since $\Lambda \subset \bar\Lambda^\star \subseteq \Lambda^\star$, by definition of $\breve\Lambda$ we conclude that

$$\#\left(\breve\Lambda\backslash\Lambda\right) \leq \#\left(\bar\Lambda^\star\backslash\Lambda\right) \lesssim \#\left(\bar\Lambda\backslash\Lambda\right) \leq \#\Lambda' \lesssim \|\mathbf{u}-\mathbf{u}_\Lambda\|^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}. \qquad \blacksquare$$

Let a map $\mathcal{V}$ satisfying the conditions of the preceding lemma is given. Then for some constant $\alpha \in (0,\eta\kappa(\mathbf{A})^{-\frac12})$ and for $i \in \mathbb{N}_0$, we define $\Lambda_i^\star := \mathcal{V}(\Lambda_i,\nabla)$, where $\Lambda_0 := \nabla_0$ and Λ_{i+1} is a smallest tree with $\|\mathbf{P}_{\Lambda_{i+1}}\mathbf{r}_i^\star\| \geq \alpha\|\mathbf{r}_i^\star\|$, where $\mathbf{r}_i^\star := \mathbf{f}_{\Lambda_i^\star}-\mathbf{A}_{\Lambda_i^\star}\mathbf{u}_{\Lambda_i}$. From the property (6.3.2), using the estimates (2.4.5) on page 22, we get

$$\begin{aligned}\|\mathbf{r}_i^\star\| &= \|\mathbf{P}_{\Lambda_i^\star}\mathbf{A}(\mathbf{u}_{\Lambda_i^\star}-\mathbf{u}_{\Lambda_i})\| \geq \|\mathbf{A}^{-1}\|^{-\frac12}|||\mathbf{u}_{\Lambda_i^\star}-\mathbf{u}_{\Lambda_i}||| \\ &\geq \eta\|\mathbf{A}^{-1}\|^{-\frac12}|||\mathbf{u}-\mathbf{u}_{\Lambda_i}||| \geq \eta\kappa(\mathbf{A})^{-\frac12}\|\mathbf{f}-\mathbf{A}\mathbf{u}_{\Lambda_i}\|,\end{aligned}$$

so by Lemma 6.3.1 we have a fixed error reduction: $|||\mathbf{u}-\mathbf{u}_{\Lambda_{i+1}}||| \leq \rho|||\mathbf{u}-\mathbf{u}_{\Lambda_i}|||$ with a constant $\rho<1$. Now assuming that $\mathbf{u} \in \mathcal{A}^s$ with some $s>0$, by the preceding lemma and the geometric decrease of $\|\mathbf{f}-\mathbf{A}\mathbf{u}_{\Lambda_i}\| \eqsim |||\mathbf{u}-\mathbf{u}_{\Lambda_i}|||$, for $i \in \mathbb{N}_0$ we have

$$\begin{aligned}\#\Lambda_k &= \textstyle\sum_{i=0}^{k-1}\#(\Lambda_{i+1}\setminus\Lambda_i) \lesssim \sum_{i=0}^{k-1}\|\mathbf{f}-\mathbf{A}\mathbf{u}_{\Lambda_i}\|^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} \\ &\lesssim \|\mathbf{f}-\mathbf{A}\mathbf{u}_{\Lambda_{k-1}}\|^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s},\end{aligned}$$

or, $\|\mathbf{u}-\mathbf{u}_{\Lambda_k}\| \lesssim (\#\Lambda_k)^{-s}|\mathbf{u}|_{\mathcal{A}^s}$, which, in view of the assumption $\mathbf{u} \in \mathcal{A}^s$, is modulo some constant factor the best possible bound on the error.

Unfortunately, for a general right hand side $\mathbf{f}$, a mapping $\mathcal{V}$ as in Lemma 6.3.2 does not exist since for any trees $\Lambda \subset \Lambda^\star$, and for any $\mathbf{f} \in \ell_2$ with $\mathbf{f}|_{\Lambda^\star\backslash\Lambda} = (\mathbf{A}\mathbf{u}_\Lambda)|_{\Lambda^\star\backslash\Lambda}$, we have $\mathbf{u}_{\Lambda^\star} = \mathbf{u}_\Lambda$. However, by using techniques from the theory of adaptive finite element methods, we realized a mapping $\mathcal{V}$ satisfying somewhat weaker conditions than those in Lemma 6.3.2, which are nevertheless sufficient conditions for showing optimality of suitable adaptive wavelet algorithms.

6.3.2 The main result

Before stating our main result, we need to introduce a number of technical assumptions and definitions. The first assumption basically assumes the existence of the map $\mathcal{V}$, which will be confirmed for second order elliptic partial differential operators in the next section.

Assumption 6.3.3. There exist

- a subspace $Y \subseteq \ell_2(\nabla)$,
- a function $\varrho : \mathcal{T} \times Y \to [0,\infty)$ with $\varrho(\tilde{\Lambda},\cdot) \leq \varrho(\Lambda,\cdot)$ for $\tilde{\Lambda} \supset \Lambda$,
- a map $\mathcal{V} : (\Lambda, \bar{\Lambda}) \mapsto \Lambda^\star \in \tilde{\mathcal{T}}$ where $\Lambda \in \tilde{\mathcal{T}}$ and $\bar{\Lambda} \supset \Lambda$ is a tree,
- and an absolute constant $\eta > 0$,

such that for any $\mathbf{g} \in Y$, $\Lambda \in \tilde{\mathcal{T}}$, and a tree $\bar{\Lambda} \supset \Lambda$, with $\Lambda^\star := \mathcal{V}(\Lambda,\bar{\Lambda})$, $\mathbf{v}_\Lambda := \mathbf{A}_\Lambda^{-1}\mathbf{P}_\Lambda\mathbf{g}$, $\mathbf{v}_{\bar{\Lambda}} := \mathbf{A}_{\bar{\Lambda}}^{-1}\mathbf{P}_{\bar{\Lambda}}\mathbf{g}$, and $\mathbf{v}_{\Lambda^\star} := \mathbf{A}_{\Lambda^\star}^{-1}\mathbf{P}_{\Lambda^\star}\mathbf{g}$, it holds that

$$|||\mathbf{v}_{\Lambda^\star} - \mathbf{v}_\Lambda||| \geq \eta |||\mathbf{v}_{\bar{\Lambda}} - \mathbf{v}_\Lambda||| - \varrho(\Lambda,\mathbf{g}). \tag{6.3.3}$$

Moreover, we assume that the map $\mathcal{V}$ is such that for any graded tree Λ and a tree $\Lambda \subset \bar{\Lambda}$,

$$\Lambda \subset \mathcal{V}(\Lambda,\bar{\Lambda}) \subseteq \mathcal{V}(\Lambda,\nabla), \qquad \#\mathcal{V}(\Lambda,\nabla) \lesssim \#\Lambda,$$

and

$$\#\left(\mathcal{V}(\Lambda,\bar{\Lambda}) \setminus \Lambda\right) \lesssim \#(\bar{\Lambda} \setminus \Lambda),$$

for the latter assuming $\bar{\Lambda}$ is finite. Finally, for any graded tree Λ, with $\Lambda^\star := \mathcal{V}(\Lambda,\nabla)$, we assume that the minimum level difference between any index from $\Lambda^\star$ and its ancestor from Λ is uniformly bounded, and that $\Lambda^\star$ can be determined by spending a number arithmetic operations and storage locations of order $\#\Lambda$.⊘

From now on in this section we will assume Assumption 6.3.3. In Section 6.4, we will verify this assumption in the case of second order elliptic boundary value problems.

The next proposition is a generalization of Lemma 6.3.2 on page 90. In particular, we use an approximate residual and inexact solution of the Galerkin systems.

Proposition 6.3.4. *With Λ a graded tree, let $\Lambda^\star := \mathcal{V}(\Lambda,\nabla)$, and with a constant $\alpha \in (0, \eta\kappa(\mathbf{A})^{-\frac{1}{2}})$, let $\delta,\delta',\delta_\varrho > 0$ be sufficiently small constants such that $\frac{\alpha+\delta+2\delta'\eta+\delta_\varrho\|\mathbf{A}^{-1}\|^{-\frac{1}{2}}}{1-\delta} < \eta\kappa(\mathbf{A})^{-\frac{1}{2}}$. Moreover, let $\mathbf{g} \in Y$ and $\tilde{\mathbf{r}}^\star \in \ell_2(\Lambda^\star)$ be such that $\varrho(\Lambda,\mathbf{g}) \leq \delta_\varrho\|\tilde{\mathbf{r}}^\star\|$ and $\|\mathbf{r}^\star - \tilde{\mathbf{r}}^\star\| \leq \delta\|\tilde{\mathbf{r}}^\star\|$, where $\mathbf{r}^\star := \mathbf{P}_{\Lambda^\star}(\mathbf{g} - \mathbf{A}\mathbf{v}_\Lambda)$ and $\mathbf{v}_\Lambda := \mathbf{A}_\Lambda^{-1}\mathbf{P}_\Lambda\mathbf{g}$, and let $\tilde{\mathbf{v}}_\Lambda \in \ell_2(\Lambda)$ be such that $\|\mathbf{P}_\Lambda(\mathbf{g}-\mathbf{A}\tilde{\mathbf{v}}_\Lambda)\|+\|\mathbf{f}-\mathbf{g}\| \leq \delta'\|\tilde{\mathbf{r}}^\star\|$. Then, whenever $\mathbf{u} \in \mathcal{A}^s$ for some $s > 0$, a* smallest *tree $\breve{\Lambda} \supset \Lambda$ with*

$$\|\mathbf{P}_{\breve{\Lambda}}\tilde{\mathbf{r}}^\star\| \geq \alpha\|\tilde{\mathbf{r}}^\star\|$$

satisfies

$$\#\left(\breve{\Lambda}\setminus\Lambda\right)\lesssim\|\mathbf{u}-\tilde{\mathbf{v}}_\Lambda\|^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}.$$

Proof. Let $\lambda>0$ be a constant whose value will be specified later, and let Λ' be a smallest tree such that $\|\mathbf{u}-\mathbf{P}_{\Lambda'}\mathbf{u}\|\le\lambda|\!|\!|\mathbf{u}-\tilde{\mathbf{v}}_\Lambda|\!|\!|$. Since $|\!|\!|\mathbf{u}-\tilde{\mathbf{v}}_\Lambda|\!|\!|\ge\|\mathbf{A}^{-1}\|^{-\frac12}\|\mathbf{u}-\tilde{\mathbf{v}}_\Lambda\|$, we have

$$\#\Lambda'\lesssim\|\mathbf{u}-\tilde{\mathbf{v}}_\Lambda\|^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}.$$

For any tree $\Sigma\supset\Lambda$, with $\mathbf{v}_\Sigma:=\mathbf{A}_\Sigma^{-1}\mathbf{P}_\Sigma\mathbf{g}$ and $\mathbf{u}_\Sigma:=\mathbf{A}_\Sigma^{-1}\mathbf{P}_\Sigma\mathbf{f}$, we have

$$\begin{aligned}|\!|\!|(\mathbf{v}_\Sigma-\mathbf{v}_\Lambda)-(\mathbf{u}_\Sigma-\tilde{\mathbf{v}}_\Lambda)|\!|\!|&\le\|\mathbf{A}^{-1}\|^{\frac12}\left(\|\mathbf{P}_\Sigma(\mathbf{f}-\mathbf{g})\|+\|\mathbf{P}_\Lambda(\mathbf{g}-\mathbf{A}\tilde{\mathbf{v}}_\Lambda)\|\right)\\&\le\delta'\|\mathbf{A}^{-1}\|^{\frac12}\|\tilde{\mathbf{r}}^\star\|.\end{aligned}\tag{6.3.4}$$

Using this, with $\bar\Lambda:=\Lambda\cup\Lambda'$ and $\bar\Lambda^\star:=\mathcal{V}(\Lambda,\bar\Lambda)$, we infer

$$\begin{aligned}\|\mathbf{P}_{\bar\Lambda^\star}\tilde{\mathbf{r}}^\star\|&\ge\|\mathbf{P}_{\bar\Lambda^\star}\mathbf{r}^\star\|-\delta\|\tilde{\mathbf{r}}^\star\|\ge\|\mathbf{A}^{-1}\|^{-\frac12}|\!|\!|\mathbf{v}_{\bar\Lambda^\star}-\mathbf{v}_\Lambda|\!|\!|-\delta\|\tilde{\mathbf{r}}^\star\|\\&\ge\|\mathbf{A}^{-1}\|^{-\frac12}\eta|\!|\!|\mathbf{v}_{\bar\Lambda}-\mathbf{v}_\Lambda|\!|\!|-\|\mathbf{A}^{-1}\|^{-\frac12}\varrho(\Lambda,\mathbf{g})-\delta\|\tilde{\mathbf{r}}^\star\|\\&\ge\|\mathbf{A}^{-1}\|^{-\frac12}\eta|\!|\!|\mathbf{u}_{\bar\Lambda}-\tilde{\mathbf{v}}_\Lambda|\!|\!|-(\delta'\eta+\delta_\varrho\|\mathbf{A}^{-1}\|^{-\frac12}+\delta)\|\tilde{\mathbf{r}}^\star\|.\end{aligned}$$

We have $|\!|\!|\mathbf{u}-\mathbf{u}_{\bar\Lambda}|\!|\!|\le|\!|\!|\mathbf{u}-\mathbf{P}_{\Lambda'}\mathbf{u}|\!|\!|\le\|\mathbf{A}\|^{\frac12}\|\mathbf{u}-\mathbf{P}_{\Lambda'}\mathbf{u}\|\le\|\mathbf{A}\|^{\frac12}\lambda|\!|\!|\mathbf{u}-\tilde{\mathbf{v}}_\Lambda|\!|\!|$, and so by Galerkin orthogonality, $|\!|\!|\mathbf{u}_{\bar\Lambda}-\tilde{\mathbf{v}}_\Lambda|\!|\!|\ge(1-\|\mathbf{A}\|\lambda^2)^{\frac12}|\!|\!|\mathbf{u}-\tilde{\mathbf{v}}_\Lambda|\!|\!|$. On the other hand, by using (6.3.4), we have

$$\begin{aligned}|\!|\!|\mathbf{u}-\tilde{\mathbf{v}}_\Lambda|\!|\!|&\ge|\!|\!|\mathbf{v}-\tilde{\mathbf{v}}_\Lambda|\!|\!|-\delta'\|\mathbf{A}^{-1}\|^{\frac12}\|\tilde{\mathbf{r}}^\star\|\\&\ge\|\mathbf{A}\|^{-\frac12}\|\mathbf{g}-\mathbf{A}\mathbf{v}_\Lambda\|-\delta'\|\mathbf{A}^{-1}\|^{\frac12}\|\tilde{\mathbf{r}}^\star\|\\&\ge\|\mathbf{A}\|^{-\frac12}\|\mathbf{r}^\star\|-\delta'\|\mathbf{A}^{-1}\|^{\frac12}\|\tilde{\mathbf{r}}^\star\|.\end{aligned}$$

Combining all these estimates and using $\|\mathbf{r}^\star\|\ge(1-\delta)\|\tilde{\mathbf{r}}^\star\|$, we deduce that

$$\begin{aligned}&\|\mathbf{P}_{\bar\Lambda^\star}\tilde{\mathbf{r}}^\star\|\\&\quad\ge\left\{(1-\|\mathbf{A}\|\lambda^2)^{\frac12}\eta[\kappa(\mathbf{A})^{-\frac12}(1-\delta)-\delta']-\delta'\eta-\delta_\varrho\|\mathbf{A}^{-1}\|^{-\frac12}-\delta\right\}\|\tilde{\mathbf{r}}^\star\|,\end{aligned}$$

and choosing a value of λ so that the expression between the curly brackets is at least α, which is possible by hypothesis, we have $\|\mathbf{P}_{\bar\Lambda^\star}\tilde{\mathbf{r}}^\star\|\ge\alpha\|\tilde{\mathbf{r}}^\star\|$.

Since $\Lambda\subset\bar\Lambda^\star\subseteq\Lambda^\star$, by definition of $\breve\Lambda$ we conclude that

$$\#\left(\breve\Lambda\backslash\Lambda\right)\le\#\left(\bar\Lambda^\star\backslash\Lambda\right)\lesssim\#\left(\bar\Lambda\backslash\Lambda\right)\le\#\Lambda'\lesssim\|\mathbf{u}-\tilde{\mathbf{v}}_\Lambda\|^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}.$$ ∎

The following proposition extends Lemma 6.3.1 on page 89 in that approximate residuals and inexact solution of the Galerkin systems are allowed.

Proposition 6.3.5. *With Λ a graded tree, let $\Lambda^\star := \mathcal{V}(\Lambda, \nabla)$, and let $0 < \delta < \alpha < 1$, $0 < \delta' < \frac{\alpha-\delta}{1+3\kappa(\mathbf{A})^{\frac{1}{2}}}$, and $\delta_\varrho > 0$ be constants. Moreover, let $\mathbf{g} \in Y$ and $\tilde{\mathbf{r}}^\star \in \ell_2(\Lambda^\star)$ be such that $\varrho(\Lambda, \mathbf{g}) \le \delta_\varrho \|\tilde{\mathbf{r}}^\star\|$ and $\|\mathbf{r}^\star - \tilde{\mathbf{r}}^\star\| \le \delta\|\tilde{\mathbf{r}}^\star\|$, where $\mathbf{r}^\star := \mathbf{P}_{\Lambda^\star}(\mathbf{g} - \mathbf{A}\mathbf{v}_\Lambda)$ and $\mathbf{v}_\Lambda := \mathbf{A}_\Lambda^{-1}\mathbf{P}_\Lambda \mathbf{g}$, and let $\tilde{\mathbf{v}}_\Lambda \in \ell_2(\Lambda)$ be such that $\|\mathbf{P}_\Lambda(\mathbf{g} - \mathbf{A}\tilde{\mathbf{v}}_\Lambda)\| + \|\mathbf{f} - \mathbf{g}\| \le \delta'\|\tilde{\mathbf{r}}^\star\|$. Then, with $\breve{\Lambda} \supset \Lambda$ being a graded tree such that $\|\mathbf{P}_{\breve{\Lambda}}\tilde{\mathbf{r}}^\star\| \ge \alpha\|\tilde{\mathbf{r}}^\star\|$, and $\tilde{\mathbf{v}}_{\breve{\Lambda}} \in \ell_2(\breve{\Lambda})$ satisfying $\|\mathbf{P}_{\breve{\Lambda}}(\mathbf{f} - \mathbf{A}\tilde{\mathbf{v}}_{\breve{\Lambda}})\| \le \delta'\|\tilde{\mathbf{r}}^\star\|$, we have*

$$|||\mathbf{u} - \tilde{\mathbf{v}}_{\breve{\Lambda}}||| \le \left(1 - (1-\beta)(1-3\beta)\xi^2\right)^{\frac{1}{2}} |||\mathbf{u} - \tilde{\mathbf{v}}_\Lambda|||,$$

where $\beta := \frac{\delta'\kappa(\mathbf{A})^{\frac{1}{2}}}{\alpha-\delta-\delta'}$ and $\xi := \frac{(\alpha-\delta)\kappa(\mathbf{A})^{-\frac{1}{2}}-\delta'}{1+\delta+\delta'\eta+\delta_\varrho\|\mathbf{A}^{-1}\|^{-\frac{1}{2}}}\eta$. Note that $3\beta, \xi \in (0,1)$ by the conditions on the constants.

Proof. From (6.3.4) with $\Sigma := \breve{\Lambda}$, we have

$$\begin{aligned}
|||\mathbf{u}_{\breve{\Lambda}} - \tilde{\mathbf{v}}_\Lambda||| &\ge |||\mathbf{v}_{\breve{\Lambda}} - \mathbf{v}_\Lambda||| - \delta'\|\mathbf{A}^{-1}\|^{\frac{1}{2}}\|\tilde{\mathbf{r}}^\star\| \\
&\ge \|\mathbf{A}\|^{-\frac{1}{2}}\|\mathbf{P}_{\breve{\Lambda}}(\mathbf{g} - \mathbf{A}\mathbf{v}_\Lambda)\| - \delta'\|\mathbf{A}^{-1}\|^{\frac{1}{2}}\|\tilde{\mathbf{r}}^\star\| \\
&\ge \|\mathbf{A}\|^{-\frac{1}{2}}\|\mathbf{P}_{\breve{\Lambda}}\tilde{\mathbf{r}}^\star\| - (\delta\|\mathbf{A}\|^{-\frac{1}{2}} + \delta'\|\mathbf{A}^{-1}\|^{\frac{1}{2}})\|\tilde{\mathbf{r}}^\star\| \\
&\ge \alpha\|\mathbf{A}\|^{-\frac{1}{2}}\|\tilde{\mathbf{r}}^\star\| - (\delta\|\mathbf{A}\|^{-\frac{1}{2}} + \delta'\|\mathbf{A}^{-1}\|^{\frac{1}{2}})\|\tilde{\mathbf{r}}^\star\|.
\end{aligned}$$

Now combining the estimates $\|\tilde{\mathbf{r}}^\star\| \ge \|\mathbf{r}^\star\| - \delta\|\tilde{\mathbf{r}}^\star\|$, and

$$\begin{aligned}
\|\mathbf{r}^\star\| &\ge \|\mathbf{A}^{-1}\|^{-\frac{1}{2}}|||\mathbf{v}_{\Lambda^\star} - \mathbf{v}_\Lambda||| \ge \|\mathbf{A}^{-1}\|^{-\frac{1}{2}}\eta|||\mathbf{v} - \mathbf{v}_\Lambda||| - \|\mathbf{A}^{-1}\|^{-\frac{1}{2}}\varrho(\Lambda, \mathbf{g}) \\
&\ge \|\mathbf{A}^{-1}\|^{-\frac{1}{2}}\eta|||\mathbf{u} - \tilde{\mathbf{v}}_\Lambda||| - \delta\eta\|\tilde{\mathbf{r}}^\star\| - \delta_\varrho\|\mathbf{A}^{-1}\|^{-\frac{1}{2}}\|\tilde{\mathbf{r}}^\star\|,
\end{aligned}$$

we get $\|\mathbf{A}^{-1}\|^{-\frac{1}{2}}\eta|||\mathbf{u} - \tilde{\mathbf{v}}_\Lambda||| \le (1 + \delta + \delta'\eta + \delta_\varrho\|\mathbf{A}^{-1}\|^{-\frac{1}{2}})\|\tilde{\mathbf{r}}^\star\|$. In view of (37), this gives $|||\mathbf{u}_{\breve{\Lambda}} - \tilde{\mathbf{v}}_\Lambda||| \ge \xi|||\mathbf{u} - \tilde{\mathbf{v}}_\Lambda|||$, and by Galerkin orthogonality, we conclude that $|||\mathbf{u} - \mathbf{u}_{\breve{\Lambda}}||| \le (1 - \xi^2)^{\frac{1}{2}}|||\mathbf{u} - \tilde{\mathbf{v}}_\Lambda|||$.

One can simply estimate $|||\mathbf{u} - \tilde{\mathbf{v}}_{\breve{\Lambda}}||| \le |||\mathbf{u} - \mathbf{u}_{\breve{\Lambda}}||| + |||\mathbf{u}_{\breve{\Lambda}} - \tilde{\mathbf{v}}_{\breve{\Lambda}}|||$, but a sharper result can be derived by using that $\mathbf{u} - \tilde{\mathbf{v}}_{\breve{\Lambda}}$ is nearly $\langle\langle\cdot,\cdot\rangle\rangle$-orthogonal to $\ell_2(\breve{\Lambda})$, with $\langle\langle\cdot,\cdot\rangle\rangle := \langle\mathbf{A}\cdot,\cdot\rangle$. We have $|||\mathbf{u}_{\breve{\Lambda}} - \tilde{\mathbf{v}}_{\breve{\Lambda}}||| \le \|\mathbf{A}^{-1}\|^{\frac{1}{2}}\|\mathbf{P}_{\breve{\Lambda}}(\mathbf{f}_\Lambda - \mathbf{A}\tilde{\mathbf{v}}_{\breve{\Lambda}})\| \le \|\mathbf{A}^{-1}\|^{\frac{1}{2}}\delta'\|\tilde{\mathbf{r}}^\star\|$, and

$$\alpha\|\tilde{\mathbf{r}}^\star\| \le \|\mathbf{P}_{\breve{\Lambda}}\tilde{\mathbf{r}}^\star\| \le \|\mathbf{P}_{\breve{\Lambda}}\mathbf{r}^\star\| + \delta\|\tilde{\mathbf{r}}^\star\| \le \|\mathbf{P}_{\breve{\Lambda}}(\mathbf{f} - \mathbf{A}\mathbf{v}_\Lambda)\| + (\delta' + \delta)\|\tilde{\mathbf{r}}^\star\|,$$

so that $|||\mathbf{u}_{\breve{\Lambda}} - \tilde{\mathbf{v}}_{\breve{\Lambda}}||| \le \beta|||\mathbf{u}_{\breve{\Lambda}} - \tilde{\mathbf{v}}_\Lambda|||$.

The rest of the proof is equivalent to the corresponding part in the proof of Proposition 3.2.2 on page 46, but we reproduce it here for the reader's convenience. Using the Galerkin orthogonality $\mathbf{u}-\mathbf{u}_{\breve{\Lambda}} \perp_{\langle\!\langle\,,\rangle\!\rangle} \ell_2(\breve{\Lambda})$, we have

$$\begin{aligned}\langle\!\langle \mathbf{u}-\tilde{\mathbf{v}}_{\breve{\Lambda}}, \tilde{\mathbf{v}}_{\breve{\Lambda}}-\tilde{\mathbf{v}}_{\Lambda}\rangle\!\rangle &= \langle\!\langle \mathbf{u}_{\breve{\Lambda}}-\tilde{\mathbf{v}}_{\breve{\Lambda}}, \tilde{\mathbf{v}}_{\breve{\Lambda}}-\tilde{\mathbf{v}}_{\Lambda}\rangle\!\rangle \\ &\leq |\!|\!| \mathbf{u}_{\breve{\Lambda}}-\tilde{\mathbf{v}}_{\breve{\Lambda}} |\!|\!| \, |\!|\!| \tilde{\mathbf{v}}_{\breve{\Lambda}}-\tilde{\mathbf{v}}_{\Lambda} |\!|\!| \leq \beta |\!|\!| \mathbf{u}_{\breve{\Lambda}}-\tilde{\mathbf{v}}_{\Lambda} |\!|\!| \, |\!|\!| \tilde{\mathbf{v}}_{\breve{\Lambda}}-\tilde{\mathbf{v}}_{\Lambda} |\!|\!|.\end{aligned}$$

Now by writing

$$|\!|\!| \mathbf{u}-\tilde{\mathbf{v}}_{\Lambda} |\!|\!|^2 = |\!|\!| \mathbf{u}-\tilde{\mathbf{v}}_{\breve{\Lambda}} |\!|\!|^2 + |\!|\!| \tilde{\mathbf{v}}_{\breve{\Lambda}}-\tilde{\mathbf{v}}_{\Lambda} |\!|\!|^2 + 2\langle\!\langle \mathbf{u}-\tilde{\mathbf{v}}_{\breve{\Lambda}}, \tilde{\mathbf{v}}_{\breve{\Lambda}}-\tilde{\mathbf{v}}_{\Lambda}\rangle\!\rangle,$$

and, for obtaining the second line in the following multi-line formula, twice applying

$$|\!|\!| \tilde{\mathbf{v}}_{\breve{\Lambda}}-\tilde{\mathbf{v}}_{\Lambda} |\!|\!| \geq |\!|\!| \mathbf{u}_{\breve{\Lambda}}-\tilde{\mathbf{v}}_{\Lambda} |\!|\!| - |\!|\!| \tilde{\mathbf{v}}_{\breve{\Lambda}}-\mathbf{u}_{\breve{\Lambda}} |\!|\!| \geq (1-\beta) |\!|\!| \mathbf{u}_{\breve{\Lambda}}-\tilde{\mathbf{v}}_{\Lambda} |\!|\!|,$$

and for the third line, using $|\!|\!| \mathbf{u}_{\breve{\Lambda}}-\tilde{\mathbf{v}}_{\Lambda} |\!|\!| \geq \xi |\!|\!| \mathbf{u}-\tilde{\mathbf{v}}_{\Lambda} |\!|\!|$, we find that

$$\begin{aligned}|\!|\!| \mathbf{u}-\tilde{\mathbf{v}}_{\Lambda} |\!|\!|^2 &\geq |\!|\!| \mathbf{u}-\tilde{\mathbf{v}}_{\breve{\Lambda}} |\!|\!|^2 + |\!|\!| \tilde{\mathbf{v}}_{\breve{\Lambda}}-\tilde{\mathbf{v}}_{\Lambda} |\!|\!| \big(|\!|\!| \tilde{\mathbf{v}}_{\breve{\Lambda}}-\tilde{\mathbf{v}}_{\Lambda} |\!|\!| - 2\beta |\!|\!| \mathbf{u}_{\breve{\Lambda}}-\tilde{\mathbf{v}}_{\Lambda} |\!|\!| \big) \\ &\geq |\!|\!| \mathbf{u}-\tilde{\mathbf{v}}_{\breve{\Lambda}} |\!|\!|^2 + (1-\beta)(1-3\beta) |\!|\!| \mathbf{u}_{\breve{\Lambda}}-\tilde{\mathbf{v}}_{\Lambda} |\!|\!|^2 \\ &\geq |\!|\!| \mathbf{u}-\tilde{\mathbf{v}}_{\breve{\Lambda}} |\!|\!|^2 + (1-\beta)(1-3\beta)\xi^2 |\!|\!| \mathbf{u}-\tilde{\mathbf{v}}_{\Lambda} |\!|\!|^2,\end{aligned}$$

which completes the proof. ∎

Now we will assume the availability of some subroutines, from which we will assemble our adaptive wavelet solver. In conjunction with formulating the requirements for those subroutines conveniently, we state the following assumption.

Assumption 6.3.6. It holds that $\mathbf{u} \in \mathcal{A}^s$ for some $s > 0$.

The following subroutine provides a means to extract information from the right hand side $\mathbf{f}$. The availability of this subroutine requires that the subspace Y is dense in ℓ_2, and that for $\mathbf{g} \in Y$, $\varrho(\Lambda, \mathbf{g})$ can be made arbitrarily small by choosing Λ sufficiently large.

Algorithm 6.3.7 Algorithm template $\mathbf{TRHS}[\Lambda, \varepsilon] \to [\mathbf{g}, \breve{\Lambda}]$

Input: $\Lambda \in \mathcal{T}$ and $\varepsilon > 0$.

Output: $\mathbf{g} \in Y$, $\Lambda \subseteq \breve{\Lambda} \in \mathcal{T}$, such that $\|\mathbf{f}-\mathbf{g}\| + \varrho(\breve{\Lambda}, \mathbf{g}) \leq \varepsilon$. Moreover, we have $\#\breve{\Lambda} - \#\Lambda \lesssim \varepsilon^{-1/s} c_{\mathbf{f}}$ for some constant $c_{\mathbf{f}}$ only dependent of $\mathbf{f}$, and the number of arithmetic operations required for this call is bounded by an absolute multiple of $\#\breve{\Lambda}$. Furthermore, for any $\tilde{\Lambda} \in \tilde{\mathcal{T}}$, the computation of $\mathbf{P}_{\tilde{\Lambda}}\mathbf{g}$ takes the order of $\#\tilde{\Lambda}$ arithmetic operations.

Analogously to the above, the following subroutine will be the device with which the solver will perceive the matrix $\mathbf{A}$. Note that, with the subroutine **APPLY** from Algorithm 2.7.9 on page 33, $\mathbf{P}_{\tilde{\Lambda}}\,(\mathbf{APPLY}[\mathbf{A}, \mathbf{v}, \varepsilon])$ has all the required properties, thus defining a valid routine.

Algorithm 6.3.8 Algorithm template $\mathbf{TAPPLY}[\tilde{\Lambda}, \mathbf{v}, \varepsilon] \to \mathbf{w}_{\tilde{\Lambda}}$

Input: $\varepsilon > 0$, and $\mathbf{v} \in \ell_2(\Lambda)$ with $\Lambda, \tilde{\Lambda} \in \tilde{\mathcal{T}}$, $\Lambda \subseteq \tilde{\Lambda}$, and $\#\tilde{\Lambda} \lesssim \#\Lambda$.

Output: $\mathbf{w}_{\tilde{\Lambda}} \in \ell_2(\tilde{\Lambda})$ and $\|\mathbf{A}_{\tilde{\Lambda}}\mathbf{v} - \mathbf{w}_{\tilde{\Lambda}}\| \leq \varepsilon$. Moreover, the number of arithmetic operations and storage locations required by the call is bounded by some absolute multiple of $\varepsilon^{-1/s}|\mathbf{v}|_{\mathcal{A}^s}^{1/s} + \#\Lambda + 1$.

For $\Lambda \in \tilde{\mathcal{T}}$ and $\mathbf{g}_\Lambda \in \ell_2(\Lambda)$, we will use the following subroutine to approximately solve the Galerkin system $\mathbf{A}_\Lambda \mathbf{v}_\Lambda = \mathbf{g}_\Lambda$. Note that the subroutine **GALSOLVE** from Algorithm 3.2.3 on page 47 defines a valid routine.

Algorithm 6.3.9 Algorithm template $\mathbf{TGALSOLVE}[\Lambda, \mathbf{g}_\Lambda, \mathbf{v}_\Lambda, \nu, \varepsilon] \to \mathbf{w}_\Lambda$

Input: $\varepsilon > 0$, $\Lambda \in \tilde{\mathcal{T}}$, and $\mathbf{g}_\Lambda, \mathbf{v}_\Lambda \in \ell_2(\Lambda)$ such that $\|\mathbf{g}_\Lambda - \mathbf{A}_\Lambda \mathbf{v}_\Lambda\| \leq \nu$.

Output: $\mathbf{w}_\Lambda \in \ell_2(\Lambda)$ and $\|\mathbf{g}_\Lambda - \mathbf{A}_\Lambda \mathbf{w}_\Lambda\| \leq \varepsilon$. Moreover, the number of arithmetic operations and storage locations required by the call is bounded by some absolute multiple of $\varepsilon^{-1/s}|\mathbf{v}_\Lambda|_{\mathcal{A}^s}^{1/s} + c(\varepsilon^{-1}\|\mathbf{g}_\Lambda - \mathbf{A}_\Lambda \mathbf{v}_\Lambda\|)\#\Lambda\}$, where $c : [0, \infty) \to [1, \infty)$ is some non-decreasing function.

In view of (6.2.4) on page 88, we assume the following subroutine.

Algorithm 6.3.10 Algorithm template $\mathbf{COMPLETE}[\Lambda] \to \tilde{\Lambda}$

Input: Let $\Lambda \in \mathcal{T}$.

Output: $\Lambda \subseteq \tilde{\Lambda} \in \tilde{\mathcal{T}}$ with $\#\tilde{\Lambda} \lesssim \#\Lambda$. The number of arithmetic operations and storage locations required by this call is bounded by an absolute multiple of $\#\tilde{\Lambda}$. Moreover, for the inputs $\Lambda_1 \subset \Lambda_2$, the corresponding outputs satisfy $\tilde{\Lambda}_1 \subset \tilde{\Lambda}_2$.

Now in view of Propositions 6.3.4 and 6.3.5 on pages 92–94, by using the above subroutines and the map $\mathcal{V}$ from Assumption 6.3.3, we construct a subroutine that approximates the residual of a Galerkin solution.

Algorithm 6.3.11 Computation of truncated Galerkin residual $\mathbf{TGALRES}[\Lambda_0, \mathbf{w}_0, \nu_0, \varepsilon] \to [\mathbf{r}_k, \Lambda_k, \tilde{\Lambda}_k, \mathbf{w}_k, \nu_k]$

Parameters: Let $\omega, \gamma, \gamma_r, \gamma_a > 0$, and $\theta > 0$ be constants with $(\gamma_r + \gamma)\omega < 1$.
Input: Let $\Lambda_0 \in \mathcal{T}$, $\mathbf{w}_0 \in \ell_2(\Lambda_0)$, $\nu_0 \geq \|\mathbf{f} - \mathbf{A}\mathbf{w}_0\|$, and $\varepsilon > 0$.
Output: $\Lambda_k \in \mathcal{T}$, $\tilde{\Lambda}_k \in \tilde{\mathcal{T}}$, $\mathbf{w}_k \in \ell_2(\tilde{\Lambda}_k)$ with $\|\mathbf{f} - \mathbf{A}\mathbf{w}_k\| \leq \nu_k$, and $\mathbf{r}_k \in P$.

$k := 0$, $\zeta_0 := \theta\nu_0$;
repeat
  $k := k + 1$, $\zeta_k := \zeta_{k-1}/2$;
  $[\mathbf{g}_k, \Lambda_k] := \mathbf{TRHS}[\Lambda_{k-1}, \gamma_r\zeta_k]$;
  $\tilde{\Lambda}_k := \mathbf{COMPLETE}[\Lambda_k]$;
  $\mathbf{w}_k := \mathbf{TGALSOLVE}[\tilde{\Lambda}_k, \mathbf{P}_{\tilde{\Lambda}_k}\mathbf{g}_k, \mathbf{w}_{k-1}, \nu_{k-1} + \gamma_r\zeta_k, \gamma\zeta_k]$;
  $\Lambda_k^\star := \mathcal{V}(\tilde{\Lambda}_k, \nabla)$;
  $\mathbf{r}_k := \mathbf{P}_{\Lambda_k^\star}\mathbf{g}_k - \mathbf{TAPPLY}[\Lambda_k^\star, \mathbf{w}_k, \gamma_a\zeta_k]$;
until $\nu_k := \kappa(\mathbf{A})^{\frac{1}{2}}\left[\eta^{-1}\|\mathbf{r}_k\| + (\eta^{-1}(\gamma_r + \gamma_a) + \gamma_r + \gamma)\,\zeta_k\right] \leq \varepsilon$ or $\zeta_k \leq \omega\|\mathbf{r}_k\|$.

Proposition 6.3.12. *With valid inputs, the subroutine* $[\mathbf{r}, \bar{\Lambda}, \tilde{\Lambda}, \mathbf{w}, \nu] := \mathbf{TGALRES}[\Lambda, \mathbf{w}_0, \nu_0, \varepsilon]$ *terminates with* $\mathbf{w} \in \ell_2(\tilde{\Lambda})$, $\|\mathbf{f} - \mathbf{A}\mathbf{w}\| \leq \nu$, *and* $\|\mathbf{P}_{\tilde{\Lambda}}(\mathbf{f} - \mathbf{A}\mathbf{w})\| \leq \nu_0\theta(\gamma + \gamma_r)/2$. *Moreover, we have* $\nu \gtrsim \min\{\nu_0, \varepsilon\}$, $\bar{\Lambda} \in \mathcal{T}$, $\tilde{\Lambda} \in \tilde{\mathcal{T}}$, $\#\tilde{\Lambda} \lesssim \#\bar{\Lambda}$, *and* $\#\bar{\Lambda} - \#\Lambda \lesssim c_{\mathbf{f}}\nu^{-1/s}$.

If the subroutine terminates with $\nu > \varepsilon$, *then* $\nu \lesssim \|\mathbf{f} - \mathbf{A}\mathbf{w}\|$, *and with* $\Lambda^\star := \mathcal{V}(\tilde{\Lambda}, \nabla)$, $\mathbf{r} \in \ell_2(\Lambda^\star)$, *there exists* $\mathbf{g} \in Y$ *such that*

$$\|\mathbf{f} - \mathbf{g}\| + \varrho(\mathbf{g}, \bar{\Lambda}) \leq \gamma_r\omega\|\mathbf{r}\|, \tag{6.3.5}$$

$$\|\mathbf{P}_{\tilde{\Lambda}}(\mathbf{g} - \mathbf{A}\mathbf{w})\| \leq \gamma\omega\|\mathbf{r}\|. \tag{6.3.6}$$

and with $\mathbf{v}_{\tilde{\Lambda}} := \mathbf{A}_{\tilde{\Lambda}}^{-1}\mathbf{g}$,

$$\|\mathbf{P}_{\Lambda^\star}(\mathbf{g} - \mathbf{A}\mathbf{v}_{\tilde{\Lambda}}) - \mathbf{r}\| \leq [\gamma_a + \gamma\kappa(\mathbf{A})^{\frac{1}{2}}]\omega\|\mathbf{r}\|, \tag{6.3.7}$$

Furthermore, the number of arithmetic operations and storage locations required by the call is bounded by some absolute multiple of

$$\nu^{-1/s}(|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + c_{\mathbf{f}}) + (\nu_0/\nu)^{1/s}(\#\Lambda + 1).$$

Proof. If at evaluation of the until-clause for the k-th iteration, $\zeta_k > \omega\|\mathbf{r}_k\|$, then $\rho_k = \|\mathbf{r}_k\| + (\gamma_r + \gamma_a)\zeta_k < (\omega^{-1} + \gamma_r + \gamma_a)\zeta_k$. Since ζ_k is halved in each iteration, we infer that, if not by $\zeta_k \leq \omega\|\mathbf{r}_k\|$, the loop will terminate by $\nu_k \leq \varepsilon$.

Let K be the value of k at the termination of the loop. Then for $1 \le k \le K$, we have

$$\begin{aligned}\|\mathbf{P}_{\tilde{\Lambda}_k}(\mathbf{f}-\mathbf{A}\mathbf{w}_k)\| &\le \|\mathbf{P}_{\tilde{\Lambda}_k}(\mathbf{g}_i-\mathbf{A}\mathbf{w}_k)\| + \|\mathbf{f}-\mathbf{g}_k\| \\ &\le (\gamma+\gamma_r)\zeta_k \le \zeta_1 = \nu_0\theta(\gamma+\gamma_r)/2.\end{aligned}$$

Let $1 \le k \le K$, and assume that $\nu_{k-1} \ge \|\mathbf{f}-\mathbf{A}\mathbf{w}_{k-1}\|$, which is true for $k=1$ by the condition on the inputs. Then it holds that

$$\|\mathbf{P}_{\tilde{\Lambda}_k}(\mathbf{g}_k-\mathbf{A}\mathbf{w}_{k-1})\| \le \|\mathbf{f}-\mathbf{g}\| + \|\mathbf{f}-\mathbf{A}\mathbf{w}_{k-1}\| \le \gamma_r\zeta_k + \nu_{k-1},$$

meaning that in the k-th iteration, the subroutine **TGALSOLVE** is called with a valid parameter. With $\mathbf{v}_k := \mathbf{A}^{-1}_{\tilde{\Lambda}_k}\mathbf{g}_k$ and $\mathbf{v}_k^\star := \mathbf{A}^{-1}_{\Lambda_k^\star}\mathbf{g}_k$, we have

$$\begin{aligned}\|\mathbf{P}_{\Lambda_k^\star}(\mathbf{g}_k-\mathbf{A}\mathbf{v})\| &\ge \|\mathbf{A}^{-1}\|^{-\frac{1}{2}}\|\!|\mathbf{v}_k^\star-\mathbf{v}_k|\!\| \\ &\ge \eta\|\mathbf{A}^{-1}\|^{-\frac{1}{2}}\|\!|\mathbf{A}^{-1}\mathbf{g}_k-\mathbf{v}_k|\!\| - \|\mathbf{A}^{-1}\|^{-\frac{1}{2}}\varrho(\tilde{\Lambda}_k,\mathbf{g}_k) \\ &\ge \eta\|\mathbf{A}^{-1}\|^{-\frac{1}{2}}\|\!|\mathbf{u}-\mathbf{w}_k|\!\| - (\gamma_r+\gamma)\eta\zeta_k \\ &\ge \eta\kappa(\mathbf{A})^{-\frac{1}{2}}\|\mathbf{f}-\mathbf{A}\mathbf{w}_k\| - (\gamma_r+\gamma)\eta\zeta_k,\end{aligned}$$

where in the third line we used the first inequality in 6.3.4 on page 93 with $\Sigma = \nabla$. Now using that $\|\mathbf{P}_{\Lambda_k^\star}(\mathbf{f}-\mathbf{A}\mathbf{w}_k)-\mathbf{r}_k\| \le (\gamma_r+\gamma)\zeta_k$, we infer $\nu_k \ge \|\mathbf{f}-\mathbf{A}\mathbf{w}_k\|$.

If the loop terminates in the first iteration, or terminates with $\nu_K > \varepsilon$, then $\nu_K \gtrsim \min\{\nu_0,\varepsilon\}$. In the other case, we have $A\|\mathbf{r}_{K-1}\| + B\zeta_K > \varepsilon$ with some fixed constants $A, B > 0$, and $2\zeta_K > \omega\|\mathbf{r}_{K-1}\|$, so that $\nu_K \gtrsim \zeta_K > \frac{A\|\mathbf{r}_{K-1}\|+B\zeta_K}{2A/\omega+B} \gtrsim \varepsilon$. From $\zeta_K \le \omega\|\mathbf{r}_K\|$ and the definition of ν_K we have $\nu_K \lesssim \|\mathbf{r}_K\|$ and $\|\mathbf{r}_K\| \le \|\mathbf{f}-\mathbf{A}\mathbf{w}_K\| + (\gamma_r+\gamma)\omega\|\mathbf{r}_K\|$, so that $\nu_K \lesssim \|\mathbf{f}-\mathbf{A}\mathbf{w}_K\|$ by $(\gamma_r+\gamma)\omega < 1$.

From the properties of **COMPLETE** we have $\#\tilde{\Lambda}_K \lesssim \#\Lambda_K$, and from the properties of **TRHS** and geometric decrease of ζ_k, we infer that $\#\Lambda_K - \#\Lambda_0 \lesssim c_{\mathbf{f}}\zeta_K^{-1/s}$. Now we will show that $\zeta_K \gtrsim \nu_K$. For $1 \le k \le K$, we have

$$\begin{aligned}\|\mathbf{A}(\mathbf{w}_k-\mathbf{w}_{k-1})\| &\le \|\mathbf{A}\|^{\frac{1}{2}}\|\!|\mathbf{w}_k-\mathbf{w}_{k-1}|\!\| \le \kappa(\mathbf{A})^{\frac{1}{2}}\|\mathbf{P}_{\tilde{\Lambda}_k}\mathbf{A}(\mathbf{w}_k-\mathbf{w}_{k-1})\| \\ &\le \kappa(\mathbf{A})^{\frac{1}{2}}\|\mathbf{P}_{\tilde{\Lambda}_k}(\mathbf{g}_k-\mathbf{A}\mathbf{w}_k)\| + \kappa(\mathbf{A})^{\frac{1}{2}}\|\mathbf{P}_{\tilde{\Lambda}_k}(\mathbf{g}_k-\mathbf{A}\mathbf{w}_{k-1})\| \\ &\le \kappa(\mathbf{A})^{\frac{1}{2}}\gamma\zeta_k + \kappa(\mathbf{A})^{\frac{1}{2}}\|\mathbf{g}_k-\mathbf{A}\mathbf{w}_{k-1}\|,\end{aligned}$$

and $\|\mathbf{g}_k-\mathbf{A}\mathbf{w}_{k-1}\| \le \nu_{k-1}+\gamma_r\zeta_k$. Using these estimates, we infer

$$\|\mathbf{g}_k-\mathbf{A}\mathbf{w}_k\| \le \|\mathbf{g}_k-\mathbf{A}\mathbf{w}_{k-1}\| + \|\mathbf{A}(\mathbf{w}_k-\mathbf{w}_{k-1})\| \lesssim \nu_{k-1}+\zeta_k,$$

implying that

$$\nu_k \lesssim \|\mathbf{r}_k\| + \zeta_k \le \|\mathbf{P}_{\Lambda_k^\star}(\mathbf{g}_k-\mathbf{A}\mathbf{w}_k)\| + \gamma_a\zeta_k + \zeta_k \lesssim \nu_{k-1}+\zeta_k.$$

We have $\nu_0 = \zeta_0$, and for $k > 1$, $\nu_{k-1} \lesssim \zeta_{k-1}$ by $\omega\|\mathbf{r}_{k-1}\| > \zeta_{k-1}$, so $\nu_k \lesssim \nu_{k-1} + \zeta_k \lesssim \zeta_{k-1} + \zeta_k \lesssim \zeta_k$, proving the first part of the proposition.

The inequalities (6.3.6) and (6.3.7) are immediate consequences of the properties of **TRHS** and **TGALSOLVE**, and the condition $\zeta_K \leq \omega\|\mathbf{r}_K\|$. One can prove (6.3.5) by using

$$\|\mathbf{P}_{\Lambda_K^\star}(\mathbf{g}_K - \mathbf{A}\mathbf{w}_K)\| \leq \|\mathbf{A}\|^{\frac{1}{2}}|\!|\!|\mathbf{v}_K - \mathbf{w}_K|\!|\!| \leq \kappa(\mathbf{A})^{\frac{1}{2}}\gamma\zeta_K.$$

The properties of the subroutines and the map $\mathcal{V}$ imply that the cost of k-th iteration can be bounded by some multiple of $\zeta_k^{-1/s}(|\mathbf{w}_k|_{\mathcal{A}^s}^{1/s} + |\mathbf{w}_{k-1}|_{\mathcal{A}^s}^{1/s}) + c(\frac{\nu_{k-1}}{\zeta_k})\#\Lambda_k + \#\Lambda_k + 1$, where $c(\cdot)$ is the non-decreasing function as described in the subroutine **TGALSOLVE** (Algorithm 6.3.9 on page 96). Since any vector $\mathbf{w}_k$ determined inside the algorithm satisfies $\|\mathbf{u} - \mathbf{w}_k\| \lesssim \nu_k$, from Remark 6.2.2, we infer that $|\mathbf{w}_k|_{\mathcal{A}^s} \lesssim |\mathbf{u}|_{\mathcal{A}^s} + (\#\Lambda_k)^s\nu_k$. At any iteration the ratio $\frac{\nu_{k-1}}{\zeta_k}$ is uniformly bounded, and $\nu_{k-1} \lesssim \zeta_{k-1} \lesssim \zeta_k$, so the cost of k-th iteration can be bounded by some multiple of $\zeta_k^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + \#\Lambda_k + 1$. Moreover, we have $\#\Lambda_k \lesssim \#\Lambda_0 + \zeta_k^{-1/s}c_{\mathbf{f}}$. By the geometric decrease of ζ_k inside the loop, the above considerations imply that the total cost of the algorithm can be bounded by some multiple of $\zeta_K^{-1/s}(|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + c_{\mathbf{f}}) + K(\#\Lambda_0 + 1)$. Taking into account the value of ζ_0, and the geometric decrease of ζ_k inside the loop, we have $K(\#\Lambda_0+1) = K\nu_0^{-1/s}\nu_0^{1/s}(\#\Lambda_0+1) \lesssim \zeta_K^{-1/s}\nu_0^{1/s}(\#\Lambda_0 + 1)$, and the proof is completed by $\zeta_K \gtrsim \nu_K$. ∎

Finally, we are ready to present our adaptive wavelet solver. Note that we employ the subroutine **RESTRICT** as in Algorithm 3.3.2 on page 53.

Algorithm 6.3.13 Adaptive Galerkin method **SOLVE**$[\varepsilon] \to \mathbf{w}_i$

Parameters: Let $\alpha \in (0, 1)$ be a constant.
Input: $\varepsilon > 0$.
Output: $\mathbf{w}_i \in P$ such that $\|\mathbf{f} - \mathbf{A}\mathbf{w}_i\| \leq \varepsilon$.

$i := 0$, $\mathbf{w}_0 := 0$, $\nu_0 := \|\mathbf{f}\|$, $\Lambda_1 := \nabla_0$;
loop
  $i := i + 1$;
  $[\mathbf{r}_i, \bar{\Lambda}_i, \tilde{\Lambda}_i, \mathbf{w}_i, \nu_i] :=$ **TGALRES**$[\Lambda_i, \mathbf{w}_{i-1}, \nu_{i-1}, \varepsilon]$;
  if $\nu_i \leq \varepsilon$ **then**
    Terminate the routine.
  end if
  $\Lambda_{i+1} :=$ **RESTRICT**$[\tilde{\Lambda}_i, \mathbf{r}_i, \alpha]$;
  Complete Λ_{i+1} to a tree by iteratively adding the parents of the indices whose parent is not in Λ_{i+1};
end loop

We need the following assumption on the subroutine **COMPLETE** to prove the optimality of the adaptive algorithm. This assumption will be verified for some important examples, cf. Example 6.4.2 on page 103.

Assumption 6.3.14. Let $\tilde{\Lambda}_0 := \nabla_0$, and for $i \in \mathbb{N}$, let $\Lambda_i \supset \tilde{\Lambda}_{i-1}$ be a tree, and let $\tilde{\Lambda}_i := \mathbf{COMPLETE}[\Lambda_i]$. Then for $k \in \mathbb{N}$, we have

$$\#\tilde{\Lambda}_k - \#\nabla_0 \lesssim \sum_{i=0}^{k-1} \#\Lambda_{i+1} - \#\tilde{\Lambda}_i.$$

Theorem 6.3.15. *Inside* **SOLVE** *and* **TGALRES**, *let the products* $\gamma\omega$, $\gamma_r\omega$, *and* $\gamma_a\omega$ *be small enough such that* $(\gamma_r+\gamma)\omega < \frac{\alpha-(\gamma_a+\gamma\kappa(\mathbf{A})^{\frac{1}{2}})\omega}{1+3\kappa(\mathbf{A})^{\frac{1}{2}}}$, *and let* θ *be such that* $\theta \leq \frac{2\omega}{\kappa(\mathbf{A})^{\frac{1}{2}}[\eta^{-1}+(\eta^{-1}(\gamma_r+\gamma_a)+\gamma_r+\gamma)\omega]}$. *Then* $\mathbf{u}_\varepsilon := \mathbf{SOLVE}[\varepsilon]$ *terminates with* $\|\mathbf{f}-\mathbf{A}\mathbf{u}_\varepsilon\| \leq \varepsilon$. *In addition, let* $\alpha \in (0, \eta\kappa(\mathbf{A})^{-\frac{1}{2}})$, *let the products* $\gamma\omega, \gamma_r\omega$, *and* $\gamma_a\omega$ *be small enough such that*

$$\frac{\alpha+[\gamma_a+\gamma\kappa(\mathbf{A})^{\frac{1}{2}}]\omega+2\eta(\gamma_r+\gamma)\omega+\gamma_r\|\mathbf{A}^{-1}\|^{-\frac{1}{2}}\omega}{1-(\gamma_a+\gamma\kappa(\mathbf{A})^{\frac{1}{2}})\omega} < \eta\kappa(\mathbf{A})^{-\frac{1}{2}},$$

and let $\varepsilon \lesssim \|\mathbf{f}\|$. *Then, we have* $\#\operatorname{supp}\mathbf{u}_\varepsilon \lesssim \varepsilon^{-1/s}(c_{\mathbf{f}} + |\mathbf{u}|_{\mathcal{A}^s}^{1/s})$ *and the number of arithmetic operations and storage locations required by the call is bounded by some absolute multiple of the same expression.*

Proof. Taking into account the conditions on the parameters, from Propositions 6.3.5 and 6.3.12, it is immediate that as long as $\nu_i > \varepsilon$, $|||\mathbf{u}-\mathbf{w}_{i+1}||| \leq \rho|||\mathbf{u}-\mathbf{w}_i|||$ with some fixed constant $\rho < 1$. Therefore the loop terminates say, directly after the K-th call of **TGALRES**.

By Assumption 6.3.14, for $1 \leq k \leq K$ we have

$$\begin{aligned}
\#\tilde{\Lambda}_k - \#\nabla_0 &\lesssim \#\bar{\Lambda}_1 - \#\nabla_0 + \sum_{i=1}^{k-1} \#\bar{\Lambda}_{i+1} - \#\tilde{\Lambda}_i \\
&\lesssim \sum_{i=1}^{k} \#\bar{\Lambda}_i - \#\Lambda_i + \sum_{i=1}^{k-1} \#\Lambda_{i+1} - \#\tilde{\Lambda}_i \\
&\lesssim \sum_{i=1}^{k} c_{\mathbf{f}}\nu_i^{-1/s} + \sum_{i=1}^{k-1} \|\mathbf{f}-\mathbf{A}\mathbf{w}_i\|^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s} \\
&\lesssim \nu_k^{-1/s}(c_{\mathbf{f}} + |\mathbf{u}|_{\mathcal{A}^s}^{1/s}).
\end{aligned}$$

From $\|\mathbf{f}\| \leq |\mathbf{f}|_{\mathcal{A}^s} \lesssim |\mathbf{u}|_{\mathcal{A}^s}$ and $\nu_k \lesssim \nu_0 = \|\mathbf{f}\|$, we have $\nabla_0 \lesssim 1 \lesssim \nu_k^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}$, implying that

$$\#\Lambda_k \lesssim \nu_k^{-1/s}(c_{\mathbf{f}} + |\mathbf{u}|_{\mathcal{A}^s}^{1/s}). \tag{6.3.8}$$

We have $\nu_K \gtrsim \min\{\nu_{K-1}, \varepsilon\} \gtrsim \varepsilon$, so the bound on $\#\operatorname{supp}\mathbf{w}$ follows.

By Proposition 6.3.12, and Lemma 3.3.3 on page 53, the cost of the i-th iteration can be bounded by an absolute multiple of

$$\nu_i^{-1/s}(|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + c_{\mathbf{f}}) + (\nu_{i-1}/\nu_i)^{1/s}(\#\Lambda_i + 1) + \#\tilde{\Lambda}_i + \#\operatorname{supp}\mathbf{r}_i + 1 \lesssim \nu_i^{-1/s}(|\mathbf{u}|_{\mathcal{A}^s}^{1/s} + c_{\mathbf{f}}) + (\nu_{i-1}/\nu_i)^{1/s}(\#\Lambda_i + 1).$$

We have $1 \lesssim \nu_{i-1}^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}$, and

$$\#\Lambda_i = \#\tilde{\Lambda}_{i-1} + \#\Lambda_i - \#\tilde{\Lambda}_{i-1} \lesssim \nu_{i-1}^{-1/s}(c_{\mathbf{f}} + |\mathbf{u}|_{\mathcal{A}^s}^{1/s}) + \nu_{i-1}^{-1/s}|\mathbf{u}|_{\mathcal{A}^s}^{1/s}.$$

Taking into account these bounds, by geometric decrease of ν_i inside the loop and $\nu_K \gtrsim \varepsilon$, we complete the proof. ∎

6.4 Elliptic boundary value problems

In this section, we will verify Assumptions 6.3.3 and 6.3.14 for the case of second order elliptic boundary value problems.

6.4.1 The wavelet setting

Let $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain and let Ψ be a Riesz basis for $H := H_0^1(\Omega)$ of wavelet type. Let $\langle\cdot,\cdot\rangle_*$ be an inner product on $L_2(\Omega)$ such that

$$\langle v, w\rangle_* \lesssim \|v\|_{L_2(\operatorname{supp} w)}\|w\|_{L_2} \quad \text{for } v, w \in L_2(\Omega).$$

We embed $L_2(\Omega)$ into H' by using this inner product: $g \in L_2(\Omega)$ is identified with the functional $\langle g,\cdot\rangle_*$ in H'. We assume that the dual basis $\tilde{\Psi}$ (cf. §2.2) of Ψ is in $L_2(\Omega)$. Moreover, we assume that the both bases are *local*, i.e., with $\Omega_\lambda := \operatorname{supp}\psi_\lambda$ and $\tilde{\Omega}_\lambda := \operatorname{supp}\tilde{\psi}_\lambda$,

$$\operatorname{diam}\Omega_\lambda, \operatorname{diam}\tilde{\Omega}_\lambda \lesssim 2^{-|\lambda|}, \qquad \lambda \in \nabla,$$

and

$$\sup_{x\in\Omega, j\in\mathbb{N}_0} \#\{|\lambda| = j : B(x, 2^{-j}) \cap \Omega_\lambda \neq \varnothing\} < \infty,$$

where $B(x,r)$ is the n-ball with radius $r > 0$ and centered at $x \in \mathbb{R}^n$. We also assume that Ω_λ contains a ball $B(x,r)$ with $r \gtrsim 2^{-|\lambda|}$ and that for $s \in \{0,1\}$,

$$\|\psi_\lambda\|_{H^s} \lesssim 2^{|\lambda|(s-1)}, \quad \text{and} \quad \|\tilde{\psi}_\lambda\|_{L_2} \lesssim 2^{|\lambda|}, \qquad \lambda \in \nabla. \tag{6.4.1}$$

For $j \in \mathbb{N}_0$, let $X_j := \operatorname{span}\{\psi_\lambda : \lambda \in \nabla_j\}$ and $\tilde{X}_j := \operatorname{span}\{\tilde{\psi}_\lambda : \lambda \in \nabla_j\}$ with $\nabla_j := \{\lambda \in \nabla : |\lambda| \leq j\}$. Then we assume the existence of biorthogonal bases, called *single scale* bases, $\Phi_j = \{\phi_{j,\lambda} : \lambda \in \nabla_j\}$ and $\tilde{\Phi}_j = \{\tilde{\phi}_{j,\lambda} : \lambda \in \nabla_j\}$ of X_j and $\tilde{X}_j$, respectively. Moreover, we assume that the bases are *local* in the sense that

$$\operatorname{diam}(\operatorname{supp}\phi_{j,\lambda}), \operatorname{diam}(\operatorname{supp}\tilde{\phi}_{j,\lambda}) \lesssim 2^{-j}, \qquad \lambda \in \nabla_j,\ j \in \mathbb{N}_0,$$

and

$$\sup_{x\in\Omega, j\in\mathbb{N}_0} \#\{\lambda \in \nabla_j : B(x, 2^{-j}) \cap \operatorname{supp}\phi_{j,\lambda} \neq \varnothing\} < \infty.$$

We also assume that for $s \in \{0, 1\}$,

$$\|\phi_{j,\lambda}\|_{H^s} \lesssim 2^{j(s-1)}, \quad \text{and} \quad \|\tilde{\phi}_{j,\lambda}\|_{L_2} \lesssim 2^j, \qquad \lambda \in \nabla_j,\ j \in \mathbb{N}_0. \tag{6.4.2}$$

It is obvious that $X_j \subset X_{j+1}$ for $j \in \mathbb{N}_0$. In addition, we assume that there exists a subspace $\Pi \subseteq X_0$ such that for any non-degenerate star-shaped domain $D \subseteq \Omega$,

$$\inf_{q\in\Pi} \|v - q\|_{L_2(D)} \lesssim (\operatorname{diam} D)\|v\|_{H^1_{0,\partial D\cap\partial\Omega}(D)} \qquad v \in H^1(D). \tag{6.4.3}$$

Remark 6.4.1. An example of wavelets satisfying all these assumptions is locally supported, piecewise polynomial biorthogonal wavelets on finite element meshes, from [84]. Another example is given by wavelets constructed via domain decomposition into smooth parametric images of cubes and tensor products of locally supported biorthogonal spline wavelets on interval, e.g. from [14, 33, 55, 56, 85]. Note that the condition (6.4.3) is satisfied when the space $\Pi \subseteq X_0$ contains all polynomials up to first order or piecewise smooth parametric images of all such polynomials.

In the following, we introduce a notion of mesh for spaces spanned by wavelets. For any given finite index set $\Lambda \subset \nabla$, let $X_\Lambda := \operatorname{span}\{\psi_\lambda : \lambda \in \Lambda\}$, and let $\mathcal{D}_\Lambda$ be a subdivision of Ω such that $\cup_{D\in\mathcal{D}_\Lambda} \overline{D} = \overline{\Omega}$, $D \cap D' = \varnothing$ for $D, D' \in \mathcal{D}_\Lambda$ with $D \neq D'$, and such that for $D \in \mathcal{D}_\Lambda$, ∂D is a piecewise smooth manifold, and

$$X_\Lambda|_{\overline{D}} \subset C^1(\overline{D}).$$

We assume that for finite subsets $\Lambda \subseteq \tilde{\Lambda} \subset \nabla$, and for $D \in \mathcal{D}_\Lambda$ and $\tilde{D} \in \mathcal{D}_{\tilde{\Lambda}}$, it holds that either $D \supseteq \tilde{D}$ or $D \cap \tilde{D} = \varnothing$. Moreover, we assume that the domains $D \in \mathcal{D}_\Lambda$ are uniformly Lipschitz and that

$$\operatorname{diam} D \lesssim 2^{-j_\Lambda(D)} \quad \text{and} \quad \operatorname{vol} D \gtrsim 2^{-nj_\Lambda(D)} \qquad D \in \mathcal{D}_\Lambda,$$

where $j_\Lambda : \mathcal{D}_\Lambda \to \mathbb{N}_0$ is defined by

$$j_\Lambda(D) = \max\{|\lambda| : \lambda \in \Lambda,\ \mathrm{vol}(D \cap \Omega_\lambda) > 0\} \qquad D \in \mathcal{D}_\Lambda.$$

We define the set $\mathcal{F}_\Lambda$ by collecting the interiors of all nonempty intersections $\partial D \cap \partial D'$ with dimension $n-1$ for all $D, D' \in \mathcal{D}_\Lambda \cup \{\mathbb{R}^n \setminus \overline{\Omega}\}$ with $D \neq D'$. We assume that $F \in \mathcal{F}_\Lambda$ is simply connected. Then one can verify that for finite subsets $\Lambda \subseteq \tilde{\Lambda} \subset \nabla$, and for $F \in \mathcal{F}_\Lambda$ and $\tilde{F} \in \mathcal{F}_{\tilde{\Lambda}}$, either $F \supseteq \tilde{F}$ or $F \cap \tilde{F} = \varnothing$. For $F \in \mathcal{F}_\Lambda$, we set $\mathcal{D}_\Lambda(F) := \{D \in \mathcal{D}_\Lambda : F \cap \overline{D} \neq \varnothing\}$. It is obvious that $\#\mathcal{D}_\Lambda(F) \leq 2$ for any $F \in \mathcal{F}_\Lambda$, and that $\operatorname{diam} F \lesssim \operatorname{diam} D$ for $D \in \mathcal{D}_\Lambda(F)$. Then we assume that each $F \in \mathcal{F}_\Lambda$ can be extended to the boundary $\partial\Omega_F \supset F$ of a uniformly Lipschitz domain Ω_F such that for some $\tilde{\nu} \in C^\infty(\overline{\Omega}_F, \mathbb{R}^n)$ with a uniformly bounded $\|\tilde{\nu}\|_{C^1}$ and for a uniformly bounded $\delta > 0$,

$$\tilde{\nu} \cdot \nu \geq \delta^{-1} \qquad \textit{a.e. on } \partial\Omega_F, \tag{6.4.4}$$

where ν is the unit outward normal of $\partial\Omega_F$ and $\tilde{\nu}\cdot\nu$ is the canonical scalar product in $\mathbb{R}^n$, and that

$$\operatorname{diam} F \eqsim 2^{-j_\Lambda(F)} \qquad F \in \mathcal{F}_\Lambda,$$

where $j_\Lambda : \mathcal{F}_\Lambda \to \mathbb{N}_0$ is defined by

$$j_\Lambda(F) = \max\{|\lambda| : \lambda \in \Lambda,\ \mathrm{vol}_{n-1}(F \cap \operatorname{int}\Omega_\lambda) > 0\} \qquad F \in \mathcal{F}_\Lambda.$$

Note that $j_\Lambda(F) \leq \max_{D\in\mathcal{D}_\Lambda(F)} j_\Lambda(D)$, since if F intersects with Ω_λ then the union $\cup_{D\in\mathcal{D}_\Lambda(F)} D$ also intersects with Ω_λ.

Furthermore, we assume that there exists a constant $N \in \mathbb{N}$ such that if $\Lambda \in \tilde{\mathcal{T}}$ is a graded tree and $\mu \in \Lambda$ is any of its elements, then, for $0 \leq j \leq |\mu| - N$,

$$\{\lambda \in \nabla_j : \mathrm{vol}(\Omega_\mu \cap \Omega_\lambda) > 0 \text{ or } \mathrm{vol}(\Omega_\mu \cap \tilde{\Omega}_\lambda) > 0\} \subset \Lambda. \tag{6.4.5}$$

In particular, this implies that for any $\lambda \in \mathcal{L}(\Lambda)$, there is no $\mu \in \Lambda$ with $\mathrm{vol}(\Omega_\mu \cap \Omega_\lambda) > 0$ and $|\mu| \geq |\lambda| + N$. In addition, we assume that for graded trees $\Lambda \in \tilde{\mathcal{T}}$, and for domains Ξ such that $\overline{\Xi} = \cup_{D\in\mathcal{D}} \overline{D}$ with $\mathcal{D} \subseteq \mathcal{D}_\Lambda$,

$$\|w\|_{H^1(\Xi)} \lesssim \left\{\min_{D\in\mathcal{D}} (\operatorname{diam} D)\right\}^{-1} \|w\|_{L_2(\Xi)} \qquad w \in X_\Lambda. \tag{6.4.6}$$

Example 6.4.2. Assuming that $\Omega \subset \mathbb{R}^n$ is a polyhedron, let $\mathcal{D}_0$ be a conforming subdivision of Ω into n-simplexes, and for $j \in \mathbb{N}$, let $\mathcal{D}_j$ be a dyadic refinement of $\mathcal{D}_{j-1}$. We define the finite element spaces by

$$X_j = \{v \in C(\overline{\Omega}) \cap H : v|_D \in P_{d-1} \text{ for } D \in \mathcal{D}_j\}, \qquad j \in \mathbb{N}_0,$$

where P_{d-1} is the space of polynomials with degree less than d. Let Φ_j be the standard nodal basis of X_j. Then there exist locally supported wavelet bases $\Psi = \{\psi_\lambda : \lambda \in \nabla\}$ and $\tilde{\Psi} = \{\tilde{\psi}_\lambda : \lambda \in \nabla\}$ of H such that $\langle \Psi, \tilde{\Psi} \rangle_{L_2} = \mathbf{I}$, and that for any $j \in \mathbb{N}_0$ there exists $\nabla_j \subset \nabla_{j+1} \subset \nabla$ such that $\{\psi_\lambda : \lambda \in \nabla_j\}$ is a basis of X_j, cf. [84]. Moreover, there exists a locally supported single scale basis $\tilde{\Phi}_j$ of $\tilde{X}_j = \operatorname{span}\{\tilde{\psi}_\lambda : \lambda \in \nabla_j\}$ such that $\langle \Phi_j, \tilde{\Phi}_j \rangle_{L_2} = \mathbf{I}$. The level number for an index $\lambda \in \nabla$ is given by $|\lambda| = \min\{j \in \mathbb{N}_0 : \lambda \in \nabla_j\}$. For a given finite subset $\Lambda \subset \nabla$ the subdivision $\mathcal{D}_\Lambda$ can be defined by the following process.

- Set $\mathcal{D}_\Lambda := \mathcal{D}_0$;
- For $j = 1, \ldots,$ and for $D \in \mathcal{D}_\Lambda$, if there is $\lambda \in \Lambda \cap \nabla_j$ such that $\operatorname{vol}(D \cap \operatorname{supp} \psi_\lambda) > 0$, then replace D in $\mathcal{D}_\Lambda$ by the union of all $D' \in \mathcal{D}_j$ that constitute $\operatorname{supp} \psi_\lambda$.

It is reasonable to assume the existence of a parent-child relation on ∇ such that if $\lambda \in \nabla$ is a child of $\mu \in \nabla$, then $|\lambda| = |\mu| + 1$ and $\operatorname{supp} \psi_\lambda \subset \operatorname{supp} \psi_\mu$. With the root ∇_0 and this parent-child relation we have a notion of *tree* structure on the subsets of ∇. Note that for any finite tree Λ, $\#\mathcal{D}_\Lambda \lesssim \#\partial\Lambda \lesssim \#\Lambda$. We call a tree $\Lambda \supseteq \nabla_0$ satisfying (6.4.5) a *graded tree*. Note that while (6.4.5) is a condition that should be satisfied for graded trees in the abstract setting, we use (6.4.5) to define the notion of graded tree itself in the context of this example. For any tree $\Lambda' \supseteq \nabla_0$, one can get a graded tree by applying the following algorithm iteratively for all $\mu \in \Lambda' \setminus \nabla_0$, starting off with $\Lambda = \nabla_0$.

Algorithm 6.4.3 Graded tree node insertion $\mathbf{APPEND}[\Lambda, \mu] \to \Lambda$

Input: Λ is a graded tree and $\mu \in \mathcal{L}(\Lambda)$.
Output: Λ is a graded tree with $\mu \in \Lambda$.

```
if |μ| < N then
    Terminate the subroutine.
end if
for all λ ∈ ∇_{|μ|−N} \ Λ such that vol(Ω_λ, Ω_μ) > 0 or vol(Ω̃_λ ∩ Ω_μ) > 0 do
    Λ := APPEND[Λ, λ];
end for
Λ := Λ ∪ {μ}.
```

With $\mu' \in \Lambda$ being the parent of μ, we have $\Omega_\mu \subset \Omega_{\mu'}$, therefore the condition (6.4.5) may be violated only for $j = |\mu| - N$. Each recursive call of **APPEND** is called with $\lambda \in \mathcal{L}(\Lambda)$, because the parent λ' of λ satisfies $\operatorname{vol}(\Omega_{\lambda'} \cap \Omega_{\mu'}) > 0$ and $|\lambda'| = |\mu'| - N$, meaning that $\lambda' \in \Lambda$. Since the number of iterations in the **for all** loop is uniformly bounded and the value of $|\lambda|$ is reduced by $N > 0$ in

each recursive call, the algorithm terminates in a finite time. By construction, the output tree Λ fulfils (6.4.5).

By using the result from Section 6.5, which is *independent* of any section in this chapter, we will now verify the condition (6.2.4) on page 88 and Assumption 6.3.14 on page 100 in the setting of this example. The functions $d(\lambda) = \operatorname{diam} \Omega_\lambda$ and $d(\lambda, \mu) = \operatorname{dist}(\Omega_\lambda, \Omega_\mu)$, $\lambda, \mu \in \nabla$, satisfy the conditions (i)-(iv) from Section 6.5, with $\chi = 1$, and the map defined by $\mathcal{R}(\Lambda, \mu) = \mathbf{APPEND}[\Lambda, \mu] \setminus \Lambda$ satisfies (6.5.1) on page 115, with $L_{\mathcal{R}} = 0$. Now Theorem 6.5.5 on page 116 implies that the above notion of graded tree complies with (6.2.4). Furthermore, the abstract subroutine **COMPLETE** that was described in Algorithm 6.3.10 on page 96 can be realized by employing the subroutine **APPEND**, and then Theorem 6.5.5 verifies Assumption 6.3.14. ⊘

In the rest of this subsection, we will prove two preliminary lemmata.

Lemma 6.4.4. *Let $\Lambda \in \tilde{\mathcal{T}}$ be a graded tree. Then, the conditions $D, D' \in \mathcal{D}_\Lambda$ and $\operatorname{dist}(D, D') \lesssim \operatorname{diam} D$ imply that $\operatorname{diam} D' \eqsim \operatorname{diam} D$ and so $\operatorname{vol} D' \eqsim \operatorname{vol} D$.*

Proof. Recall that for $\lambda \in \nabla$, the support Ω_λ contains a ball $B(x, r)$ with radius $r \geq C2^{-|\lambda|}$ with an absolute constant $C > 0$. In view of (6.4.5), if $\operatorname{dist}(D, D') \leq C2^{-\ell}$ for $\ell \leq j_\Lambda(D') - N$, then we have $j_\Lambda(D) \geq \ell$. So $\operatorname{dist}(D, D') \leq C2^{N+K}2^{-j_\Lambda(D')}$ with a constant $K \geq 0$ implies $j_\Lambda(D) \geq j_\Lambda(D') - N - K$, that is, $\operatorname{diam} D \lesssim 2^K \operatorname{diam} D'$.

On the other hand, if $\operatorname{dist}(D, D') \leq C2^{-\ell}$ for $\ell \leq j_\Lambda(D) - N$, then we have $j_\Lambda(D') \geq \ell$. This implies $\operatorname{diam} D' \lesssim 2^K \operatorname{diam} D$, and the rest of the proof is straightforward. ∎

The following lemma shows the existence of a mapping that realizes a quasi-optimal local polynomial approximation. The proof is inspired by the proof of [35, Lemma 3.3], and it exploits the gradedness of the index trees and the locality of the dual wavelets. For a different approach that makes use of special properties of splines, see [9].

Lemma 6.4.5. *In the above setting, for any graded tree $\Lambda \in \tilde{\mathcal{T}}$, there exists a mapping $Q_\Lambda : L_2(\Omega) \to X_\Lambda$ such that for $D \in \mathcal{D}_\Lambda$,*

$$\|v - Q_\Lambda v\|_{L_2(D)} \lesssim \inf_{p \in \Pi} \|v - p\|_{L_2(D^*)}.$$

with an n-ball $D^* \supset D$ satisfying $\operatorname{diam} D^* \lesssim \operatorname{diam} D$, and for $F \in \mathcal{F}_\Lambda$,

$$\|v - Q_\Lambda v\|_{L_2(F)} \lesssim (\operatorname{diam} F)^{-\frac{1}{2}} \inf_{p \in \Pi} \|v - p\|_{L_2(F^*)},$$

with an n-ball $F^* \supset F$ satisfying $\operatorname{diam} F^* \lesssim \operatorname{diam} F$.

Moreover, for any tree $\bar\Lambda \supset \Lambda$ and $v \in X_{\bar\Lambda}$, with

$$\mathcal{D}_{\Lambda,\bar\Lambda} := \{D \in \mathcal{D}_\Lambda : \operatorname{vol}(D \cap \Omega_\lambda) > 0 \text{ for some } \lambda \in \bar\Lambda \setminus \Lambda\},$$

we have $(v - Q_\Lambda v)|_D = 0$ when $D \notin \mathcal{D}_{\Lambda,\bar\Lambda}$, and with

$$\mathcal{F}_{\Lambda,\bar\Lambda} := \{F \in \mathcal{F}_\Lambda : \operatorname{vol}_{n-1}(F \cap \operatorname{int}\Omega_\lambda) > 0 \text{ for some } \lambda \in \bar\Lambda \setminus \Lambda\},$$

we have $(v - Q_\Lambda v)|_F = 0$ when $F \notin \mathcal{F}_{\Lambda,\bar\Lambda}$.

Proof. Let $Q_\Lambda v := \sum_{\lambda\in\Lambda}\langle v, \tilde\psi_\lambda\rangle_*\psi_\lambda$ and let $Q_j := Q_{\nabla_j}$ for $j \in \mathbb{N}_0$. Then the last statement of the lemma is trivially true. With $j := j_\Lambda(D)$, we have

$$(v - Q_\Lambda v)|_D = (v - Q_j v)|_D + \sum_{\lambda\in\Lambda^-(D)} \langle v, \tilde\psi_\lambda\rangle_*\psi_\lambda, \tag{6.4.7}$$

with

$$\Lambda^-(D) := \{\lambda \in \nabla \setminus \Lambda : |\lambda| \le j_\Lambda(D),\ \operatorname{vol}(D \cap \Omega_\lambda) > 0\}.$$

The condition (6.4.5) immediately implies that $\#\Lambda^-(D) \lesssim 1$ and $j_\Lambda(D) - |\lambda| \lesssim 1$ for $\lambda \in \Lambda^-(D)$. Now we will estimate the L_2-norms of the two terms in the right hand side separately. For the last term we have

$$\begin{aligned}|\langle v, \tilde\psi_\lambda\rangle_*| \cdot \|\psi_\lambda\|_{L_2(D)} &= |\langle v - p, \tilde\psi_\lambda\rangle_*| \cdot \|\psi_\lambda\|_{L_2(D)} \\ &\lesssim \|v - p\|_{L_2(\tilde\Omega_\lambda)}\|\tilde\psi_\lambda\|_{L_2}\|\psi_\lambda\|_{L_2} \lesssim \|v - p\|_{L_2(\tilde\Omega_\lambda)},\end{aligned}$$

which, together with the condition on $\Lambda^-(D)$, implies that

$$\left\|\sum_{\lambda\in\Lambda^-(D)} \langle v, \tilde\psi_\lambda\rangle_*\psi_\lambda\right\|_{L_2(D)} \lesssim \sum_{\lambda\in\Lambda^-(D)} \|v - p\|_{L_2(\tilde\Omega_\lambda)} \lesssim \|v - p\|_{L_2(D^*)},$$

where we assumed that

$$\bigcup_{\lambda\in\Lambda^-(D)} \tilde\Omega_\lambda \subseteq D^*. \tag{6.4.8}$$

For the first term in the right hand side of (6.4.7), we have

$$\|v - Q_j v\|_{L_2(D)} \le \|v - p\|_{L_2(D)} + \|Q_j v - p\|_{L_2(D)}.$$

Using the single scale basis, we get

$$\|Q_j v - p\|_{L_2(D)} = \left\| \sum_{\lambda \in \Lambda^\circ(D)} \langle v - p, \tilde{\phi}_{j,\lambda} \rangle_* \phi_{j,\lambda} \right\|_{L_2(D)} \leq \sum_{\lambda \in \Lambda^\circ(D)} |\langle v - p, \tilde{\phi}_{j,\lambda} \rangle_*| \cdot \|\phi_{j,\lambda}\|_{L_2(D)} \lesssim \|v - p\|_{L_2(D^*)}.$$

with

$$\Lambda^\circ(D) := \{\lambda \in \Lambda \cap \nabla_j,\ D \cap \operatorname{supp} \phi_{j,\lambda} \neq \varnothing\},$$

and with the assumption

$$\bigcup_{\lambda \in \Lambda^\circ(D)} \operatorname{supp} \tilde{\phi}_{j,\lambda} \subseteq D^*. \tag{6.4.9}$$

By the locality of ψ_λ and $\phi_{j,\lambda}$, and the properties of $\Lambda^-(D)$, we conclude that there is a ball D^* satisfying (6.4.8), (6.4.9) and $\operatorname{diam} D^* \lesssim \operatorname{diam} D$.

Now we will prove the second part of the lemma. With $j := j_\Lambda(F)$, similarly to the previous case, we have

$$(v - Q_\Lambda v)|_F = (v - Q_j v)|_F + \sum_{\lambda \in \Lambda^-(F)} \langle v, \tilde{\psi}_\lambda \rangle_* \psi_\lambda, \tag{6.4.10}$$

where $\Lambda^-(F) := \{\lambda \in \nabla \setminus \Lambda : |\lambda| \leq j_\Lambda(F),\ F \cap \operatorname{int} \Omega_\lambda \neq \varnothing\}$, with $\operatorname{int} \Omega_\lambda$ denoting the interior of Ω_λ. Since $\Lambda^-(F) \subseteq \cup_{D \in \mathcal{D}_\Lambda(F)} \Lambda^-(D)$, it holds that $\#\Lambda^-(F) \lesssim 1$ and that $j_\Lambda(F) - |\lambda| \lesssim 1$ for $\lambda \in \Lambda^-(F)$. The rest of the proof is completely analogous to the previous case except we use [50, Theorem 1.5.1.10] with the help of the assumption (6.4.4) to estimate L_2-norms on F. For instance, for the second term in the right hand side of (6.4.10) we have

$$\begin{aligned} |\langle v, \tilde{\psi}_\lambda \rangle_*| \cdot \|\psi_\lambda\|_{L_2(F)} &= |\langle v - p, \tilde{\psi}_\lambda \rangle_*| \|\psi_\lambda\|_{L_2(F)} \leq |\langle v - p, \tilde{\psi}_\lambda \rangle_*| \|\psi_\lambda\|_{L_2(\partial\Omega_F)} \\ &\lesssim \|v - p\|_{L_2(\tilde{\Omega}_\lambda)} \|\tilde{\psi}_\lambda\|_{L_2(\tilde{\Omega}_\lambda)} \|\psi_\lambda\|^{1/2}_{L_2(\Omega_F)} \|\psi_\lambda\|^{1/2}_{H^1(\Omega_F)} \\ &\lesssim 2^{|\lambda|/2} \|v - p\|_{L_2(\tilde{\Omega}_\lambda)}. \end{aligned}$$

■

6.4.2 Differential operators

Let

$$a(v, w) := \int_\Omega \left(\sum_{j,k=1}^n a_{jk} \partial_k v \partial_j w + \sum_{j=1}^n b_j \partial_j v w + c v w \right), \tag{6.4.11}$$

be bounded and coercive bilinear form on $H \times H$, i.e., it satisfies

$$a(v,w) \lesssim \|v\|_H \|w\|_H \quad \text{and} \quad a(v,v) \gtrsim \|v\|_H \qquad v,w \in H.$$

Then the operator $A : H \to H'$ defined by

$$\langle Av, w\rangle := a(v,w) \qquad \text{for } v,w \in H,$$

is bounded and H-elliptic.

We assume that $a_{jk}|_D \in H^1(D)$ for any $D \in \mathcal{D}_{\nabla_0}$ and $b_j, c \in L_2(\Omega)$.

6.4.3 Verification of Assumption 6.3.3

Let $X \subseteq H$ be a linear subspace, $g \in L_2(\Omega)$, and let $v \in X$ be the solution of the Galerkin problem

$$a(v,w) = \langle g, w\rangle_{L_2} \qquad w \in X. \tag{6.4.12}$$

Note that taking $X = H$ yields $v = A^{-1}g$. We have for $\tilde{v} \in X_\Lambda$ and $w \in X$,

$$\begin{aligned} a(v-\tilde{v},w) &= \int_\Omega \Big(gw - \textstyle\sum_{j,k} a_{jk}\partial_k\tilde{v}\partial_j w - \sum_j b_j \partial_j \tilde{v} w - c\tilde{v}w \Big) \\ &= \sum_{D\in\mathcal{D}_\Lambda} \Big\{ \int_D \Big(gw + \textstyle\sum_{j,k} \partial_j a_{jk}\partial_k\tilde{v}w - \sum_j b_j\partial_j\tilde{v}w - c\tilde{v}w \Big) \\ &\quad - \int_{\partial D} \textstyle\sum_{j,k} \nu_j a_{jk}\partial_k\tilde{v}w \Big\} \\ &=: \sum_{D\in\mathcal{D}_\Lambda} \int_D R_D(g,\tilde{v})w + \sum_{F\in\mathcal{F}_\Lambda} \int_F R_F(\tilde{v})w, \end{aligned} \tag{6.4.13}$$

where ν_j is the j-th component of the outward unit normal of ∂D. We have

$$R_F(\tilde{v}) = \sum_{j,k} \nu_j(F) \left\{ (a_{jk}\partial_k\tilde{v})_+ - (a_{jk}\partial_k\tilde{v})_- \right\},$$

where $\nu(F)$ with the components $\nu_j(F)$ is a unit normal of F, and $(\cdot)_\pm$ refers to the value in the positive (or negative) side of F with respect to $\nu(F)$. Note that $R_F(\tilde{v})$ does not depend on the orientation of $\nu(F)$. From the conditions on the coefficients and because F is piecewise smooth, we infer that $R_D(g,\tilde{v}) \in L_2(D)$ and $R_F(\tilde{v}) \in L_2(F)$.

For any graded tree $\Lambda \in \tilde{\mathcal{T}}$, a tree $\bar{\Lambda} \supset \Lambda$, and functions $g \in L_2(\Omega)$ and

$\tilde{v} \in X_\Lambda$, we define an error estimator by

$$E_{\Lambda,\bar{\Lambda}}(g,\tilde{v}) := \left\{ \sum_{D\in\mathcal{D}_{\Lambda,\bar{\Lambda}}} (\operatorname{diam} D)^2 \|R_D(g,\tilde{v})\|^2_{L_2(D)} + \sum_{F\in\mathcal{F}_{\Lambda,\bar{\Lambda}}} (\operatorname{diam} F)\|R_F(\tilde{v})\|^2_{L_2(F)} \right\}^{\frac{1}{2}}, \tag{6.4.14}$$

where $\mathcal{D}_{\Lambda,\bar{\Lambda}}$ and $\mathcal{F}_{\Lambda,\bar{\Lambda}}$ are as in Lemma 6.4.5 on page 105.

The following result shows that $E_{\Lambda,\bar{\Lambda}}(g, v_\Lambda)$ is an upper bound on the difference between the Galerkin solutions on X_Λ and on $X_{\bar{\Lambda}}$. Given the result of Lemma 6.4.5, the proof follows the standard techniques, cf. [89], but we include it here for the reader's convenience.

Theorem 6.4.6. *Let $\Lambda \in \tilde{\mathcal{T}}$, $g \in L_2(\Omega)$, and let $v_\Lambda \in X_\Lambda$ and $v_{\bar{\Lambda}} \in X_{\bar{\Lambda}}$ be the solutions of the Galerkin problem* (6.4.12) *with $X = X_\Lambda$ and $X = X_{\bar{\Lambda}}$, respectively. Then we have*

$$\|v_{\bar{\Lambda}} - v_\Lambda\|_{H^1} \lesssim E_{\Lambda,\bar{\Lambda}}(g, v_\Lambda).$$

Proof. Since $a(v_{\bar{\Lambda}}, w) = \langle g, w\rangle_{L_2}$ for $w \in X_{\bar{\Lambda}}$, we have $a(v_{\bar{\Lambda}} - v_\Lambda, w) = 0$ for $w \in X_\Lambda$. Using this, the definition (6.4.13), and applying the Cauchy-Bunyakovsky-Schwarz (CBS) inequality, Lemma 6.4.5 and (6.4.3) on page 102, and again the CBS inequality, for $w \in X_{\bar{\Lambda}}$, we infer

$$\begin{aligned}
a(v_{\bar{\Lambda}} - v_\Lambda, w) &= a(v_{\bar{\Lambda}} - v_\Lambda, w - Q_\Lambda w) \\
&= \sum_{D\in\mathcal{D}_\Lambda} \int_D R_D(g,\tilde{v})(w - Q_\Lambda w) + \sum_{F\in\mathcal{F}_\Lambda} \int_F R_F(\tilde{v})(w - Q_\Lambda w) \\
&\le \sum_{D\in\mathcal{D}_{\Lambda,\bar{\Lambda}}} \|R_D(g,\tilde{v})\|_{L_2(D)} \|w - Q_\Lambda w\|_{L_2(D)} \\
&\quad + \sum_{F\in\mathcal{F}_{\Lambda,\bar{\Lambda}}} \|R_F(\tilde{v})\|_{L_2(F)} \|w - Q_\Lambda w\|_{L_2(F)} \\
&\lesssim \sum_{D\in\mathcal{D}_{\Lambda,\bar{\Lambda}}} \|R_D(g,\tilde{v})\|_{L_2(D)} (\operatorname{diam} D) \|w\|_{H^1(D^*)} \\
&\quad + \sum_{F\in\mathcal{F}_{\Lambda,\bar{\Lambda}}} \|R_F(\tilde{v})\|_{L_2(F)} (\operatorname{diam} F)^{\frac{1}{2}} \|w\|_{H^1(F^*)}
\end{aligned}$$

$$\lesssim \Bigg\{ \sum_{D\in\mathcal{D}_{\Lambda,\bar\Lambda}} (\operatorname{diam} D)^2 \|R_D(g,\tilde v)\|^2_{L_2(D)} + \sum_{F\in\mathcal{F}_{\Lambda,\bar\Lambda}} (\operatorname{diam} F)\|R_F(\tilde v)\|^2_{L_2(F)} \Bigg\}^{\frac12} \|w\|_{H^1(\Omega)}.$$

Now using that $\|\cdot\|_{H^1} \lesssim \sup_{w\in H} \frac{a(\cdot,w)}{\|w\|_{H^1}}$, we finish the proof. ∎

The following result shows that $E_{\Lambda,\bar\Lambda}(g,v_\Lambda)$ is also a lower bound on the difference between two Galerkin solutions. Although the proof follows the standard techniques, cf. [62, 89], we include it here for the reader's convenience since the setting here is somewhat different than the usual finite element setting.

Theorem 6.4.7. *Let $\Lambda,\Lambda^\star\in\tilde{\mathcal{T}}$ and $\bar\Lambda\in\mathcal{T}$ be such that $\Lambda\subset\Lambda^\star\subset\bar\Lambda$ and that*

$$\min_{\{D^\star\in\mathcal{D}_{\Lambda^\star}:D^\star\subseteq D\}} \operatorname{diam} D^\star \gtrsim \operatorname{diam} D \qquad D\in\mathcal{D}_{\Lambda,\bar\Lambda}.$$

For $D\in\mathcal{D}_{\Lambda,\bar\Lambda}$, let $\Pi(D)\subseteq L_2(D)$, and let $\vartheta_D:\Pi(D)\to X_{\Lambda^\star}$ be uniformly bounded in the standard metric on $L_2(D)\to L_2(D)$ and such that for $p\in\Pi(D)$

$$\operatorname{supp}\vartheta_D p\subseteq\overline{D} \qquad \text{and} \qquad \|p\|^2_{L_2(D)}\lesssim\int_D p\vartheta_D p. \tag{6.4.15}$$

For $F\in\mathcal{F}_{\Lambda,\bar\Lambda}$, let $\Pi(F)\subseteq L_2(F)$, and let $\vartheta_F:\Pi(F)\to X_{\Lambda^\star}$ be uniformly bounded in the standard metric on $L_2(F)\to L_2(F)$ and such that for $p\in\Pi(F)$

$$\operatorname{supp}\vartheta_F p\subseteq\bigcup_{D\in\mathcal{D}_\Lambda(F)}\overline{D}, \qquad \|p\|^2_{L_2(F)}\lesssim\int_F p\vartheta_F p \tag{6.4.16}$$

and

$$\|\vartheta_F p\|_{L_2(D)}\lesssim(\operatorname{diam} F)^{\frac12}\|p\|_{L_2(F)}. \tag{6.4.17}$$

Moreover, let $g\in L_2(\Omega)$, and let $v_\Lambda\in X_\Lambda$ and $v_{\Lambda^\star}\in X_{\Lambda^\star}$ be the solutions to the Galerkin problem (6.4.12) with $X=X_\Lambda$ and $X=X_{\Lambda^\star}$, respectively. Then, there exists a function $\rho:\mathcal{T}\times L_2(\Omega)\to[0,\infty)$ such that

$$E_{\Lambda,\bar\Lambda}(g,v_\Lambda)\lesssim\|v_{\Lambda^\star}-v_\Lambda\|_{H^1}+\rho(\Lambda,g)$$

Proof. For $D\in\mathcal{D}_{\Lambda,\bar\Lambda}$, set $R_D=R_D(g,v_\Lambda)$ and let $\overline{R}_D\in\Pi(D)$. Then, with $w:=\vartheta_D\overline{R}_D\in X_{\Lambda^\star}$, using the second estimate in (6.4.15), taking into account the

definition (6.4.13) and the fact that $\operatorname{supp} w \subseteq \overline{D}$, and finally applying the CBS inequality and the inverse inequality (6.4.6), we have

$$\begin{aligned}\|\overline{R}_D\|_{L_2(D)}^2 &\lesssim \int_D \overline{R}_D w = a(v_{\Lambda^\star} - v_\Lambda, w) + \int_D (\overline{R}_D - R_D)w \\ &\lesssim \|v_{\Lambda^\star} - v_\Lambda\|_{H^1(D)}\|w\|_{H^1(D)} + \|\overline{R}_D - R_D\|_{L_2(D)}\|w\|_{L_2(D)} \\ &\lesssim \left\{(\operatorname{diam} D)^{-1}\|v_{\Lambda^\star} - v_\Lambda\|_{H^1(D)} + \|\overline{R}_D - R_D\|_{L_2(D)}\right\}\|w\|_{L_2(D)}.\end{aligned}$$

Now using the uniform boundedness of $\vartheta_D : L_2(D) \to L_2(D)$ and the triangle inequality, we infer

$$\|R_D\|_{L_2(D)} \lesssim (\operatorname{diam} D)^{-1}\|v_{\Lambda^\star} - v_\Lambda\|_{H^1(D)} + \|\overline{R}_D - R_D\|_{L_2(D)}. \tag{6.4.18}$$

For $F \in \mathcal{F}_{\Lambda,\bar{\Lambda}}$, let $\overline{R}_F \in \Pi(F)$ and set $w := \vartheta_F \overline{R}_F \in X_{\Lambda^\star}$ and $R_F = R_F(v_\Lambda)$. Then, similarly to the above, we get

$$\begin{aligned}\|\overline{R}_F\|_{L_2(F)}^2 &\lesssim \int_F \overline{R}_F w \\ &= a(v_{\Lambda^\star} - v_\Lambda, w) + \int_F (\overline{R}_F - R_F)w - \sum_{D\in\mathcal{D}_\Lambda(F)} \int_D R_D w \\ &\lesssim \sum_{D\in\mathcal{D}_\Lambda(F)} \|v_{\Lambda^\star} - v_\Lambda\|_{H^1(D)}\|w\|_{H^1(D)} + \|\overline{R}_F - R_F\|_{L_2(F)}\|w\|_{L_2(F)} \\ &\quad + \sum_{D\in\mathcal{D}_\Lambda(F)} \left\{(\operatorname{diam} D)^{-1}\|v_{\Lambda^\star} - v_\Lambda\|_{H^1(D)} + \|\overline{R}_D - R_D\|_{L_2(D)}\right\}\|w\|_{L_2(D)} \\ &\lesssim \sum_{D\in\mathcal{D}_\Lambda(F)} \left\{(\operatorname{diam} D)^{-1}\|v_{\Lambda^\star} - v_\Lambda\|_{H^1(D)} + \|\overline{R}_D - R_D\|_{L_2(D)}\right\}\|w\|_{L_2(D)} \\ &\quad + \|\overline{R}_F - R_F\|_{L_2(F)}\|w\|_{L_2(F)},\end{aligned}$$

where we have used (6.4.18) in the third line. By using (6.4.17), the uniform boundedness of $\vartheta_F : L_2(F) \to L_2(F)$, and the triangle inequality, we have

$$\begin{aligned}\|R_F\|_{L_2(F)} &\lesssim \|\overline{R}_F - R_F\|_{L_2(F)} \\ &\quad + \sum_{D\in\mathcal{D}_\Lambda(F)} \left\{(\operatorname{diam} D)^{-\frac{1}{2}}\|v_{\Lambda^\star} - v_\Lambda\|_{H^1(D)} + (\operatorname{diam} F)^{\frac{1}{2}}\|\overline{R}_D - R_D\|_{L_2(D)}\right\}.\end{aligned} \tag{6.4.19}$$

Whenever $F \in \mathcal{F}_{\Lambda,\bar{\Lambda}}$ and $D \in \mathcal{D}_\Lambda(F)$, we have $D \in \mathcal{D}_{\Lambda,\bar{\Lambda}}$. Then in view of the definition (6.4.14), the estimates (6.4.18) and (6.4.19) show that

$$\begin{aligned}\left[E_{\Lambda,\bar{\Lambda}}(g, v_\Lambda)\right]^2 &\lesssim \sum_{D\in\mathcal{D}_{\Lambda,\bar{\Lambda}}} \|v_{\Lambda^\star} - v_\Lambda\|_{H^1(D)}^2 \\ &\quad + \sum_{D\in\mathcal{D}_{\Lambda,\bar{\Lambda}}} (\operatorname{diam} D)^2\|\overline{R}_D - R_D\|_{L_2(D)}^2 + \sum_{F\in\mathcal{F}_{\Lambda,\bar{\Lambda}}} (\operatorname{diam} F)\|\overline{R}_F - R_F\|_{L_2(F)}^2.\end{aligned}$$

Now the proof is obtained with the function

$$\rho(\Lambda, g) := \inf_{\{\overline{R}_D \in \Pi(D),\, \overline{R}_F \in \Pi(F)\}} \left\{ \sum_{D \in \mathcal{D}_\Lambda} (\operatorname{diam} D)^2 \|\overline{R}_D - R_D\|^2_{L_2(D)} \right. \tag{6.4.20}$$

$$\left. + \sum_{F \in \mathcal{F}_\Lambda} (\operatorname{diam} F) \|\overline{R}_F - R_F\|^2_{L_2(F)} \right\}^{\frac{1}{2}}. \qquad \blacksquare$$

Corollary 6.4.8. *Let $\mathcal{V} : (\Lambda, \bar{\Lambda}) \mapsto \Lambda^\star \in \tilde{\mathcal{T}}$ be a mapping such that all the conditions of Theorem 6.4.7 are satisfied for any $\Lambda \in \tilde{\mathcal{T}}$, $\bar{\Lambda} \supset \Lambda$ a tree, and $\Lambda^\star := \mathcal{V}(\Lambda, \bar{\Lambda})$. Then, the condition* (6.3.3) *in Assumption 6.3.3 on page 92 is valid with $Y = \{\mathbf{g} \in \ell_2 : \mathbf{g}^T \tilde{\Psi} \in L_2(\Omega)\}$ and $\varrho(\Lambda, \mathbf{g}) \eqsim \rho(\Lambda, \mathbf{g}^T \tilde{\Psi})$, where $\rho(\cdot, \cdot)$ is as in Theorem 6.4.7.*

Example 6.4.9. Here we return to Example 6.4.2 of finite element wavelets. Let $\Lambda \subset \Lambda' \subseteq \Lambda^\star$ be graded trees such that each $D \in \mathcal{D}_{\Lambda,\bar{\Lambda}}$ contains in the interior a vertex from $\mathcal{D}_{\Lambda'}$ and each $F \in \mathcal{F}_{\Lambda,\bar{\Lambda}}$ contains in the interior a vertex from $\mathcal{F}_{\Lambda'}$. We denote these vertices by V_D and V_F, respectively, and for $D \in \mathcal{D}_\Lambda$ and $F \in \mathcal{F}_\Lambda$, define the bubble functions b_D and b_F such that

- both b_D and b_F are nonnegative and piecewise linear w.r.t. $\mathcal{D}_{\Lambda'}$,
- $b_D(V_D) = 1$ and $b_F(V_F) = 1$,
- $\operatorname{supp} b_D \subseteq \overline{D}$ and $\operatorname{supp} b_F \subseteq \cup_{\{D' \in \mathcal{D}_\Lambda : \overline{D'} \cap F \neq \varnothing\}} \overline{D'}$.

Then, we take $\Pi(D) := P_{d-2}(D)$ for $D \in \mathcal{D}_\Lambda$, and $\Pi(F) := P_{d-2}(F)$ for $F \in \mathcal{F}_\Lambda$, and define $\vartheta_D : \Pi(D) \ni p \mapsto b_D p$ and $\vartheta_F(p)$ for $p \in \Pi(F)$ by extending p constantly along a transversal to F and multiplying it with b_F. Here by a transversal to F we mean a vector whose angle with F is uniformly bounded away from 0. Now the maps ϑ_D and ϑ_F satisfy the conditions of Theorem 6.4.7, cf. [62], provided that for $D \in \mathcal{D}_{\Lambda,\bar{\Lambda}}$ and for $F \in \mathcal{F}_{\Lambda,\bar{\Lambda}}$, the space $X_{\Lambda^\star}$ contains $\vartheta_D \Pi(D)$ and $\vartheta_F \Pi(F)$, respectively.

In view of the above considerations, we introduce an algorithm for constructing $\Lambda^\star$ for given graded tree Λ and tree $\bar{\Lambda} \supset \Lambda$.

Algorithm 6.4.10 Realization of the mapping $\mathcal{V} : (\Lambda, \bar{\Lambda}) \mapsto \Lambda^\star$

Input: Let Λ be a graded tree, and let $\bar{\Lambda} \supset \Lambda$ be a tree.
Output: $\Lambda^\star \in \tilde{\mathcal{T}}$ with $\Lambda \subset \Lambda^\star \subset \bar{\Lambda}$.

1: $\Lambda^\star := \Lambda$;
2: **for all** $E \in \mathcal{D}_{\Lambda,\bar{\Lambda}} \cup \mathcal{F}_{\Lambda,\bar{\Lambda}}$ **do**
3: $\quad \Lambda' := \Lambda$;
4: $\quad$ Add to Λ' all necessary indices $\lambda \in \nabla \setminus \Lambda$, so that E contains a vertex V_E from $\mathcal{D}_{\Lambda'}$;
5: $\quad$ Construct the function b_E;
6: $\quad$ Add to $\Lambda^\star$ all indices $\lambda \in \nabla \setminus \Lambda^\star$ for which $\operatorname{supp} \tilde{\psi}_\lambda$ intersects with $\operatorname{sing\,supp} b_E$;
7: $\quad \Lambda^\star := \Lambda^\star \cup \Lambda'$;
8: **end for**
9: Complete $\Lambda^\star$ to a tree by iteratively adding the parents of the indices whose parent is not in $\Lambda^\star$;
10: Complete $\Lambda^\star$ to a graded tree by iteratively applying **APPEND** for the indices in $\Lambda^\star \setminus \Lambda$.

Note that the set of vertices of $\mathcal{D}_{\Lambda'}$ is the same as the set of vertices of $\mathcal{F}_{\Lambda'}$. By the condition (6.4.5) and the locality of the wavelets, the number of indices added to $\Lambda^\star$ in an iteration of the **for all** loop is uniformly bounded, meaning that with $\Lambda_1^\star$ denoting the value of $\Lambda^\star$ just after this loop, we have $\#\Lambda_1^\star - \#\Lambda \lesssim \#\mathcal{D}_{\Lambda,\bar{\Lambda}} + \#\mathcal{F}_{\Lambda,\bar{\Lambda}} \lesssim \#\mathcal{D}_{\Lambda,\bar{\Lambda}}$. Moreover, denoting by $\Lambda_2^\star$ the value of $\Lambda^\star$ just after the evaluation of the statement in Line 9, we have $\#\Lambda_2^\star - \#\Lambda \lesssim \#\Lambda_1^\star - \#\Lambda$ since the minimum level difference between any index from $\Lambda_1^\star$ and its ancestor from Λ is uniformly bounded. As noted earlier, the condition (6.4.5) implies that for any $\lambda \in \mathcal{L}(\Lambda)$, there is no $\mu \in \Lambda$ with $\operatorname{vol}(\Omega_\mu \cap \Omega_\lambda) > 0$ and $|\mu| \geq |\lambda| + N$. Since the minimum level difference between any index from $\Lambda_2^\star$ and its ancestor from Λ is uniformly bounded, each application of **APPEND** adds a uniformly bounded number of indices to $\Lambda^\star$, implying that $\#\Lambda^\star - \#\Lambda \lesssim \#\mathcal{D}_{\Lambda,\bar{\Lambda}}$. It is obvious that $\#\mathcal{D}_{\Lambda,\bar{\Lambda}} \lesssim \#\Lambda$. Moreover, we have $\mathcal{D}_{\Lambda,\bar{\Lambda}} = \{D \in \mathcal{D}_\Lambda : \operatorname{vol}(D \cap \Omega_\lambda) > 0 \text{ for some } \lambda \in \bar{\Lambda} \cap \mathcal{L}(\Lambda)\}$, and for $D \in \mathcal{D}_\Lambda$ and $\lambda \in \mathcal{L}(\Lambda)$ with $\operatorname{vol}(D \cap \Omega_\lambda) > 0$, we have $\operatorname{vol}(D) \gtrsim 2^{-n|\lambda|}$. We end this example by deducing that, for finite trees $\bar{\Lambda}$, $\mathcal{D}_{\Lambda,\bar{\Lambda}} \lesssim \#\left(\bar{\Lambda} \cap \mathcal{L}(\Lambda)\right) \lesssim \#(\bar{\Lambda} \setminus \Lambda)$. ⊘

Remark 6.4.11. For graded trees, one can perform a transformation into a local scaling function representation in linear time, cf. [35, §5.3]. So since Λ and $\tilde{\Lambda}$ are graded trees and $\#\tilde{\Lambda} \lesssim \#\Lambda$ in **TAPPLY** (Algorithm 6.3.8 on page 96), we can design a valid subroutine **TAPPLY** using these local transforms and the stiffness matrix in the local scaling function representation, which is sparse for

differential operators. We remark that in the truncated residuals approach, at least for differential operators the compressibility of the infinite stiffness matrix is not necessary.

6.5 Completion of tree

Let ∇ be a countable set, and let a parent-child relation be defined on ∇. Note that the main result of this section is *independent* of the results from the previous sections, so in particular, the set ∇ is an abstract set, not necessarily being the index set we considered in the previous sections. We assume that every element $\lambda \in \nabla$ has a uniformly bounded number of children, and has at most one parent. We say that $\lambda \in \nabla$ is a *descendant* of $\mu \in \nabla$ and write $\lambda \succ \mu$ if λ is a child of a descendant of *or* is a child of μ. The relations $\prec$ (ascendant of), $\succeq$ (descendant of or equal to), and $\preceq$ (ascendant of or equal to) are defined accordingly. The *level* or *generation* of an element $\lambda \in \nabla$, denoted by $|\lambda| \in \mathbb{N}_0$, is the number of its ascendants. Obviously, $\lambda \succ \mu$ implies $|\lambda| > |\mu|$. We call the set $\nabla_0 := \{\lambda \in \nabla : |\lambda| = 0\}$ the *root*, and assume that $\#\nabla_0 < \infty$.

A subset $\Lambda \subseteq \nabla$ is said to be a *tree* if with every member $\lambda \in \Lambda$ all its ascendants are included in Λ. For a tree Λ, those $\lambda \in \Lambda$ whose children are not contained in Λ are called *leaves* of Λ, and the set of all leaves of Λ is denoted by $\partial\Lambda$. Similarly, those $\lambda \notin \Lambda$ whose parent belongs to Λ is called *outer leaves* of Λ and the set of all outer leaves of Λ is denoted by $\mathcal{L}(\Lambda)$.

We assume that there are functions $d : \nabla \to \mathbb{R}$ and $d : \nabla \times \nabla \to \mathbb{R}$ satisfying the following conditions:

(i) For any $\lambda \in \nabla$, with some absolute constants $C_d, \chi \geq 0$, it holds that

$$0 < d(\lambda) \leq C_d 2^{-\chi|\lambda|};$$

(ii) For any $\lambda, \mu \in \nabla$, we have $d(\lambda, \mu) = d(\mu, \lambda) \geq 0$, and $d(\lambda, \mu) = 0$ if $\lambda \succeq \mu$;

(iii) For any $\lambda, \mu, \nu \in \nabla$, there holds a triangle inequality:

$$d(\lambda, \nu) \leq d(\lambda, \mu) + d(\mu) + d(\mu, \nu);$$

(iv) Let $L \in \mathbb{N}_0$ and $C > 0$ be arbitrary but fixed constants. Then for any fixed $\mu \in \nabla$, $\ell \in \mathbb{N}_0$ with $\ell \leq |\mu| + L$, there exists a uniformly bounded number of $\lambda \in \nabla$ with $d(\lambda, \mu) \leq C 2^{-\chi\ell}$.

Example 6.5.1. In the situation of Example 6.4.2 on page 103, let $d(\lambda) = \operatorname{diam} \Omega_\lambda$ and $d(\lambda, \mu) = \operatorname{dist}(\Omega_\lambda, \Omega_\mu)$, $\lambda, \mu \in \nabla$. Then these functions satisfy the above conditions (i)-(iv) with $\chi = 1$. ⊘

Example 6.5.2. Let $\Omega \subset \mathbb{R}^n$ be some polyhedral domain and let it be subdivided into finitely many pairwise disjoint n-simplices. We denote by ∇_0 the set of these n-simplices, and form the set ∇ by collecting all n-simplices created by a (possibly trivial) finite sequence of dyadic refinements of an initial simplex $\lambda \in \nabla_0$. The parent-child relation on ∇ is defined by saying that $\lambda \in \nabla$ is a child of $\mu \in \nabla$ if λ is created by one elementary dyadic refinement of μ. Then the functions $d(\cdot) := \operatorname{diam}(\cdot)$ and $d(\cdot,\cdot) := \operatorname{dist}(\cdot,\cdot)$ satisfy the above conditions with $\chi = 1$. ⊘

Let $\mathcal{T}$ denote the set of all finite trees, and let $\tilde{\mathcal{T}} \subseteq \mathcal{T}$ be a subset such that $\nabla_0 \in \tilde{\mathcal{T}}$. Then we introduce a map $\mathcal{R}$ that sends the pair of a tree $\Lambda \in \tilde{\mathcal{T}}$ and any of its outer leaves $\mu \in \mathcal{L}(\Lambda)$ to a set $\mu \in \mathcal{R}(\Lambda,\mu) \subset \nabla$ such that $\mathcal{R}(\Lambda,\mu) \cap \Lambda = \varnothing$, and $\mathcal{R}(\Lambda,\mu) \cup \Lambda$ is a tree in $\tilde{\mathcal{T}}$. We assume that for any $\lambda \in \mathcal{R}(\Lambda,\mu)$ it holds that

$$d(\lambda,\mu) \le C_{\mathcal{R}} 2^{-\chi|\lambda|}, \quad \text{and} \quad |\lambda| \le |\mu| + L_{\mathcal{R}}, \tag{6.5.1}$$

where $C_{\mathcal{R}} \in \mathbb{R}_+$ and $L_{\mathcal{R}} \in \mathbb{N}_0$ are constants.

Example 6.5.3. In the setting of Example 6.4.2 on page 103, let $\mathcal{R}(\Lambda,\mu) = \mathbf{APPEND}[\Lambda,\mu] \setminus \Lambda$. Then this map satisfies the above condition (6.5.1) with $\chi = 1$ and $L_{\mathcal{R}} = 0$. ⊘

We can apply the map $\mathcal{R}$ iteratively on some tree $\Lambda \in \tilde{\mathcal{T}}$ and get bigger and bigger trees in $\tilde{\mathcal{T}}$. What is interesting to us here is that choosing the map $\mathcal{R}$ (and so $\tilde{\mathcal{T}}$) appropriately we can impose special structures on the resulting tree, while keeping the size reasonably small. For instance, it is possible to grow any tree to a graded tree using this approach such that the result is optimal in some sense. To this end, let us study the following algorithm.

Algorithm 6.5.4 Tree completion

```
Λ := ∇_0;
for i = 1 to K do
  Let M̄_i ⊆ L(Λ);
  for all μ ∈ M̄_i do
    if μ ∉ Λ then
      Λ := Λ ∪ R(Λ, μ);
    end if
  end for
end for.
```

The following theorem is an easy extension of [87, Theorem 6.1] and [8, Theorem 2.4], and since the setting here is somewhat more general, we include the proof for reader's convenience.

Theorem 6.5.5. *Let M be the set of elements μ for which the map $\mathcal{R}$ is applied in the above algorithm, which set is thus contained in $\cup_i \bar{M}_i$. Then for the output tree $\Lambda \in \tilde{\mathcal{T}}$ we have $\#(\Lambda \setminus \nabla_0) \lesssim \#M$ uniformly in K.*

Proof. The proof will closely follow the proof of Theorem 6.1 in [87]. Let $a : \mathbb{N}_0 \cup \{-1, \dots, -L_{\mathcal{R}}\} \to (0, \infty)$ and $b : \mathbb{N} \to (1, \infty)$ be some sequences with $\sum_p a(p) < \infty$, $\sum_p b(p) 2^{-\chi p} < \infty$, and $\inf_{p \geq 1} [b(p) - 1] a(p) > 0$. For instance, $a(p) = (p + L_{\mathcal{R}} + 1)^{-1}$ and $b(p) = 1 + 2^{\kappa p}$ with a constant $\kappa \in (0, \chi)$ satisfy these conditions.

With $A := C_{\mathcal{R}} + (2^{\chi} C_{\mathcal{R}} + 2^{\chi} C_d + C_d) \sum_p b(p) 2^{-\chi p}$, we define the function $f : \Lambda \times M \to \mathbb{R}$ by

$$f(\lambda, \mu) = \begin{cases} a(|\mu| - |\lambda|) & \text{if } d(\lambda, \mu) < A 2^{-\chi |\lambda|} \text{ and } |\mu| - |\lambda| \geq -L_{\mathcal{R}}, \\ 0 & \text{otherwise.} \end{cases}$$

From condition (iv), for any $\mu \in M$ we have

$$\sum_{\lambda \in \Lambda} f(\lambda, \mu) = \sum_{\ell=0}^{|\mu| + L_{\mathcal{R}}} \sum_{|\lambda| = \ell} f(\lambda, \mu) \lesssim \sum_{\ell=0}^{|\mu| + L_{\mathcal{R}}} a(|\mu| - \ell) \leq \sum_p a(p) \lesssim 1,$$

implying that $\sum_{\mu \in M} \sum_{\lambda \in \Lambda} f(\lambda, \mu) \lesssim \#M$.

We claim that for any $\lambda \in \Lambda \setminus \nabla_0$,

$$\sum_{\mu \in M} f(\lambda, \mu) \gtrsim 1,$$

so that

$$\#(\Lambda \setminus \nabla_0) \lesssim \sum_{\lambda \in \Lambda \setminus \nabla_0} \sum_{\mu \in M} f(\lambda, \mu) \leq \sum_{\mu \in M} \sum_{\lambda \in \Lambda} f(\lambda, \mu) \lesssim \#M,$$

as required. Now we will prove this claim.

The claim is true for $\lambda \in M$ since $f(\lambda, \lambda) = a(0) \gtrsim 1$. Let $\lambda_0 \in \Lambda \setminus (M \cup \nabla_0)$. For $j \geq 0$, assume that λ_j has been defined and let λ'_j be the parent of λ_j for $j \geq 1$, and $\lambda'_0 := \lambda_0$. Then we define $\lambda_{j+1} \in M$ such that $\lambda'_j \in \mathcal{R}(\Lambda', \lambda_{j+1})$ with some tree Λ'. Let s be the smallest positive integer such that $|\lambda_s| \in I := \{|\lambda_0| - L_{\mathcal{R}}, \dots, |\lambda_0|\}$. Note that such an s exists. Indeed, the sequence $\{\lambda_j\}$ ends with some $\lambda_J \in \mathcal{L}(\nabla_0)$ thus with $|\lambda_J| = 1 \leq |\lambda_0|$, and from the properties of $\mathcal{R}$ we have $|\lambda'_j| \leq |\lambda_{j+1}| + L_{\mathcal{R}}$ or $|\lambda_{j+1}| \geq |\lambda_j| - L_{\mathcal{R}} - 1$ for $j \geq 1$ and $|\lambda_1| \geq |\lambda_0| - L_{\mathcal{R}}$, meaning that if not $|\lambda_1| \in I$, we have $|\lambda_1| > |\lambda_0|$. Therefore the interval I can

not be skipped by $j \mapsto \lambda_j$. For $1 \le j \le s$ we have

$$\begin{aligned} d(\lambda_0,\lambda_j) &\le d(\lambda_0,\lambda_1) + d(\lambda_1) + d(\lambda_1,\lambda_j) \le \sum_{k=1}^{j} d(\lambda_{k-1},\lambda_k) + \sum_{k=1}^{j-1} d(\lambda_k) \\ &\le \sum_{k=1}^{j} d(\lambda'_{k-1},\lambda_k) + \sum_{k=1}^{j-1} d(\lambda'_k) + d(\lambda_k) \\ &\le C_{\mathcal{R}} 2^{-\chi|\lambda_0|} + C_{\mathcal{R}} \sum_{k=1}^{j-1} 2^{-\chi|\lambda'_k|} + C_d \sum_{k=1}^{j-1} 2^{-\chi|\lambda'_k|} + 2^{-\chi|\lambda_k|} \\ &\le C_{\mathcal{R}} 2^{-\chi|\lambda_0|} + (2^{\chi} C_{\mathcal{R}} + 2^{\chi} C_d + C_d) \sum_{k=1}^{j-1} 2^{-\chi|\lambda_k|} \\ &= C_{\mathcal{R}} 2^{-\chi|\lambda_0|} + (2^{\chi} C_{\mathcal{R}} + 2^{\chi} C_d + C_d) \sum_{p=1}^{\infty} m(p,j) 2^{-\chi(|\lambda_0|+p)}, \end{aligned}$$

where $m(p,j)$ denotes the number of $k \in \{1,\dots,j-1\}$ with $|\lambda_k| = |\lambda_0| + p$. Note that $m(p,1) = 0$ for any p.

In case $m(p,s) \le b(p)$ for all $p \ge 1$, then by the definition of the constant A we have $d(\lambda_0,\lambda_s) < A 2^{-\chi|\lambda_0|}$. Since $-L_{\mathcal{R}} \le |\lambda_s| - |\lambda_0| \le 0$, we have $f(\lambda_0,\lambda_s) = a(|\lambda_s| - |\lambda_0|) \gtrsim 1$, which proves the claim.

Otherwise, there exist p with $m(p,s) > b(p)$. For each of those p, there exists a smallest $j = j(p)$ with $m(p,j(p)) > b(p)$ because $m(p,j) \ge m(p,j-1)$. With $j^* := \min_{p\ge 1} j(p)$, let p^* be such that $j(p^*) = j^*$. So we have $m(p,j^*-1) \le b(p)$ for all $p \ge 1$, and $m(p^*,j^*-1) \ge m(p^*,j^*) - 1 > b(p^*) - 1 > 0$. This implies that $j^* - 1 \ge 1$. As in the above case, we find that for all $1 \le k \le j^* - 1$, $d(\lambda_0,\lambda_k) < A 2^{-\chi|\lambda_0|}$ and $f(\lambda_0,\lambda_k) = a(|\lambda_k| - |\lambda_0|)$. Finally by using the definition of $m(\cdot,\cdot)$ we have

$$\begin{aligned} \sum_{\{1\le k\le j^*-1 : |\lambda_k| = \lambda_0|+p^*\}} f(\lambda_0,\lambda_k) &= m(p^*,j^*-1) a(p^*) \\ &> [b(p^*) - 1] a(p^*) \ge \inf_{p\ge 0} [b(p) - 1] a(p) \gtrsim 1, \end{aligned}$$

which proves the claim. ∎

Chapter 7

Computability of differential operators

7.1 Introduction

For a boundedly invertible $\mathbf{M} : \ell_2 \to \ell_2$, and $\mathbf{g} \in \ell_2$, we consider the problem of finding the solution $\mathbf{u} \in \ell_2$ of

$$\mathbf{M}\mathbf{u} = \mathbf{g}.$$

One can apply the adaptive algorithms from the preceding chapters, thereby e.g. Theorem 5.3.9 and Theorem 3.3.5 now say that if $\mathbf{u} \in \mathcal{A}^s$ for some s, *and* $\mathbf{M}$ is s^*-computable for an $s^* > s$, then the number of arithmetic operations and storage locations used by the adaptive wavelet algorithm for computing an approximation for $\mathbf{u}$ within tolerance ε is of the order $\varepsilon^{-1/s}$. Since in view of (2.3.3) the same order of storage locations is generally needed to approximate $\mathbf{u}$ within this tolerance using best N-term approximations, assuming these would be available, this result shows that the solution methods achieve the *optimal computational complexity* for the given problem.

To conclude optimality of the adaptive wavelet method, it is necessary to show that $\mathbf{M}$ is s^*-computable for some $s^* > \frac{d-t}{n}$, since otherwise for a solution u that has sufficient Besov regularity, the computability will be the limiting factor. On the other hand, since, for wavelets of order d, by imposing whatever smoothness conditions $\mathbf{u} \in \mathcal{A}^s$ can only be guaranteed for $s \leq \frac{d-t}{n}$, showing s^*-computability for some $s^* > \frac{d-t}{n}$ is also a sufficient condition for optimality of the adaptive wavelet method.

The work in this chapter is a joint work with Rob Stevenson, see Section 1.2

On the other hand, s^*-compressibility for some $s^* > \frac{d-t}{n}$ has been demonstrated in [86] for both differential and singular integral operators, and piecewise polynomial wavelets that are sufficiently smooth and have sufficiently many vanishing moments.

Only in the special case of a differential operator with constant coefficients, entries of $\mathbf{M}$ can be computed exactly, in $\mathcal{O}(1)$ operations, so that s^*-compressibility immediately implies s^*-computability. In general, numerical quadrature is required to approximate the entries. In this chapter, considering differential operators, we will show that $\mathbf{M}$ is s^*-computable for the same value of s^* as is was shown to be s^*-compressible. The case of singular integral operators will be treated in the next chapter. We split the task into two parts. First we derive a criterion on the accuracy-work balance of a numerical quadrature scheme to approximate any entry of $\mathbf{M}$, such that, for a suitable choice of the work invested in approximating the entries of the compressed matrix $\mathbf{M}_j$ as function of both wavelets involved, we obtain an approximation $\mathbf{M}_j^*$ of which the computation of each column requires $\mathcal{O}(2^j)$ operations, and $\|\mathbf{M}_j - \mathbf{M}_j^*\| \leq 2^{-js^*}$, meaning that, on account of Lemma 2.7.12, $\mathbf{M}$ is s^*-computable. Second, we show that we can fulfill above criterion by the application of standard composite quadrature rules of a fixed, sufficiently high order.

This chapter is organized as follows. We collect some error estimates for numerical quadrature in Section 7.2. In Section 7.3, assumptions are formulated on the boundary value problem and the wavelets, and the result concerning s^*-compressibility is recalled from [86]. In Section 7.4, rules for the numerical approximation of the entries of the stiffness matrix are derived, with which s^*-computability for some $s^* > \frac{d-t}{n}$ will be demonstrated.

At the end of this introduction, we fix a few more notations. A monomial of n variables is conveniently written using a *multi-index* $\alpha \in \mathbb{N}_0^n$ as $x^\alpha := x_1^{\alpha_1} \dots x_n^{\alpha_n}$. Likewise we write partial differentiation operators, that is, $\partial^\alpha := \partial_1^{\alpha_1} \dots \partial_n^{\alpha_n}$. We set $|\alpha| := \alpha_1 + \ldots + \alpha_n$, and the relation $\alpha \leq \beta$ is defined as $\alpha_i \leq \beta_i$ for all $i \in \overline{1,n}$. We have $|\alpha \pm \beta| = |\alpha| \pm |\beta|$ provided that $\alpha - \beta \in \mathbb{N}_0^n$ in case of subtraction.

7.2 Error estimates for numerical quadrature

We start with deriving an error bound in L_∞-norm for polynomial approximation, which improves upon available results (e.g. in [38, Theorem 1.1]) in the sense that our upper bound does not contain an unspecified constant that may vary as function of the polynomial order p. This latter fact will be particularly important for analyzing the errors of quadrature schemes with varying orders as we will apply

in the next chapter. We define the *radius* of a star-shaped domain Ω by

$$\mathrm{rad}(\Omega) := \min_{y \in S(\Omega)} \max_{x \in \partial\Omega} |x - y|, \tag{7.2.1}$$

where $S(\Omega) := \mathrm{clos}\{y \in \Omega : \Omega \text{ is star-shaped w.r.t. } y\}$. Apparently, we always have $\mathrm{rad}(\Omega) \leq \mathrm{diam}(\Omega)$, and the radius of a convex domain equals the radius of its smallest circumscribed sphere.

Lemma 7.2.1. *Let $\Omega \subset \mathbb{R}^n$ be a star-shaped domain and let $f \in W^p_\infty(\Omega)$, $p \in \mathbb{N}$. Then there exists a polynomial $g \in P_{p-1}$ on Ω for which*

$$\|f - g\|_{L_\infty(\Omega)} \leq \frac{n^p}{p!} \cdot \mathrm{rad}(\Omega)^p \cdot |f|_{W^p_\infty(\Omega)}. \tag{7.2.2}$$

Proof. We first assume that $f \in C^\infty(\Omega) \cap W^p_\infty(\Omega)$. Let a point $y \in S(\Omega)$ be such that $\max_{x \in \partial\Omega} |x - y| = \mathrm{rad}(\Omega)$. Let g be the Taylor polynomial of order p at the point y, i.e.,

$$g(x) = \sum_{|\alpha| < p} \frac{(x-y)^\alpha}{\alpha!} (\partial^\alpha f)(y). \tag{7.2.3}$$

Then the Taylor remainder is given by

$$f(x) - g(x) = p \sum_{|\alpha| = p} \frac{(x-y)^\alpha}{\alpha!} \int_0^1 s^{p-1} (\partial^\alpha f)(x + (y-x)s) ds.$$

Using

$$\left| \int_0^1 s^{p-1} (\partial^\alpha f)(x + (y-x)s) ds \right| \leq \int_0^1 s^{p-1} ds \cdot |f|_{W^p_\infty(\Omega)} = \frac{1}{p} \cdot |f|_{W^p_\infty(\Omega)},$$

and

$$|(x-y)^\alpha| = |x_1 - y_1|^{\alpha_1} \dots |x_n - y_n|^{\alpha_n} \leq \mathrm{rad}(\Omega)^p,$$

we have

$$|f(x) - g(x)| \leq \sum_{|\alpha| = p} \frac{1}{\alpha!} \cdot \mathrm{rad}(\Omega)^p \cdot |f|_{W^p_\infty(\Omega)}.$$

Then by applying the identity

$$\sum_{|\alpha| = p} \frac{1}{\alpha!} = \frac{n^p}{p!}$$

we get (7.2.2) for $f \in C^\infty(\Omega) \cap W^p_\infty(\Omega)$.

To complete the proof we use a density argument that was proven in [12]. For any $f \in W^p_\infty(\Omega)$, there exist functions $f_k \in C^\infty(\Omega) \cap W^p_\infty(\Omega)$, $k \in \mathbb{N}$, such that $f_k \to f$ in $W^{p-1}_\infty(\Omega)$, and $\|f_k\|_{W^p_\infty(\Omega)} \to \|f\|_{W^p_\infty(\Omega)}$ as $k \to \infty$. With this result, for each $k \in \mathbb{N}$ let us denote by $g_k \in P_{p-1}$ the Taylor polynomial (7.2.3) corresponding to f_k. Then, since for any $k, j \in \mathbb{N}$ and $|\alpha| < p$ we have

$$\begin{aligned} |(\partial^\alpha f_k)(y) - (\partial^\alpha f_j)(y)| &\leq \|\partial^\alpha f_k - \partial^\alpha f_j\|_{L_\infty(\Omega)} \\ &\leq \|\partial^\alpha f_k - \partial^\alpha f\|_{L_\infty(\Omega)} + \|\partial^\alpha f_j - \partial^\alpha f\|_{L_\infty(\Omega)}, \end{aligned}$$

where the right-hand side tends to zero as $j, k \to \infty$, we infer that there is a $g \in P_{p-1}$ such that $g_k \to g$ in $L_\infty(\Omega)$. Writing

$$\|f - g\|_{L_\infty(\Omega)} \leq \|f - f_k\|_{L_\infty(\Omega)} + \|g - g_k\|_{L_\infty(\Omega)} + \|f_k - g_k\|_{L_\infty(\Omega)},$$

and by taking the limit $k \to \infty$, the proof is completed. ∎

On a star-shaped domain Ω, let us now consider quadrature rules of the form $Q : f \mapsto \sum_j w_j f(x_j)$ to approximate $I : f \mapsto \int_\Omega f$. We will only consider rules that are *internal* meaning that all $x_j \in \operatorname{clos}\Omega$. The quadrature *error functional* is defined as $E := I - Q$.

Proposition 7.2.2. *For a rule Q of* order p, *meaning that $E(f) = 0$ for all $f \in P_{p-1}(\Omega)$, and any $f \in W^p_\infty(\Omega)$ we have*

$$|E(f)| \leq \left(1 + \frac{\sum_j |w_j|}{\operatorname{vol}(\Omega)}\right) \cdot \frac{n^p}{p!} \cdot \operatorname{rad}(\Omega)^p \cdot \operatorname{vol}(\Omega) \cdot |f|_{W^p_\infty(\Omega)}. \tag{7.2.4}$$

Proof. Taking g as in Lemma 7.2.1, the proof is an easy consequence of that lemma and the estimate

$$|I(f) - Q(f)| = |I(f) - Q(f) + Q(g) - I(g)| \leq |I(f - g)| + |Q(g - f)|. \quad ∎$$

Note that for a rule that is *positive*, meaning that all $w_j > 0$, and that has order $p > 0$, we have $\frac{\sum_j |w_j|}{\operatorname{vol}(\Omega)} = 1$.

Let us now consider a collection $\mathcal{O}$ of disjoint star-shaped Lipschitz subdomains $\Omega' \subset \Omega$, the latter not necessarily being star-shaped, such that $\operatorname{clos}\Omega = \cup_{\Omega' \in \mathcal{O}} \operatorname{clos}\Omega'$, which collection we will refer to as being a *quadrature mesh*. Writing $I(f)$ as $\sum_{\Omega' \in \mathcal{O}} \int_{\Omega'} f$, on each subdomain Ω' we employ a quadrature rule $Q_{\Omega'}(f) = \sum_j w_j^{\Omega'} f(x_j^{\Omega'})$ of order p, defining a *composite* quadrature rule Q of *rank* $N := \#\mathcal{O}$ (and order p) by $Q(f) := \sum_{\Omega' \in \mathcal{O}} Q_{\Omega'}(f)$.

Proposition 7.2.3. *For the error functional $E = I - Q$ of this composite quadrature rule, and $f \in W^p_\infty(\Omega)$ we have*

$$|E(f)| \leq \left(1 + \sup_{\Omega' \in \mathcal{O}} \frac{\sum_j |w_j^{\Omega'}|}{\operatorname{vol}(\Omega')}\right) \cdot \sup_{\Omega' \in \mathcal{O}} \left(\frac{N^{1/n} \operatorname{rad}(\Omega')}{\operatorname{diam}(\Omega)}\right)^p \times N^{-p/n} \cdot \frac{n^p}{p!} \cdot \operatorname{diam}(\Omega)^p \cdot \operatorname{vol}(\Omega) \cdot |f|_{W^p_\infty(\Omega)}.$$

Proof. Writing $\operatorname{rad}(\Omega') = \frac{N^{1/n}\operatorname{rad}(\Omega')}{\operatorname{diam}(\Omega)} N^{-1/n} \operatorname{diam}(\Omega)$, and using that $\sum_{\Omega' \in \mathcal{O}} \operatorname{vol}(\Omega') = \operatorname{vol}(\Omega)$, the result follows from Proposition 7.2.2. ∎

In view of above estimate, as well as to control the number of function evaluations that are required, in this chapter we will consider families $(\mathcal{O}_\ell)_{\ell \in \mathbb{N}}$ of quadrature meshes and corresponding families of composite quadrature rules $Q_\ell : f \mapsto \sum_{Q' \in \mathcal{O}_\ell} \sum_j w_j^{\Omega'} f(x_j^{\Omega'})$ of rank $N_\ell := \#\mathcal{O}_\ell$ and *fixed order* p that are *admissible* meaning that they satisfy

$$\sup_{\ell \in \mathbb{N}, \Omega' \in \mathcal{O}_\ell} \max \left\{ \frac{\sum_j |w_j^{\Omega'}|}{\operatorname{vol}(\Omega')}, \frac{N_\ell^{1/n} \operatorname{rad}(\Omega')}{\operatorname{diam}(\Omega)}, \#x_j^{\Omega'} \right\} < \infty.$$

Note that the bound on the number of abscissae in each subdomain is reasonable because the space of polynomials of total degree $p-1$ has $\binom{p-1+n}{n} \leq p^n \lesssim 1$ degrees of freedom.

Finally in this section, we consider *product quadrature rules* which are generally applied on Cartesian product domains. Let A and B be domains of possibly different dimensions, equipped with the quadrature rules $Q^{(A)} : g \mapsto \sum_j w_j g(x_j)$ and $Q^{(B)} : h \mapsto \sum_k v_k h(y_k)$ to approximate $I^{(A)} : g \mapsto \int_A g$ and $I^{(B)} : h \mapsto \int_B h$, respectively. For simplicity, in this setting we will always assume that these rules are *positive* and have *strictly positive orders*. Now with the product rule $Q^{(A)} \times Q^{(B)}$ we mean the mapping $f \mapsto \sum_{jk} w_j v_k f(x_j, y_k)$ to approximate $I : f \mapsto \int_{A \times B} f$.

Lemma 7.2.4. *With error functionals $E^{(A)} := I^{(A)} - Q^{(A)}$ and $E^{(B)} := I^{(B)} - Q^{(B)}$, the product rule $Q := Q^{(A)} \times Q^{(B)}$ satisfies*

$$|I(f) - Q(f)| \leq \operatorname{vol}(A) \sup_{x \in A} |E^{(B)}(f(x, \cdot))| + \operatorname{vol}(B) \sup_{y \in B} |E^{(A)}(f(\cdot, y))|, \quad (7.2.5)$$

as long as both $E^{(A)}(f(\cdot, y))$ and $E^{(B)}(f(x, \cdot))$ make sense for all $y \in B$ and $x \in A$, respectively.

Proof. We have

$$\begin{aligned} I(f) - Q(f) &= \int_{A\times B} f(x,y)dxdy - \sum_{j,k} w_j v_k f(x_j, y_k) \\ &= \int_B \left(\int_A f(x,y)dx - \sum_j w_j f(x_j, y) \right) dy \\ &\quad + \sum_j w_j \left(\int_B f(x_j, y)dy - \sum_k v_k f(x_j, y_k) \right). \end{aligned}$$

The proof is completed by taking absolute values and using that $\sum_j w_j = \mathrm{vol}(A)$. ∎

As an application of this lemma, we have the following result for product quadrature rules on rectangular domains.

Proposition 7.2.5. *Consider the rectangular domain $\square := (0, l_1) \times \ldots \times (0, l_n)$. For the i-th coordinate direction, let $Q^{(i)}_{N_i}$ be a composite quadrature rule of order p with respect to a quadrature mesh on $(0, l_i)$ of N_i equally sized subintervals. Then for the product quadrature rule $Q := Q^{(1)}_{N_1} \times \ldots \times Q^{(n)}_{N_n}$ to approximate $I : f \mapsto \int_\square f$, and f such that $\partial_i^p f \in L_\infty(\square)$, $i \in \overline{1,n}$, we have*

$$|I(f) - Q(f)| \leq \frac{2^{1-p}}{p!}\ \mathrm{vol}(\square) \cdot \sum_{i=1}^{n} l_i^p N_i^{-p} \cdot \max_{i\in\overline{1,n}} \|\partial_i^p f\|_{L_\infty(\square)}. \tag{7.2.6}$$

In particular, this quadrature rule is exact on $Q_{p-1}(\square) := P_{p-1}(0, l_1) \times \ldots \times P_{p-1}(0, l_n)$.

Proof. Using that $\mathrm{rad}(0, l_i) = l_i/2$, Proposition 7.2.3 shows that for each i,

$$\left| \int_0^{l_i} g - Q^{(i)}_{N_i}(g) \right| \leq \frac{2^{1-p}}{p!} N_i^{-p} l_i^{p+1} |g|_{W^p_\infty(0,l_i)}.$$

Using Lemma 7.2.4 we arrive at the claim by induction. ∎

Corollary 7.2.6. *For the special case $N_1 = \ldots = N_n = N^{1/n}$, with $l := \max_i l_i$ we have*

$$|I(f) - Q(f)| \leq n\frac{2^{1-p}}{p!}\ N^{-p/n} \cdot l^{n+p} \cdot \max_{i\in\overline{1,n}} \|\partial_i^p f\|_{L_\infty(\square)}. \tag{7.2.7}$$

7.3 Compressibility

For some domain $\Omega \subset \mathbb{R}^n$, $t \in \mathbb{N}_0$ and $\Gamma_D \subset \partial\Omega$, possibly with $\Gamma_D = \varnothing$, let

$$H^t_{0,\Gamma^D}(\Omega) = \mathrm{clos}_{H^t(\Omega)}\{u \in H^t(\Omega) \cap C^\infty(\Omega) : \mathrm{supp}\, u \cap \Gamma^D = \varnothing\},$$

and let $L : H^t_{0,\Gamma^D}(\Omega) \to (H^t_{0,\Gamma^D}(\Omega))'$ be defined by

$$\langle u, Lv\rangle = \sum_{|\alpha|,|\beta|\le t} \langle \partial^\alpha u, a_{\alpha\beta}\partial^\beta v\rangle,$$

where $a_{\alpha\beta} \in L_\infty(\Omega)$ so that L is bounded. Obviously L has an extension, that we will also denote by L, as a bounded operator from $H^t(\Omega) \to H^{-t}(\Omega)$. For completeness, $H^s(\Omega)$ for $s < 0$ denotes the dual of $H^{-s}(\Omega)$.

We assume that there exists a $\sigma > 0$, such that

$$L, L' : H^{t+\sigma}(\Omega) \to H^{-t+\sigma}(\Omega) \quad \text{are bounded.} \tag{7.3.1}$$

Sufficient is that for arbitrary $\varepsilon > 0$, and all α, β with $\min\{|\alpha|, |\beta|\} > t - \sigma$, it holds that

$$a_{\alpha\beta} \in \begin{cases} W_\infty^{\sigma-t+\min\{|\alpha|,|\beta|\}}(\Omega) & \text{when } \sigma \in \mathbb{N}, \\ C^{\sigma-t+\min\{|\alpha|,|\beta|\}+\varepsilon}(\Omega) & \text{when } \sigma \notin \mathbb{N}. \end{cases}$$

In addition, we assume that the coefficients $a_{\alpha\beta}$ are *piecewise smooth*, in the sense that there exist M disjoint Lipschitz domains Ω_q, $q \in \overline{1, M}$, such that $a_{\alpha\beta}$ is smooth on each Ω_q, and $\mathrm{clos}\,\Omega = \cup_q \mathrm{clos}\,\Omega_q$.

Let

$$\Psi = \{\psi_\lambda : \lambda \in \Lambda\}$$

be a *Riesz basis for* $H^t_{0,\Gamma^D}(\Omega)$ of wavelet type. The index λ encodes both the level, denoted by $|\lambda| \in \mathbb{N}_0$, and the location of the wavelet ψ_λ. We will assume that the wavelets are *local* and *piecewise smooth* with respect to nested subdivisions in the following sense: We assume that there exists a sequence $(\mathcal{O}_\ell)_{\ell\in\mathbb{N}_0}$ of collections $\mathcal{O}_\ell = \{\Omega_i^\ell : i \in J^\ell\}$ of disjoint "uniformly" (in i and ℓ) Lipschitz domains Ω_i^ℓ, with $\mathrm{clos}\,\Omega = \cup_{i\in J^\ell} \mathrm{clos}\,\Omega_i^\ell$ and

$$\mathrm{diam}(\Omega_i^\ell) \eqsim 2^{-\ell} \qquad \text{and} \qquad \mathrm{vol}(\Omega_i^\ell) \eqsim 2^{-n\ell}, \tag{7.3.2}$$

where each Ω_i^ℓ is contained in some Ω_q, and its closure is the union of the closures of a uniformly bounded number of subdomains from $\mathcal{O}_{\ell+1}$. We assume that for each $\lambda \in \Lambda$ there exists a $J_\lambda \subset J^{|\lambda|}$ with

$$\sup_{\lambda\in\Lambda} \#J_\lambda < \infty \quad \text{and} \quad \sup_{\ell\in\mathbb{N}_0, i\in J^\ell} \#\{\lambda : |\lambda| = \ell, i \in J_\lambda\} < \infty,$$

such that $\operatorname{supp}\psi_\lambda = \cup_{i\in J_\lambda} \operatorname{clos}\Omega_i^{|\lambda|}$, being a connected set, and that on each $\Omega_i^{|\lambda|}$, ψ_λ is smooth with

$$\sup_{x\in\Omega_i^{|\lambda|}} |\partial^\beta \psi_\lambda(x)| \lesssim 2^{(|\beta|+\frac{n}{2}-t)|\lambda|} \qquad \text{for } \beta \in \mathbb{N}_0^n. \tag{7.3.3}$$

Examples of such wavelets are (the images under smooth mappings of) tensor products of univariate spline wavelets, or finite element wavelets subordinate to a subdivision of the domain into n-simplices.

Remark 7.3.1. Precisely, we call a collection of domains $\{A_\nu\} \subset \mathbb{R}^n$ uniformly Lipschitz domains when there exist affine mappings B_ν with $|DB_\nu| \lesssim \operatorname{vol}(A_\nu)^{-1}$ and $|(DB_\nu)^{-1}| \lesssim \operatorname{vol}(A_\nu)$ such that the sets $B_\nu(A_\nu)$ satisfy the condition of *minimal smoothness* in the sense of Stein (cf. [82, §VI.3]), with uniform parameters ε, N and M.

A minimally smooth domain in $\mathbb{R}^n$, in the sense of Stein, is an open set for which there is a number $\varepsilon > 0$ and open sets U_i, $i = 1, 2, \ldots$, such that: (i) for each $x \in \partial\Omega$, the ball $B(x, \varepsilon)$ is contained in one of U_i; (ii) a point $x \in \mathbb{R}^n$ is in at most N of the sets U_i where N is an absolute constant; (iii) for each i, $U_i \cap \Omega = U_i \cap \Omega_i$ for some domain Ω_i which is the rotation of a Lipschitz graph domain with Lipschitz constant M independent of i. ⊘

Furthermore, we assume that there exist $\gamma > t$, $\tilde{d} > -t$ such that for $r \in [-\tilde{d}, \gamma)$, $s < \gamma$,

$$\|\cdot\|_{H^r(\Omega)} \lesssim 2^{\ell(r-s)} \|\cdot\|_{H^s(\Omega)}, \qquad \text{on } W_\ell := \operatorname{span}\{\psi_\lambda : |\lambda| = \ell\}. \tag{7.3.4}$$

For $r > s$, this is the well-known inverse inequality. For $r < s$, (7.3.4) is a consequence of the property of wavelets of having *vanishing moments*, or, more generally, *cancellation properties*.

Remark 7.3.2. It is known that the above wavelet assumptions are satisfied by biorthogonal wavelets when the primal and dual spaces have regularity indices $\gamma > t$, $\tilde{\gamma} > 0$ and orders $d > \gamma$, $\tilde{d} > \tilde{\gamma}$ respectively (cf. [29, 36]), the primal spaces consist of "piecewise" smooth functions, and finally, no boundary conditions are imposed on the dual spaces (cf. [32]). In particular, (7.3.4) for $r \in [-\tilde{d}, -\tilde{\gamma}]$ can be deduced from the lines following (A.2) in [36]. In case homogeneous boundary conditions are incorporated in the dual spaces, slightly weaker statements can be proven, see [86, Remark 2.5].

We recall here the main result on compressibility for differential operators from [86].

Theorem 7.3.3. *Let* $\mathbf{M} = \langle \Psi, L\Psi \rangle$. *Choose* κ *satisfying*

$$\begin{aligned} &\kappa = \frac{1}{n-1} && \text{when } n > 1, \qquad (7.3.5)\\ &\kappa > \frac{\min\{t+\tilde{d}, \sigma\}}{\gamma - t} \quad \text{and} \quad \kappa \geq 1 && \text{when } n = 1. \end{aligned}$$

For $j \in \mathbb{N}$, *define the infinite matrix* $\mathbf{M}_j$ *by replacing all entries* $M_{\lambda\lambda'} = \langle \psi_\lambda, L\psi_{\lambda'} \rangle$ *by zeros when*

$$\big||\lambda| - |\lambda'|\big| > j\kappa, \quad \text{or} \qquad (7.3.6)$$

$$\big||\lambda| - |\lambda'|\big| > j/n \quad \text{and} \quad \begin{cases} \exists i' \in J_{\lambda'},\ \operatorname{supp}\psi_\lambda \subseteq \operatorname{clos}\Omega_{i'}^{|\lambda'|} & \text{when } |\lambda| > |\lambda'|, \\ \exists i \in J_\lambda,\ \operatorname{supp}\psi_{\lambda'} \subseteq \operatorname{clos}\Omega_i^{|\lambda|} & \text{when } |\lambda| < |\lambda'|. \end{cases} \qquad (7.3.7)$$

Then the number of non-zero entries in each column of $\mathbf{M}_j$ *is of order* 2^j, *and for any*

$$s \leq \min\{\tfrac{t+\tilde{d}}{n}, \tfrac{\sigma}{n}\}, \ \text{with } s < \tfrac{\gamma - t}{n-1} \text{ when } n > 1,$$

it holds that $\|\mathbf{M} - \mathbf{M}_j\| \lesssim 2^{-js}$. *We conclude that* $\mathbf{M}$ *is* s^**-compressible, as defined in Definition 2.7.11, with* $s^* = \min\{\tfrac{t+\tilde{d}}{n}, \tfrac{\sigma}{n}, \tfrac{\gamma-t}{n-1}\}$ *when* $n > 1$, *and* $s^* = \min\{t + \tilde{d}, \sigma\}$ *when* $n = 1$.

From this theorem we infer that if $\tilde{d} > d - 2t$, $\sigma > d - t$ and, when $n > 1$, $\frac{\gamma-t}{n-1} > \frac{d-t}{n}$, then $s^* > \frac{d-t}{n}$ as required. For $n > 1$, the condition involving γ is satisfied for instance for spline wavelets, where $\gamma = d - \frac{1}{2}$, in case $\frac{d-t}{n} > \frac{1}{2}$.

If each entry of $\mathbf{M}$ can be exactly computed in $\mathcal{O}(1)$ operations, then s^*-compressibility implies s^*-computability, as defined in Definition 2.7.8, and so, when indeed $s^* > \frac{d-t}{n}$, it implies the optimal computational complexity of the adaptive wavelet scheme from the preceding chapters. This assumption on the computation of the entries is realistic when both the coefficients $a_{\alpha\beta}$ of the differential operator and the wavelets are piecewise polynomials. In general, however, numerical quadrature will be needed to approximate the entries of $\mathbf{M}_j$. Then the question arises how to realize a sufficient accuracy of these approximations such that the additional error has, qualitatively, the same upper bound as $\|\mathbf{M} - \mathbf{M}_j\|$, where in each column the average work per entry is $\mathcal{O}(1)$, in which case s^*-compressibility implies s^*-computability. In the next section, additionally assuming that the wavelets are essentially *piecewise polynomials*, we will see that it is possible to select quadrature rules with which this is realized.

7.4 Computability

Let us denote by $\mathbf{M}_j^*$ the matrix, with elements $M_{j,\lambda\lambda'}^*$, obtained by approximating the entries of $\mathbf{M}_j$ using some numerical scheme dependent on j. The following theorem defines a criterion on the computational cost in relation to the accuracy for computing individual entries of $\mathbf{M}$ so that s^*-compressibility implies s^*-computability.

Theorem 7.4.1. *Let $\mathbf{M}$, $\mathbf{M}_j$ and s^* be as in Theorem 7.3.3. Assume that for some $d^* \in \mathbb{R}$ and p with*

$$p > s^*n + d^* \quad \text{and} \quad p \geq s^*n, \tag{7.4.1}$$

an approximation $M_{\lambda\lambda'}^*$ of $M_{\lambda\lambda'}$ can be computed in $\mathcal{O}(N)$ operations, having an error

$$|M_{\lambda\lambda'} - M_{\lambda\lambda'}^*| \lesssim N^{-p/n} 2^{-||\lambda|-|\lambda'||(n/2+p-d^*)}. \tag{7.4.2}$$

Then for parameters θ and ϱ with

$$\theta \leq 1 \quad \text{and} \quad s^*n/p \leq \theta \leq \varrho < 1 - d^*/p, \tag{7.4.3}$$

by spending the number of

$$N_{j,\lambda\lambda'} \eqsim \max\{1, 2^{j\theta - ||\lambda|-|\lambda'||n\varrho}\} \tag{7.4.4}$$

arithmetical operations to the computation of $M_{j,\lambda\lambda'}^*$, one has $\|\mathbf{M}_j - \mathbf{M}_j^*\| \lesssim 2^{-js^*}$, and the work for computing each column of $\mathbf{M}_j^*$ is of order 2^j.

Since the conditions (7.4.1) *and* (7.4.3) *define a nonempty set in the $\theta - \varrho$ plane, we conclude that $\mathbf{M}$ is s^*-computable.*

The proof will use Schur's lemma that we recall here for the reader's convenience.

Schur's lemma. *If for a matrix $\mathbf{A} = (a_{\lambda,\lambda'})_{\lambda,\lambda' \in \Lambda}$, there is a sequence $\omega_\lambda > 0$, $\lambda \in \Lambda$, and a constant C such that*

$$\sum_{\lambda' \in \Lambda} \omega'_\lambda |a_{\lambda\,\lambda'}| \leq \omega_\lambda C, \quad (\lambda \in \Lambda), \quad \text{and} \quad \sum_{\lambda \in \Lambda} \omega_\lambda |a_{\lambda\,\lambda'}| \leq \omega'_\lambda C, \quad (\lambda' \in \Lambda),$$

then $\|\mathbf{A}\| \leq C$.

Proof (Proof of Theorem 7.4.1). Denoting the (λ, λ')-th entry of the error matrix $\mathbf{M}_j - \mathbf{M}_j^*$ by $\varepsilon_{j,\lambda\lambda'}$, from (7.4.2) and (7.4.4) we have

$$\begin{aligned} \varepsilon_{j,\lambda\lambda'} &\lesssim N_{j,\lambda\lambda'}^{-p/n} 2^{-||\lambda|-|\lambda'||(n/2+p-d^*)} \\ &\lesssim 2^{-||\lambda|-|\lambda'||(n/2+p-\varrho p-d^*)} 2^{-j\theta p/n}. \end{aligned} \tag{7.4.5}$$

We have $\sigma := n/2+p-\varrho p-d^* = n/2+p(1-\varrho-d^*/p) > n/2$ from (7.4.3). Let λ be some given index. The locality assumptions on the wavelets show that for fixed $\lambda \in \Lambda$, the number of indices λ' with fixed $|\lambda'|$ with $\mathrm{vol}(\operatorname{supp}\psi_{\lambda'} \cap \operatorname{supp}\psi_\lambda) > 0$ is of order $\max\{1, 2^{(|\lambda'|-|\lambda|)n}\}$. With weights $\omega_{\lambda'} = 2^{-|\lambda'|n/2}$, we find

$$\begin{aligned}\omega_\lambda^{-1}\sum_{\lambda'}\omega_{\lambda'}|\varepsilon_{j,\lambda\lambda'}| &\lesssim 2^{|\lambda|n/2}\sum_{0\le|\lambda'|\le|\lambda|} 2^{-|\lambda'|n/2}2^{-(|\lambda|-|\lambda'|)\sigma}2^{-j\theta p/n}\cdot 1\\ &\quad + 2^{|\lambda|n/2}\sum_{|\lambda'|>|\lambda|} 2^{-|\lambda'|n/2}2^{-(|\lambda'|-|\lambda|)\sigma}2^{-j\theta p/n}\cdot 2^{(|\lambda'|-|\lambda|)n}\\ &\lesssim 2^{-j\theta p/n}.\end{aligned}$$

By the symmetry of the estimate (7.4.5) in λ and λ', from Schur's lemma we conclude that

$$\|\mathbf{M}_j - \mathbf{M}_j^*\| \lesssim 2^{-j\theta p/n} \le 2^{-js^*},$$

because $\theta \ge s^*n/p$.

Denoting by $\Lambda_{j,\lambda}$ the set of row-indices of nonzero entries in the λ-th column of $\mathbf{M}_j$, the computational work $W_{j,\lambda}$ for this column is

$$\begin{aligned}W_{j,\lambda} = \sum_{\lambda'\in\Lambda_{j,\lambda}} N_{j,\lambda\lambda'} &\lesssim \sum_{\lambda'\in\Lambda_{j,\lambda}} \max\{1, 2^{j\theta-||\lambda|-|\lambda'||n\varrho}\}\\ &\lesssim 2^j + \sum_{\{\lambda'\in\Lambda_{j,\lambda}:||\lambda|-|\lambda'||\le j/n\}} 2^{j\theta-||\lambda|-|\lambda'||n\varrho},\end{aligned}$$

where we used the fact that, since $\varrho \ge \theta$, $2^{j\theta-||\lambda|-|\lambda'||n\varrho} < 1$ for $||\lambda| - |\lambda'|| > j/n$, and that the number of nonzero entries in each column of $\mathbf{M}_j$ is $\mathcal{O}(2^j)$. The second term can be bounded by a constant multiple of

$$\begin{aligned}\sum_{-j/n\le|\lambda'|-|\lambda|\le 0} 2^{j\theta-(|\lambda|-|\lambda'|)n\varrho}\cdot 1 + \sum_{0<|\lambda'|-|\lambda|\le j/n} 2^{j\theta-(|\lambda'|-|\lambda|)n\varrho}\cdot 2^{(|\lambda'|-|\lambda|)n}&\\ \lesssim 2^{j\theta}2^{j\max\{0,1-\varrho\}}.&\end{aligned}$$

From (7.4.3) we have $\theta \le 1$ and $\theta \le \varrho$, and so $1-\varrho+\theta \le 1$, from which we conclude that $W_{j,\lambda} = \mathcal{O}(2^j)$. ∎

By applying the error estimates from Section 7.2, we will now show how numerical quadrature schemes satisfying (7.4.2) can be realized. An entry of the matrix can be rewritten as

$$M_{\lambda\lambda'} = \sum_{|\alpha|,|\beta|\le t}\int_{\operatorname{supp}\psi_\lambda\cap\operatorname{supp}\psi_{\lambda'}} a_{\alpha\beta}\partial^\alpha\psi_\lambda\partial^\beta\psi_{\lambda'}.$$

Without loss of generality, in the remainder of this section we assume that

$$|\lambda| \geq |\lambda'|.$$

Then, it is clear that the intersection $\operatorname{supp}\psi_\lambda \cap \operatorname{supp}\psi_{\lambda'}$ is the union of sets $\Omega_i^{|\lambda|}$, $i \in J_{\lambda\lambda'}$, for some $J_{\lambda\lambda'} \subseteq J_\lambda$. Therefore we can expand the integral into integrals of smooth functions

$$M_{\lambda\lambda'} = \sum_{i \in J_{\lambda\lambda'}} I_{\lambda\lambda',i}, \tag{7.4.6}$$

where

$$I_{\lambda\lambda',i} := \sum_{|\alpha|,|\beta|\leq t} \int_{\Omega_i^{|\lambda|}} a_{\alpha\beta}\partial^\alpha\psi_\lambda\partial^\beta\psi_{\lambda'}. \tag{7.4.7}$$

Recall that for each $\ell \in \mathbb{N}_0$, $i \in J^\ell$, there is a $q = q(\ell, i) \in \overline{1, M}$ with $\Omega_i^\ell \subset \Omega_q$, and furthermore that all $a_{\alpha\beta}$ are smooth on any Ω_q. In the following, we assume that there is a constant $e \in \mathbb{N}$, and that there are smooth regular mappings $\kappa_q : \mathbb{R}^n \to \mathbb{R}^n$, such that for each $\lambda \in \Lambda$, $i \in J_\lambda$, and $q = q(|\lambda|, i)$,

$$(\psi_\lambda \circ \kappa_q)|_{\kappa_q^{-1}(\Omega_i^{|\lambda|})} \in P_{e-1}. \tag{7.4.8}$$

With the commonly used approaches to constructed wavelets on non-trivial geometries via domain decomposition techniques ([14, 22, 33, 34, 85]), above assumption is valid when one starts from a piecewise polynomial multiresolution analysis on the corresponding reference domain. Note that the smallest e for which (7.4.8) holds satisfies $e \geq d$ $(\geq t + 1)$.

Since by a transformation of coordinates, $\sum_{|\alpha|,|\beta|\leq t} \int_{\Omega_i^{|\lambda|}} a_{\alpha\beta}\partial^\alpha\psi_\lambda\partial^\beta\psi_{\lambda'}$ can be written as $\sum_{|\alpha|,|\beta|\leq t} \int_{\kappa_q^{-1}(\Omega_i^{|\lambda|})} \tilde{a}_{\alpha\beta}\partial^\alpha(\psi_\lambda \circ \kappa_q)\partial^\beta(\psi_{\lambda'} \circ \kappa_q)$, where, as $a_{\alpha\beta}$, $\tilde{a}_{\alpha\beta}$ is a function that is smooth, in the following without loss of generality we may assume that $\kappa_q = \mathrm{id}$.

Proposition 7.4.2. *Consider a composite quadrature rule from an admissible family (uniformly in $\lambda \in \Lambda$ and $i \in J_\lambda$) of fixed order p and rank N to approximate each of the integrals from* (7.4.7)*, where $\psi_\lambda|_{\Omega_i^{|\lambda|}} \in P_{e-1}$. Then, with*

$$d^* := e - 1 - t, \tag{7.4.9}$$

the error of this numerical integration is bounded by

$$|M_{\lambda\lambda'} - M^*_{\lambda\lambda'}| \lesssim N^{-p/n} 2^{-||\lambda|-|\lambda'||(n/2+p-d^*)}, \tag{7.4.10}$$

Taking $p > s^*n + d^*$, we conclude that the criterion for s^*-computability from Theorem 7.4.1 is satisfied.

Proof. In view of Proposition 7.2.3 we have to bound $\partial^\zeta(a_{\alpha\beta}\partial^\alpha\psi_\lambda\partial^\beta\psi_{\lambda'})$ for $|\zeta| = p$, or $\partial^\eta a_{\alpha\beta}\partial^{\alpha+\theta}\psi_\lambda\partial^{\beta+\xi}\psi_{\lambda'}$ for $|\eta+\theta+\xi| = p$. Since $a_{\alpha\beta}$ is smooth, $|\lambda| \geq |\lambda'|$, and $\partial^{\alpha+\theta}\psi_\lambda$ vanishes when $|\alpha+\theta| \geq e$, by invoking (7.3.3) we see that the worst case occurs when $\eta = 0$, $|\alpha+\theta| = r := \min\{e-1, |\alpha|+p\}$, and thus $|\xi| = p-r+|\alpha|$, yielding

$$\begin{aligned}|a_{\alpha\beta}\partial^\alpha\psi_\lambda\partial^\beta\psi_{\lambda'}|_{W^p_\infty(\Omega_i^{|\lambda|})} &\lesssim 2^{(r+n/2-t)|\lambda|}2^{(p-r+|\alpha|+|\beta|+n/2-t)|\lambda'|}\\ &\leq 2^{(e-1+n/2-t)|\lambda|}2^{(p-e+1+|\alpha|+|\beta|+n/2-t)|\lambda'|}.\end{aligned}$$

Now using that $\operatorname{diam}(\Omega_i^{|\lambda|}) \eqsim 2^{-|\lambda|}$ and $|\alpha|, |\beta| \leq t$, Proposition 7.2.3 shows that

$$\begin{aligned}|M_{\lambda\lambda'} - M^*_{\lambda\lambda'}| &\lesssim N^{-p/n}2^{-|\lambda|(n+p)}2^{(e-1+n/2-t)|\lambda|}2^{(p-e+1+t+n/2)|\lambda'|}\\ &= N^{-p/n}2^{-||\lambda|-|\lambda'||(n/2+p-d^*)}.\end{aligned}$$

■

In the case of tensor product constructions yielding wavelets that are piecewise in Q_{d-1}, the (piecewise) polynomial order e is $n(d-1)+1$, so that d^* from Proposition 7.4.2 is equal to $n(d-1)-t$ ($\geq (n-1)d$). In the next proposition, we will see that for such wavelets the application of product quadrature rules gives rise to smaller d^*, and so allows for smaller quadrature orders p.

Proposition 7.4.3. *Suppose that $\Omega_i^{|\lambda|}$ is an n-rectangle, that a product composite quadrature rule of order p and rank N as in Corollary 7.2.6 is applied to approximate each of the integrals from* (7.4.7)*, and that $\psi_\lambda|_{\Omega_i^{|\lambda|}} \in Q_{d-1}(\Omega_i^{|\lambda|})$. Then, with*

$$d^* := d-1, \tag{7.4.11}$$

the error of the numerical integration is bounded by

$$|M_{\lambda\lambda'} - M^*_{\lambda\lambda'}| \lesssim N^{-p/n}2^{-||\lambda|-|\lambda'||(n/2+p-d^*)}. \tag{7.4.12}$$

Taking $p > s^*n + d^*$, we conclude that the criterion for s^*-computability from Theorem 7.4.1 is satisfied.

Proof. Without loss of generality, we may assume that the n-rectangle $\Omega_i^{|\lambda|} \subset \mathbb{R}^n$ is aligned with the Cartesian coordinates. In view of Corollary 7.2.6, for any $i \in \overline{1,n}$ we have to bound $\partial_i^p(a_{\alpha\beta}\partial^\alpha\psi_\lambda\partial^\beta\psi_{\lambda'})$, or $\partial_i^k a_{\alpha\beta}\partial_i^l\partial^\alpha\psi_\lambda\partial_i^m\partial^\beta\psi_{\lambda'}$ for $k+l+m = p$. Since $a_{\alpha\beta}$ is smooth, $|\lambda| \geq |\lambda'|$, and $\partial_i^l\partial^\alpha\psi_\lambda$ vanishes when $\alpha_i + l \geq d$, by invoking (7.3.3) we see that the worst case occurs when $k = 0$, $\alpha_i + l = r := \min\{d-1, \alpha_i+p\}$, and thus $m = p-r+\alpha_i$, yielding

$$\begin{aligned}\left|\partial_i^p\left(a_{\alpha\beta}\partial^\alpha\psi_\lambda\partial^\beta\psi_{\lambda'}\right)\right| &\lesssim 2^{(|\alpha|-\alpha_i+r+n/2-t)|\lambda|}2^{(p-r+\alpha_i+|\beta|+n/2-t)|\lambda'|}\\ &\lesssim 2^{(|\alpha|+d-1+n/2-t)|\lambda|}2^{(p-d+1+|\beta|+n/2-t)|\lambda'|}.\end{aligned}$$

Since $\operatorname{diam}(\Omega_i^{|\lambda|}) \eqsim 2^{-|\lambda|}$ and $|\alpha|, |\beta| \leq t$, an application of Corollary 7.2.6 shows that

$$\begin{aligned} |M_{\lambda\lambda'} - M^*_{\lambda\lambda'}| &\lesssim N^{-p/n} 2^{-|\lambda|(n+p)} 2^{(d-1+n/2)|\lambda|} 2^{(p-d+1+n/2)|\lambda'|} \\ &= N^{-p/n} 2^{-||\lambda|-|\lambda'||(n/2+p-d^*)}. \end{aligned}$$

■

Chapter 8

Computability of singular integral operators

8.1 Introduction

Boundary integral methods reduce elliptic boundary value problems in domains to integral equations formulated on the boundary of the domain. Although the dimension of the underlying manifold decreases by one, the finite element discretization of the resulting boundary integral equations gives densely populated stiffness matrices, causing serious obstructions to accurate numerical solution processes. In order to overcome this difficulty, various successful approaches for approximating the stiffness matrix by sparse ones have been developed, such as multipole expansions, panel clustering, and wavelet compression, see e.g. [2, 51]. We will restrict ourselves here to the latter approach.

In [7], it was first observed that wavelet bases give rise to almost sparse stiffness matrices for the Galerkin discretization of singular integral operators, meaning that the stiffness matrix has many small entries that can be discarded without reducing the order of convergence of the resulting solution. This result ignited the development of efficient compression techniques for boundary integral equations based upon wavelets. In [37, 67, 78] it was shown that for a wide class of boundary integral operators a wavelet basis can be chosen so that the full accuracy of the Galerkin discretization can be retained at a computational work of order N (possibly with a logarithmic factor in some studies), where N is the number of degrees of freedom used in the discretization. First nontrivial implementations of

The work in this chapter is a joint work with Rob Stevenson, see Section 1.2

these algorithms and performance tests were reported in [53, 59].

The main reason why a stiffness matrix entry is small is that the kernel of the involved integral operator is increasingly smooth away from its diagonal, and that the wavelets have vanishing moments, meaning that they are L_2-orthogonal to all polynomials up to a certain degree. Another advantage of a wavelet-Galerkin discretization is that the diagonally scaled stiffness matrices are well-conditioned uniformly in their sizes, guaranteeing a uniform convergence rate of iterative methods for the linear systems. Finally, as we have seen in the foregoing chapters, recent developments suggest a natural use of wavelets in adaptive discretization methods that approximate the solution using, up to a constant factor, as few degrees of freedom as possible.

Let $H^t(\Gamma)$ be the usual Sobolev space defined on a sufficiently smooth n-dimensional manifold $\Gamma \subset \mathbb{R}^{n+1}$, and let $H^{-t}(\Gamma)$ be its dual space. Then we consider the problem of finding the solution $u \in H^t(\Gamma)$ of

$$Lu = g,$$

where $L : H^t(\Gamma) \to H^{-t}(\Gamma)$ is a boundedly invertible linear operator, and $g \in H^{-t}(\Gamma)$. We will think of this problem as being the result of a variational formulation of a strongly elliptic boundary integral equation of order $2t$. With Ψ being a Riesz basis for $H^t(\Gamma)$, we can transform it into an equivalent infinite matrix-vector problem

$$\mathbf{Mu} = \mathbf{g},$$

where $\mathbf{M} : \ell_2 \to \ell_2$ is boundedly invertible, and $\mathbf{g}, \mathbf{u} \in \ell_2$.

Now the discussion in Introduction of the preceding chapter applies: One requires $\mathbf{M}$ to be s^*-computable for some $s^* > \frac{d-t}{n}$.

As we indicated in the preceding chapter, s^*-compressibility for some $s^* > \frac{d-t}{n}$ has been demonstrated in [86] for both differential and singular integral operators, and piecewise polynomial wavelets that are sufficiently smooth and have sufficiently many vanishing moments.

Only in the special case of a differential operator with constant coefficients, entries of $\mathbf{M}$ can be computed exactly, in $\mathcal{O}(1)$ operations, so that s^*-compressibility immediately implies s^*-computability. In general, numerical quadrature is required to approximate the entries. In the present chapter, considering singular integral operators resulting from the boundary integral method, we will show that $\mathbf{M}$ *is s^*-computable for the same value of s^* as it was shown to be s^*-compressible.* Summarizing, this result shows that using the routine **APPLY** as in Algorithm 2.7.9, the compression rules from [86] (recalled in Theorem 8.2.4), and the quadrature schemes derived in this paper to approximately compute the remaining entries, the adaptive wavelet methods from e.g. Chapter 2, 3, and

5 now define *fully discrete* algorithms that achieve the optimal computational complexity for the given problem.

We split our task into two parts. First we derive a criterion on the accuracy-work balance of a numerical quadrature scheme to approximate any entry of $\mathbf{M}$, such that, for a suitable choice of the work invested in approximating the entries of the compressed matrix $\mathbf{M}_j$ as function of both wavelets involved, we obtain an approximation $\mathbf{M}_j^*$ of which the computation of each column requires $\mathcal{O}(j^c 2^j)$ operations with a fixed constant c, and $\|\mathbf{M}_j - \mathbf{M}_j^*\| \leq 2^{-js^*}$, meaning that, on account of Lemma 2.7.12 on page 34 with a slight adjustment, $\mathbf{M}$ is s^*-computable. Second, we show that for any desired $s^* > 0$ we can fulfill the above criterion by the application of certain quadrature rules of variable order.

In view of Proposition 7.2.3 on page 123, as well as to control the number of function evaluations that are required, in this paper we will consider families $(Q_p)_{p\in\mathbb{N}}$ of composite quadrature rules $Q_p : f \mapsto \sum_{\Omega'\in\mathcal{O}} \sum_j w_j^{p,\Omega'} f(x_j^{p,\Omega'})$ of order p with a fixed mesh $\mathcal{O}$, that are *admissible* meaning that they satisfy

$$\sup_{p\in\mathbb{N},\Omega'\in\mathcal{O}} \max\left\{ \frac{\sum_j |w_j^{p,\Omega'}|}{\mathrm{vol}(\Omega')}, \frac{\#x_j^{p,\Omega'}}{p^n} \right\} < \infty.$$

Note that the bound on the number of abscissae in each subdomain is reasonable because the space of polynomials of total degree $p-1$ has $\binom{p-1+n}{n} \leq p^n$ degrees of freedom. Moreover, for a quadrature mesh $\mathcal{O}$ we define the following quantity

$$C_{\mathcal{O}} := \sup_{\Omega'\in\mathcal{O}} \frac{(\#\mathcal{O})^{1/n}\mathrm{rad}(\Omega')}{\mathrm{diam}(\Omega)}. \tag{8.1.1}$$

This chapter is organized as follows. In Section 8.2, assumptions are formulated on the singular integral operator and the wavelets, and the result concerning s^*-compressibility is recalled from [86]. Then in Section 8.3, rules for the numerical approximation of the entries of the stiffness matrix are derived, with which s^*-computability for some $s^* > \frac{d-t}{n}$ will be demonstrated.

At the end of this introduction, we fix a few more notations. A monomial of n variables is conveniently written using a *multi-index* $\alpha \in \mathbb{N}_0^n$ as $x^\alpha := x_1^{\alpha_1} \dots x_n^{\alpha_n}$. Likewise we write partial differentiation operators, that is, $\partial^\alpha := \partial_1^{\alpha_1} \dots \partial_n^{\alpha_n}$. We set $|\alpha| := \alpha_1 + \ldots + \alpha_n$, and the relation $\alpha \leq \beta$ is defined as $\alpha_i \leq \beta_i$ for all $i \in \overline{1,n}$. We have $|\alpha \pm \beta| = |\alpha| \pm |\beta|$ provided that $\alpha - \beta \in \mathbb{N}_0^n$ in case of subtraction. Binomial coefficients are naturally defined as $\binom{\alpha}{\beta} := \binom{\alpha_1}{\beta_1} \dots \binom{\alpha_n}{\beta_n}$.

8.2 Compressibility

For some $\mu \in \mathbb{N}$, let Γ be a patchwise smooth, compact n-dimensional, globally $C^{\mu-1,1}$ manifold in $\mathbb{R}^{n+1}$. Following [34], we assume that $\Gamma = \cup_{q=1}^M \overline{\Gamma_q}$, with

$\Gamma_q \cap \Gamma_{q'} = \varnothing$ when $q \neq q'$, and that for each $1 \leq q \leq M$, there exists

- a domain $\Omega_q \subset \mathbb{R}^n$, and a C^∞-parametrization $\kappa_q : \mathbb{R}^n \to \mathbb{R}^{n+1}$ with $\mathrm{Im}(\kappa_q|_{\Omega_q}) = \Gamma_q$,
- a domain $\mathbb{R}^n \supset \hat{\Omega}_q \supset\supset \Omega_q$, and an extension of $\kappa_q|_{\Omega_q}$ to a $C^{\mu-1,1}$ parametrization $\hat{\kappa}_q : \hat{\Omega}_q \to \mathrm{Im}(\hat{\kappa}_q) \subset \Gamma$.

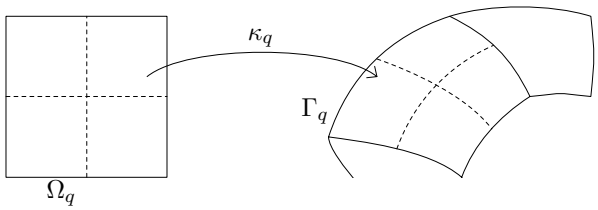

Figure 8.1: *Parametrization of the manifold.*

Formally supposing that the domains Ω_q are pairwise disjoint, for notational convenience we introduce the invertible mapping $\kappa : \cup_q \Omega_q \to \cup_q \Gamma_q \subset \Gamma$ via

$$\kappa(x) := \kappa_q(x) \quad \text{with } q \text{ such that } x \in \Omega_q.$$

For $|s| \leq \mu$, the Sobolev spaces $H^s(\Gamma)$ are well-defined, where for $s < 0$, $H^s(\Gamma)$ is the dual of $H^{-s}(\Gamma)$. Let

$$\Psi = \{\psi_\lambda : \lambda \in \Lambda\}$$

be a *Riesz basis for* $H^t(\Gamma)$ of wavelet type. The index λ encodes both the level, denoted by $|\lambda| \in \mathbb{N}_0$, and the location of the wavelet ψ_λ. We will assume that the wavelets are *local* and *piecewise smooth* with respect to nested subdivisions in the following sense. We assume that there exists a sequence $(\mathcal{O}_\ell)_{\ell \in \mathbb{N}_0}$ of collections $\mathcal{O}_\ell$ of disjoint uniformly Lipschitz domains $\Theta \in \mathcal{O}_\ell$, with

$$\mathrm{diam}(\Theta) \eqsim 2^{-\ell} \qquad \text{and} \qquad \mathrm{vol}(\Theta) \eqsim 2^{-n\ell}, \tag{8.2.1}$$

and where each $\Theta \in \mathcal{O}_\ell$ is contained in some Ω_q, and its closure is the union of the closures of a uniformly bounded number of subdomains from $\mathcal{O}_{\ell+1}$. For a precise definition of a collection of sets to be a collection of uniformly Lipschitz domains, we refer to Remark 7.3.1. Defining the collections of *panels*

$$\mathcal{G}_\ell := \{\kappa(\Theta) : \Theta \in \mathcal{O}_\ell\}, \qquad (\ell \in \mathbb{N}_0),$$

we assume that $\Gamma = \cup_{\Pi\in\mathcal{G}_\ell}\overline{\Pi}$, $(\ell \in \mathbb{N}_0)$, and that for each $\lambda \in \Lambda$ there exists a subcollection $\mathcal{G}_\lambda \subset \mathcal{G}_{|\lambda|}$ with

$$\sup_{\lambda\in\Lambda} \#\mathcal{G}_\lambda < \infty \quad \text{and} \quad \sup_{\ell\in\mathbb{N}_0,\Pi\in\mathcal{G}_\ell} \#\{\lambda : |\lambda| = \ell,\, \Pi \in \mathcal{G}_\lambda\} < \infty,$$

such that $\operatorname{supp}\psi_\lambda = \cup_{\Pi\in\mathcal{G}_\lambda} \operatorname{clos}\Pi$, being a connected set, and that on each $\Theta \in \kappa^{-1}(\mathcal{G}_\lambda)$, the pull-back $\hat{\psi}_{\lambda,\Theta} := (\psi_\lambda \circ \kappa)|_\Theta$ is smooth with

$$\sup_{x\in\Theta} |\partial^\beta \hat{\psi}_{\lambda,\Theta}(x)| \lesssim 2^{(|\beta|+\frac{n}{2}-t)|\lambda|} \qquad \text{for } \beta \in N_0^n. \tag{8.2.2}$$

We assume that the wavelets have the so-called *cancellation property of order* $\tilde{d} \in \mathbb{N}$, saying that there exists a constant $\eta > 0$, such that for any $p \in [1,\infty]$, for all continuous, patchwise smooth functions v and $\lambda \in \Lambda$,

$$|\langle v, \psi_\lambda\rangle| \lesssim 2^{-|\lambda|(\frac{n}{2}-\frac{n}{p}+t+\tilde{d})} \max_{1\le q\le M} |v|_{W_p^{\tilde{d}}(B(\operatorname{supp}\psi_\lambda;2^{-|\lambda|}\eta)\cap\Gamma_q)}, \tag{8.2.3}$$

where for $A \subset \mathbb{R}^{n+1}$ and $\varepsilon > 0$, $B(A;\varepsilon) := \{y \in \mathbb{R}^{n+1} : \operatorname{dist}(A,y) < \varepsilon\}$.

Furthermore, for some $k \in \mathbb{N}_0 \cup \{-1\}$, with $k < \mu$ and

$$\gamma := k + \tfrac{3}{2} > t, \tag{8.2.4}$$

we assume that all $\psi_\lambda \in C^k(\Gamma)$, where $k = -1$ means no global continuity condition, and that for all $r \in [-\tilde{d},\gamma)$, $s < \gamma$, necessarily with $|s|, |r| \le \mu$,

$$\|\cdot\|_{H^r(\Gamma)} \lesssim 2^{\ell(r-s)} \|\cdot\|_{H^s(\Gamma)} \qquad \text{on } W_\ell := \operatorname{span}\{\psi_\lambda : |\lambda| = \ell\}. \tag{8.2.5}$$

Inside a patch, a similar property can be required for larger ranges: For all $q \in \overline{1,M}$, and $r \in [-\tilde{d},\gamma)$, $s < \gamma$, we assume that

$$\|\cdot\|_{H^r(\Gamma_q)} \lesssim 2^{\ell(r-s)} \|\cdot\|_{H^s(\Gamma_q)} \quad \text{on } \operatorname{span}\{\psi_\lambda : |\lambda| = \ell, B(\operatorname{supp}\psi_\lambda; 2^{-\ell}\eta) \subset \overline{\Gamma_q}\}. \tag{8.2.6}$$

Remark 8.2.1. Wavelets that satisfy the assumptions in principle for any d, $\tilde{d}$ and smoothness permitted by both d and the regularity of the manifold were constructed in [34]. Apart from this construction, all known approaches based on non-overlapping domain decompositions yield wavelets which over the interfaces between patches are only continuous. With the constructions from [14, 22, 33], biorthogonality was realized with respect to a modified $L_2(\Gamma)$-scalar product. As a consequence, with the interpretation of functions as functionals via the Riesz mapping with respect to the standard $L_2(\Gamma)$ scalar product, for negative t the wavelets only generate a Riesz basis for $H^t(\Gamma)$ when $t > -\frac{1}{2}$, and likewise wavelets with supports that extend to more than one patch generally have no cancellation

properties in the sense of (8.2.3). Recently in [85], this difficulty was overcame, and wavelets were constructed that all have the cancellation property of the full order, and that generate Riesz bases for the full range of Sobolev spaces $H^t(\Gamma)$ that is allowed by continuous gluing of functions over the patch interfaces *and* the regularity of the manifold.

For some $|t| \leq \mu$, let L be a bounded operator from $H^t(\Gamma) \to H^{-t}(\Gamma)$, where we have in mind a singular integral operator of order $2t$. We assume that the operator L is defined by

$$Lu(z) = \int_\Gamma K(z, z')u(z')d\Gamma_{z'}, \quad (z \in \Gamma), \tag{8.2.7}$$

and that its *local kernel function*

$$\hat{K}(x, x') := K(\kappa(x), \kappa(x')) \cdot |\partial\kappa(x)| \cdot |\partial\kappa(x')|$$

satisfies for all $x, x' \in \cup_{1\leq q\leq M}\Omega_q$, and $\alpha, \beta \in \mathbb{N}_0^n$,

$$|\partial_x^\alpha \partial_{x'}^\beta \hat{K}(x, x')| \lesssim \frac{|\alpha + \beta|!}{\varsigma^{|\alpha+\beta|}} \cdot \mathrm{dist}(\kappa(x), \kappa(x'))^{-(n+2t+|\alpha+\beta|)}, \tag{8.2.8}$$

with a constant $\varsigma > 0$ (cf. [37, 53]), provided that $n + 2t + |\alpha + \beta| > 0$. If the kernel function $K(z, z')$ contains non-integrable singularities, the integral (8.2.7) has to be understood in the *finite part* sense of Hadamard, see e.g. [73, 80]. Following [37], we emphasize that (8.2.8) requires patchwise smoothness but no global smoothness of Γ. Only assuming global Lipschitz continuity of Γ, the local kernel of any standard boundary integral operator of order $2t$ can be shown to satisfy (8.2.8).

We assume that for some $\sigma \in (0, \mu - |t|]$, both L and its adjoint L' are bounded from $H^{t+\sigma}(\Gamma) \to H^{-t+\sigma}(\Gamma)$.

Remark 8.2.2. If Γ is a C^∞-manifold, then these boundary integral operators are known to be pseudo-differential operators, meaning that for any $\sigma \in \mathbb{R}$ they define bounded mappings from $H^{t+\sigma}(\Gamma) \to H^{-t+\sigma}(\Gamma)$. For Γ being only Lipschitz continuous, for the classical boundary integral equations it is known that $L : H^{t+\sigma}(\Gamma) \to H^{-t+\sigma}(\Gamma)$ is bounded for the maximum possible value $\sigma = 1 - |t|$ (cf. [23]). With increasing smoothness of Γ one may expect this boundedness for larger values of σ. Results in this direction can be found in [60].

Furthermore, with $\tilde{H}^s(\Gamma_q) := \begin{cases} H^s(\Gamma_q) & \text{when } s \geq 0, \\ (H_0^{-s}(\Gamma_q))' & \text{when } s < 0, \end{cases}$ we assume that there exists a $\tau \in (0, \mu - |t|]$ such that

$$L : H^{t+\tau}(\Gamma) \to \tilde{H}^{-t+\tau}(\Gamma_q) \quad \text{is bounded for all } 1 \leq q \leq M. \tag{8.2.9}$$

Remark 8.2.3. Since for any $|s| \leq \mu$, the restriction of functions on Γ to Γ_q is a bounded mapping from $H^s(\Gamma)$ to $\tilde{H}^s(\Gamma_q)$, from the boundedness of $L : H^{t+\sigma}(\Gamma) \to H^{-t+\sigma}(\Gamma)$, it follows that in any case (8.2.9) is valid for $\tau = \sigma$. So for example for Γ being a C^∞-manifold, (8.2.9) is valid for any $\tau \in \mathbb{R}$. Yet, in particular when $t < 0$, for Γ being less smooth it might happen that (8.2.9) is valid for a τ that is strictly larger than any σ for which $L : H^{t+\sigma}(\Gamma) \to H^{-t+\sigma}(\Gamma)$ is bounded.

In the following theorem, we recall the main result on compressibility for boundary integral operators from [86].

Theorem 8.2.4. *For Ψ being a Riesz basis for $H^t(\Gamma)$ as described above with $t + \tilde{d} > 0$, and $\tilde{d} > \gamma - 2t$, let $\mathbf{M} = \langle \Psi, L\Psi \rangle$.*

Let $\alpha \in (\frac{1}{2}, 1)$ and $b_i := (1+i)^{-1-\varepsilon}$ for some $\varepsilon > 0$. Choose k satisfying

$$\begin{aligned} &k = \frac{1}{n-1} && \text{when } n > 1, \\ &k > \frac{\min\{t+\tilde{d}, \tau\}}{\gamma - t} \quad \text{and} \quad k \geq \max\{1, \frac{\min\{t+\tilde{d}, \tau\}}{\min\{t+\mu, \sigma\}}\} && \text{when } n = 1. \end{aligned} \tag{8.2.10}$$

We define the infinite matrix $\mathbf{M}_j$ for $j \in \mathbb{N}$ by replacing all entries $\mathbf{M}_{\lambda,\lambda'} = \langle \psi_\lambda, L\psi_{\lambda'} \rangle$ by zeros when

$$\big||\lambda| - |\lambda'|\big| > jk, \quad \text{or} \tag{8.2.11}$$

$$\big||\lambda| - |\lambda'|\big| \leq j/n \quad \text{and} \quad \delta(\lambda, \lambda') \geq \max\{3\eta, 2^{\alpha(j/n - ||\lambda|-|\lambda'||)}\}, \quad \text{or} \tag{8.2.12}$$

$$\begin{aligned} &\big||\lambda| - |\lambda'|\big| > j/n \quad \text{and} \\ &\qquad \tilde{\delta}(\lambda, \lambda') \geq \max\{2^{n(j/n - ||\lambda|-|\lambda'||)} b_{||\lambda|-|\lambda'||-j/n}, 2\eta 2^{-||\lambda|-|\lambda'||}\}, \end{aligned} \tag{8.2.13}$$

where

$$\delta(\lambda, \lambda') := 2^{\min\{|\lambda|,|\lambda'|\}} \operatorname{dist}(\operatorname{supp} \psi_\lambda, \operatorname{supp} \psi_{\lambda'}), \tag{8.2.14}$$

and

$$\tilde{\delta}(\lambda, \lambda') := 2^{\min\{|\lambda|,|\lambda'|\}} \times \begin{cases} \operatorname{dist}(\operatorname{supp} \psi_\lambda, \operatorname{sing\,supp} \psi_{\lambda'}) & \text{when } |\lambda| > |\lambda'|, \\ \operatorname{dist}(\operatorname{sing\,supp} \psi_\lambda, \operatorname{supp} \psi_{\lambda'}) & \text{when } |\lambda| < |\lambda'|, \end{cases}$$

and η is from (8.2.3).

Then the number of non-zero entries in each column of $\mathbf{M}_j$ is of order 2^j, and for any

$$s \leq \min\left\{\tfrac{t+\tilde{d}}{n}, \tfrac{\tau}{n}\right\}, \text{ with } s < \tfrac{\gamma - t}{n-1}, \ s \leq \tfrac{\sigma}{n-1} \text{ and } s \leq \tfrac{\mu+t}{n-1} \text{ when } n > 1,$$

it holds that $\|\mathbf{M} - \mathbf{M}_j\| \lesssim 2^{-js}$. We conclude that $\mathbf{M}$ is s^*-compressible, as defined in Definition 2.7.11, with $s^* = \min\{\frac{t+\tilde{d}}{n}, \frac{\tau}{n}, \frac{\sigma}{n-1}, \frac{\gamma-t}{n-1}, \frac{\mu+t}{n-1}\}$ when $n > 1$, and $s^* = \min\{t + \tilde{d}, \tau\}$ when $n = 1$.

From this theorem we infer that if $\tilde{d} > d - 2t$, $\tau > d - t$ and, when $n > 1$, $\frac{\min\{\gamma - t, \sigma, t + \mu\}}{n-1} > \frac{d-t}{n}$, then $s^* > \frac{d-t}{n}$ as required. For $n > 1$, the condition involving γ is satisfied for instance for spline wavelets, where $\gamma = d - \frac{1}{2}$, in case $\frac{d-t}{n} > \frac{1}{2}$.

If each entry of $\mathbf{M}$ can be exactly computed in $\mathcal{O}(1)$ operations, then s^*-compressibility implies s^*-computability, as defined in Definition 2.7.8, and so, when indeed $s^* > \frac{d-t}{n}$, it implies the optimal computational complexity of the adaptive wavelet schemes from the earlier chapters. In general, one is not able to compute the matrix entries exactly. What is more, it is far from obvious how to compute the entries of $\mathbf{M}_j$ sufficiently accurate while keeping the average computational expense per entry in each column uniformly bounded. In the next section, additionally assuming that the wavelets are essentially *piecewise polynomials*, we will show that it is possible to arrange quadrature schemes which admit s^*-computability of $\mathbf{M}$.

8.3 Computability

In this section, we will present a numerical integration scheme which computes an approximation $\mathbf{M}_j^*$ of $\mathbf{M}_j$ such that, for some specified constant c, by spending $\mathcal{O}(j^c 2^j)$ computational work per column of $\mathbf{M}_j^*$, the approximation error satisfies $\|\mathbf{M}_j - \mathbf{M}_j^*\| \lesssim 2^{-js^*}$ with s^* given by Theorem 8.2.4, implying that $\mathbf{M}$ is s^*-computable.

Let us consider the computation of individual entries

$$\mathbf{M}_{\lambda,\lambda'} = \int_\Gamma \psi_\lambda(z) \left(\int_\Gamma K(z, z') \psi_{\lambda'}(z') d\Gamma_{z'} \right) d\Gamma_z \tag{8.3.1}$$

of $\mathbf{M}$. Unless explicitly stated otherwise, throughout this section we assume that

$$|\lambda| \geq |\lambda'|.$$

We start with an assumption.

Assumption 8.3.1. For any $\Xi \in \mathcal{G}_\lambda$, $\Xi' \in \mathcal{G}_{|\lambda|}$ with $\Xi' \subset \operatorname{supp} \psi_{\lambda'}$, in the following we assume that the integral

$$\int_\Xi \int_{\Xi'} K(z, z') \psi_\lambda(z) \psi_{\lambda'}(z') d\Gamma_z d\Gamma_{z'}$$

is well-defined.

This assumption obviously holds in case of proper or improper integrals. However, it requires an appropriate interpretation of the integrals in case of strongly-

or hyper-singular kernels. For strongly singular kernels on surfaces in $\mathbb{R}^3$ the assumption was confirmed in [52].

As a consequence of the assumption, we may write

$$\mathbf{M}_{\lambda,\lambda'} = \sum_{\Pi\in\mathcal{G}_\lambda}\sum_{\Pi'\in\mathcal{G}_{\lambda'}} I_{\lambda\lambda'}(\Pi,\Pi'), \tag{8.3.2}$$

with, for $\Pi \in \mathcal{G}_\lambda$ and $\Pi' \in \mathcal{G}_{\lambda'}$,

$$I_{\lambda\lambda'}(\Pi,\Pi') := \sum_{\{\Xi'\in\mathcal{G}_{|\lambda|}:\Xi'\subset\Pi'\}} \int_\Pi\int_{\Xi'} K(z,z')\psi_\lambda(z)\psi_{\lambda'}(z')d\Gamma_z d\Gamma_{z'}. \tag{8.3.3}$$

We assume that for each $\Pi \in \mathcal{G}_\lambda$, $\Pi' \in \mathcal{G}_{\lambda'}$ an approximation of the integral $I_{\lambda\lambda'}(\Pi,\Pi')$ is obtained by some numerical scheme dependent on j, and using (8.3.2), that these approximations are used to assemble the matrix $\mathbf{M}_j^*$. The following theorem defines a criterion on the computational cost in relation to the accuracy of computing the integrals $I_{\lambda\lambda'}(\Pi,\Pi')$ so that s^*-compressibility implies s^*-computability.

Theorem 8.3.2. *Let $s^* > 0$ be any given constant, and $\mathbf{M}$, $\mathbf{M}_j$ be as in Theorem 8.2.4. Let $\sigma : \cup_\ell \mathcal{G}_\ell \to \mathbb{R}$ be some fixed function such that*

$$\sigma(\Xi) \eqsim \operatorname{diam}(\Xi) \qquad \text{for } \Xi \in \cup_\ell \mathcal{G}_\ell, \tag{8.3.4}$$

and let $d^*, e^* \in \mathbb{R}$ and $\varrho > 1$ be fixed constants. Assume that for any $p \in \mathbb{N}$, an approximation $I^*_{\lambda\lambda'}(\Pi,\Pi')$ of the integral $I_{\lambda\lambda'}(\Pi,\Pi')$ can be computed such that by spending the number of

$$W \lesssim p^{2n}(1 + ||\lambda| - |\lambda'||) \tag{8.3.5}$$

arithmetical operations, the error satisfies

$$|E_{\lambda\lambda'}(\Pi,\Pi')| \lesssim \varrho^{-p} 2^{||\lambda|-|\lambda'||d^*} \max\left\{1, \frac{\operatorname{dist}(\Pi,\Pi')}{\varrho\max\{\sigma(\Pi),\sigma(\Pi')\}}\right\}^{e^*-p}. \tag{8.3.6}$$

Then for any fixed $\vartheta \geq 0$, and for parameters θ and τ with

$$\theta \geq s^*/\log_2\varrho \quad \text{and} \quad \tau > (n/2 + d^*)/\log_2\varrho, \tag{8.3.7}$$

by choosing p for the computation of $I^*_{\lambda\lambda'}(\Pi,\Pi')$ as the smallest positive integer satisfying

$$p > e^* + n \quad \text{and} \quad p \geq j\theta + \tau||\lambda| - |\lambda'|| - \vartheta, \tag{8.3.8}$$

the so computed approximation $\mathbf{M}_j^*$ of $\mathbf{M}_j$ satisfies $\|\mathbf{M}_j - \mathbf{M}_j^*\| \lesssim 2^{-js^*}$, where the work for computing each column of $\mathbf{M}_j^*$ is $\mathcal{O}(j^{2n+1}2^j)$.

By taking s^* as given in Theorem 8.2.4, we conclude that the matrix $\mathbf{M}$ is s^*-computable for the same value of s^* as it was shown to be s^*-compressible.

The proof will use Schur's lemma that we recall here for the reader's convenience.

Schur's lemma. *If for a matrix* $\mathbf{A} = (a_{\lambda,\lambda'})_{\lambda,\lambda'\in\Lambda}$*, there is a sequence* $w_\lambda > 0$*,* $\lambda \in \Lambda$*, and a constant* C *such that*

$$\sum_{\lambda'\in\Lambda} w_{\lambda'}|a_{\lambda\,\lambda'}| \le w_\lambda C, \quad (\lambda\in\Lambda), \quad \textit{and} \quad \sum_{\lambda\in\Lambda} w_\lambda|a_{\lambda\,\lambda'}| \le w_{\lambda'}C, \quad (\lambda'\in\Lambda),$$

then $\|\mathbf{A}\| \le C$*.*

Proof (Proof of Theorem 8.3.2). Since $\#\mathcal{G}_\lambda$, $\#\mathcal{G}_{\lambda'} \lesssim 1$, it is sufficient to give the proof pretending that $\#\mathcal{G}_\lambda = \#\mathcal{G}_{\lambda'} = 1$.

With the matrix $(\Delta_{\lambda,\lambda'})_{\lambda,\lambda'\in\Lambda}$ defined by

$$\Delta_{\lambda,\lambda'} := \max\left\{1, \frac{\operatorname{dist}(\Pi,\Pi')}{\varrho\max\{\sigma(\Pi),\sigma(\Pi')\}}\right\}, \quad \Pi\in\mathcal{G}_\lambda,\ \Pi'\in\mathcal{G}_{\lambda'},$$

for each $\lambda\in\Lambda$, $\ell'\in\mathbb{N}_0$, and $\beta > n$, we can verify that

$$\sum_{|\lambda'|=\ell'} \Delta_{\lambda,\lambda'}^{-\beta} \lesssim 2^{n\max\{0,\ell'-|\lambda|\}}, \tag{8.3.9}$$

using the locality of the wavelets and the fact that $\sigma(\Pi') \eqsim \operatorname{diam}(\Pi') \eqsim 2^{-|\lambda'|}$ and that $\operatorname{vol}(\Pi') \eqsim 2^{-|\lambda'|n}$.

Denoting the entry (λ,λ') of the error matrix $\mathbf{M}_j - \mathbf{M}_j^*$ by $\varepsilon_{j,\lambda\lambda'}$, and by substituting $p \ge j\theta + \tau||\lambda| - |\lambda'|| - \vartheta$ into (8.3.6), we infer that

$$\varepsilon_{j,\lambda\lambda'} \lesssim 2^{-j\theta\log_2\varrho}2^{-||\lambda|-|\lambda'||(\tau\log_2\varrho - d^*)}\Delta_{\lambda,\lambda'}^{-(p-e^*)}. \tag{8.3.10}$$

Recall that $\sigma := \tau\log_2\varrho - d^* > n/2$ and $p - e^* > n$. Applying Schur's lemma to the error matrix $\mathbf{M}_j - \mathbf{M}_j^*$ with weights $w_\lambda = 2^{-|\lambda|n/2}$, we have

$$\begin{aligned}
w_\lambda^{-1}\sum_{\lambda'} w_{\lambda'}|\varepsilon_{j,\lambda\lambda'}| &\lesssim 2^{-j\theta\log_2\varrho}2^{|\lambda|n/2}\sum_{\ell'\ge 0} 2^{-\ell' n/2}2^{-(|\lambda|-\ell')\sigma}\cdot\sum_{|\lambda'|=\ell'}\Delta_{\lambda,\lambda'}^{-(p-e^*)}\\
&\lesssim 2^{-j\theta\log_2\varrho}2^{|\lambda|n/2}\sum_{0\le\ell'\le|\lambda|} 2^{-\ell' n/2}2^{-(|\lambda|-\ell')\sigma}\cdot 1\\
&\quad + 2^{-j\theta\log_2\varrho}2^{|\lambda|n/2}\sum_{\ell'>|\lambda|} 2^{-\ell' n/2}2^{-(\ell'-|\lambda|)\sigma}\cdot 2^{(\ell'-|\lambda|)n}\\
&\lesssim 2^{-j\theta\log_2\varrho},
\end{aligned}$$

where we used (8.3.9) in the second step. Now by the symmetry of the estimate (8.3.10) in λ and λ', we conclude that the error in the computed matrix $\mathbf{M}_j^*$ satisfies

$$\|\mathbf{M}_j - \mathbf{M}_j^*\| \lesssim 2^{-j\theta\log_2\varrho} \le 2^{-js^*}.$$

The work for computing the entry $(\mathbf{M}_j^*)_{\lambda,\lambda'}$ is of order

$$p(j,\lambda,\lambda')^{2n}(1+||\lambda|-|\lambda'||) \lesssim (j\theta+\tau||\lambda|-|\lambda'||)^{2n}(1+||\lambda|-|\lambda'||).$$

Since $\mathbf{M}_j^*$ contains nonzero entries only for $||\lambda|-|\lambda'|| \le jk$, we can bound the work for computing each element $(\mathbf{M}_j^*)_{\lambda,\lambda'}$ by a constant multiple of j^{2n+1}. Now using the fact that each column of $\mathbf{M}_j$ contains $\mathcal{O}(2^j)$ nonzero entries, we conclude the computational work per column is $\mathcal{O}(j^{2n+1}2^j)$. ∎

By applying the error estimates from Section 7.2, we will now show how numerical quadrature schemes satisfying (8.3.5) and (8.3.6) can be realized. We will consider variable order quadrature rules, meaning that constants absorbed by the "$\lesssim$" symbol will not depend on the quadrature order. To this end, we consider a general finite subdivision $\Upsilon \subset (\cup_\ell \mathcal{G}_\ell)^2$ of the integration domain $\Pi\times\Pi'$ such that $\{\Xi\times\Xi' \in \Upsilon : \operatorname{dist}(\Xi,\Xi')=0\} \subset \mathcal{G}^2_{|\lambda|}$. Then in view of Assumption 8.3.1, we can split the integral (8.3.3) as

$$I_{\lambda\lambda'}(\Pi,\Pi') = \sum_{\Xi\times\Xi'\in\Upsilon} I_{\lambda\lambda'}(\Xi,\Xi'), \tag{8.3.11}$$

with

$$I_{\lambda\lambda'}(\Xi,\Xi') := \int_\Xi\int_{\Xi'} K(z,z')\psi_\lambda(z)\psi_{\lambda'}(z')d\Gamma_z d\Gamma_{z'}.$$

First we will study the numerical evaluation of an individual integral $I(\Xi,\Xi')$ for the case that $\operatorname{dist}(\Xi,\Xi')>0$. We can write the integral $I(\Xi,\Xi')$ in local coordinates

$$I_{\lambda\lambda'}(\Xi,\Xi') = \int_\Theta\int_{\Theta'} \hat{K}(x,x')\hat{\psi}_{\lambda,\kappa^{-1}(\Pi)}(x)\hat{\psi}_{\lambda',\kappa^{-1}(\Pi')}(x')dxdx', \tag{8.3.12}$$

where $\Theta=\kappa^{-1}(\Xi)$ and $\Theta'=\kappa^{-1}(\Xi')$.

Definition 8.3.3. The wavelet basis Ψ is said to be of *P-type of order e* when for all $\lambda\in\Lambda$ and $\Theta\in\mathcal{O}_{|\lambda|}$, $\hat{\psi}_{\lambda,\Theta}\in P_{e-1}(\Theta)$. Similarly, Ψ is of *Q-type of order e* when for all $\lambda\in\Lambda$ and $\Theta\in\mathcal{O}_{|\lambda|}$, Θ is an n-rectangle and $\hat{\psi}_{\lambda,\Theta}\in Q_{e-1}(\Theta)$. ⊘

Lemma 8.3.4. *Assume that the wavelet basis Ψ is of P-type of order e and that* $\operatorname{dist}(\kappa(\Theta),\kappa(\Theta'))>0$. *For the domains Θ and Θ', we employ composite quadrature rules from admissible families (uniformly in Θ,Θ') of orders p and fixed ranks N, and apply the product of these quadrature rules to approximate the non-singular integral $I_{\lambda\lambda'}(\kappa(\Theta),\kappa(\Theta'))$ from* (8.3.12). *We define*

$$\sigma(\kappa(\tilde{\Theta})) := \frac{nC}{\varsigma N^{1/n}}\operatorname{diam}(\tilde{\Theta}) \qquad \text{for all } \tilde{\Theta}\in\cup_\ell\mathcal{O}_\ell, \tag{8.3.13}$$

where $\varsigma > 0$ is the constant involved in the Calderon-Zygmund estimate (8.2.8)*, and C is an upper bound on the quantity* (8.1.1) *for quadrature meshes on $\tilde{\Theta} \in \cup_\ell \mathcal{O}_\ell$. Then with*

$$\omega := \frac{\operatorname{dist}(\kappa(\Theta), \kappa(\Theta'))}{\max\{\sigma(\kappa(\Theta)), \sigma(\kappa(\Theta'))\}}, \tag{8.3.14}$$

for any $p \geq \max\{e-2t-n, e-1\}$, the quadrature error $E(\kappa(\Theta), \kappa(\Theta'))$ satisfies

$$\begin{aligned} |E(\Xi, \Xi')| \lesssim\ & 2^{||\lambda|-|\lambda'||(n/2-t)} \omega^{-(n+p)} \max\{1, \omega\}^{e-1} \\ & \times \min\{\sigma(\kappa(\Theta)), \sigma(\kappa(\Theta'))\}^n \operatorname{dist}(\kappa(\Theta), \kappa(\Theta'))^{-2t}. \end{aligned} \tag{8.3.15}$$

Proof. Since there will be no risk of confusion, we will write $\hat{\psi}_\lambda$ and $\hat{\psi}_{\lambda'}$ instead of $\hat{\psi}_{\lambda,\kappa^{-1}(\Pi)}$ and $\hat{\psi}_{\lambda',\kappa^{-1}(\Pi')}$, respectively. By Lemma 7.2.4, the error of the product quadrature is

$$|E(\kappa(\Theta), \kappa(\Theta'))| \leq \operatorname{vol}(\Theta') \cdot \sup_{x' \in \Theta'} |E(x')| + \operatorname{vol}(\Theta) \cdot \sup_{x \in \Theta} |E'(x)|, \tag{8.3.16}$$

where we denoted by $E(x')$ the error of the quadrature over the domain Θ with the integrand $x \mapsto \hat{K}(x, x')\hat{\psi}_\lambda(x)\hat{\psi}_{\lambda'}(x')$. Analogously $E'(x)$ denotes the error of the quadrature over Θ'. Using Proposition 7.2.3 to bound $E(x')$, we have

$$|E(x')| \lesssim \frac{n^p}{p!} C^p N^{-p/n} \operatorname{vol}(\Theta) \cdot \operatorname{diam}(\Theta)^p \cdot |\hat{\psi}_{\lambda'}(x')| \cdot |\hat{K}(\cdot, x')\hat{\psi}_\lambda|_{W^p_\infty(\Theta)}. \tag{8.3.17}$$

The partial derivatives with $|\eta| = p$, satisfy

$$\begin{aligned} \left|\partial_x^\eta \left(\hat{K}(x, x')\hat{\psi}_\lambda(x)\right)\right| &= \left|\sum_{\xi \leq \eta} \binom{\eta}{\xi} \partial_x^{\eta-\xi} \hat{K}(x, x') \partial_x^\xi \hat{\psi}_\lambda(x)\right| \\ &\leq \sum_{\{\xi \leq \eta : |\xi| \leq e-1\}} \binom{\eta}{\xi} \left|\partial_x^{\eta-\xi} \hat{K}(x, x') \partial_x^\xi \hat{\psi}_\lambda(x)\right|, \end{aligned}$$

since $\partial^\xi \hat{\psi}_\lambda$ can only be nonzero when $|\xi| \leq e-1$ because $\hat{\psi}_\lambda \in P_{e-1}$. Applying the estimates (8.2.2) and (8.2.8) we have, with $\delta := \operatorname{dist}(\kappa(\Theta), \kappa(\Theta'))$

$$\begin{aligned} &|\hat{K}(\cdot, x')\hat{\psi}|_{W^p_\infty(\Theta)} \\ &\qquad \lesssim \max_{|\eta|=p} \sum_{\{\xi \leq \eta : |\xi| \leq e-1\}} \binom{\eta}{\xi} \frac{(p-|\xi|)!}{\varsigma^{p-|\xi|}} \delta^{-(n+2t+p-|\xi|)} 2^{(|\xi|+n/2-t)|\lambda|} \\ &\qquad \lesssim 2^{|\lambda|(n/2-t)} \delta^{-(n+2t+p)} \max_{|\eta|=p} \sum_{\{\xi \leq \eta : |\xi| \leq e-1\}} \binom{\eta}{\xi} \frac{(p-|\xi|)!}{\varsigma^{p-|\xi|}} \left(2^{|\lambda|}\delta\right)^{|\xi|} \\ &\qquad \lesssim \frac{p!}{\varsigma^p} \cdot 2^{|\lambda|(n/2-t)} \delta^{-(n+2t+p)} \cdot \max\{1, 2^{|\lambda|}\delta\}^{e-1}, \end{aligned}$$

where $\binom{\eta}{\xi}(p-|\xi|)! \le p!$ was used. By substituting this result into (8.3.17), setting $c := nC/(\varsigma N^{1/n})$, and using $\mathrm{vol}(\Theta) \lesssim \mathrm{diam}(\Theta)^n$, $\mathrm{vol}(\Theta') \lesssim \mathrm{diam}(\Theta')^n$, and again (8.2.2), we get

$$\begin{aligned}
\mathrm{vol}(\Theta') \sup_{x' \in \Theta'} |E(x')| &\lesssim \mathrm{diam}(\Theta')^n c^p \,\mathrm{diam}(\Theta)^{n+p} \cdot 2^{(|\lambda|+|\lambda'|)(n/2-t)} \\
&\quad \times \delta^{-(n+2t+p)} \max\{1, 2^{|\lambda|}\delta\}^{e-1} \\
&= \mathrm{diam}(\Theta')^n \,\mathrm{diam}(\Theta)^{n+p} \cdot 2^{(|\lambda|+|\lambda'|)(n/2-t)} c^{-n} \delta^{-2t} \omega^{-n-p} \\
&\quad \times \max\{\mathrm{diam}(\Theta), \mathrm{diam}(\Theta')\}^{-n-p} \max\{1, 2^{|\lambda|}\delta\}^{e-1} \\
&= c^{-n} 2^{(|\lambda|+|\lambda'|)(n/2-t)} \delta^{-2t} \omega^{-n-p} \min\{\mathrm{diam}(\Theta), \mathrm{diam}(\Theta')\}^n \\
&\quad \times \left(\frac{\mathrm{diam}(\Theta)}{\max\{\mathrm{diam}(\Theta), \mathrm{diam}(\Theta')\}} \right)^p \max\{1, 2^{|\lambda|}\delta\}^{e-1},
\end{aligned}$$

by definition of ω. For the expression in the last row, employing the inequalities

$$\left(\frac{\mathrm{diam}(\Theta)}{\max\{\mathrm{diam}(\Theta), \mathrm{diam}(\Theta')\}} \right)^p \le 1,$$

and

$$\begin{aligned}
&\left(\frac{\mathrm{diam}(\Theta)}{\max\{\mathrm{diam}(\Theta), \mathrm{diam}(\Theta')\}} \right)^p \left(2^{|\lambda|}\delta\right)^{e-1} = \left(\frac{\mathrm{diam}(\Theta)}{2^{-|\lambda|}} \right)^{e-1} \\
&\quad \times \left(\frac{\delta}{\max\{\mathrm{diam}(\Theta), \mathrm{diam}(\Theta')\}} \right)^{e-1} \left(\frac{\mathrm{diam}(\Theta)}{\max\{\mathrm{diam}(\Theta), \mathrm{diam}(\Theta')\}} \right)^{p-e+1} \\
&\lesssim \omega^{e-1},
\end{aligned}$$

and taking the maximum over these two, the assertion of the lemma is proven for the first term in (8.3.16). The remaining second term in (8.3.16) can be estimated exactly in the same fashion by interchanging the roles of λ and λ'. ∎

Obviously, if Ψ is of Q-type of order e, then it is also of P-type of order $n(e-1)+1$. In the next lemma, however, we will see that product quadrature rules are quantitatively more efficient for Q-type wavelets.

Lemma 8.3.5. *Assume that the wavelet basis Ψ is of Q-type of order e and that $\mathrm{dist}(\kappa(\Theta), \kappa(\Theta')) > 0$. For the domains Θ and Θ', we employ composite product quadrature rules of orders p and fixed ranks N as in Corollary 7.2.6, and apply the product of these quadrature rules to approximate the non-singular integral $I_{\lambda\lambda'}(\kappa(\Theta), \kappa(\Theta'))$ from (8.3.12). We define*

$$\sigma(\kappa(\tilde{\Theta})) := \frac{1}{2\varsigma N^{1/n}} \tilde{l} \qquad \text{for all } \tilde{\Theta} \in \cup_\ell \mathcal{O}_\ell, \tag{8.3.18}$$

where $\tilde{l}$ is the maximum edge length of $\tilde{\Theta}$, and ς is the constant involved in the Calderon-Zygmund estimate (8.2.8)*. Then with*

$$\omega := \frac{\operatorname{dist}(\kappa(\Theta), \kappa(\Theta'))}{\max\{\sigma(\kappa(\Theta)), \sigma(\kappa(\Theta'))\}}, \tag{8.3.19}$$

for any $p \geq \max\{e-2t-n, e-1\}$, the quadrature error $E(\kappa(\Theta), \kappa(\Theta'))$ satisfies

$$\begin{aligned} |E(\Xi, \Xi')| \lesssim\ & 2^{||\lambda|-|\lambda'||(n/2-t)} \omega^{-(n+p)} \max\{1, \omega\}^{e-1} \\ & \times \min\{\sigma(\kappa(\Theta)), \sigma(\kappa(\Theta'))\}^n \operatorname{dist}(\kappa(\Theta), \kappa(\Theta'))^{-2t}. \end{aligned} \tag{8.3.20}$$

Proof. Adopting the notations from the previous proof, we use Corollary 7.2.6 to estimate $E(x')$.

$$|E(x')| \leq n \frac{2^{1-p}}{p!} N^{-p/n} l^{n+p} \cdot |\hat{\psi}_{\lambda'}(x')| \cdot \max_{j=\overline{1,n}} \left\| \partial_{x_j}^p \left(\hat{K}(x, x') \hat{\psi}_\lambda(x) \right) \right\|_{L_\infty(\Theta)}.$$

The partial derivative of order p along the j-th coordinate direction satisfies

$$\begin{aligned} \left| \partial_{x_j}^p \left(\hat{K}(x, x') \hat{\psi}_\lambda(x) \right) \right| &= \left| \sum_{k=0}^{p} \binom{p}{k} \partial_{x_j}^{p-k} \hat{K}(x, x') \partial_{x_j}^k \hat{\psi}_\lambda(x) \right| \\ &\leq \sum_{k=0}^{\min\{p, e-1\}} \binom{p}{k} \left| \partial_{x_j}^{p-k} \hat{K}(x, x') \partial_{x_j}^k \hat{\psi}_\lambda(x) \right|, \end{aligned}$$

since $\partial_{x_j}^k \hat{\psi}_\lambda(x)$ can only be nonzero when $k \leq e-1$ because $\hat{\psi}_\lambda \in Q_{e-1}$. Applying the estimates (8.2.2) and (8.2.8) we have, with $\delta := \operatorname{dist}(\kappa(\Theta), \kappa(\Theta))$

$$\begin{aligned} &\max_{j=\overline{1,n}} \|\hat{K}(\cdot, x') \hat{\psi}\|_{L_\infty(\Theta)} \\ &\qquad \lesssim 2^{|\lambda|(n/2-t)} \delta^{-(n+2t+p)} \sum_{k=0}^{\min\{p, e-1\}} \binom{p}{k} \frac{(p-k)!}{\varsigma^{p-k}} \left(2^{|\lambda|} \delta\right)^k \\ &\qquad \lesssim \frac{p!}{\varsigma^p} \cdot 2^{|\lambda|(n/2-t)} \delta^{-(n+2t+p)} \cdot \max\{1, 2^{|\lambda|} \delta\}^{e-1}. \end{aligned}$$

Further we can proceed as in the preceding proof. ∎

We now turn back to the computation of the integral $I_{\lambda\lambda'}(\Pi, \Pi')$ in (8.3.3). From Lemmata 8.3.4 and 8.3.5, we see that convergence of the quadrature rule as a function of the order p depends on the quantity ω, which is in essence the distance between the panels in terms of the size of the bigger panel. For panels Π and Π' that have a sufficiently large mutual distance, namely, when $\operatorname{dist}(\Pi, \Pi') >$

$\max\{\sigma(\Pi), \sigma(\Pi')\}$ and thus $\omega > 1$, it makes sense to apply quadrature directly on the domain $\Pi \times \Pi'$, that is, not to apply a further splitting as in (8.3.11).

For the integrals with $0 < \operatorname{dist}(\Pi, \Pi') \leq \max\{\sigma(\Pi), \sigma(\Pi')\}$, however, the subdivision Υ has to be nontrivial. By subdividing the integration domain $\Pi \times \Pi'$ in such a way that $\omega > 1$ for all individual integrals $I_{\lambda\lambda'}(\Xi, \Xi')$, we will ensure convergence of the numerical integration also for these integrals.

Finally, for the case that $\operatorname{dist}(\Pi, \Pi') = 0$, quadrature methods developed for standard Galerkin boundary elements cannot be applied directly in the wavelet setting, because the panels Π and Π' can have very different sizes. Therefore, our strategy here will be to split the bigger panel into smaller panels such that the resulting singular integrals are over panels of the same level, *and* such that the nonsingular integrals are arranged so that $\omega > 1$ for each of them. In view of these considerations, we consider Algorithm 8.3.6 for producing a subdivision of the product domain $\Pi \times \Pi'$.

Algorithm 8.3.6 Nonuniform subdivision of the product domain $\Pi \times \Pi'$

Parameters: Let $\rho > 0$ be given, and $\sigma : \cup_l \mathcal{G}_l \to \mathbb{R}$ be a function satisfying

$$\sigma(\Xi) \eqsim \operatorname{diam}(\Xi) \quad \text{uniformly in } \Xi \in \cup_l \mathcal{G}_l. \tag{8.3.21}$$

Input: $\Pi \in \mathcal{G}_\ell$ and $\Pi' \in \mathcal{G}_{\ell'}$ with $\ell \geq \ell'$.
Output: $\Upsilon \subset (\cup_l \mathcal{G}_l)^2$.

1: Set $\Upsilon := \varnothing$, $\Xi := \Pi$, $\Xi' := \Pi'$, and $\tilde{\ell} := \ell$, $\tilde{\ell}' := \ell'$;
2: If the pair Ξ and Ξ' does satisfy one of the conditions

$$\operatorname{dist}(\Xi, \Xi') \geq \rho \cdot \max\{\sigma(\Xi), \sigma(\Xi')\}, \tag{8.3.22}$$

or

$$\operatorname{dist}(\Xi, \Xi') = 0 \quad \text{and} \quad \Xi = \Pi,\ \Xi' \in \mathcal{G}_\ell, \tag{8.3.23}$$

accept the pair: $\Upsilon := \Upsilon \cup \{\Xi \times \Xi'\}$;If not, go to either step 3 or 4;
3: If $\tilde{\ell}' \leq \tilde{\ell}$, subdivide Ξ' into next level elements $\Xi'_i \in \mathcal{G}_{\tilde{\ell}'+1}$, and perform step 2 with $\tilde{\ell}' = \tilde{\ell}' + 1$, $\Xi' = \Xi'_i$ for each Ξ'_i;
4: If $\tilde{\ell}' > \tilde{\ell}$, subdivide Ξ into next level elements $\Xi_i \in \mathcal{G}_{\tilde{\ell}+1}$, and perform step 2 with $\tilde{\ell} = \tilde{\ell} + 1$, $\Xi = \Xi_i$ for each Ξ_i.

Remark 8.3.7. Algorithm 8.3.6 can already be found in, e.g., [53, 59, 67] with $\sigma(\Xi) = 2^{-\ell}$ for $\Xi \in \mathcal{G}_\ell$. This nonuniform subdivision effectively distributes the "strength" of the nearly singular behavior of the integrand over individual subdomains. In [59, 67] the value of ρ is fixed independent of the user and the subdivision Υ is of type $\Upsilon = \Xi \times \Upsilon'$, where Υ' is a subdivision of Ξ'. For the

algorithm in [53], as is the case for the version herein, the parameter ρ can be chosen by the user, and therefore the subdivision Υ is needed to be more general. Later we will see that the parameter ρ can be used to control the convergence rate of quadrature schemes based on the subdivision generated by Algorithm 8.3.6.

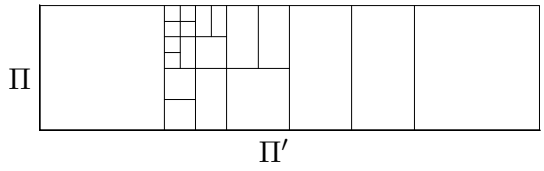

Figure 8.2: *A possible subdivision of* $\Pi \times \Pi'$ *generated by Algorithm 8.3.6:* $n = 1$, $\operatorname{dist}(\Pi, \Pi') = 0$ *and* $\Pi \cap \Pi' = \varnothing$.

Remark 8.3.8. Since the manifold is Lipschitz, and the subdivisions are nested and satisfy (8.2.1), one can verify that for any pair $\Xi, \Xi' \in \cup_\ell \mathcal{G}_\ell$ such that $\operatorname{dist}(\Xi, \Xi') > 0$,

$$\operatorname{dist}(\Xi, \Xi') \geq c_\Gamma \min\{\operatorname{diam}\Xi, \operatorname{diam}\Xi'\},$$

with the constant c_Γ depending only on the manifold Γ and its parametrization.

Theorem 8.3.9. *For any* $\Pi \times \Pi' \in \mathcal{G}_\ell \times \mathcal{G}_{\ell'}$ *with* $\ell \geq \ell'$*, Algorithm 8.3.6 terminates. We have* $\cup_{\Xi \times \Xi' \in \Upsilon} \Xi \times \Xi' = \Pi \times \Pi'$ *and the number of elements in* Υ *can be bounded by*

$$\#\Upsilon \lesssim (\rho^n + 1)(\ell - \ell') + \rho^{2n} + 1, \tag{8.3.24}$$

with the constant absorbed by the "$\lesssim$*" symbol not depending on* Π, Π', *and* ρ.

Proof. In each two successive subdivisions the maximum diameter of the "current" panels decreases by a constant factor, while the minimum distance between the "current" pairs does not decrease. Furthermore, thinking of a pair of panels that have distance zero, if the panels of a current pair live on different levels, then the difference in levels is decreased by a subdivision. Therefore the conditions (8.3.22) or (8.3.23) will eventually be satisfied starting from any pair, implying that the algorithm will terminate.

To avoid some technicalities, we prove here the estimate (8.3.24) for the simple case that the manifold Γ is $\mathbb{R}^n$, and that $\sigma(\tilde{\Xi}) = \operatorname{diam}(\tilde{\Xi}) = 2^{-\tilde{\ell}}$ for all $\tilde{\Xi} \in \mathcal{G}_{\tilde{\ell}}$, $\tilde{\ell} \in \mathbb{N}_0$. For the general case an analogous proof is obtained by using the fact that Γ is Lipschitz and that $\sigma(\tilde{\Xi}) \eqsim \operatorname{diam}(\tilde{\Xi}) \eqsim 2^{-\tilde{\ell}}$ for all $\tilde{\Xi} \in \mathcal{G}_{\tilde{\ell}}$, $\tilde{\ell} \in \mathbb{N}_0$.

Let $N_{\tilde{\ell}}$ denote the number of pairs $\Xi \times \Xi' \in \Upsilon$ such that $\Xi' \in \mathcal{G}_{\tilde{\ell}}$. Then we can estimate the total number of pairs by estimating the numbers $N_{\tilde{\ell}}$ and summing

over all $\tilde{\ell}$. It is obvious that if $\operatorname{dist}(\Pi, \Pi') > 0$, the number of pairs $\Xi \times \Xi' \in \Upsilon$ that satisfy (8.3.23) is zero, and if $\operatorname{dist}(\Pi, \Pi') = 0$, this number is uniformly bounded. Since in (8.3.24) this number is absorbed by the term 1 at the right hand side, in the remainder we will only count pairs of type (8.3.22).

In case $\tilde{\ell} \leq \ell$, we have $\Xi = \Pi$ for any $\Xi' \in \mathcal{G}_{\tilde{\ell}}$ with $\Xi \times \Xi' \in \Upsilon$. When, moreover $\tilde{\ell} > \ell'$ we have $\operatorname{dist}(\Pi, \Xi') \leq (2\rho + 2)2^{-\tilde{\ell}}$. Indeed, if not, then the "parent" $\Xi'' \in \mathcal{G}_{\tilde{\ell}-1}$ of Ξ' would have satisfied $\operatorname{dist}(\Pi, \Xi'') > 2\rho 2^{-\tilde{\ell}} = \max\{\sigma(\Pi), \sigma(\Xi'')\}$ and so Ξ' would never have been created by the algorithm. We conclude that for $\ell' < \tilde{\ell} \leq \ell$, $N_{\tilde{\ell}} \lesssim \left((2\rho + 2)2^{-\tilde{\ell}} + 2^{-\ell}\right)^n / 2^{-\tilde{\ell} n} \lesssim \rho^n + 1$.

Now we consider $\Xi \times \Xi' \in \Upsilon$ with $\Xi' \in \mathcal{G}_{\tilde{\ell}}$ and $\tilde{\ell} > \ell$ (and such that $\Xi \times \Xi'$ satisfies (8.3.22)). By construction of the algorithm, we have either $\Xi \in \mathcal{G}_{\tilde{\ell}}$ or $\Xi \in \mathcal{G}_{\tilde{\ell}-1}$. Similar arguments as have been used above show that for fixed Ξ, the number of such pairs is bounded by a constant multiple of $\rho^n + 1$. Since the number of such Ξ is bounded by a constant multiple of $2^{(\tilde{\ell}-\ell)n}$, we conclude that for $\tilde{\ell} > \ell$, $N_{\tilde{\ell}} \lesssim (\rho^n + 1)2^{(\tilde{\ell}-\ell)n}$.

In light of Remark 8.3.8, it is easy to see that the smallest subelements generated by this algorithm will belong to the level $\ell_{\max}$ with $\rho 2^{-\ell_{\max}} \gtrsim 2^{-\ell}$, implying that $2^{(\ell_{\max}-\ell)n} \lesssim \rho^n$. Therefore, we conclude that the number of elements in the subdivision Υ is bounded by a constant multiple of

$$\begin{aligned} 1 + \sum_{\tilde{\ell}=\ell'+1}^{\ell_{\max}} N_{\tilde{\ell}} &\lesssim 1 + \sum_{\tilde{\ell}=\ell'+1}^{\ell} (\rho^n + 1) + \sum_{\tilde{\ell}=\ell+1}^{\ell_{\max}} (\rho^n + 1)2^{(\tilde{\ell}-\ell)n} \\ &\lesssim (\rho^n + 1)(\ell - \ell') + \rho^{2n} + 1. \end{aligned}$$

■

From the condition (8.3.23), we have that the singular integrals corresponding to the subdivision Υ are always over pairs of panels on the same level. In this paper, we make the following Assumption 8.3.10 on quadrature schemes for computing those singular integrals. For completeness, in Section 8.4 we confirm this assumption for the simple case of the single layer kernel on polyhedral surfaces in $\mathbb{R}^3$. In any case for weakly- and strongly singular integrals, using the quadrature schemes from e.g. [72, 74], we expect that Assumption 8.3.10 can be verified generally.

Assumption 8.3.10. We assume that there exist $d_0^* \in \mathbb{R}$ and $\varrho_0 > 1$ such that for any $\lambda, \lambda' \in \Lambda$ with $|\lambda| \geq |\lambda'|$, $\Xi, \Xi' \in \mathcal{G}_{|\lambda|}$ with $\operatorname{dist}(\Xi, \Xi') = 0$, and for any order $p \in \mathbb{N}$, an approximation $I^*_{\lambda\lambda'}(\Xi, \Xi')$ of $I_{\lambda\lambda'}(\Xi, \Xi')$ can be computed within $W \lesssim p^{2n}$ arithmetical operations, having an error

$$|I_{\lambda\lambda'}(\Xi, \Xi') - I^*_{\lambda\lambda'}(\Xi, \Xi)'| \lesssim \varrho_0^{-p} 2^{||\lambda|-|\lambda'||d_0^*}. \tag{8.3.25}$$

Now we are ready to present an algorithm how to compute the integral (8.3.11) with the help of a generally non-uniform subdivision of the integration domain $\Pi \times \Pi'$.

Algorithm 8.3.11 Computation of the integral $I_{\lambda\lambda'}(\Pi, \Pi')$

Parameters: Let Ψ be of P-type of order e, and let $p \in \mathbb{N}$ and $\rho > 1$ be given. Choose the function $\sigma(\cdot)$ as in Lemma 8.3.4.

Input: $\lambda, \lambda' \in \Lambda$, and $\Pi \in \mathcal{G}_\lambda$, and $\Pi' \in \mathcal{G}_{\lambda'}$.

Output: $I^*_{\lambda\lambda'}(\Pi, \Pi') \in \mathbb{R}$.

1: Apply Algorithm 8.3.6 with the above ρ and $\sigma(\cdot)$ to get the subdivision Υ of $\Pi \times \Pi'$;
2: For each subdomain $\Xi \times \Xi' \in \Upsilon$ apply either step 4 or 5; and sum the results as in (8.3.11), to get $I^*_{\lambda\lambda'}(\Pi, \Pi')$;
3: If $\operatorname{dist}(\Xi, \Xi') > 0$, apply the quadrature scheme of order p from Lemma 8.3.4;
4: If $\operatorname{dist}(\Xi, \Xi') = 0$, apply the computational scheme of order p from Assumption 8.3.10.

Remark 8.3.12. For Q-type wavelets, Algorithm 8.3.11 can be redefined by replacing "Lemma 8.3.4" by "Lemma 8.3.5".

Theorem 8.3.13. *Let Ψ be of P-type of order e, and assume that an approximation $I^*_{\lambda\lambda'}(\Pi, \Pi')$ of $I_{\lambda\lambda'}(\Pi, \Pi')$ is computed by using Algorithm 8.3.11. Assume that $n \geq 2t$. Then, in case that*

$$\operatorname{dist}(\Pi, \Pi') \geq \rho \max\{\sigma(\Pi), \sigma(\Pi')\}, \tag{8.3.26}$$

with $e^* = e - 1 - 2t - n$, the error of the numerical integration satisfies

$$|E_{\lambda\lambda'}(\Pi, \Pi')| \lesssim \rho^{-p} 2^{-||\lambda|-|\lambda'||(t+n/2)} \left(\frac{\operatorname{dist}(\Pi, \Pi')}{\rho \max\{\sigma(\Pi), \sigma(\Pi')\}} \right)^{e^*-p}, \tag{8.3.27}$$

and the work for computing $I^*_{\lambda\lambda'}(\Pi, \Pi')$ is bounded by a constant multiple of p^{2n}, provided that $p \geq \max\{e-1, e^*+1\}$. In case that (8.3.26) does not hold, for any $d_1^* \geq |t| - n/2$, with $d_1^* > -n/2$ when $t = 0$, the error satisfies

$$|E_{\lambda\lambda'}(\Pi, \Pi')| \lesssim \rho^{-p} 2^{||\lambda|-|\lambda'||d_1^*} + \varrho_0^{-p} 2^{||\lambda|-|\lambda'||d_0^*}, \tag{8.3.28}$$

and the work is bounded by a constant multiple of $p^{2n}(1 + ||\lambda| - |\lambda'||)$, provided that $p \geq \max\{e-1, e^*+1\}$. In view of Remark 8.3.12, these results also hold for Q-type wavelets of order e.

By taking $\varrho := \min\{\varrho_0, \rho\}$ and $d^* := \max\{d_0^*, d_1^*\}$, we conclude that the criteria (8.3.5) and (8.3.6) for s^*-computability from Theorem 8.3.2 are satisfied.

Proof. Without loss of generality, we assume that $|\lambda| \geq |\lambda'|$. First, we will consider the case that (8.3.26) holds. In this case, we have the subdivision $\Upsilon = \{\Pi \times \Pi'\}$, and so the computational work is of order of p^{2n}. Applying Lemma 8.3.4 with $\Theta = \kappa^{-1}(\Pi)$ and $\Theta' = \kappa^{-1}(\Pi')$, taking into account the definition of ω, and using the fact that $\omega \geq \rho > 1$ and that $\min\{\sigma(\Pi), \sigma(\Pi')\} \lesssim 2^{-|\lambda|}$, we get

$$\begin{aligned} |E_{\lambda\lambda'}(\Pi, \Pi')| &\lesssim 2^{(|\lambda|-|\lambda'|)(n/2-t)}\omega^{-(n+p)} \max\{1, \omega\}^{e-1} \min\{\sigma(\Pi), \sigma(\Pi')\}^n \\ &\quad \times \operatorname{dist}(\Pi, \Pi')^{-2t} \\ &\lesssim 2^{-|\lambda|(t+n/2)+|\lambda'|(t-n/2)}\omega^{e-1-n-p}\omega^{-2t} \max\{\sigma(\Pi), \sigma(\Pi')\}^{-2t}. \end{aligned}$$

Now using the estimate $\max\{\sigma(\Pi), \sigma(\Pi')\} \eqsim 2^{-|\lambda'|}$ and $n \geq 2t$, we have

$$\begin{aligned} |E_{\lambda\lambda'}(\Pi, \Pi')| &\lesssim 2^{-(|\lambda|-|\lambda'|)(t+n/2)-|\lambda'|(n-2t)}\omega^{e^*-p} \\ &\lesssim \rho^{-p}2^{-(|\lambda|-|\lambda'|)(t+n/2)}(\omega/\rho)^{e^*-p}, \end{aligned}$$

proving the first part of the theorem.

Let us now consider the case that (8.3.26) does not hold. Since ρ is fixed, the number of subdomains of the subdivision Υ is of order $1 + ||\lambda| - |\lambda'||$, and thus we get the work bound. By Assumption 8.3.10, the sum of the errors made in the approximations for $I_{\lambda\lambda'}(\Xi, \Xi')$ with $\Xi \times \Xi' \in \Upsilon$ and $\operatorname{dist}(\Xi, \Xi') = 0$ is responsible for the last term in (8.3.28).

We need to estimate the portion of the total error $E_{\lambda\lambda'}(\Pi, \Pi')$ that corresponds to the integrals $I_{\lambda\lambda'}(\Xi, \Xi')$ with $\Xi \times \Xi' \in \Upsilon$ and $\operatorname{dist}(\Xi, \Xi') > 0$. We denote by I_1 the sum of all these integrals arising from the subdivision Υ, and by I_1^* the computed approximation for I_1. Since by construction for any $\Xi \times \Xi' \in \Upsilon$ with $\operatorname{dist}(\Xi, \Xi') > 0$ it holds that $\frac{\operatorname{dist}(\Xi,\Xi')}{\max\{\sigma(\Xi),\sigma(\Xi')\}} \geq \rho > 1$, Lemma 8.3.4 gives

$$\begin{aligned} |I_1 - I_1^*| \lesssim & \sum_{\{\Xi\times\Xi'\in\Upsilon:\operatorname{dist}(\Xi,\Xi')>0\}} 2^{(|\lambda|-|\lambda'|)(n/2-t)}\rho^{e-1-n-p} \\ & \times \min\{\sigma(\Xi), \sigma(\Xi')\}^n \operatorname{dist}(\Xi, \Xi')^{-2t} \\ \lesssim & \rho^{-p}2^{-|\lambda|(t+n/2)+|\lambda'|(t-n/2)} \sum_{\{\Xi\times\Xi'\in\Upsilon:\operatorname{dist}(\Xi,\Xi')>0\}} \operatorname{dist}(\Xi, \Xi')^{-2t}, \end{aligned} \tag{8.3.29}$$

where we have used that $\min\{\sigma(\Xi), \sigma(\Xi')\} \lesssim 2^{-|\lambda|}$.

From the proof of Lemma 8.3.9, recall that for the number $N_{\tilde{\ell}}$ of $\Xi \times \Xi' \in \Upsilon$ with $\Xi' \in \mathcal{G}_{\tilde{\ell}}$, we have $N_{\tilde{\ell}} = 0$ for $\tilde{\ell} > \ell_{\max}$ where, since ρ is a fixed constant, $\ell_{\max} - |\lambda| \lesssim 1$, and furthermore $N_{\tilde{\ell}} \lesssim 1$ for $|\lambda'| \leq \tilde{\ell} \leq \ell_{\max}$. Since for $\Xi \times \Xi' \in \Upsilon$ with $\operatorname{dist}(\Xi, \Xi') > 0$ and $\Xi' \in \mathcal{G}_{\tilde{\ell}}$, $\operatorname{dist}(\Xi, \Xi') \eqsim 2^{-\tilde{\ell}}$, we may bound the sum in the

last row of (8.3.29) on a constant multiple of

$$\sum_{\tilde{\ell}=|\lambda'|}^{\ell_{\max}} 2^{\tilde{\ell}\cdot 2t} \lesssim \begin{cases} 1+||\lambda|-|\lambda'|| & \text{if } t=0, \\ 2^{|\lambda'|\cdot 2t} & \text{if } t<0, \\ 2^{|\lambda|\cdot 2t} & \text{if } t>0. \end{cases}$$

By substituting this result into (8.3.29), the proof is completed. ∎

8.4 Quadrature for singular integrals

In this section, we confirm Assumption 8.3.10 for the simple case of the single layer kernel on polyhedral surfaces in $\mathbb{R}^3$.

We assume that the manifold Γ is the surface of a three dimensional polyhedron, and that the subdivisions $\mathcal{G}_\ell$, ($\ell \in \mathbb{N}$), are generated by dyadic refinements of $\mathcal{G}_0$, being an initial conforming triangulation of Γ.

We take the operator L to be the single layer operator (thus $t = -\frac{1}{2}$) having the kernel

$$K(z,z') = \frac{1}{4\pi|z-z'|} \qquad z \neq z', \tag{8.4.1}$$

and assume that the wavelet basis Ψ is of P-type of order e. Let $\lambda, \lambda' \in \Lambda$ be indices with $|\lambda| \geq |\lambda'|$. Then in view of Assumption 8.3.10, we are ultimately interested in computing the integral

$$I := \int_{\Xi}\int_{\Xi'} K(z,z')\psi_\lambda(z)\psi_{\lambda'}(z')d\Gamma_z d\Gamma_{z'}, \tag{8.4.2}$$

where $\Xi, \Xi' \in \mathcal{G}_{|\lambda|}$ and $\operatorname{dist}(\Xi,\Xi') = 0$. With

$$T := \{(x_1,x_2) \in \mathbb{R}^2 : 0 < x_2 < x_1 < 1\},$$

we can find affine bijections $\chi_\Xi : T \to \Xi$, and $\chi_{\Xi'} : T \to \Xi'$, thus with Jacobians $J_\Xi := |\partial\chi_\Xi| \eqsim 2^{-2|\lambda|}$, and $J_{\Xi'} := |\partial\chi_{\Xi'}| \eqsim 2^{-2|\lambda|}$, such that

$$I = \int_T\int_T \frac{g(x,y)}{|r(x,y)|}dxdy, \tag{8.4.3}$$

where $g(x,y) := (4\pi)^{-1}J_\Xi J_{\Xi'}\psi_\lambda(\chi_\Xi(x))\psi_{\lambda'}(\chi_{\Xi'}(y))$ and $r(x,y) := \chi_{\Xi'}(y) - \chi_\Xi(x)$. Taking into account that $n = 2$ and $t = -\frac{1}{2}$, from (8.2.2) we derive the following estimates for $\beta \in \mathbb{N}_0^2$

$$|\partial_x^\beta g| \lesssim 2^{-\frac{5}{2}|\lambda|+\frac{3}{2}|\lambda'|} \quad \text{and} \quad |\partial_y^\beta g| \lesssim 2^{-\frac{5}{2}|\lambda|+\frac{3}{2}|\lambda'|}2^{(|\lambda'|-|\lambda|)|\beta|}. \tag{8.4.4}$$

We present here a slight variation of the quadrature scheme developed in e.g. [67, 72, 74], see also [73]. The idea is to apply a degenerate coordinate transformation which is a generalization of the so called *Duffy's triangular coordinates*, effectively removing the singularity of the integrand while preserving a polyhedral shape of the integration domain. The coordinate transformations introduced here are somewhat simpler than the ones in the above mentioned papers, and we expect that the presentation is geometrically more intuitive.

To this end, we need to partition the integration domain $T\times T$ into several pyramides, which is necessary for us to use Duffy's transformations in order to remove the singularities, cf. [67, 72]. Denote the vertices of the triangle T by $A_0 = (0,0)$, $A_1 = (0,1)$, and $A_2 = (1,1)$. Then obviously, $T\times T$ has nine vertices $A_{ik} := A_i \times A_k$ for $i,k = 0,1,2$. Note that $A_{00} = O$.

We break $T\times T$ up into two pyramides $D_1 := \{(x,y) \in T\times T : x_1 > y_1\}$ and $D_2 := \{(x,y) \in T\times T : x_1 < y_1\}$. One can verify that D_1 is the pyramid with vertex O and base $B_1 = A_{10}A_{11}A_{12}A_{20}A_{21}A_{22}$, being a triangular prism, and that D_2 is the pyramid with vertex O and base B_2 $B_2 = A_{01}A_{11}A_{21}A_{02}A_{12}A_{22}$, being also a triangular prism. Moreover, these prisms can be described as $B_1 = \{1\}\times(0,1)\times T$ and $B_2 = T\times\{1\}\times(0,1)$. Introducing the reflection with respect to the plane $x = y$ by $\mathcal{R} : (x,y) \mapsto (y,x)$, we notice the symmetry $B_2 = \mathcal{R}B_1$ and so $D_2 = \mathcal{R}D_1$.

By subdividing the prism B_1 into tetrahedra, we can get a simplicial partitioning of $T\times T$, because any simplicial partitioning of B_1 induces a simplicial partitioning of D_1, and by taking the image under the mapping $\mathcal{R}$, a simplicial partitioning of D_2. Our choice of such a partitioning is depicted in Figure 8.3.

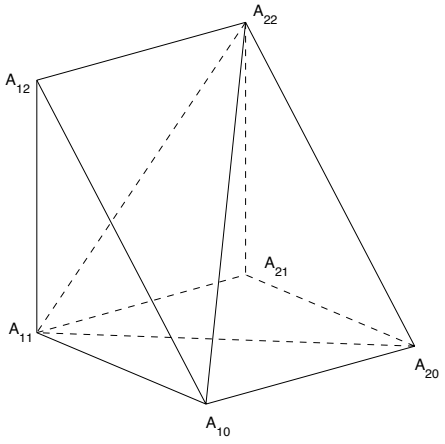

Figure 8.3: *A simplicial partitioning of the prism B_1.*

Consequently, the domain $T\times T$ is subdivided into the following simplices

described by their vertices.

$$D_1 \begin{cases} S_1 = OA_{10}A_{11}A_{12}A_{22}, \\ S_2 = OA_{10}A_{11}A_{20}A_{22}, \\ S_3 = OA_{11}A_{20}A_{21}A_{22}, \end{cases} \quad \text{and} \quad D_2 \begin{cases} S_4 = OA_{01}A_{11}A_{21}A_{22}, \\ S_5 = OA_{01}A_{11}A_{02}A_{22}, \\ S_6 = OA_{11}A_{02}A_{12}A_{22}. \end{cases}$$

We notice the symmetry $S_i = \mathcal{R}S_{i+3}$ for $i = 1, 2, 3$. The above partitionings of $T \times T$ will be used in quadrature schemes for the integral (8.4.3).

In the following we will distinguish three basic cases:

- *Coincident panels:* $\Xi = \Xi'$, that is, the case of identical panels;
- *Edge adjacent panels:* $\overline{\Xi}$ and $\overline{\Xi'}$ share one common edge;
- *Vertex adjacent panels:* $\overline{\Xi}$ and $\overline{\Xi'}$ share one common vertex.

In view of (8.4.3), we need to integrate a singular function over a four dimensional polyhedral domain $T \times T$. The singularity of the function is located on different dimensional sets in different situations: whereas the singularity occurs at a point for vertex adjacent integrals, it occurs all along an edge in case the integral is edge adjacent, and for coincident integrals, the singularity is on a two dimensional "diagonal" of the domain. Therefore in each of the three cases, we first characterize the singularity in terms of the distance to the singularity set, and then introduce special coordinate transformations that annihilate the singularity.

Case of identical panels

First we will discuss the case of identical panels $\Xi = \Xi'$. In this case, the difference $r = \chi_\Xi(y) - \chi_\Xi(x)$ is zero if and only if $t := y - x = 0$. Since χ_Ξ is affine, we can write

$$r = 2^{-|\lambda|} l_1(t) = 2^{-|\lambda|} l_1(y_1 - x_1, y_2 - x_2),$$

where $l_1 : \mathbb{R}^2 \mapsto \mathbb{R}^3$ is a linear function depending only on the shape of Ξ. Noting that any panel Ξ is similar to a panel from the initial triangulation, we only have to deal with finitely many functions l_1. Introducing polar coordinates (ρ, θ) in $\mathbb{R}^2$ by $\rho = |t|$ and $\theta = t/|t| \in S^1$, being the unit circle in $\mathbb{R}^2$, this difference r reads as

$$r = 2^{-|\lambda|} \rho l_1(\theta).$$

Our goal is now to obtain an expression for $|r|^{-1}$, because this quantity essentially determines the singular behavior of the local kernel. Since r is defined on some

complete neighborhood of $t = 0$, the function $l_1(\theta)$ has to be nonzero for any $\theta \in S^1$, and so we have

$$|r|^{-1} = 2^{|\lambda|}\rho^{-1}a(\theta)$$

with $a(\theta) := |l_1(\theta)|^{-1}$ which is analytic in a neighborhood of S^1. Now the integrand of (8.4.3) can be written as

$$|r(x,y)|^{-1}g(x,y) = 2^{|\lambda|}\rho^{-1}a(\theta)g(x,y). \tag{8.4.5}$$

It is time to use the above described simplicial partitioning of the integration domain $T \times T$, in combination with special coordinate transformations for the purpose of removing the singularity of the integrand. Introducing the notation $\mathcal{P} := T\times(0,1)\times(0,1)$, we define the transformations $\phi_i : \mathcal{P}\to S_i : (\eta,\zeta,\xi) \mapsto (x,y)$ for $i \in \overline{1,6}$.

$$\phi_1(\eta,\zeta,\xi) = \begin{pmatrix} (1-\xi)\eta_1 + \xi \\ (1-\xi)\eta_2 \\ (1-\xi)\eta_1 + \xi\zeta \\ (1-\xi)\eta_2 + \xi\zeta \end{pmatrix}, \qquad \phi_2(\eta,\zeta,\xi) = \begin{pmatrix} (1-\xi)\eta_1 + \xi \\ (1-\xi)\eta_2 + \xi\zeta \\ (1-\xi)\eta_1 \\ (1-\xi)\eta_2 \end{pmatrix},$$

$$\phi_3(\eta,\zeta,\xi) = \begin{pmatrix} (1-\xi)\eta_1 + \xi \\ (1-\xi)\eta_2 + \xi \\ (1-\xi)\eta_1 + \xi\zeta \\ (1-\xi)\eta_2 \end{pmatrix}, \tag{8.4.6}$$

and $\phi_{i+3} := \mathcal{R}\circ\phi_i$ for $i = 1,2,3$. The Jacobian of each transformation ϕ_i is given by $\xi(1-\xi)^2$. Recall that ρ^{-1} characterizes the singularity of the integrand (8.4.5). In this regard, for each transformation ϕ_i one can show that

$$\rho = \xi f_i(\zeta), \quad \text{with an analytic } f_i(\zeta)\geq\tfrac{1}{\sqrt{2}} \quad \text{for any } \zeta \in [0,1].$$

For instance, for ϕ_1 we have

$$\rho^2 = \xi^2(\zeta^2 + (1-\zeta)^2) \geq \xi^2\cdot\tfrac{1}{2},$$

since $\zeta^2 + (1-\zeta)^2\geq\frac{1}{2}$ for any $\zeta \in \mathbb{R}$. Moreover, for each ϕ_i one can verify that $\theta = \vartheta_i(\zeta)$ for some analytic function $\vartheta_i : [0,1] \to S^1$.

In all, the Jacobian of the mapping ϕ_i annihilates the singularity in the integrand (8.4.5), meaning that the integral I in (8.4.3) now can be written as the following proper integral

$$\begin{aligned} I &= \int_0^1\int_0^1\int_T \xi(1-\xi)^2 \sum_{i=1}^{6} \frac{g(\phi_i(\eta,\zeta,\xi))}{|r(\phi_i(\eta,\zeta,\xi))|} d\eta d\zeta d\xi \\ &= 2^{|\lambda|}\int_0^1\int_0^1\int_T (1-\xi)^2 \sum_{i=1}^{6} \frac{a(\vartheta_i(\zeta))g(\phi_i(\eta,\zeta,\xi))}{f_i(\zeta)} d\eta d\zeta d\xi. \end{aligned} \tag{8.4.7}$$

Therefore we will be able to use standard quadrature schemes to approximate the integral I. Note that in numerical quadrature we can use the first expression in (8.4.7) for the integral I. The functions f_i and $a \circ \vartheta_i$ are introduced here merely for the analysis purpose.

Since the integrand in (8.4.7) is polynomial with respect to the variables ξ and η, we can always choose exact quadrature rules for integrations over those variables.

Proposition 8.4.1. *Approximate the integral* (8.4.7) *by a product quadrature rule $Q_\xi \times Q_\zeta \times Q_\eta$, where Q_ξ and Q_η are quadrature rules exact for the integration over the variables $\xi \in (0,1)$ and $\eta \in T$, respectively, and Q_ζ is a composite quadrature rule for the integration over $\zeta \in (0,1)$ of varying order p and fixed rank N. Then there exist a constant $\delta > 0$ such that the quadrature error satisfies*

$$|E(\Xi, \Xi')| \lesssim 2^{-\frac{3}{2}(|\lambda|-|\lambda'|)}(\delta N)^{-p}. \tag{8.4.8}$$

Choosing N such that $\delta N > 1$, we conclude that in this case Assumption 8.3.10 is fulfilled with $d_0^* = -\frac{3}{2}$.

Proof. In view of Lemma 7.2.4, it suffices to consider the integration over ζ. Using the analyticity of $\zeta \mapsto \frac{a(\vartheta_i(\zeta))}{f_i(\zeta)}$ one derives

$$\sup_{\zeta\in[0,1]} \left| \partial_\zeta^k \frac{a(\vartheta_i(\zeta))}{f_i(\zeta)} \right| \lesssim \frac{k!}{\delta^k} \quad \text{for } k \in \mathbb{N}_0,\ i \in \overline{1,6},$$

for some constant $\delta > 0$. From (8.4.4) and (8.4.6) we have for each $i \in \overline{1,6}$ that $g \circ \phi_i$ is a polynomial of order e and

$$|\partial_\zeta^k (g \circ \phi_i)| \lesssim 2^{-\frac{5}{2}|\lambda|+\frac{3}{2}|\lambda'|} \quad \text{for } k \in \overline{1, e-1}.$$

Now using Proposition 7.2.6 the proof is obtained. ∎

Case of edge adjacent panels

Now we will discuss the case when $\overline{\Xi}$ and $\overline{\Xi'}$ share exactly one common edge. Without loss of generality, we assume that $\chi_\Xi(x) = \chi_{\Xi'}(x)$ for all $x \in (0,1) \times \{0\}$. Then, the difference $r = \chi_{\Xi'}(y) - \chi_\Xi(x)$ is zero if and only if

$$t = (t_1, t_2, t_3) := (y_1 - x_1, x_2, y_2)$$

equals zero. Since χ_Ξ and $\chi_{\Xi'}$ are affine, we can write

$$r = \chi_{\Xi'}(x_1 + t_1, t_3) - \chi_\Xi(x_1, t_2) = 2^{-|\lambda|} l_1(t),$$

where $l_1 : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear function depending only on the shapes of Ξ and Ξ'. Introducing polar coordinates (ρ, θ) in $\mathbb{R}^3$ by $\rho = |t|$ and $\theta = t/|t| \in S^2$, being the unit sphere in $\mathbb{R}^3$, this difference r reads as

$$r = r(\rho, \theta) = 2^{-|\lambda|}\rho l_1(\theta).$$

Since r is defined on a complete neighborhood of $t = 0$ in $\mathbb{R} \times \mathbb{R}^2_{\geq 0}$, the function $l_1(\theta) \neq 0$ for any $\theta \in S^2$ with $\theta_2, \theta_3 \geq 0$, allowing us to write

$$|r|^{-1} = 2^{|\lambda|}\rho^{-1}b(\theta)$$

with $b(\theta) := |l_1(\theta)|^{-1}$ which is analytic in a neighborhood of $S^2 \cap \left(\mathbb{R} \times \mathbb{R}^2_{\geq 0}\right)$. Then the integrand of (8.4.3) can be written as

$$|r(x, y)|^{-1} g(x, y) = 2^{|\lambda|}\rho^{-1} b(\theta) g(x, y). \tag{8.4.9}$$

Now we define the transformations $\phi_i : \mathcal{P} \to S_i : (\eta, \zeta, \xi) \mapsto (x, y)$ for $i \in \overline{1, 6}$.

$$\phi_1(\eta, \zeta, \xi) = \begin{pmatrix} (1-\xi)\zeta + \xi \\ \xi\eta_2 \\ (1-\xi)\zeta + \xi\eta_1 \\ \xi\eta_1 \end{pmatrix}, \qquad \phi_2(\eta, \zeta, \xi) = \begin{pmatrix} (1-\xi)\zeta + \xi \\ \xi\eta_1 \\ (1-\xi)\zeta + \xi\eta_2 \\ \xi\eta_2 \end{pmatrix},$$

$$\phi_3(\eta, \zeta, \xi) = \begin{pmatrix} (1-\xi)\zeta + \xi \\ \xi \\ (1-\xi)\zeta + \xi\eta_1 \\ \xi\eta_2 \end{pmatrix}, \tag{8.4.10}$$

and $\phi_{i+3} := \mathcal{R} \circ \phi_i$ for $i = 1, 2, 3$. For each transformation ϕ_i one can show that the Jacobian equals $\xi^2(1 - \xi)$, and that

$$\rho = \xi f_i(\eta), \quad \text{with an analytic } f_i(\eta) \geq \tfrac{1}{\sqrt{2}} \quad \text{for any } \eta \in \overline{T}.$$

For instance, for ϕ_1 we have

$$\rho^2 = \xi^2(\eta_1^2 + (1 - \eta_1)^2 + \eta_2^2) \geq \xi^2 \cdot \tfrac{1}{2}.$$

Moreover, for each ϕ_i one can verify that $\theta = \vartheta_i(\eta)$ with some analytic function $\vartheta_i : \overline{T} \to S^2$.

In all, the Jacobian of the mapping ϕ_i annihilates the singularity in the integrand (8.4.9), meaning that the integral I in (8.4.3) now can be written as the

following proper integral

$$\begin{aligned} I &= \int_0^1 \int_0^1 \int_T \xi^2(1-\xi) \sum_{i=1}^{6} \frac{g(\phi_i(\eta,\zeta,\xi))}{|r(\phi_i(\eta,\zeta,\xi))|} d\eta d\zeta d\xi \\ &= 2^{|\lambda|} \int_0^1 \int_0^1 \int_T \xi(1-\xi) \sum_{i=1}^{6} \frac{b(\vartheta_i(\eta)) g(\phi_i(\eta,\zeta,\xi))}{f_i(\eta)} d\eta d\zeta d\xi, \end{aligned} \tag{8.4.11}$$

and thus the standard quadrature schemes on $\mathcal{P}$ can be applied.

Proposition 8.4.2. *Approximate the integral* (8.4.11) *by a product quadrature rule $Q_\xi \times Q_\zeta \times Q_\eta$, where Q_ξ and Q_ζ are quadrature rules exact for the integration over the variables $\xi, \zeta \in (0,1)$, respectively, and Q_η is a composite quadrature rule for the integration over $\eta \in T$ of varying order p and fixed rank N. Then there exist a constant $\delta > 0$ such that the quadrature error satisfies*

$$|E(\Xi, \Xi')| \lesssim 2^{-\frac{3}{2}(|\lambda|-|\lambda'|)} (\delta N)^{-p}. \tag{8.4.12}$$

Choosing N such that $\delta N > 1$, we conclude that in this case Assumption 8.3.10 is fulfilled with $d_0^* = -\frac{3}{2}$.

The proof is obtained similarly to the proof of Proposition 8.4.1.

Case of vertex adjacent panels

Let $\overline{\Xi}$ and $\overline{\Xi'}$ share exactly one common point. Without loss of generality, we assume that $\chi_\Xi(0) = \overline{\Xi} \cap \overline{\Xi'} = \chi_{\Xi'}(0)$. Then obviously, the difference $r = r(x,y) = \chi_{\Xi'}(y) - \chi_\Xi(x)$ is zero if and only if $t := (x,y)$ equals zero. Since χ_Ξ and $\chi_{\Xi'}$ are affine, we can write

$$r(x,y) = 2^{-|\lambda|} l_1(x,y),$$

where $l_1 : \mathbb{R}^4 \to \mathbb{R}^3$ is a linear function depending only on the shapes of Ξ and Ξ'. Introducing polar coordinates (ρ, θ) in $\mathbb{R}^4$ by $\rho = |t|$ and $\theta = t/|t| \in S^3$, being the unit sphere in $\mathbb{R}^4$, this difference r reads as

$$r = r(\rho, \theta) = 2^{-|\lambda|} \rho l_1(\theta).$$

Since r is defined on a complete neighborhood of $t = 0$ in $\{t \in \mathbb{R}^4 : t_1 \geq t_2 \geq 0,\, t_3 \geq t_4 \geq 0\}$, the function $l_1(\theta) \neq 0$ for any $\theta \in S^3$ with $\theta_1 \geq \theta_2 \geq 0$ and $\theta_3 \geq \theta_4 \geq 0$, allowing us to write

$$|r|^{-1} = 2^{|\lambda|} \rho^{-1} c(\theta)$$

with $c(\theta) := |l_1(\theta)|^{-1}$ which is analytic in a neighborhood of $\{\theta \in S^3 : \theta_1 \geq \theta_2 \geq 0,\ \theta_3 \geq \theta_4 \geq 0\}$. Then the integrand of (8.4.3) can be written as

$$|r(x,y)|^{-1} g(x,y) = 2^{|\lambda|} \rho^{-1} c(\theta) g(x,y). \tag{8.4.13}$$

We define the transformations ϕ_1 and ϕ_2 that map the coordinates $(\eta, \zeta, \xi) \in \mathcal{P}$ onto the four dimensional pyramides D_1 and D_2 respectively.

$$\phi_1(\eta,\zeta,\xi) = \xi(1,\zeta,\eta_1,\eta_2), \quad \text{and} \quad \phi_2(\eta,\zeta,\xi) = \xi(\eta_1,\eta_2,1,\zeta). \tag{8.4.14}$$

Notice that $\phi_1 = \mathcal{R} \circ \phi_2$ with $\mathcal{R}$ being the reflection $x \leftrightarrow y$. For both of the transformations, the Jacobian equals ξ^3, and we have

$$\rho = \xi f(\eta,\zeta) \quad \text{with } f(\eta,\zeta) = \sqrt{1 + {\eta_1}^2 + {\eta_2}^2 + \zeta^2}.$$

Moreover, we have $\theta = \vartheta_1(\eta,\zeta) := f(\eta,\zeta)^{-1}(1,\zeta,\eta_1,\eta_2)$ for the transformation ϕ_1, and $\theta = \vartheta_2(\eta,\zeta) := \mathcal{R}\vartheta_1(\eta,\zeta)$ for the transformation ϕ_2.

In all, the Jacobian of the mapping ϕ_i annihilates the singularity in the integrand (8.4.13), meaning that the integral I in (8.4.3) now can be written as the following proper integral

$$\begin{aligned} I &= \int_0^1 \int_0^1 \int_T \xi^3 \sum_{i=1}^{2} \frac{g(\phi_i(\eta,\zeta,\xi))}{|r(\phi_i(\eta,\zeta,\xi))|} d\eta d\zeta d\xi \\ &= 2^{|\lambda|} \int_0^1 \int_0^h \int_T \xi^2 \sum_{i=1}^{2} \frac{c(\vartheta_i(\eta,\zeta)) g(\phi_i(\eta,\zeta,\xi))}{f(\eta,\xi)} d\eta d\zeta d\xi, \end{aligned} \tag{8.4.15}$$

and thus the standard quadrature schemes on $\mathcal{P}$ can be applied.

Proposition 8.4.3. *Approximate the integral* (8.4.15) *by a product quadrature rule $Q_\xi \times Q_\zeta \times Q_\eta$, where Q_ξ is a quadrature rule exact for the integration over $\xi \in (0,1)$, and Q_ζ and Q_η are composite quadrature rules for the integration over $\zeta \in (0,1)$ and $\eta \in T$, respectively, of varying order p and fixed rank N. Then there exist a constant $\delta > 0$ such that the quadrature error satisfies*

$$|E(\Xi,\Xi')| \lesssim 2^{-\frac{3}{2}(|\lambda|-|\lambda'|)} (\delta N)^{-p}. \tag{8.4.16}$$

Choosing N such that $\delta N > 1$, we conclude that in this case Assumption 8.3.10 is fulfilled with $d_0^* = -\frac{3}{2}$.

Chapter 9

Conclusion

9.1 Discussion

In [17, 18], Cohen, Dahmen and DeVore introduced adaptive wavelet paradigms for solving operator equations. A number of algorithms with asymptotically optimal computational complexity were developed, among others, under the assumption that on average, an individual entry of the stiffness matrix can be computed at unit cost. Although it has been indicated that this assumption is realistic, it is far from obvious.

The work presented in this thesis shows that the average unit cost assumption is valid for both differential and singular integral operators, at least when the wavalets are piecewise polynomials (Chapters 7 and 8). As a consequence, we can conclude that the "fully discrete" adaptive wavelet algorithm has optimal computational complexity.

A crucial ingredient for proving the optimal complexity of the adaptive wavelet algorithms was the *coarsening* step that was applied after every fixed number of iterations. As we have shown in Chapter 3, it turns out that coarsening is unnecessary for proving optimal computational complexity of algorithms of the type considered in [17]. Since with the new method no information is deleted that has been created by a sequence of computations, we expect that it is more efficient. The algorithm from Chapter 3 can be applied directly with minor modifications to a larger class of problems (Chapter 5). We also investigated the possibility of using polynomial preconditioners in the context of adaptive wavelet methods (Chapter 4).

In [5, 54], adaptive wavelet methods with "truncated" residuals were introduced which are modifications of the methods from [17], and convincing numerical experiments were reported showing that the methods are comparatively efficient. We developed a theoretical framework that could be used to prove optimal com-

putational complexity of the methods with truncated residuals, and for elliptic boundary value problems, a complete proof of optimality is given (Chapter 6).

9.2 Future work

There are many interesting directions in which future research on adaptive wavelet algorithms can be taken.

Proving optimality of adaptive wavelet BEMs with truncated residuals is an interesting and important open issue. Supposing that the proof would be similar to our proof in the case of methods for boundary value problems (Chapter 6), one would need efficient and reliable *a posteriori* error estimators for BEMs. To our knowledge, the only known such estimators are those developed by Birgit Faermann, cf. [42, 43, 44]. Then, the so-called local discrete lower bound for those estimators seems to be far from obvious. This direction could also be a way to approach the convergence and complexity analyses of adaptive BEMs.

Another interesting topic is the use of anisotropically supported wavelets for adaptive algorithms. When isotropically supported wavelets are employed, the convergence rate grows with the regularity of the solution in terms of (isotropic) Besov spaces (cf. Chapter 2), and decreases with increasing space dimension. The latter fact is an instance of the so called *curse of dimensionality*. Fortunately, high dimensional problems are usually simple and formulated on tensor product domains, and this exceptionally symmetric structure seems to give rise to a certain regular behaviour of the solution. Recently, Nitsche [65] identified certain anisotropic Besov spaces as the natural smoothness spaces for measuring the regularity of the solution when anisotropically supported tensor product wavelets are employed. It is also shown that in two and three dimensions, solutions to elliptic PDE's exhibit arbitrarily high regularity measured in terms of these spaces, while it is not known whether the same holds in the isotropic setting. Moreover, adaptive wavelet algorithms applied with anisotropically supported tensor product wavelets are proven to display asymptotically optimal convergence rates independent of space dimension, only restricted by the anisotropic Besov regularity of the solution and certain properties of the wavelets which can be bettered at will by choosing appropriate wavelets, cf. [65, 79]. So above all, it seems that the curse of dimensionality can be avoided. Yet, there still remains a few technical details. One has to be sure that the constants in the asymptotic estimates do not blow up with increasing dimension. Although it is generally expected, a sufficient regularity of the PDE's in more than three dimension needs to be verified. Furthermore, appropriate data structures for the implementation of the adaptive wavelet methods for high dimensions should be identified.

Anisotropically supported wavelets could also pay off even in low dimensions

when a strong singularity along a layer is present, for instance, when boundary integral equations on polyhedra, or singularly perturbed boundary value problems are considered. There have appeared some interesting results in the direction of using (isotropic as well as anisotropic) wavelets for stabilizing singularly perturbed boundary value problems, cf. [6, 13].

Furthermore, there are many related active research areas that are not touched upon in this thesis. Those include the issues of adaptive wavelet methods for nonlinear variational problems (cf. [3, 19, 20, 35]), using wavelets for goal oriented adaptivity (cf. [31]), using wavelet frames instead of a basis (cf. [27, 28, 83]), and adaptive wavelet methods for eigenvalue computation.

Bibliography

[1] S. F. Ashby, T. A. Manteuffel, and J. S. Otto, *A comparison of adaptive Chebyshev and least squares polynomial preconditioning for Hermitian positive definite linear systems*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 1–29.

[2] K. Atkinson, *The numerical solution of boundary integral equations*, in The State of the Art in Numerical Analysis, I. Duff and G. Watson, eds., Clarendon Press, Oxford, 1997, pp. 223–259.

[3] A. Barinka, *Fast computation tools for adaptive wavelet schemes*, PhD thesis, RWTH Aachen, Germany, March 2005.

[4] J. Bergh and J. Löfström, *Interpolation spaces: An introduction*, vol. 223 of Grundlehren der mathematischen Wissenschaften, Springer-Verlag, Berlin, Heidelberg, New York, 1976.

[5] S. Berrone and T. Kozubek, *An adaptive WEM algorithm for solving elliptic boundary problems in fairly general domains*, Preprint 38, Politecnico di Torino, Italy, 2004. Submitted.

[6] S. Bertoluzza, C. Canuto, and A. Tabacco, *Stable discretizations of convection-diffusion problems via computable negative-order inner products*, SIAM J. Numer. Anal., 38 (2000), pp. 1034–1055 (electronic).

[7] G. Beylkin, R. Coifman, and V. Rokhlin, *The fast wavelet transform and numerical algorithms*, Comm. Pure & Appl. Math., 44 (1991), pp. 141–183.

[8] P. Binev, W. Dahmen, and R. DeVore, *Adaptive finite element methods with convergence rates*, Numer. Math., 97(2) (2004), pp. 219–268.

[9] K. Bittner and K. Urban, *Adaptive wavelet methods using semiorthogonal spline wavelets: Sparse evaluation of nonlinear functions*, preprint, Universität Ulm, Germany, 2004.

[10] J. Bramble and J. Pasciak, *A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems*, Math. Comp., 50(181) (1988), pp. 1–17.

[11] S. C. Brenner and R. L. Scott, *The mathematical theory of finite element methods*, Springer Verlag, New York, 1994.

[12] V. I. Burenkov, *Sobolev spaces on domains*, Teubner Verlag, Stuttgart, Leipzig, 1998.

[13] C. Canuto and A. Tabacco, *Anisotropic wavelets along vector fields and applications to PDE's*, Arab. J. Sci. Eng. Sect. C Theme Issues, 28 (2003), pp. 89–105. Wavelet and fractal methods in science and engineering, Part I.

[14] C. Canuto, A. Tabacco, and K. Urban, *The wavelet element method part I: Construction and analysis*, Appl. Comput. Harmon. Anal., 6 (1999), pp. 1–52.

[15] C. Canuto and K. Urban, *Adaptive optimization in convex Banach spaces*, SIAM J. Numer. Anal., 42 (2005), pp. 2043–2075.

[16] A. Cohen, *Wavelet methods in numerical analysis*, in Handbook of Numerical Analysis. Vol. VII, P. Ciarlet and J. L. Lions, eds., North-Holland, Amsterdam, 2000, pp. 417–711.

[17] A. Cohen, W. Dahmen, and R. DeVore, *Adaptive wavelet schemes for elliptic operator equations – Convergence rates*, Math. Comp., 70 (2001), pp. 27–75.

[18] ——, *Adaptive wavelet methods II – Beyond the elliptic case*, Found. Comput. Math., 2 (2002), pp. 203–245.

[19] ——, *Adaptive wavelet schemes for nonlinear variational problems*, SIAM J. Numer. Anal., 41 (2003), pp. 1785–1823.

[20] ——, *Sparse evaluation of compositions of functions using multiscale expansions*, SIAM J. Math. Anal., 35(2) (2003), pp. 279–303.

[21] A. Cohen, I. Daubechies, and J. C. Feauveau, *Biorthogonal bases of compactly supported wavelets*, Comm. Pure & Appl. Math., 45 (1992), pp. 485–560.

[22] A. Cohen and R. Masson, *Wavelet adaptive method for second order elliptic problems: Boundary conditions and domain decomposition*, Numer. Math., 86 (2000), pp. 193–238.

[23] M. Costabel, *Boundary integral operators on Lipschitz domains: Elementary results*, SIAM J. Numer. Anal., 19(3) (1988), pp. 613–626.

[24] M. Costabel and W. L. Wendland, *Strong ellipticity of boundary integral operators*, J. Reine. Angew. Math., 372 (1986), pp. 34–63.

[25] S. Dahlke, W. Dahmen, and K. Urban, *Adaptive wavelet methods for saddle point probelms – Optimal convergence rates*, SIAM J. Numer. Anal., 40 (2002), pp. 1230–1262.

[26] S. Dahlke and R. DeVore, *Besov regularity for elliptic boundary value problems*, Comm. Part. Diff. Eqs., 22(1&2) (1997), pp. 1–16.

[27] S. Dahlke, M. Fornasier, and T. Raasch, *Adaptive frame methods for elliptic operator equations*, Bericht 3, Philipps-Universität Marburg, Germany, 2004.

[28] S. Dahlke, M. Fornasier, T. Raasch, R. P. Stevenson, and M. Werner, *Adaptive frame methods for elliptic operator equations: The steepest descent approach*, Technical Report 1347, Utrecht University, The Netherlands, February 2006. Submitted.

[29] W. Dahmen, *Stability of multiscale transformations*, J. Fourier Anal. & Appl., 2 (1996), pp. 341–361.

[30] W. Dahmen, H. Harbrecht, and R. Schneider, *Adaptive methods for boundary integral equations – Complexity and convergence estimates*, IGPM report 250, RWTH Aachen, Germany, March 2005.

[31] W. Dahmen, A. Kunoth, and J. Vorloeper, *Convergence of adaptive wavelet methods for goal-oriented error estimation*, IGPM report, RWTH Aachen, 2006. To appear in Proceedings of ENUMATH 05.

[32] W. Dahmen and R. Schneider, *Wavelets with complementary boundary conditions – Function spaces on the cube*, Results in Math., 34 (1998), pp. 255–293.

[33] ——, *Composite wavelet bases for operator equations*, Math. Comp., 68 (1999), pp. 1533–1567.

[34] ——, *Wavelets on manifolds I: Construction and domain decomposition*, SIAM J. Math. Anal., 31 (1999), pp. 184–230.

[35] W. Dahmen, R. Schneider, and Y. Xu, *Nonlinear functionals of wavelet expansions - adaptive reconstruction and fast evaluation*, Numer. Math., 86 (2000), pp. 49–101.

[36] W. Dahmen and R. P. Stevenson, *Element-by-element construction of wavelets satisfying stability and moment conditions*, SIAM J. Numer. Anal, 37 (1999), pp. 319–352.

[37] W. A. Dahmen, H. Harbrecht, and R. Schneider, *Compression techniques for boundary integral equations - Optimal complexity estimates*, SIAM J. Numer. Anal., 43 (2006), pp. 2251–2271.

[38] S. Dekel and D. Leviatan, *The Bramble-Hilbert lemma for convex domains*, SIAM J. Numer. Anal, 35 (2004), pp. 1203–1212.

[39] R. DeVore, *Nonlinear approximation*, Acta Numerica, 7 (1998), pp. 51–150.

[40] S. C. Eisenstat, H. C. Elman, and M. H. Schultz, *Variational iterative methods for nonsymmetric systems of linear equations*, SIAM J. Numer. Anal., 20(2) (1983), pp. 345–357.

[41] J. v. d. Eshof and G. L. Sleijpen, *Inexact Krylov subspace methods for linear systems*, SIAM J. Matrix Anal. Appl., 26 (2004), pp. 125–153.

[42] B. Faermann, *Local a-posteriori error indicators for the Galerkin discretization of boundary integral equations*, Numer. Math., 79 (1998), pp. 43–76.

[43] ——, *Localization of the Aronszajn-Slobodeckij norm and application to adaptive boundary element methods. Part I. The two-dimensional case*, IMA J. Numer. Anal., 20 (2000), pp. 203–234.

[44] ——, *Localization of the Aronszajn-Slobodeckij norm and application to adaptive boundary element methods. Part II. The three-dimensional case*, Numer. Math., 92 (2002), pp. 467–499.

[45] Ts. Gantumur, *An optimal adaptive wavelet method for nonsymmetric and indefinite elliptic problems*, Technical Report 1343, Utrecht University, The Netherlands, January 2006. Submitted.

[46] Ts. Gantumur, H. Harbrecht, and R. P. Stevenson, *An optimal adaptive wavelet method without coarsening of the iterands*, Technical Report 1325, Utrecht University, The Netherlands, March 2005. To appear in *Math. Comp.*

[47] Ts. Gantumur and R. P. Stevenson, *Computation of differential operators in wavelet coordinates*, Math. Comp., 75 (2006), pp. 697–709.

[48] ——, *Computation of singular integral operators in wavelet coordinates*, Computing, 76 (2006), pp. 77–107.

[49] G. H. Golub and M. L. Overton, *The convergence of inexact Chebyshev and Richardson iterative methods for solving linear systems*, Numer. Math., 53 (1988), pp. 571–593.

[50] P. Grisvard, *Elliptic problems in nonsmooth domains*, vol. 24 of Monographs and Studies in Mathematics, Pitman, Boston, London, Melbourne, 1985.

[51] W. Hackbusch, *Integral equations. Theory and numerical treatment*, ISNM, Birkhauser Verlag, Basel, Boston, Berlin, 1995.

[52] W. Hackbusch and S. A. Sauter, *On the efficient use of the Galerkin method to solve Fredholm integral equations*, Appl. Math., 38 (1993), pp. 301–322.

[53] H. Harbrecht, *Wavelet Galerkin schemes for the boundary element method in three dimensions*, PhD thesis, TU Chemnitz, Germany, 2001.

[54] H. Harbrecht and R. Schneider, *Adaptive wavelet Galerkin BEM*, in Computational Fluid and Solid Mechanics 2003, K.-J. Bathe, ed., Elsevier, Amsterdam, Boston, 2003, pp. 1982–1986.

[55] ——, *Biorthogonal wavelet bases for the boundary element method*, Math. Nachr., 269-270 (2004), pp. 167–188.

[56] H. Harbrecht and R. P. Stevenson, *Wavelets with patchwise cancellation properties*, Technical Report 1311, Utrecht University, The Netherlands, October 2004. To appear in *Math. Comp.*

[57] G. Hsiao and W. Wendland, *Boundary element methods: foundation and error analysis*, in Encyclopedia of Computational Mechanics, E. Stein, R. de Borst, and T. J. Hughes, eds., John Wiley & Sons Ltd, New York, 2004.

[58] O. G. Johnson, C. A. Micchelli, and G. Paul, *Polynomial preconditioners for conjugate gradient calculations*, SIAM J. Numer. Anal., 20 (1983), pp. 362–376.

[59] Ch. Lage and Ch. Schwab, *Wavelet Galerkin algorithms for boundary integral equations*, SIAM J. Sci. Comput., 20 (1999), pp. 2195–2222.

[60] V. Maz'ya and T. Shaposhnikova, *Higher regularity in the classical layer potential theory for Lipschitz domains*, Indiana Univ. Math. J, 54 (2005), pp. 99–142.

[61] W. C. McLean, *Strongly elliptic systems and boundary integral equations*, Cambridge University Press, Cambridge, New York, 2000.

[62] K. Mekchay and R. Nochetto, *Convergence of an adaptive finite element method for general second order linear elliptic PDE*, preprint, University of Maryland, 2004.

[63] A. A. R. Metselaar, *Handling wavelet expansions in numerical analysis*, PhD thesis, Universiteit Twente, The Netherlands, June 2002.

[64] H. Nguyen, *Finite element wavelets for solving partial differential equations*, PhD thesis, Utrecht University, The Netherlands, April 2005.

[65] A. Nitsche, *Sparse tensor product approximation of elliptic problems*, PhD thesis, ETH Zürich, Switzerland, October 2004.

[66] W. M. Patterson, *Iterative methods for the solution of a linear operator equation in Hilbert space – A survey*, no. 394 in Lecture Notes in Mathematics, Springer-Verlag, Berlin, Heidelberg, New York, 1974.

[67] T. v. Petersdorff and Ch. Schwab, *Fully discrete multiscale Galerkin BEM*, in Multiscale wavelet methods for partial differential equations, W. A. Dahmen, P. Kurdila, and P. Oswald, eds., Wavelet analysis and its applications, San Diego, 1997, Academic Press, pp. 287–346.

[68] S. Rolewicz, *Metric linear spaces*, D. Reidel Publishing Co., Dordrecht, 1984.

[69] W. Rudin, *Functional analysis*, International Series in Pure and Applied Mathematics, McGraw-Hill, Inc., New York, second ed., 1991.

[70] Y. Saad, *Practical use of polynomial preconditionings for the conjugate gradient method*, SIAM J. Sci. Statist. Comput., 6 (1985), pp. 865–881.

[71] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 856–869.

[72] S. A. Sauter, *Cubature techniques for 3-d Galerkin BEM*, in Boundary elements: Implementation and analysis of advanced algorithms, W. Hackbusch and G. Wittum, eds., Notes on numerical fluid mechanics 54, Braunschweig, 1996, Vieweg Verlag, pp. 29–44.

[73] S. A. Sauter and Ch. Lage, *Transformation of hypersingular integrals and black-box cubature*, Math. Comp., 70 (2000), pp. 223–250.

[74] S. A. Sauter and Ch. Schwab, *Randelement-methoden. Analyse, Numerik und Implementierung schneller Algorithmen*, B.G.Teubner, Stuttgart, Leipzig, Wiesbaden, 2004.

[75] G. Savaré, *Regularity results for elliptic equations in Lipschitz domains*, J. Funct. Anal., 152 (1998), pp. 176–201.

[76] A. Schatz, *An observation concerning Ritz-Galerkin methods with indefinite bilinear forms*, Math. Comp., 28 (1974), pp. 959–962.

[77] M. Schechter, *Principles of Functional Analysis*, vol. 36 of Graduate Studies in Mathematics, Americal Mathematical Society, Providence, Rhode Island, 2001.

[78] R. Schneider, *Multiskalen- und Wavelet-Matrixkompression: Analysisbasierte Methoden zur Lösung großer vollbesetzter Gleigungssysteme*, Advances in Numerical Mathematics, Teubner, Stuttgart, 1998.

[79] Ch. Schwab and R. P. Stevenson, *Adaptive wavelet algorithms for PDE's on product domains*, Technical Report 1353, Utrecht University, The Netherlands, 2006.

[80] Ch. Schwab and W. L. Wendland, *Kernel properties and representations of boundary integral operators*, Math. Nachr., 156 (1992), pp. 187–218.

[81] G. Stampacchia, *Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus*, Ann. Inst. Fourier (Grenoble), 15 (1965), pp. 189–258.

[82] E. Stein, *Singular integrals and differentiability properties of functions*, Princeton University Press, Princeton, NJ, 1970.

[83] R. P. Stevenson, *Adaptive solution of operator equations using wavelet frames*, SIAM J. Numer. Anal, 41(3) (2003), pp. 1074–1100.

[84] ——, *Locally supported, piecewise polynomial biorthogonal wavelets on non-uniform meshes*, Constr. Approx., 19 (2003), pp. 477–508.

[85] ——, *Composite wavelet bases with extended stability and cancellation properties*, Technical Report 1304, Utrecht University, The Netherlands, July 2004. Submitted.

[86] ——, *On the compressibility of operators in wavelet coordinates*, SIAM J. Math. Anal, 35(5) (2004), pp. 1110–1132.

[87] ——, *The completion of locally refined simplicial partitions created by bisection*, Technical Report 1336, Utrecht University, The Netherlands, September 2005. Submitted.

[88] ——, *Optimality of a standard adaptive finite element method*, Technical Report 1329, Utrecht University, The Netherlands, May 2005. Submitted.

[89] R. Verfürth, *A review of a posteriori error estimation and adaptive mesh-refinement techniques*, Wiley-Teubner, Chichester, 1996.