Zentrum für Technomathematik
Fachbereich 3 – Mathematik und Informatik

Adaptive Wavelet Methods for Saddle Point Problems - Optimal Convergence Rates

Stephan Dahlke **Wolfgang Dahmen**
Karsten Urban

Report 01–07

Berichte aus der Technomathematik

Report 01–07 **Juli 2001**

ADAPTIVE WAVELET METHODS FOR SADDLE POINT PROBLEMS — OPTIMAL CONVERGENCE RATES*

STEPHAN DAHLKE†, WOLFGANG DAHMEN‡, AND KARSTEN URBAN‡

Abstract. In this paper an adaptive wavelet scheme for saddle point problems is developed and analysed. Under the assumption that the underlying continuous problem satisfies the inf-sup condition it is shown in the first part under which circumstances the scheme exhibits asymptotically optimal complexity. This means that within a certain range the convergence rate which relates the achieved accuracy to the number of involved degrees of freedom is asymptotically the same as the best wavelet N-term approximation of the solution with respect to the relevant norms. Moreover, the computational work needed to compute the approximate solution stays proportional to the number of degrees of freedom. It is remarkable that compatibility constraints on the trial spaces such as the Ladyshenskaja-Babuška-Brezzi (LBB) condition do not arise. In the second part the general results are applied to the Stokes problem. Aside from the verification of those requirements on the algorithmic ingredients the theoretical analysis had been based upon, the regularity of the solutions in certain Besov scales is analyzed. These results reveal under which circumstances the work/accuracy balance of the adaptive scheme is even asymptotically better than that resulting from preassigned uniform refinements. This in turn is used to select and interpret some first numerical experiments that are to quantitatively complement the theoretical results for the Stokes problem.

Key Words: Saddle point problems, wavelet bases, norm equivalences, adaptive refinements, fast approximate operator application, Uzawa iteration.

1. Introduction. This paper draws on two major sources of motivation. First, it has recently been shown in [8] that certain adaptive wavelet schemes are asymptotically optimal for a wide class of selfadjoint elliptic operator equations. This means that the achieved accuray in the energy norm expressed in terms of the numbers of involved degrees of freedom is asymptotically the same as the rate of the *best N-term approximation*, i.e., the minimal number of basis functions needed to approximate the solution within the given accuracy tolerance. Moreover, (up to additional log-factors in sorting operations, see also Remark 4.9 below) it was shown that the computational work needed to compute the approximate solution stays proportional to the number of degrees of freedom. While the class of operator equations covers boundary value problems for partial differential equations as well as singular integral equations, symmetry did play a crucial role in the analysis and design of the scheme. These techniques have meanwhile been extended to non-coercive problems through *wavelet least squares formulations* [9].

Second, in [14] the results of a predecessor [13] of [8] also for the symmetric elliptic case have been extended to saddle point problems. The key idea there was to use an outer Uzawa iteration and to solve the interior symmetric positive definite problems by a scheme of the type considered in [13]. However, no statements about the efficiency of such schemes in terms of convergence rates and work count was made in [14].

In this paper we also consider saddle point problems actually under slightly weaker assumptions than in [14] and propose an adaptive wavelet scheme for their numerical solution. In order to avoid (among other things) the squaring of condition numbers, it is based as in [14] on an outer Uzawa iteration although it differs from the scheme in [14] in several essential ways. It draws on detailed algorithmic ingredients from [8] which allow one to quantify concrete computational steps and estimate their complexity which results in a somewhat different balance of accuracies. It also applies when the symmetric bilinear form is only elliptic on the kernel of the constraint operator.

On a more fundamental level, in the same spirit as in [8, 9], there are two essential features that distinguish the present approach from [13, 14] and more so from classical discretizations. The first one is that through appropriate wavelet bases the original continuous problem is transformed right from the beginning into an *equivalent* problem which is *well-posed in the Euclidean metric*. All essential computational steps refer then to approximation in ℓ_2 and therefore bear a great potential of being portable to other problem classes. In fact, many of the basic routines developed in [2, 8] in the context of elliptic problems can be used here as well. The second important point is that the wavelet representation allows us to think of performing, up to a controlled perturbation, an iteration on the full infinite dimensional

*This work has been supported in part by the Deutsche Forschungsgemeinschaft grants Da 117/8–2, SFB 401, and the TMR Network "Wavelets in Numerical Simulation" funded by the European Commission.

†Universität Bremen, Fachbereich 3, ZeTem, Postfach 330440, 28359 Bremen, Germany, dahlke@math.uni-bremen.de

‡RWTH Aachen, Institut für Geometrie und Praktische Mathematik, Templergraben 55, 52056 Aachen, Germany, {dahmen,urban}@igpm.rwth-aachen.de

problem realized through the *adaptive approximate application* of the full infinite dimensional operators. The tolerances have to be chosen so that the convergence speed of the perturbed realizable iteration is indeed governed by the properties of the ideal infinite dimensional iteration.

This offers, in particular, a first intuitive explanation for the following fact which at the first glance strikes one as a paradoxon, namely compatibility constraints on the choice of trial functions such as the *LBB condition* do *not* arise. In fact, recall that even when the infinite dimensional saddle point problem is well posed and hence satisfies an inf-sup condition inappropriate choices of finite dimensional trial spaces could lead to discrete problems with poor stability properties, that is the inverses of the corresponding system matrices may have arbitrarily large norm. This fact is relevant whenever linear systems are too be solved for any such given pair of trial spaces. In the present context this situation will never arise. Instead an iterative process is conceptually applied to the *full infinite dimensional problem* where each iteration involves an adaptive application of the underlying infinite dimensional operators within a certain stage dependent dynamic accuracy tolerance. This process is inherently nonlinear. Roughly speaking proper adaptation in the above sense inherits the stability of the infinite dimensional problem. In this sense adaptation not only reduces complexity but also *stabilizes* the computation *automatically*.

The paper is organized as follows. After formulating the problem in Section 2 we describe and analyse an adaptive method in Section 3. It will be shown in Section 4 under which conditions on the algorithmic ingredients it exhibits an *asymptotically optimal accuracy/work balance* in the following sense. Whenever the exact solution has, within a certain range of exponents s, an error of *best N-term approximation* with respect to an underlying wavelet basis decaying like N^{-s}, then the error achieved by the adaptive scheme also decays like N^{-s} where N is the number of used degrees of freedom. Moreover, the computational work stays proportional to N. A key role in this context is played by the *compressibility range* of the involved operators in wavelet coordinates. Given this property one can apply a certain adaptive scheme for applying the operator to any finitely supported vector with optimal accuracy/work balance [8].

In Section 5 the general results are applied to the *Stokes problem*. Specifically, we investigate in Section 5.3 the compressibility range of the wavelet representation of the Stokes operator for a certain family of wavelet bases and derive sharp estimates for this range. This identifies the range of decay rates for which the general results from the preceding sections apply.

It should be stressed that the scheme works without *any* a-priori assumptions on the solution while its complexity is analysed under the assumption that the solution has a certain order of best N-term approximation and the involved operators in wavelet coordinates have a certain compressibility range (see Section 4). Certain rates of best N-term approximation, in turn, are (almost) equivalent to a certain regularity of the solution in a *Besov scale*. Roughly speaking, when the Sobolev regularity of the solution is lower than its Besov regularity, the adaptive scheme is expected to offer even an *asymptotically better* accuracy/work balance than linear schemes. To see whether or under which circumstances the adaptive scheme can be rigorously proven to offer even an asymptotically better accuracy/work balance than schemes based on uniform preassigned mesh refinements, we investigate in Section 5.4 the Besov regularity of singularity solutions for the Stokes problem. The results show that in two spatial dimensions sufficiently high order wavelet bases would give rise to adaptive schemes with *arbitrarily* high convergence rates.

Finally in Section 6 we present some numerical experiments essentially guided by the above mentioned theoretical considerations. Here we make use of the software developed in [2] as well as in [25]. The results confirm that the adaptive scheme performs essentially independently of the pairing of trial functions for velocities and pressure. For instance, the rate of best N-term approximation is met within a factor two when both velocities and pressure are approximated by piecewise linear trial functions.

After completion of this work we became aware of related investigations in [4] pursuing similar ideas in a finite element context. There convergence in the sense of [14] is proven for a similar Uzawa technique without establishing, however, rigorous estimates for the corresponding work/accuracy balance.

2. Saddle Point Problems.

2.1. The Setting. Let X, M denote Hilbert spaces with norms $\|\cdot\|_X, \|\cdot\|_M$, respectively. Dual pairings on $X \times X'$ and $M \times M'$ (X', M' denoting the duals of X, M, respectively) will always be denoted by $\langle \cdot, \cdot \rangle$. It will be clear from the context which spaces are referred to. Suppose that $a(\cdot, \cdot)$ is a continuous

symmetric bilinear form on $X \times X$ and that $b(\cdot,\cdot)$ is a continuos bilinear form on $X \times M$, i.e.,

$$|a(v,w)| \lesssim \|v\|_X \|w\|_X, \quad |b(q,v)| \lesssim \|v\|_X \|q\|_M.$$

Moreover, denoting by $B : X \to M'$ the operator induced by $b(p,v) = \langle p, Bv\rangle$ and setting $V := \ker B$, assume that $a(\cdot,\cdot)$ is elliptic on V and $b(\cdot,\cdot)$ satisfies the *inf-sup condition*

$$a(v,v) \geq \alpha \|v\|_X^2, \quad v \in V, \qquad \inf_{q\in M} \sup_{v\in X} \frac{b(v,q)}{\|v\|_X \|q\|_M} > \beta. \tag{2.1.1}$$

It is well known that then the variational problem has for any $f \in X'$, $g \in M'$

$$\begin{array}{rcccl} a(u,v) & + & b(p,v) & = & \langle f, v\rangle \quad \forall\, v \in X, \\ b(q,u) & & & = & \langle q, g\rangle \quad \forall\, q \in M, \end{array} \tag{2.1.2}$$

a unique solution $U = (u,p) \in X \times M$, see e.g. [5]. Defining $A : X \to X'$ by $a(v,w) = \langle v, Aw\rangle$, $v \in X$, (2.1.2) is equivalent to the 2×2 block operator equation

$$\mathcal{L}U := \begin{pmatrix} A & B' \\ B & 0 \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix} =: F, \tag{2.1.3}$$

where $\mathcal{L}$ is an isomorphism from $X \times M$ into its dual $X' \times M'$, i.e. there exist positive constants $c_{\mathcal{L}}, C_{\mathcal{L}}$ such that

$$c_{\mathcal{L}} \left(\|v\|_X^2 + \|q\|_M^2\right)^{1/2} \leq \left\| \mathcal{L} \begin{pmatrix} v \\ q \end{pmatrix} \right\|_{X' \times M'} \leq C_{\mathcal{L}} \left(\|v\|_X^2 + \|q\|_M^2\right)^{1/2}. \tag{2.1.4}$$

Classical examples are mixed formulations of second order elliptic boundary value problems, the Stokes problem or the system obtained when appending essential boundary conditions by Lagrange multipliers.

2.2. Wavelet Coordinates. Now suppose that we have wavelet bases $\Psi_X = \{\psi_{X,\lambda} : \lambda \in \mathcal{J}_X\}$, $\Psi_M = \{\psi_{M,\lambda} : \lambda \in \mathcal{J}_M\}$ for X and M at our disposal such that for suitable diagonal matrices $\mathbf{D}_X$, $\mathbf{D}_M$ and constants c_X, C_X, c_M, C_M one has

$$c_X \|\mathbf{v}\|_{\ell_2(\mathcal{J}_X)} \leq \|\mathbf{v}^T \mathbf{D}_X^{-1} \Psi_X\|_X \leq C_X \|\mathbf{v}\|_{\ell_2(\mathcal{J}_X)}, \tag{2.2.1}$$

and likewise

$$c_M \|\mathbf{q}\|_{\ell_2(\mathcal{J}_M)} \leq \|\mathbf{q}^T \mathbf{D}_M^{-1} \Psi_M\|_M \leq C_M \|\mathbf{q}\|_{\ell_2(\mathcal{J}_M)}, \tag{2.2.2}$$

where $\mathbf{v}^T \mathbf{D}_X^{-1} \Psi_X := \sum_{\lambda\in\mathcal{J}_X} d_{X,\lambda}^{-1} v_\lambda \psi_{X,\lambda}$. The validity of such norm equivalences will be crucial in what follows. Note that often M is a closed subspace of finite codimension in a larger Hilbert space $\hat{M}$ for which (2.2.2) holds. For instance, in the case of the Stokes problem M is the space of all L_2 functions with zero mean. Thus the arrays of wavelet coefficients of elements in M will in general form a closed subspace $\ell_{2,0}(\mathcal{J}_M)$ of finite codimension in $\ell_2(\mathcal{J}_M)$.

At this point we dispense with any additional technical details about the precise nature of the basis functions but refer to [7, 15] for surveys and further references, see also the comments in connection with numerical realizations below. A further important property is the *cancellation property* which entails near sparseness of wavelet representations for many operators. This will also be detailed when necessity arises.

Defining now for any two countable arrays Θ, Φ and some inner product $c(\cdot,\cdot)$ the matrix $c(\Theta,\Phi) := (c(\theta,\phi))_{\theta\in\Theta,\phi\in\Phi}$, consider as usual the scaled wavelet representations

$$\mathbf{A} := a(\mathbf{D}_X^{-1}\Psi, \mathbf{D}_X^{-1}\Psi), \quad \mathbf{B} := b(\mathbf{D}_M^{-1}\Psi_M, \mathbf{D}_X^{-1}\Psi_X), \tag{2.2.3}$$

as well as the arrays $\mathbf{f} := \mathbf{D}_X^{-1}\langle \Psi_X, f\rangle$, $\mathbf{g} := \mathbf{D}_M^{-1}\langle \Psi_M, g\rangle$ and $\mathbf{F} := (\mathbf{f}^T, \mathbf{g}^T)^T$. Then (2.1.2) or (2.1.3) is equivalent to the following two by two block matrix system

$$\begin{pmatrix} \mathbf{A} & \mathbf{B}^T \\ \mathbf{B} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \mathbf{g} \end{pmatrix}. \tag{2.2.4}$$

It will make things much more transparent when working from now on exclusively in the ℓ_2 setting.

2.3. Well-Posedness in ℓ_2. It follows from (2.2.1) and (2.2.2) together with (2.1.4) that the operator

$$\mathbf{L} := \begin{pmatrix} \mathbf{A} & \mathbf{B}^T \\ \mathbf{B} & \mathbf{0} \end{pmatrix} : \ell_2(\mathcal{J}) := \ell_2(\mathcal{J}_X) \times \ell_{2,0}(\mathcal{J}_M) \to \ell_2(\mathcal{J}), \quad \mathcal{J} := \mathcal{J}_X \times \mathcal{J}_M,$$

is an isomorphism, i.e., there exist positive constants c_L, C_L such that for $\mathbf{V} := (\mathbf{v}^T, \mathbf{q}^T)^T \in \ell_2(\mathcal{J})$, $\|\mathbf{V}\|^2_{\ell_2(\mathcal{J})} = \|\mathbf{v}\|^2_{\ell_2(\mathcal{J}_X)} + \|\mathbf{q}\|^2_{\ell_2(\mathcal{J}_M)}$

$$c_L\|\mathbf{V}\|_{\ell_2(\mathcal{J})} \le \|\mathbf{L}\mathbf{V}\|_{\ell_2(\mathcal{J})} \le C_L\|\mathbf{V}\|_{\ell_2(\mathcal{J})}, \quad \mathbf{V} \in \ell_2(\mathcal{J}), \tag{2.3.1}$$

see e.g., [15, 22] for further details. Clearly c_L, C_L can be expressed in terms of the constants $c_{\mathcal{L}}, C_{\mathcal{L}}$, c_Y, C_Y for $Y \in \{X, M\}$. Furthermore there exist constants C_B, C'_A such that

$$\|\mathbf{B}\mathbf{v}\|_{\ell_2(\mathcal{J}_M)} \le C_B\|\mathbf{v}\|_{\ell_2(\mathcal{J}_X)}, \quad \|\mathbf{B}^T\mathbf{q}\|_{\ell_2(\mathcal{J}_X)} \le C_B\|\mathbf{q}\|_{\ell_2(\mathcal{J}_M)}. \tag{2.3.2}$$

and

$$\|\mathbf{A}\mathbf{v}\|_{\ell_2(\mathcal{J}_X)} \le C'_A\|\mathbf{v}\|_{\ell_2(\mathcal{J}_X)}. \tag{2.3.3}$$

2.4. The Schur Complement. In many cases a somewhat stronger property than the first relation in (2.1.1) is valid, namely that

$$a(v, v) \sim \|v\|^2_X, \qquad v \in X, \tag{2.4.1}$$

which, of course means that $\mathbf{A}$ is invertible on all of $\ell_2(\mathcal{J}_X)$. In this case block elimination reduces (2.2.4) to the so called reduced system

$$\mathbf{S}\mathbf{p} = \mathbf{B}\mathbf{A}^{-1}\mathbf{f} - \mathbf{g}, \tag{2.4.2}$$

involving the (infinite dimesnional) *Schur complement*

$$\mathbf{S} := \mathbf{B}\mathbf{A}^{-1}\mathbf{B}^T : \ell_{2,0}(\mathcal{J}_M) \to \ell_{2,0}(\mathcal{J}_M) \tag{2.4.3}$$

which is symmetric positive definite and, under the above assumptions, in fact an automorphism on $\ell_{2,0}(\mathcal{J}_M)$, i.e., there exist positive constants c_S, C_S such that

$$c_S\|\mathbf{q}\|_{\ell_2(\mathcal{J}_M)} \le \|\mathbf{S}\mathbf{q}\|_{\ell_2(\mathcal{J}_M)} \le C_S\|\mathbf{q}\|_{\ell_2(\mathcal{J}_M)}, \quad \mathbf{q} \in \ell_{2,0}(\mathcal{J}_M). \tag{2.4.4}$$

Once $\mathbf{p}$ has been determined from (2.4.2) it remains to solve the positive definite problem

$$\mathbf{A}\mathbf{u} = \mathbf{f} - \mathbf{B}^T\mathbf{p}. \tag{2.4.5}$$

However, under the weaker assumption (2.1.1) on the bilinear form $a(\cdot, \cdot)$ one has to take first a precaution whose variational counterpart is sometimes referred to as *augmented Lagrangian method*. In the present setting it boils down to considering the matrix

$$\hat{\mathbf{A}} := \mathbf{A} + c\mathbf{B}^T\mathbf{B}, \tag{2.4.6}$$

where c is some sufficiently large but fixed positive constant.

REMARK 2.1. *Under the assumption (2.1.1) the matrix* $\hat{\mathbf{A}}$ *is an automorphism on* $\ell_2(\mathcal{J}_X)$*, i.e., there exist positive constants* c_A, C_A *such that*

$$c_A\|\mathbf{v}\|_{\ell_2(\mathcal{J}_X)} \le \|\hat{\mathbf{A}}\mathbf{v}\|_{\ell_2(\mathcal{J}_X)} \le C_A\|\mathbf{v}\|_{\ell_2(\mathcal{J}_X)}, \quad \mathbf{v} \in \ell_2(\mathcal{J}_X). \tag{2.4.7}$$

Proof: It follows from (2.3.2) and (2.3.3) that $\hat{\mathbf{A}}$ is bounded on $\ell_2(\mathcal{J}_X)$. Moreover, by (2.3.1) the matrix $\mathbf{L}^T\mathbf{L} = \mathbf{L}^2$ is positive definite on $\ell_2(\mathcal{J})$. Since $\mathbf{B}\mathbf{B}^T$ is a principal block of $\mathbf{L}^2$ it is positive definite on $\ell_{2,0}(\mathcal{J}_M)$. This entails that $\hat{\mathbf{A}}$ is also injective on $\ell_2(\mathcal{J}_X)$. To see this note that by the first relation in (2.1.1), $\mathbf{v}^T\hat{\mathbf{A}} \neq \mathbf{0}$ for $\mathbf{v} \in \ker\mathbf{B}$. On the other hand, when $\mathbf{v}$ is in the range of $\mathbf{B}^T$, i.e., $\mathbf{v} = \mathbf{B}^T\mathbf{q}$ for some $\mathbf{q} \in \ell_{2,0}(\mathcal{J}_M)$, then one has

$$\mathbf{v}^T\hat{\mathbf{A}}\mathbf{v} = \mathbf{q}^T\mathbf{B}\mathbf{A}\mathbf{B}^T\mathbf{q} + c\|\mathbf{B}\mathbf{B}^T\mathbf{q}\|^2_{\ell_2(\mathcal{J}_M)} \tag{2.4.8}$$

which, by the previous remark, is strictly positive whenever $\mathbf{p} \neq \mathbf{0}$, confirming injectivity of $\hat{\mathbf{A}}$ on $\ell_2(\mathcal{J}_X)$. By symmetry (2.4.8) also implies surjectivity. Due to the boundedness of $\hat{\mathbf{A}}$, the claim follows now from the Inverse Mapping Theorem. ∎

Now multiply (2.2.4) from the left by the $\ell_2(\mathcal{J})$-isomorphism

$$\begin{pmatrix} \mathbf{id} & c\mathbf{B}^T \\ \mathbf{B} & \mathbf{0} \end{pmatrix}, \tag{2.4.9}$$

which yields the equivalent system

$$\begin{pmatrix} \hat{\mathbf{A}} & \mathbf{B}^T \\ \mathbf{B} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{f}} \\ \mathbf{g} \end{pmatrix}, \tag{2.4.10}$$

for some $c > 0$, where for $\hat{\mathbf{A}}$ given by (2.4.6)

$$\hat{\mathbf{f}} := \mathbf{f} + c\mathbf{B}^T\mathbf{g}.$$

By Remark 2.1 block elimination can be applied to this new system (2.4.10) which then reduces to the coupled systems (2.4.2), (2.4.5) with $\mathbf{A}$ and $\mathbf{f}$ replaced by $\hat{\mathbf{A}}$, respectively $\hat{\mathbf{f}}$.

To simplify notation we will use the following convention throughout the remainder of the paper. We will always set

$$\mathbf{A} := \mathbf{D}_X^{-1}a(\Psi_X, \Psi_X)\mathbf{D}_X^{-1} + c\mathbf{B}^T\mathbf{B}, \quad \mathbf{f} := \mathbf{D}_X^{-1}\langle\Psi_X, f\rangle + c\mathbf{B}^T\mathbf{g}, \tag{2.4.11}$$

with $\mathbf{B} := \mathbf{D}_M^{-1}b(\Psi_M, \Psi_X)\mathbf{D}_X^{-1}$ as in (2.2.3) and $\mathbf{g} := \mathbf{D}_M^{-1}\langle\Psi_M, g\rangle$. When the bilinear form $a(\cdot,\cdot)$ satisfies the stronger assumption (2.4.1) the constant c in (2.4.11) can be chosen to be zero. Otherwise, c is any fixed positive number. Thus without loss of generality we can always make use of the reduced systems (2.4.3), (2.4.5) with a proper interpretation of the matrix $\mathbf{A}$ according to the above convention. Consequently, $\mathbf{A}$ satisfies in this sense (2.4.7).

A standard way of formulating finite dimensional problems is to take Galerkin discretizations for (2.1.2). As soon as one *fixes* a pair of finite dimensional trial spaces in X and M, for instance, spanned by collections of wavelets, the corresponding Galerkin discretization gives rise to a finite dimensional linear system, e.g. in terms of a principal finite submatrix of (2.2.4). However, it is well-known that stability of the infinite dimensional problem does not guarantee the finite dimensional problems to be uniformly stable as well. Compatibility constraints in terms of the LBB condition come into play. It will be seen that this will *not* be the case in the following adaptive framework.

3. An Adaptive Uzawa-Strategy.

3.1. Infinite Dimensional Uzawa Iteration. The idea is to use a stationary iterative scheme for the solution of the reduced system (2.4.2) which is essentially the Uzawa strategy proposed in [14]. In contrast, we formulate it here directly for the discrete infinite dimensional ℓ_2-problem (2.2.4). To this end, we have to address first an issue which is somewhat hidden in the ℓ_2-setting. The Spaces X, M are always function spaces on some domain Ω. As will be explained in more detail later the wavelet bases Ψ_X and Ψ_M are then typically constructed as *Riesz bases* for the corresponding spaces $L_2(\Omega)$, i.e, in addition to the norm equivalences (2.2.1), (2.2.2) one also has

$$\|\mathbf{v}\|_{\ell_2(\mathcal{J}_X)} \sim \|\mathbf{v}^T\Psi_X\|_{L_2(\Omega)}, \quad \|\mathbf{q}\|_{\ell_2(\mathcal{J}_M)} \sim \|\mathbf{q}^T\Psi_M\|_{L_2(\Omega)}. \tag{3.1.1}$$

This means that there exist *dual* bases $\tilde{\Psi}_X$, $\tilde{\Psi}_M$ in $L_2(\Omega)$ which are also Riesz bases and satisfy

$$(\Psi_X, \tilde{\Psi}_X) = \mathbf{id}, \quad (\Psi_M, \tilde{\Psi}_M) = \mathbf{id}, \tag{3.1.2}$$

where $(\cdot,\cdot)$ denotes the standard inner product in $L_2(\Omega)$. In full agreement with the fact that the operator B maps X into M' one observes that for $v = \mathbf{v}^T\mathbf{D}_X^{-1}\Psi_X$ the array $\mathbf{Bv}$ represents expansion coefficients of Bv with respect to the *dual basis* $\tilde{\Psi}_M$. In fact,

$$\begin{aligned}(\mathbf{Bv})^T\mathbf{D}_M\tilde{\Psi}_M &= \mathbf{v}^T\langle B\mathbf{D}_X^{-1}\Psi_X, \Psi_M\rangle\mathbf{D}_M^{-1}\mathbf{D}_M\tilde{\Psi}_M = \mathbf{v}^T\langle B\mathbf{D}_X^{-1}\Psi_X, \Psi_M\rangle\tilde{\Psi}_M \\ &= B(\mathbf{v}^T\mathbf{D}_X^{-1}\Psi_X) = Bv.\end{aligned}$$

Likewise the array $\mathbf{g}$ consists by definition of the wavelet coefficients with respect to the dual basis $\tilde{\Psi}_M$. On the other hand, the unknown array $\mathbf{q}$ in the reduced system (2.4.2) contains coefficients with respect to the primal basis Ψ_M. Now, as mentioned before, in some cases the space M is actually a closed subspace of a somewhat larger Hilbert space characterized by Ψ_M. Therefore the wavelet coefficients of elements of M with respect to Ψ_M (or $\mathbf{D}_M^{-1}\Psi_M$) satisfy certain constraints which generally depend on the particular wavelet basis. To change representations if necessary, observe that, in view of (3.1.2) $\tilde{\Psi}_M = (\tilde{\Psi}_M, \tilde{\Psi}_M)\Psi_M$, so that such a change of bases is realized by the matrix

$$\mathbf{R} := (\tilde{\Psi}_M, \tilde{\Psi}_M) \tag{3.1.3}$$

because

$$\tilde{\mathbf{p}}^T\tilde{\Psi}_M = \tilde{\mathbf{p}}^T\mathbf{R}\Psi_M = (\mathbf{R}\tilde{\mathbf{p}})^T\Psi_M.$$

It immediately follows from (3.1.1) that both $\mathbf{R}$ and $\mathbf{R}^{-1} = (\Psi_M, \Psi_M)$ are bounded on $\ell_2(\mathcal{J}_M)$,

$$\|\mathbf{R}\|_{\ell_2(\mathcal{J}_M)\to\ell_2(\mathcal{J}_M)} \le C_R. \tag{3.1.4}$$

Since $\mathbf{S}$ is positive definite and satisfies (2.4.4) there exists therefore some positive ω (e.g. $\omega < 2C_SC_R$) such that

$$\rho := \|\mathbf{id} - \omega\mathbf{RS}\|_{\ell_2(\mathcal{J}_M)\to\ell_2(\mathcal{J}_M)} < 1. \tag{3.1.5}$$

Then the infinite dimensional version of the Uzawa scheme reads as follows.

UZAWA: *Given any* $\mathbf{p}_0 \in \ell_{2,0}(\mathcal{J}_M)$, *compute for* $i = 1, 2, \ldots$

$$\mathbf{Au}_i = \mathbf{f} - \mathbf{B}^T\mathbf{p}_{i-1}, \tag{3.1.6}$$

$$\mathbf{p}_i = \mathbf{p}_{i-1} + \omega\mathbf{R}(\mathbf{Bu}_i - \mathbf{g}). \tag{3.1.7}$$

This is known to converge when $\rho < 1$. In fact, since $\mathbf{u} = \mathbf{A}^{-1}(\mathbf{f} - \mathbf{B}^T\mathbf{p})$ it is easy to see that

$$\mathbf{p} - \mathbf{p}_i = (\mathbf{id} - \omega\mathbf{RS})(\mathbf{p} - \mathbf{p}_{i-1}),$$

so that

$$\|\mathbf{p} - \mathbf{p}_i\|_{\ell_2(\mathcal{J}_M)} \le \rho^i\|\mathbf{p} - \mathbf{p}_0\|_{\ell_2(\mathcal{J}_M)}. \tag{3.1.8}$$

Moreover, it has been shown in [14] that for $\mathbf{p}_0 = \mathbf{0}$ one has

$$\|\mathbf{p} - \mathbf{p}_i\|_{\ell_2(\mathcal{J}_M)} \le \|\mathbf{A}^{-1}\mathbf{f}\|_{\ell_2(\mathcal{J}_X)}\|\omega\mathbf{RSB}\|_{\ell_2(\mathcal{J}_X)\to\ell_2(\mathcal{J}_M)}\frac{\rho^i}{1-\rho}. \tag{3.1.9}$$

3.2. The Adaptive Scheme. As in [9] the key idea is to apply the above Uzawa iteration to the ininite dimensional problem. In view of (3.1.6) and (3.1.7), this involves three tasks, namely adding sequences with generally infinite support such as the data $\mathbf{f}$ and $\mathbf{g}$, the application of *infinite* matrices like $\mathbf{B}$ or $\mathbf{B}^T$ to *finitely supported* vectors, as well as the solution of elliptic problems involving the infinite matrix $\mathbf{A}$. Of course, in practice neither one of these tasks can be performed *exactly*. Therefore one has to employ suitable approximations whose accuracy will depend on the current stage of the algorithm and which will be described next.

To this end, we shall not distinguish formally between finitely supported vectors and infinite sequences in $\ell_2(\mathcal{J}')$ where in the sequel $\mathcal{J}' \in \{\mathcal{J}_X, \mathcal{J}_M\}$, but will rather view both quantities as sequences (expanded by zero entries if necessary).

The first basic ingredient is the routine

NCOARSE $[\eta, \mathbf{v}] \to (\bar{\mathbf{v}}, \Lambda)$ which determines for a given *finitely supported* vector $\mathbf{v}$ a vector $\bar{\mathbf{v}}$ with smallest possible support Λ such that

$$\|\mathbf{v} - \bar{\mathbf{v}}\|_{\ell_2(\mathcal{J})} \leq \eta. \tag{3.2.1}$$

In particular, **NCOARSE** will be used to approximate the arrays $\mathbf{f}^\circ := \mathbf{D}_X^{-1}\langle \Psi_X, f\rangle_X$ and $\mathbf{g}$ of given data by finitely supported vectors. The way how to think about **NCOARSE** in this context can be formulated as the following

Assumption f: In a preprocessing step for a given target accuracy sufficiently many (wavelet) coefficients in the arrays $\mathbf{f}^\circ$ *and* $\mathbf{g}$ *are made available and ordered by size.*

In many applications f and g are simple and, as model data given by the user, are considered here as completely accessible. Coarser approximations of the data are then obtained by applying **NCOARSE** to these preprocessed finite arrays (see Section 6.1 in [8] for a more detailed discussion).

The second basic ingredient is an approximate application of an infinite matrix to a finitely supported vector. Given an infinite matrix $\mathbf{C}$ (as a mapping from $\ell_2(\mathcal{J}'')$ to $\ell_2(\mathcal{J}')$ for any pair $(\mathcal{J}', \mathcal{J}'') \in \{\mathcal{J}_X, \mathcal{J}_M\}^2$), the scheme

APPLY $[\eta, \mathbf{C}, \mathbf{v}] \to (\mathbf{w}, \Lambda)$ produces for any finitely supported input vector $\mathbf{v}$ a vector $\mathbf{w}$ with finite support $\Lambda \subset \mathcal{J}'$ such that

$$\|\mathbf{C}\mathbf{v} - \mathbf{w}\|_{\ell_2(\mathcal{J}')} \leq \eta. \tag{3.2.2}$$

A scheme with this property has been developed in [8]. We postpone a quick description of the relevant features along with estimates for its computational cost to a later section.

Note that, in particular, the routines **APPLY** and **NCOARSE** allow us to approximately evaluate the right-hand sides of (3.1.6) and (3.1.7).

So the remaining task in an approximate Uzawa iteration of the form (3.1.6), (3.1.7) is to solve the operator equation (3.1.6) with system matrix $\mathbf{A}$. This is an elliptic problem in the sense of [8] and we will make heavy use of the results obtained there, see also [2] for implementations and numerical tests. The scheme from [8] is also built solely on the above routines **NCOARSE** and **APPLY**. There are, however, two minor points that need to be briefly addressed. First in [8] the matrix $\mathbf{A}$ is just the wavelet representation of the underlying elliptic operator while in the present situation $\mathbf{A}$ has the form (2.4.11) for some positive constant c when $a(\cdot,\cdot)$ is not elliptic on all of X. Nevertheless, once a scheme **APPLY** for wavelet representations is availabele a scheme for applying matrices of the form (2.4.11) with $c \neq 0$ is easily obtained from such a building block as follows. To simplify notation we set $\mathbf{A}^\circ := \mathbf{D}_X^{-1} a(\Psi_X, \Psi_X)\mathbf{D}_X^{-1}$:

APPLY$^*[\eta, \mathbf{A}, \mathbf{v}] \to (\mathbf{w}, \Lambda)$
(i) **APPLY** $[\eta/2, \mathbf{A}^\circ, \mathbf{v}] \to (\mathbf{w}_1, \Lambda_1)$;

(ii) **APPLY** $[\eta/4cC_B, \mathbf{B}, \mathbf{w}_1] \to (\mathbf{w}_2, \Lambda_2)$;

(iii) APPLY $[\eta/4, c\mathbf{B}^T, \mathbf{w}_2] \to (\mathbf{w}_3, \Lambda_3)$ and set

$$\mathbf{w} := \mathbf{w}_1 + \mathbf{w}_3, \quad \Lambda := \Lambda_1 \cup \Lambda_3.$$

REMARK 3.1. *One easily derives from (3.2.2) that the output* $\mathbf{w}$ *produced by* **APPLY**$^*[\eta, \mathbf{A}, \mathbf{v}]$ *satisfies for* $\mathbf{A}$ *given by (2.4.11)*

$$\|\mathbf{A}\mathbf{v} - \mathbf{w}\|_{\ell_2(\mathcal{J}_X)} \le \eta. \tag{3.2.3}$$

Moreover, it is also clear that up to a uniform constant the work/accuracy balance for **APPLY*** *is the same as that for* **APPLY**. *Note that the matrix* $\mathbf{B}^T\mathbf{B}$ *is, of course, never computed.*

We will extract now from the results in [8] a version for the treatment of (3.1.6) (with **APPLY** replaced by **APPLY*** if necessary) that suits the present needs best. To this end, consider for $\mathbf{A}$ as above the elliptic problem

$$\mathbf{A}\mathbf{u} = \mathbf{h} \tag{3.2.4}$$

for some $\mathbf{h} \in \ell_2(\mathcal{J}_X)$ with exact solution $\hat{\mathbf{u}}$.

ELLSOLVE $[\varepsilon, \mathbf{A}, \mathbf{v}, \mathbf{h}] \to (\bar{\mathbf{u}}, \Lambda)$
Given $\varepsilon > 0$ *and an approximate solution* $\mathbf{v}$ *to (3.2.4), then the output* $\bar{\mathbf{u}}$ *with finite support* Λ *satisfies*

$$\|\hat{\mathbf{u}} - \bar{\mathbf{u}}\|_{\ell_2(\mathcal{J}_X)} \le \varepsilon. \tag{3.2.5}$$

The second point is that in [8] the right-hand data are assumed to be a given array of wavelet coefficients as explained above that can be preprocessed. In the present situation the right-hand data are composed of such preprocessable data like $\mathbf{f}$ and an additional matrix/vector product involving dynamically updated entities. We therefore have to approximate these data by finitely supported vectors that can then be processed as in Sections 7.2, 7.3 of [8]. The corresponding perturbations can be estimated as follows.

REMARK 3.2. *Consider again (3.2.4) and suppose that approximate finitely supported right-hand side data* $\mathbf{h}_\eta \in \ell_2(\mathcal{J}_X)$ *are given such that*

$$\|\mathbf{h} - \mathbf{h}_\eta\|_{\ell_2(\mathcal{J}_X)} \le \eta. \tag{3.2.6}$$

Then the output $\bar{\mathbf{u}}$ *of* **ELLSOLVE** $[\eta, \mathbf{A}, \mathbf{v}, \mathbf{h}_\eta]$ *satisfies*

$$\|\hat{\mathbf{u}} - \bar{\mathbf{u}}\|_{\ell_2(\mathcal{J}_X)} \le \varepsilon + c_A^{-1}\eta. \tag{3.2.7}$$

Proof: The claim follows from (3.2.5) combined with (2.4.7) to estimate the perturbation effect. □

We will describe next the computation of a finitely supported $\mathbf{h}_\eta$ when $\mathbf{h} = \mathbf{f} - \mathbf{B}^T\bar{\mathbf{p}}_{i-1}$, see (3.1.6). Defining $\mathbf{f}^\circ := \mathbf{D}_X^{-1}\langle\Psi_X, f\rangle$, recall from (2.4.11) that

$$\mathbf{f} - \mathbf{B}\bar{\mathbf{p}}_{i-1} = \mathbf{f}^\circ - \mathbf{B}^T(\bar{\mathbf{p}}_{i-1} - c\mathbf{g}),$$

which thus involves coarsening the given (preprocessed) data $\mathbf{f}^\circ, \mathbf{g}$ and a multiplication by $\mathbf{B}^T$. The respective concrete accuracy tolerences are given in the following routine:

RHS $[\bar{\mathbf{p}}, \eta] \to (\mathbf{h}_\eta, \Lambda^h)$
Given a finitely supported $\bar{\mathbf{p}}$ *the routine* **RHS** *computes a vector* $\mathbf{h}_\eta$ *with finite support* Λ^h *satisfying*

$$\|\mathbf{f} - \mathbf{B}^T\bar{\mathbf{p}} - \mathbf{h}_\eta\|_{\ell_2(\mathcal{J}_X)} \le \eta, \tag{3.2.8}$$

as follows:

(i) Apply **NCOARSE** $[\eta/3, \mathbf{f}^\circ] \to (\bar{\mathbf{f}}, \Lambda^f)$, **NCOARSE** $[\eta/3cC_B, \mathbf{g}] \to (\bar{\mathbf{g}}, \Lambda^g)$ *and set* $\mathbf{r} := \bar{\mathbf{g}} + \bar{\mathbf{p}}$.

(ii) **APPLY** $[\eta/3, \mathbf{B}^T, \mathbf{r}] \to (\mathbf{w}, \Lambda^w)$ *and set*

$$\mathbf{h}_\eta := \bar{\mathbf{f}} - \mathbf{w}, \quad \Lambda^h := \Lambda^f \cup \Lambda^w.$$

Since by (3.2.1) $\|(\bar{\mathbf{p}} - c\mathbf{g}) - \mathbf{r}\|_{\ell_2(\mathcal{J}_X)} \leq \eta/3C_B$ the estimate (3.2.8) indeed readily follows from (3.2.2).

Our numerical realization of the *ideal* (infinite dimensional) Uzawa scheme (3.1.6), (3.1.7) has the following structure. A fixed uniformly bounded number K, depending only on the constants associated with the wavelet bases and the mapping properties of the involved operators, of approximate applications of (3.1.6), (3.1.7) are applied which is then followed by a coarsening step before the iteration is further resumed. Such an *iteration block* will be arranged to *advance* the current approximate solutions so as to reduce the current error bounds by a fixed factor. Before giving a precise description, we would like to stress that the Uzawa scheme as a gradient method for the reduced system (2.4.2) treats in some sense $q \in M$ as the "preferred" variable. In fact, the accuracy of the approximate solution to the elliptic problem (3.1.6) need not be too accurate relative the the current accuracy of the approximation to q. In order to formulate now the basic iteration block as a concrete routine we will use the following choice for the number K of perturbed iterations before the next coarsening step. Let γ_i denote any positive summable numbers, e.g. $\gamma_i = (1+i)^{-2}$. Moreover, we need some control parameters. Set

$$C_1 := \omega(C_R C_B + 2)\gamma + 1, \tag{3.2.9}$$

where $\gamma := \sum_{i=0}^{\infty} \gamma_i$, and let K denote the smallest integer such that

$$\rho^K ((\rho c_A)^{-1} C_B C_1 + 1) \leq 1/10. \tag{3.2.10}$$

ADV $[\bar{\mathbf{u}}, \bar{\mathbf{p}}, \delta] \to (\tilde{\mathbf{u}}, \tilde{\mathbf{p}}, \Lambda_u, \Lambda_q)$
Given current approximations $\bar{\mathbf{u}}, \bar{\mathbf{p}}$ *of the solution to (2.2.4) such that*

$$\|\bar{\mathbf{u}} - \mathbf{u}\|_{\ell_2(\mathcal{J}_X)} \leq \delta, \quad \|\bar{\mathbf{p}} - \mathbf{p}\|_{\ell_2(\mathcal{J}_M)} \leq \delta, \tag{3.2.11}$$

ADV $[\bar{\mathbf{u}}, \bar{\mathbf{p}}, \delta]$ *produces new approximations* $\tilde{\mathbf{u}}, \tilde{\mathbf{p}}$ *as follows*
(i) *Set* $i = 1$, $\bar{\mathbf{p}}_0 := \bar{\mathbf{p}}$, $\bar{\mathbf{u}}_0 := \bar{\mathbf{u}}$.

(ii) *If* $i \leq K$ *go to* **(iii)**; *else*
NCOARSE $[2\delta/5, \bar{\mathbf{p}}_{i-1}] \to (\tilde{\mathbf{p}}, \Lambda_q)$;
NCOARSE $[2\delta/5, \bar{\mathbf{u}}_{i-1}] \to (\tilde{\mathbf{u}}, \Lambda_u)$;

(iii) *Apply* **RHS** $[\bar{\mathbf{p}}_{i-1}, c_A \gamma_i \rho^i \delta/2] \to (\mathbf{h}_i, \Lambda_i^h)$;

(iv) **ELLSOLVE** $[\gamma_i \rho^i \delta/2, \mathbf{A}, \bar{\mathbf{u}}_{i-1}, \mathbf{h}_i] \to (\bar{\mathbf{u}}_i, \Lambda_i^X)$.

(v) **NCOARSE** $[\gamma_i \rho^i \delta/2C_R, \mathbf{g}] \to (\hat{\mathbf{g}}_i, \hat{\Lambda}_i)$;
APPLY $[\gamma_i \rho^i \delta/2, \mathbf{R}, \hat{\mathbf{g}}_i] \to (\mathbf{g}_i, \Lambda_i^g)$;

APPLY $[\gamma_i \rho^i \delta/2C_R, \mathbf{B}, \bar{\mathbf{u}}_i] \to (\hat{\mathbf{p}}_i, \hat{\Lambda}_i)$;
APPLY $[\gamma_i \rho^i \delta/2, \mathbf{R}, \hat{\mathbf{p}}_i] \to (\mathbf{p}_i', \Lambda_i^p)$; *set*

$$\bar{\mathbf{p}}_i = \bar{\mathbf{p}}_{i-1} + \omega(\mathbf{p}_i' - \mathbf{g}_i);$$

set $i + 1 \to i$ *and go to* **(ii)**.

It will be shown later that the error bounds of the new approximations produced by **ADV** are reduced by a factor two. The role of the final application of **NCOARSE** in step **(ii)** of **ADV** will be seen later to play an important role with regard to asymptotically optimal complexity, roughly speaking, by keeping only significant coefficients.

Of course, when the characterization of the space M does not entail any constraints on the wavelet coeffcicients $\mathbf{R}$ can be replaced by the identity in (3.1.7) in which case step **(v)** of **ADV** simplifies in an obvious manner.

To formulate the main algorithm recall that by (2.3.1)

$$\|\mathbf{u}\|^2_{\ell_2(\mathcal{J}_X)} + \|\mathbf{p}\|^2_{\ell_2(\mathcal{J}_M)} \leq c_L^{-1}\left(\|\mathbf{f}^\circ\|^2_{\ell_2(\mathcal{J}_X)} + \|\mathbf{g}\|^2_{\ell_2(\mathcal{J}_M)}\right).$$

Therefore the right-hand side gives a bound for the initial error when using $\mathbf{0}$ as initial guess for $\mathbf{u}, \mathbf{p}$, respectively. The complete adaptive Uzawa iteration can be described now as follows.

UZAWA^c $[\mathbf{A}, \mathbf{B}, \mathbf{f}, \mathbf{g}, \varepsilon] \to (\mathbf{u}(\varepsilon), \mathbf{p}(\varepsilon))$:
Set $\mathbf{\Lambda}_0 := (\Lambda_0^X, \Lambda_0^M) \subset \mathcal{J} =: \mathcal{J}_X \times \mathcal{J}_M$ to be empty $\Lambda_0^M = \Lambda_0^X = \emptyset$, $\mathbf{p}_0 = \bar{\mathbf{p}}_0 = \mathbf{0}$, $\bar{\mathbf{u}} = \mathbf{0}$, $\delta_0 := c_L^{-1/2}\left(\|\mathbf{f}^\circ\|^2_{\ell_2(\mathcal{J}_X)} + \|\mathbf{g}\|^2_{\ell_2(\mathcal{J}_M)}\right)^{1/2}$, $J = 0$, choose a target accuracy ε.
(i) **ADV** $[\bar{\mathbf{u}}, \bar{\mathbf{p}}, \delta_J] \to (\tilde{\mathbf{u}}, \tilde{\mathbf{p}}, \Lambda_u, \Lambda_q)$;

(ii) Set $\delta_{J+1} := \delta_J/2$.
If $\delta_{J+1} \leq \varepsilon$, stop and accept $\mathbf{u}(\varepsilon) := \tilde{\mathbf{u}}, \mathbf{p}(\varepsilon) := \tilde{\mathbf{p}}$ as solution.
Else, set $\bar{\mathbf{u}} = \tilde{\mathbf{u}}$, $\bar{\mathbf{p}} = \tilde{\mathbf{p}}$, $J+1 \to J$ and go to **(i)**.

3.3. Convergence. The convergence of **UZAWA**^{c} relies on the error reduction caused by **ADV**.

PROPOSITION 3.3. *Given a scheme* **APPLY** *such that (3.2.2) holds then, under the above assumptions concerning* **NCOARSE** *on the data* $\mathbf{f}, \mathbf{g}$, *the vectors* $\tilde{\mathbf{u}}, \tilde{\mathbf{p}}$ *produced by* **ADV** $[\bar{\mathbf{u}}, \bar{\mathbf{p}}, \delta]$ *above satisfy*

$$\|\tilde{\mathbf{u}} - \mathbf{u}\|_{\ell_2(\mathcal{J}_X)} \leq \delta/2, \quad \|\tilde{\mathbf{p}} - \mathbf{p}\|_{\ell_2(\mathcal{J}_M)} \leq \delta/2. \tag{3.3.1}$$

Hence, after finitely many steps the scheme **UZAWA** *produces finitely supported solutions* $(\mathbf{u}(\varepsilon), \mathbf{p}(\varepsilon))$ *satisfying*

$$\|\mathbf{u} - \mathbf{u}(\varepsilon)\|_{\ell_2(\mathcal{J}_X)} \leq \varepsilon, \quad \|\mathbf{p} - \mathbf{p}(\varepsilon)\|_{\ell_2(\mathcal{J}_M)} \leq \varepsilon. \tag{3.3.2}$$

Proof: Set $\mathbf{p}_0 := \bar{\mathbf{p}}_0 = \bar{\mathbf{p}}$, $\bar{\mathbf{u}}_0 := \bar{\mathbf{u}}$ and observe that

$$\begin{aligned}
\mathbf{p}_i - \bar{\mathbf{p}}_i &= \mathbf{p}_{i-1} + \omega\mathbf{R}(\mathbf{B}\mathbf{u}_i - \mathbf{g}) - \bar{\mathbf{p}}_{i-1} - \omega(\mathbf{p}_i' - \mathbf{g}_i) \\
&= \mathbf{p}_{i-1} - \bar{\mathbf{p}}_{i-1} + \omega(\mathbf{R}\mathbf{B}\mathbf{u}_i - \mathbf{p}_i' - \mathbf{R}\mathbf{g} + \mathbf{g}_i) \\
&= (\mathbf{id} - \omega\mathbf{R}\mathbf{S})(\mathbf{p}_{i-1} - \bar{\mathbf{p}}_{i-1}) + \omega\left(\mathbf{R}(\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^T)(\mathbf{p}_{i-1} - \bar{\mathbf{p}}_{i-1}) + \mathbf{R}\mathbf{B}\mathbf{u}_i - \mathbf{p}_i' + \mathbf{g}_i - \mathbf{R}\mathbf{g}\right).
\end{aligned} \tag{3.3.3}$$

Since $\mathbf{A}\mathbf{u}_i = \mathbf{f} - \mathbf{B}^T\mathbf{p}_{i-1}$ we can replace $\mathbf{B}^T\mathbf{p}_{i-1}$ by $\mathbf{f} - \mathbf{A}\mathbf{u}_i$ to obtain

$$\begin{aligned}
&\omega\left(\mathbf{R}(\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^T)(\mathbf{p}_{i-1} - \bar{\mathbf{p}}_{i-1}) + \mathbf{R}\mathbf{B}\mathbf{u}_i - \mathbf{p}_i' + \mathbf{g}_i - \mathbf{R}\mathbf{g}\right) \\
&= \omega\left(\mathbf{R}(\mathbf{B}\mathbf{A}^{-1}\mathbf{f} - \mathbf{B}\mathbf{u}_i + \mathbf{B}\mathbf{u}_i - \mathbf{B}\mathbf{A}^{-1}\mathbf{B}^T\bar{\mathbf{p}}_{i-1}) - \mathbf{p}_i' + \mathbf{g}_i - \mathbf{R}\mathbf{g}\right) \\
&= \omega\left(\mathbf{R}\mathbf{B}\mathbf{A}^{-1}(\mathbf{f} - \mathbf{B}^T\bar{\mathbf{p}}_{i-1}) - \mathbf{p}_i' + (\mathbf{g}_i - \mathbf{R}\mathbf{g})\right)
\end{aligned} \tag{3.3.4}$$

Thus

$$\mathbf{R}\mathbf{B}\mathbf{A}^{-1}(\mathbf{f} - \mathbf{B}^T\bar{\mathbf{p}}_{i-1}) - \mathbf{p}_i' = \mathbf{R}\mathbf{B}\left(\mathbf{A}^{-1}(\mathbf{f} - \mathbf{B}^T\bar{\mathbf{p}}_{i-1}) - \bar{\mathbf{u}}_i\right) + (\mathbf{R}\mathbf{B}\bar{\mathbf{u}}_i - \mathbf{p}_i') \tag{3.3.5}$$

Hence combining (3.3.3), (3.3.4) and (3.3.5) and recalling (3.1.5), yields

$$\begin{aligned}
\|\mathbf{p}_i - \bar{\mathbf{p}}_i\|_{\ell_2(\mathcal{J}_M)} &\leq \rho\|\mathbf{p}_{i-1} - \bar{\mathbf{p}}_{i-1}\|_{\ell_2(\mathcal{J}_M)} + \omega\left(\|\mathbf{R}\mathbf{B}\left(\mathbf{A}^{-1}(\mathbf{f} - \mathbf{B}^T\bar{\mathbf{p}}_{i-1}) - \bar{\mathbf{u}}_i\right)\|_{\ell_2(\mathcal{J}_M)}\right. \\
&\quad \left. + \|\mathbf{R}\mathbf{B}\bar{\mathbf{u}}_i - \mathbf{p}_i'\|_{\ell_2(\mathcal{J}_M)} + \|\mathbf{g}_i - \mathbf{R}\mathbf{g}\|_{\ell_2(\mathcal{J}_M)}\right) \\
&\leq \rho\|\mathbf{p}_{i-1} - \bar{\mathbf{p}}_{i-1}\|_{\ell_2(\mathcal{J}_M)} + \omega C_R C_B\|\mathbf{A}^{-1}(\mathbf{f} - \mathbf{B}^T\bar{\mathbf{p}}_{i-1}) - \bar{\mathbf{u}}_i\|_{\ell_2(\mathcal{J}_X)} + 2\omega\gamma_i\rho^i\delta
\end{aligned} \tag{3.3.6}$$

where we have used the tolerances in step **(v)** of **ADV**. By (3.2.8) we have for the output $\mathbf{h}_i$ of step **(iii)** in **ADV** that $\|\mathbf{h}_i - (\mathbf{f} - \mathbf{B}^T\bar{\mathbf{p}}_{i-1})\|_{\ell_2(\mathcal{J}_X)} \leq c_A\gamma_i\rho^i\delta/2$ which, in view of the tolerances in step **(iv)** of **ADV** and (3.2.7), implies

$$\|\mathbf{A}^{-1}(\mathbf{f} - \mathbf{B}^T\bar{\mathbf{p}}_{i-1}) - \bar{\mathbf{u}}_i\|_{\ell_2(\mathcal{J}_X)} \leq \gamma_i\rho^i\delta. \tag{3.3.7}$$

Therefore we deduce from (3.3.6) that

$$\|\mathbf{p}_i - \bar{\mathbf{p}}_i\|_{\ell_2(\mathcal{J}_M)} \le \rho\|\mathbf{p}_{i-1} - \bar{\mathbf{p}}_{i-1}\|_{\ell_2(\mathcal{J}_M)} + \omega(C_R C_B + 2)\gamma_i\rho^i\delta. \tag{3.3.8}$$

Iterating this estimate and bearing in mind that $\mathbf{p}_0 = \bar{\mathbf{p}}_0$, provides

$$\|\mathbf{p}_i - \bar{\mathbf{p}}_i\|_{\ell_2(\mathcal{J}_M)} \le \omega(C_R C_B + 2)\left(\sum_{l=1}^{i}\gamma_i\right)\rho^i\delta. \tag{3.3.9}$$

Since by (3.1.8) and the assumption, $\|\mathbf{p} - \mathbf{p}_i\|_{\ell_2(\mathcal{J}_M)} \le \rho^i\|\mathbf{p} - \mathbf{p}_0\|_{\ell_2(\mathcal{J}_M)} = \rho^i\|\mathbf{p} - \bar{\mathbf{p}}\|_{\ell_2(\mathcal{J}_M)} \le \rho^i\delta$ we conclude that

$$\|\mathbf{p} - \bar{\mathbf{p}}_i\|_{\ell_2(\mathcal{J}_M)} \le \left\{\omega(C_R C_B + 2)\left(\sum_{l=1}^{i}\gamma_l\right) + 1\right\}\rho^i\delta, \tag{3.3.10}$$

which, in view of (3.2.10), gives

$$\|\mathbf{p} - \bar{\mathbf{p}}_K\|_{\ell_2(\mathcal{J}_M)} \le \delta/10. \tag{3.3.11}$$

Now recall that by step **(ii)** of **ADV** the final approximation $\tilde{\mathbf{p}}$ is obtained by coarsening $\bar{\mathbf{p}}_K$. Thus

$$\|\mathbf{p} - \tilde{\mathbf{p}}\|_{\ell_2(\mathcal{J}_M)} \le \|\mathbf{p} - \bar{\mathbf{p}}_K\|_{\ell_2(\mathcal{J}_M)} + \|\bar{\mathbf{p}}_K - \tilde{\mathbf{p}}\|_{\ell_2(\mathcal{J}_M)} \le \left(\frac{2}{5} + \frac{1}{10}\right)\delta = \frac{\delta}{2} \tag{3.3.12}$$

as claimed.

It remains to estimate the accuracy of $\bar{\mathbf{u}}_K$. Denoting by $\hat{\mathbf{u}}_i$ be the exact solution of $\mathbf{A}\hat{\mathbf{u}}_i = \mathbf{f}_i - \mathbf{B}^T\bar{\mathbf{p}}_{i-1}$, (3.3.7) and (3.2.7) say that $\|\hat{\mathbf{u}}_i - \bar{\mathbf{u}}_i\|_{\ell_2(\mathcal{J}_X)} \le \gamma_i\rho^i\delta$. Writing

$$\mathbf{u} - \bar{\mathbf{u}}_i = \mathbf{u} - \hat{\mathbf{u}}_i + \bar{\mathbf{u}}_i - \bar{\mathbf{u}}_i = \mathbf{A}^{-1}\mathbf{B}^T(\bar{\mathbf{p}}_{i-1} - \mathbf{p}) + \hat{\mathbf{u}}_i - \bar{\mathbf{u}}_i, \tag{3.3.13}$$

and defining $C_1 := \omega(C_R C_B + 2)\gamma + 1$, one obtains

$$\|\mathbf{u} - \bar{\mathbf{u}}_i\|_{\ell_2(\mathcal{J}_X)} \le (c_A\rho)^{-1}C_B C_1\rho^i\delta + \gamma_i\rho^i\delta = \left((c_A\rho)^{-1}C_B C_1 + \gamma_i\right)\rho^i\delta.$$

Again, we infer from (3.2.10) that

$$\|\mathbf{u} - \bar{\mathbf{u}}_K\|_{\ell_2(\mathcal{J}_X)} \le \delta/10, \tag{3.3.14}$$

so that by the same reasoning as in (3.3.12) $\tilde{\mathbf{u}}$ produced by $\mathbf{NCOARSE}\,[2\delta/5, \bar{\mathbf{u}}_K]$ satisfies $\|\mathbf{u} - \tilde{\mathbf{u}}\|_{\ell_2(\mathcal{J}_X)} \le \delta/2$, which completes the proof. ∎

As an immediate consequence of the norm equivalences (2.2.1), (2.2.2) one has the following fact.

COROLLARY 3.4. *Let $u = \mathbf{u}^T\mathbf{D}_X^{-1}\Psi_X$, $p = \mathbf{p}^T\mathbf{D}_M^{-1}\Psi_M$ be the exact solution of (2.1.2). Then the finite expansions $u(\varepsilon) := \mathbf{u}^T(\varepsilon)\mathbf{D}_X^{-1}\Psi_X$, $p(\varepsilon) = \mathbf{p}^T(\varepsilon)\mathbf{D}_M^{-1}\Psi_M$ with terms from the finite index sets $\Lambda_{u(\varepsilon)} \subset \mathcal{J}_X$, $\Lambda_{q(\varepsilon)} \subset \mathcal{J}_M$ satisfy*

$$\|u - u(\varepsilon)\|_X \le c\varepsilon, \qquad \|q - q(\varepsilon)\|_M \le c\varepsilon, \tag{3.3.15}$$

uniformly in ε, where c depends only on the constants in (2.1.4), (2.2.1), (2.2.2).

To keep things transparent we have based the above considerations on the simplest version (3.1.6), (3.1.7) of an Uzawa iteration. It will be seen below that already this version gives rise to asymptotically optimal convergence properties. Of course, similar results would be obtained for different accuracy tolerances as long as they differ by constants leading possibly to different values of K. Nevertheless, several more important possibilities suggest themselves to realize quantitative improvements, e.g. by replacing the Richardson iteration by a gradient or conjugate gradient iteration. This avoids the need of estimating step size parameters and should speed error reduction. Note that these variants still involve only the same algorithmic tasks namely approximate application of operators in the above sense. Furthermore, the number K of subiterations is likely to be too pessimistic. Therefore it would be preferable to monitor the error decay as follows. Note that $\mathbf{p}_i - \mathbf{g}_i$ in step **(v)** of **ADV** approximates $\mathbf{R}(\mathbf{B}\mathbf{u}_i - \mathbf{p}_i)$ and, in view of (3.1.6), (3.1.7), the residual $\mathbf{R}(\mathbf{B}\mathbf{A}^{-1}\mathbf{f} - \mathbf{g} - \mathbf{S}\mathbf{p}_{i-1})$. By (2.4.4) and the bounded invertibility of $\mathbf{R}$ this residual can be bounded from below and above by fixed constant multiples of the current error of the approximate solution to the reduced system (2.4.2). Thus monitoring $\|\mathbf{p}_i' - \mathbf{g}_i\|_{\ell_2(\mathcal{J}_M)}$ can be used as a stopping criterion. This is expected to result in frequent early termination of step (ii) in **ADV**. These points will be taken up in more detail elsewhere.

4. Complexity Analysis. Of course, the central questions now are how to come up with an **APPLY** scheme with the desired properties and what is the computational cost of $\mathbf{UZAWA}^c$ for a given target accuracy ε. In the present generality cost will be measured by storage requirements and the number of flops required by the scheme (well being aware of the fact that this is not the full story).

4.1. Best N-Term Approximation. As in [8] we will relate the performance of the adaptive scheme to what could be achieved at best namely the approximation of the solution in terms of *possibly few degrees of freedom* within the given discretization context - here determined by the underlying wavelet bases. Note that, in view of (3.3.15), it suffices to deal with the conceptually much simpler approximation in $\ell_2(\mathcal{J})$. To explain this, it is useful to recall first the following notion of *best N-term approximation* in ℓ_2:

$$\sigma_{N,\ell_2(\mathcal{J}')}(\mathbf{v}) := \inf_{\mathbf{w},\#\operatorname{supp}\mathbf{w}\le N} \|\mathbf{v}-\mathbf{w}\|_{\ell_2(\mathcal{J}')}, \tag{4.1.1}$$

where $\ell_2(\mathcal{J}')$ stands again for $\ell_2(\mathcal{J}_X)$ or $\ell_2(\mathcal{J}_M)$. Thus $\sigma_{N,\ell_2(\mathcal{J}')}(\mathbf{v})$ describes the error as a function of the number of degrees of freedom when the (possibly infinitely supported) vector is approximated by a vector with at most N nonzero entries whose value and position can be chosen freely. Thus the approximant is not taken from any fixed linear space but from the nonlinear manifold of all vectors with at most N nonzero entries. This notion is well understood for ℓ_2, see e.g. [19]. Obviously, $\sigma_{N,\ell_2(\mathcal{J}')}(\mathbf{v})$ is realized by retaining the N *largest* coefficients in $\mathbf{v}$ which are, of course, unknown when $\mathbf{v}$ is a solution of a system of equations. To understand how this error behaves denote for any $\mathbf{v}\in\ell_2(\mathcal{J}')$ by $\mathbf{v}^*=\{v_{\lambda_l}\}_{l\in\mathbb{N}}$ its *decreasing rearrangement* in the sense that $|v_{\lambda_l}|\ge|v_{\lambda_{l+1}}|$ and let

$$\Lambda(\mathbf{v},N) := \{\lambda_l : l=1,\dots,N\}, \quad \mathbf{v}_N := \mathbf{v}|_{\Lambda(\mathbf{v},N)}. \tag{4.1.2}$$

It is clear that $\mathbf{v}_N$ is a best N-term approximation of $\mathbf{v}$.

In particular, it will be important to characterize the sequences in $\ell_2(\mathcal{J}')$ whose best N-term approximation behaves like N^{-s} for some $s>0$. The following facts are well-known [8, 19]. Let for $0<\tau<2$

$$|\mathbf{v}|_{\ell_\tau^w(\mathcal{J}')} := \sup_{n\in\mathbb{N}} n^{1/\tau}|v_n^*|, \quad \|\mathbf{v}\|_{\ell_\tau^w(\mathcal{J}')} := \|\mathbf{v}\|_{\ell_2(\mathcal{J}')} + |\mathbf{v}|_{\ell_\tau^w(\mathcal{J}')}. \tag{4.1.3}$$

It is easy to see that for any $\tau<\tau'\le 2$

$$\|\mathbf{v}\|_{\ell_{\tau'}(\mathcal{J}')} \lesssim \|\mathbf{v}\|_{\ell_\tau^w(\mathcal{J}')} \le 2\|\mathbf{v}\|_{\ell_\tau(\mathcal{J}')}, \tag{4.1.4}$$

so that by Jensen's inequality, in particular, $\ell_\tau^w(\mathcal{J}')\subset\ell_2(\mathcal{J}')$.

PROPOSITION 4.1. *Let*

$$\frac{1}{\tau} = s+\frac{1}{2}, \tag{4.1.5}$$

then

$$\mathbf{v}\in\ell_\tau^w(\mathcal{J}') \iff \|\mathbf{v}-\mathbf{v}_N\|_{\ell_2(\mathcal{J}')} \lesssim N^{-s}\|\mathbf{v}\|_{\ell_\tau^w(\mathcal{J}')}. \tag{4.1.6}$$

In complete analogy one can define $\|\cdot\|_{\ell_\tau^w(\mathcal{J})}$ for $\ell_\tau^w(\mathcal{J}) := \ell_\tau^w(\mathcal{J}_X\times\mathcal{J}_M)$ by forming the rearrangements from both component vectors $\mathbf{v}\in\ell_2(\mathcal{J}_X)$, $\mathbf{p}\in\ell_{2,0}(\mathcal{J}_M)$ and regrouping the entries to both component vectors.

We will make use of the following result from [8] which interrelates best N-term approximation in ℓ_2 with the routine **NCOARSE**, see Section 6.2 in [8].

PROPOSITION 4.2. *Given $\mathbf{v}\in\ell_2(\mathcal{J}')$, a toleraqnce $\eta>0$ and a finitely supported $\mathbf{w}$ such that*

$$\|\mathbf{v}-\mathbf{w}\|_{\ell_2(\mathcal{J}')} \le \eta/5,$$

Then (as has been used before), the output $\bar{\mathbf{w}}$ of **NCOARSE**$[\mathbf{w},4\eta/5]$ *satisfies $\|\mathbf{v}-\bar{\mathbf{w}}\|_{\ell_2(\mathcal{J}')}\le\eta$. Moreover, when $\mathbf{v}\in\ell_\tau^w(\mathcal{J}')$ and $\frac{1}{\tau}=s+\frac{1}{2}$ for some $s>0$, then there exists a constant C depending only on s when s tends to infinity such that:*

$$\|\mathbf{v}-\bar{\mathbf{w}}\|_{\ell_2(\mathcal{J}')} \le C\|\mathbf{v}\|_{\ell_\tau^w(\mathcal{J}')}(\#\operatorname{supp}\bar{\mathbf{w}})^{-s}, \tag{4.1.7}$$

and

$$\|\bar{\mathbf{w}}\|_{\ell^w_\tau(\mathcal{J}')} \le C\|\mathbf{v}\|_{\ell^w_\tau(\mathcal{J}')}, \quad \#\operatorname{supp}\bar{\mathbf{w}} \le C\|\mathbf{v}\|^{1/s}\eta^{-1/s}. \tag{4.1.8}$$

Best N-term approximation will be one important ingredient in the realization of the approximate application of infinite matrices represented by **APPLY**. The other one is the (a-priori known) quasi sparseness of wavelet representations which can be formalized as follows, see [8].

DEFINITION 4.3. *A matrix* $\mathbf{C}$ *belongs to the class* $\mathcal{C}_{s^*}$ *if for every* $s < s^*$ *there exists a positive summable sequence* $(\alpha_j)_{j\ge 0}$ *and for every* $j \ge 0$ *there exists a matrix* $\mathbf{C}_j$ *with at most* $2^j\alpha_j$ *nonzero entries per row and column such that*

$$\|\mathbf{C}_j - \mathbf{C}\| \lesssim \alpha_j 2^{-sj}. \tag{4.1.9}$$

A matrix in $\mathcal{C}_{s^*}$ *is called* compressible *or sometimes* s^**-compressible.*

Compressibility of a wavelet representation of certain operators follows from the above mentioned cancellation properties of the wavelets, see [8] as well as Section 5.3 for concretizations.

Now suppose that the (possibly infinite) matrix $\mathbf{C}$ (defined on $\ell_2(\mathcal{J}')$ say) is known to be compressible in the sense of (4.1.9) for some range of $s > 0$. For any given finitely supported $\mathbf{v} \in \ell_2(\mathcal{J}')$, let $\mathbf{v}_{[j]} := \mathbf{v}_{2^j}$ denote its best 2^j-term approximation in $\ell_2(\mathcal{J}')$. We shall numerically approximate $\mathbf{Cv}$ by using the vector

$$\mathbf{w}_k := \mathbf{C}_k\mathbf{v}_{[0]} + \mathbf{C}_{k-1}(\mathbf{v}_{[1]} - \mathbf{v}_{[0]}) + \cdots + \mathbf{C}_0(\mathbf{v}_{[k]} - \mathbf{v}_{[k-1]}) \tag{4.1.10}$$

for a certain value of k determined by the desired numerical accuracy. This leads to a practical scheme **APPLY** $[\eta, \mathbf{C}, \mathbf{v}] \to (\mathbf{w}, \Lambda)$, whose detailed description is given in [8], Section 6.4, see also [2]. For later use we recall its properties, see Properties 6.4 in [8].

PROPOSITION 4.4. *Assume that* $\mathbf{C} \in \mathcal{C}_{s^*}$. *Given a tolerance* $\eta > 0$ *and a vector* $\mathbf{v}$ *with finite support, the algorithm* **APPLY** *produces a vector* $\mathbf{w} = \mathbf{w}(\mathbf{v}, \eta)$ *which satisfies (3.2.2).*

Moreover, if $\mathbf{v} \in \ell^w_\tau(\mathcal{J}')$, *with* $\tau = (s + 1/2)^{-1/2}$ *and* $0 < s < s^*$, *then the following properties hold:*

(i) The size of the output Λ *is bounded by*

$$\#(\Lambda) \le C\|\mathbf{v}\|^{1/s}_{\ell^w_\tau(\mathcal{J}')}\eta^{-1/s}, \tag{4.1.11}$$

and the number of entries of $\mathbf{C}$ *that need to be computed is* $\le C\|\mathbf{v}\|^{1/s}_{\ell^w_\tau(\mathcal{J}')}\eta^{-1/s}$.

(ii) The number of **arithmetic operations** *needed to compute* $\mathbf{w}(\mathbf{v}, \eta)$ *does not exceed* $C\eta^{-1/s}\|\mathbf{v}\|^{1/s}_{\ell^w_\tau(\mathcal{J}')} + 2N$ *with* $N := \#\operatorname{supp}\mathbf{v}$.

(iii) The number of operations for **sorting** *needed to assemble the slices* $\mathbf{v}_{[j]}$ *of* $\mathbf{w}(\mathbf{v}, \eta)$, $j = 0, 1, \cdots, \lfloor \log N \rfloor$, *does not exceed* $CN \log N$.

(iv) The output vector $\mathbf{w}$ *satisfies*

$$\|\mathbf{w}\|_{\ell^w_\tau(\mathcal{J}')} \le C\|\mathbf{v}\|_{\ell^w_\tau(\mathcal{J}')}. \tag{4.1.12}$$

As for the log-terms for sorting, see Remark 4.9 at the end of this section. We shall make use of the following fact, see [8].

REMARK 4.5. *It follows from Proposition 4.1 and Proposition 4.4 (i) that any matrix* $\mathbf{C} \in \mathcal{C}_{s^*}$ *is bounded on* ℓ^w_τ *when* τ *is related to* $s < s^*$ *by (4.1.5).*

As mentioned above, wavelet representations of differential operators are compressible. Therefore the following observation is useful.

REMARK 4.6. *When* $\mathbf{A}^\circ := \mathbf{D}_X^{-1}a(\Psi_X, \Psi_X)\mathbf{D}_X^{-1}$ *and* $\mathbf{B}$ *belong to* $\mathcal{C}_{s^*}$ *for some* $s^* > 0$, *then one easily shows that the scheme* **APPLY*** *inherits all the properties described in Proposition 4.4 above, see [8] Properties 6.4.*

The complexity estimates in *(ii)* and *(iii)* of Proposition 4.4 hold under the assumption that the entries of $\mathbf{C}$ are accessible during the calculation. In fact, the subsequent developments will always be based on the following

Assumption C: The entries of the matrices $\mathbf{A}^\circ$ *and* $\mathbf{B}$ *are accessible at unit cost.*

Using piecewise polynomial wavelets this assumption can be realized for constant coefficient operators in a relatively straightforward manner. This task becomes much more delicate under more general circumstances, e.g. when isoparametric mappings are involved in the construction of the wavelets, see Section 5.2 below. In [3] a fast evaluation scheme is developed that computes sufficiently accurate approximations to the summands on the right-hand side of (4.1.10) at a computational cost that still satisfies the bounds in *(ii)*, *(iii)* of Proposition 4.4 above. Thus *Assumption C* is justified for a wide range of practically relevant situations.

With Remark 4.6 at hand, we are now in the position for estimating the complexity analysis of **ELLSOLVE** based on the results in [8, 9] with the **APPLY** scheme for compressible matrices replaced, if necessary, by the extended version **APPLY*** introduced above. The fact that in the present context **ELLSOLVE** applies to varying auxiliary problems with little a-priori information on the corresponding intermediate solutions prevents us from applying the results from [8] directly. Nevertheless, we can extract from the analysis in [8, 9] some facts that will apply in the present situation as well. This is most transparent when considering the simplified scheme in [9] which (in the very spirit of the current approach) for the special case of an elliptic (coercive) problem is based on a simple iteration for (3.2.4) of the form

$$\hat{\mathbf{u}}^{n+1} = \hat{\mathbf{u}}^n + \bar{\omega}(\mathbf{h} - \mathbf{A}\hat{\mathbf{u}}^n). \tag{4.1.13}$$

In particular, when the right-hand sides are already finitely supported as in the present situation, the scheme consists of at most $\bar{K}$ perturbed iterations of the form (4.1.13), employing **APPLY*** and **NCOARSE** with judiceaously chosen accuracy tolerances, followed by a coarsening step so as to reduce a current error bound by a factor two, say (see the algorithm **SOLVE** in Section 4.2 of [9]). This implies the following fact.

PROPOSITION 4.7. *Consider the problem (3.2.4) and suppose that the initial approximation* $\mathbf{v}$ *used as input for* **ELLSOLVE** *satisfies*

$$\|\hat{\mathbf{u}} - \mathbf{v}\|_{\ell_2(\mathcal{J}_X)} \le \bar{\varepsilon} \tag{4.1.14}$$

for some $\bar{\varepsilon} > \varepsilon$. *Moreover assume that* s *and* τ *are related by (4.1.5) and that*

$$\varepsilon \le \bar{C}\bar{\varepsilon} \tag{4.1.15}$$

for some positive constant $\bar{C}$. *Then the output* $\bar{\mathbf{u}}$ *and* $\Lambda := \operatorname{supp} \bar{\mathbf{u}}$ *of* **ELLSOLVE** $[\varepsilon, \mathbf{A}, \mathbf{v}, \mathbf{h}]$ *satisfies*

$$\begin{aligned} \#(\Lambda) &\le \hat{C}\left(\#(\operatorname{supp}\mathbf{v}) + \left(\|\mathbf{v}\|^{1/s}_{\ell^w_\tau(\mathcal{J}_X)} + \|\mathbf{h}\|^{1/s}_{\ell^w_\tau(\mathcal{J}_X)}\right)\varepsilon^{-1/s}\right), \\ \|\bar{\mathbf{u}}\|_{\ell^w_\tau(\mathcal{J}_X)} &\le \hat{C}\left(\|\mathbf{v}\|_{\ell^w_\tau(\mathcal{J}_X)} + \|\mathbf{h}\|_{\ell^w_\tau(\mathcal{J}_X)}\right). \end{aligned} \tag{4.1.16}$$

Moreover, the number of arithmetic operations required for the computation of $\bar{\mathbf{u}}$ *remains bounded by*

$$\hat{C}\left\{\operatorname{supp}\mathbf{v} + \varepsilon^{-1/s}\left(\|\mathbf{v}\|^{1/s}_{\ell^w_\tau(\mathcal{J}_X)} + \|\mathbf{h}\|^{1/s}_{\ell^w_\tau(\mathcal{J}_X)}\right)\right\}. \tag{4.1.17}$$

An additional factor $\hat{C}\log\varepsilon^{-1}$ *is allowed for operations spent on sorting arrays (see Remark 4.9). The constant* $\hat{C}$ *depends in all cases only on the constants in (2.4.7), (2.2.1), on* s *when* s *tends to infinity, and on the constant* $\bar{C}$ *in (4.1.15).*

Proof: In view of (4.1.15) only a uniformly bounded number of blocks of perturbed iterations (4.1.13) separated by coarsening steps is needed to reduce the current error bound from $\bar{\varepsilon}$ to ε, see Proposition 4.2 in [9]. This number depends clearly on the bound $\bar{C}$ for the ratio $\bar{\varepsilon}/\varepsilon$. Each block, in turn, involves a uniformly bounded number $\bar{K}$ of perturbed applications of (4.1.13), where $\bar{K}$ depends only on the constants in (2.4.7) and (2.2.1). The claim follows now immediately from Propositions 4.2 and 4.4 (see also the proof of Theorem 5.7 in [9]). ∎

The main result can now be formulated as follows.

THEOREM 4.8. *Assume that the scaled wavelet representations* $\mathbf{A}^\circ$, $\mathbf{B}$ *in (2.2.4) and* $\mathbf{R}$ *from (3.1.3) belong to* $\mathcal{C}_{s^*}$ *for some* $s^* > 0$. *If the exact solution* (u, p) *of (2.1.2) satisfies for some* $s < s^*$

$$\inf_{\#\mathrm{supp}\mathbf{v}\le N} \|u - \mathbf{v}^T\mathbf{D}_X^{-1}\Psi_X\|_X \lesssim N^{-s}, \qquad \inf_{\#\mathrm{supp}\mathbf{q}\le N} \|p - \mathbf{q}^T\mathbf{D}_M^{-1}\Psi_M\|_M \lesssim N^{-s}, \quad N\to\infty, \tag{4.1.18}$$

then the approximations $(\mathbf{u}(\varepsilon), \mathbf{p}(\varepsilon))$ *produced by* **UZAWA**^{c} *satisfy*

$$\|u - \mathbf{u}(\varepsilon)^T\mathbf{D}_X^{-1}\Psi_X\|_X \lesssim (\#\mathrm{supp}\,\mathbf{u}(\varepsilon))^{-s}, \quad \|p - \mathbf{p}(\varepsilon)^T\mathbf{D}_M^{-1}\Psi_M\|_M \lesssim (\#\mathrm{supp}\,\mathbf{p}(\varepsilon))^{-s}. \tag{4.1.19}$$

Moreover, under assumptions f, C *(pages 7 and 14, resp.) the computational work needed to compute* $\mathbf{u}(\varepsilon), \mathbf{p}(\varepsilon)$ *is also of the order* $\varepsilon^{-1/s}$ *(except for additional* log *terms for sorting).*

Proof: First note that by (2.2.4), Proposition 4.1 and Remark 4.5, $\mathbf{u} \in \ell_\tau^w(\mathcal{J}_X)$ implies $\mathbf{g} \in \ell_\tau^w(\mathcal{J}_M)$. Since by the same argument $\mathbf{B}^T\mathbf{p}, \mathbf{A}\mathbf{u} \in \ell_\tau^w(\mathcal{J}_X)$, (2.4.5) says that also $\mathbf{f} \in \ell_\tau^w(\mathcal{J}_X)$, i.e.,

$$\|\mathbf{g}\|_{\ell_\tau^w(\mathcal{J}_M)} \lesssim \|\mathbf{u}\|_{\ell_\tau^w(\mathcal{J}_X)}, \quad \|\mathbf{f}\|_{\ell_\tau^w(\mathcal{J}_X)} \lesssim \|\mathbf{u}\|_{\ell_\tau^w(\mathcal{J}_X)} + \|\mathbf{p}\|_{\ell_\tau^w(\mathcal{J}_M)}. \tag{4.1.20}$$

We proceed now estimating the computational cost of one call of **ADV** adhering to the notation used in this context before. We will make frequent use of the fact that all accuracy tolerances appearing in **ADV** remain, in view of the uniform boundedness of K, proportional to the current accuracy $\delta = \delta_J$ in the Jth call of **ADV** in **UZAWA**^{c}. First observe that, by Proposition 4.2 and step **(ii)** in **ADV** combined with the error estimate (3.3.11), one has

$$\#(\mathrm{supp}\,\tilde{\mathbf{p}}) \le C\delta^{-1/s}\left(\|\mathbf{p}\|^{1/s}_{\ell^w_{\tau'}(\mathcal{J}_M)} + \|\bar{\mathbf{p}}\|^{1/s}_{\ell^w_{\tau'}(\mathcal{J}_M)}\right) + \#(\mathrm{supp}\,\bar{\mathbf{p}}), \tag{4.1.21}$$

and

$$\|\tilde{\mathbf{p}}\|_{\ell^w_{\tau'}(\mathcal{J}_M)} \le C\|\mathbf{p}\|_{\ell^w_{\tau'}(\mathcal{J}_M)}, \tag{4.1.22}$$

where C depends only on s when s tends to infinity.

We still have to control the computational cost of the intermediate steps in **(iv)** of **ADV** leading to the final update $\bar{\mathbf{p}}_K$ which is then subjected to the coarsening step that led to the above estimates. To this end, we infer from Remark 4.5, Propositions 4.2 and 4.4 that, since the number K of updates in step **(v)** of **ADV** is uniformly bounded, one has

$$\#(\mathrm{supp}\,\bar{\mathbf{p}}_i) \le C\delta^{-1/s}\left(\|\bar{\mathbf{u}}_i\|^{1/s}_{\ell^w_\tau(\mathcal{J}_X)} + \|\mathbf{g}\|^{1/s}_{\ell^w_\tau(\mathcal{J}_M)}\right) + \#(\mathrm{supp}\,\bar{\mathbf{p}}_{i-1}), \tag{4.1.23}$$

and

$$\|\bar{\mathbf{p}}_i\|_{\ell^w_{\tau'}(\mathcal{J}_M)} \le C\left(\|\bar{\mathbf{p}}_{i-1}\|_{\ell^w_\tau(\mathcal{J}_M)} + \|\bar{\mathbf{u}}_i\|_{\ell^w_\tau(\mathcal{J}_X)}\right), \tag{4.1.24}$$

Thus we have to estimate next the quantities $\|\bar{\mathbf{u}}_i\|_{\ell^w_\tau(\mathcal{J}_X)}$, $\mathrm{supp}\,\bar{\mathbf{u}}_i$, $i = 1, \dots, K$. Again, the coarsening step **(ii)** in **ADV** combined with the error estimate (3.3.14) ensures, in view of Proposition 4.2, that $\tilde{\mathbf{u}}$ and hence the input $\bar{\mathbf{u}} = \bar{\mathbf{u}}_0$ of **ELLSOLVE** $[\gamma_1\rho\delta/2, \mathbf{A}, \bar{\mathbf{u}}_0, \mathbf{h}_1]$ satisfies

$$\|\tilde{\mathbf{u}}\|_{\ell^w_\tau(\mathcal{J}_X)} \le C\|\mathbf{u}\|_{\ell^w_\tau(\mathcal{J}_X)}, \quad \mathrm{supp}\,\tilde{\mathbf{u}} \le C\delta^{-1/s}\|\mathbf{u}\|^{1/s}_{\ell^w_\tau(\mathcal{J}_X)}, \tag{4.1.25}$$

where $\delta = \delta_J$ is the current accuracy level in the Jth call of **ADV** in **UZAWA**^{c}. We will exploit this for the estimation of the intermediate approximations $\bar{\mathbf{u}}_i$ in a call of **ADV** by applying Proposition 4.7. To this end, we first have to determine the accuracy of $\bar{\mathbf{u}}_{i-1}$ as an initial guess for **ELLSOLVE** $[\gamma_i\rho^i\delta/2, \mathbf{A}, \bar{\mathbf{u}}_{i-1}, \mathbf{h}_i]$. In fact, a little care is needed because the right-hand sides $\mathbf{h}_i$ change. Recall that $\hat{\mathbf{u}}_i$ denotes the exact solution of $\mathbf{A}\hat{\mathbf{u}}_i = \mathbf{h}_i$, see **(iii)** in **ADV**. Then, by (3.3.13), for $\delta = \delta_J$ in the Jth call of **ADV** in **UZAWA**^{c} one obtains for some constant C

$$\begin{aligned}
\|\hat{\mathbf{u}}_i - \bar{\mathbf{u}}_{i-1}\|_{\ell_2(\mathcal{J}_X)} &\le \|\hat{\mathbf{u}}_i - \mathbf{u}\|_{\ell_2(\mathcal{J}_X)} + \|\mathbf{u} - \bar{\mathbf{u}}_{i-1}\|_{\ell_2(\mathcal{J}_X)} \\
&\le c_A^{-1}\|\mathbf{f} - \mathbf{B}^T\mathbf{p} - \mathbf{h}_{i-1}\|_{\ell_2(\mathcal{J}_X)} + C\delta \\
&\le c_A^{-1}\left(\|\mathbf{f} - \mathbf{B}^T\mathbf{p} - (\mathbf{f} - \mathbf{B}^T\bar{\mathbf{p}}_{i-1})\|_{\ell_2(\mathcal{J}_X)} + \|\mathbf{f} - \mathbf{B}^T\bar{\mathbf{p}}_{i-1} - \mathbf{h}_{i-1}\|_{\ell_2(\mathcal{J}_X)}\right) + C\delta \\
&\le c_A^{-1}C_B\|\mathbf{p} - \bar{\mathbf{p}}_{i-1}\|_{\ell_2(\mathcal{J}_M)} + \gamma_i\rho^i\delta/2 + C\delta \\
&\le C'\delta_J,
\end{aligned}$$

where we have used (3.3.10) and (3.2.8). Thus the ratio of initial and target accuracies in each call of **ELLSOLVE** remains uniformly bounded by a constant $\bar{C}$ depending on the number K in **ADV**, so that Proposition 4.7 applies. To this end, consider first $i = 1$ in step **(iv)** of **ADV**. By the above bound (4.1.22) on $\bar{\mathbf{p}}_0 = \tilde{\mathbf{p}}^{J-1}$, Remark 4.5, Propositions 4.2, 4.4 and steps **(i)**, **(ii)** in **RHS**, we conclude that

$$\|\mathbf{h}_1\|_{\ell_\tau^w(\mathcal{J}_X)} \leq C(\|\mathbf{p}\|_{\ell_\tau^w(\mathcal{J}_M)} + \|\mathbf{f}\|_{\ell_\tau^w(\mathcal{J}_X)}) \leq C(\|\mathbf{p}\|_{\ell_\tau^w(\mathcal{J}_M)} + \|\mathbf{u}\|_{\ell_\tau^w(\mathcal{J}_X)}), \tag{4.1.26}$$

where we have used (4.1.20) in the last step. Here and in the sequel, unless stated otherwise, C will be a constant (that may vary from place to place) which is independent of $\mathbf{u}, \mathbf{p}$ and at most depending on the problem constants as before. Proposition 4.7 combined with (4.1.25) implies now

$$\begin{aligned} &\|\bar{\mathbf{u}}_1\|_{\ell_\tau^w(\mathcal{J}_X)} \leq C(\|\mathbf{p}\|_{\ell_\tau^w(\mathcal{J}_M)} + \|\mathbf{u}\|_{\ell_\tau^w(\mathcal{J}_X)}) \\ &\#(\operatorname{supp} \bar{\mathbf{u}}_1) \leq C\left(\#(\operatorname{supp} \bar{\mathbf{u}}_0) + \delta^{-1/s}(\|\mathbf{p}\|^{1/s}_{\ell_\tau^w(\mathcal{J}_M)} + \|\mathbf{u}\|^{1/s}_{\ell_\tau^w(\mathcal{J}_X)})\right). \end{aligned} \tag{4.1.27}$$

Again keeping (4.1.22) in mind and substituting (4.1.27) in (4.1.24) for $i = 1$, we obtain

$$\|\bar{\mathbf{p}}_1\|_{\ell_\tau^w(\mathcal{J}_M)} \leq C(\|\mathbf{p}\|_{\ell_\tau^w(\mathcal{J}_M)} + \|\mathbf{u}\|_{\ell_\tau^w(\mathcal{J}_X)}). \tag{4.1.28}$$

We can now repeat this argument K times obtaining that for all $i \leq K$

$$\begin{aligned} &\|\bar{\mathbf{u}}_i\|_{\ell_\tau^w(\mathcal{J}_X)} \leq C(\|\mathbf{p}\|_{\ell_\tau^w(\mathcal{J}_M)} + \|\mathbf{u}\|_{\ell_\tau^w(\mathcal{J}_X)}) \\ &\#(\operatorname{supp} \bar{\mathbf{u}}_i) \leq C\left(\#(\operatorname{supp} \bar{\mathbf{u}}_0) + \delta^{-1/s}(\|\mathbf{p}\|^{1/s}_{\ell_\tau^w(\mathcal{J}_M)} + \|\mathbf{u}\|^{1/s}_{\ell_\tau^w(\mathcal{J}_X)})\right) \\ &\|\bar{\mathbf{p}}_i\|_{\ell_\tau^w(\mathcal{J}_M)} \leq C(\|\mathbf{p}\|_{\ell_\tau^w(\mathcal{J}_M)} + \|\mathbf{u}\|_{\ell_\tau^w(\mathcal{J}_X)}) \\ &\#(\operatorname{supp} \bar{\mathbf{p}}_i) \leq C\left(\#(\operatorname{supp} \bar{\mathbf{p}}_{i-1}) + \delta^{-1/s}(\|\mathbf{p}\|^{1/s}_{\ell_\tau^w(\mathcal{J}_M)} + \|\mathbf{u}\|^{1/s}_{\ell_\tau^w(\mathcal{J}_X)})\right). \end{aligned} \tag{4.1.29}$$

Of course, the constants C depend on the number of steps K and may build up. However, it is important to note that the thresholding applied by step **(ii)** in **ADV** produces a new constant that no longer depends on K and in some sense sets the estimate back. In view of the bound on the operations count given in Proposition 4.7, we conclude that under the given assumptions on the exact solutions $\mathbf{u}, \mathbf{p}$ the convergence rate N^{-s} is indeed preserved by **ALGORITHM**^{c} within the claimed bounds for the corresponding computational work. The assertion follows now directly from Corollary 3.4, (3.3.15). ∎

REMARK 4.9. *One should note that a strict ordering of the wavelet coefficients by size is actually not essential. What matters is to group the coefficients in binary bins, i.e., to collect all those coefficients whose modulus falls into* $[a2^{-j}, a2^{-j-1})$*, say. In this way one can avoid the logarithmic terms appearing in the work counts for sorting, see [1].*

5. Applications to the Stokes Problem. In this section the above developments will be applied to a classical example, namely the *Stokes problem.*

5.1. The Continuous Problem. We consider a Lipschitz domain $\Omega \subset \mathbb{R}^d$ and assume for simplicity homogeneous boundary conditions, i.e.,

$$\begin{aligned} -\Delta u + \nabla p &= f \quad \text{in } \Omega \subset \mathbb{R}^d, \quad u|_{\partial\Omega} = 0, \\ \nabla \cdot u &= 0. \end{aligned} \tag{5.1.1}$$

The standard L_2 inner product on a domain G will be denoted by $\langle v, w\rangle_G := \int_G v(x)\, w(x)\, dx$ where we will drop the subscript whenever the inner product refers to Ω. The mixed formulation takes the form (2.1.2) with

$$X = H_0^1(\Omega)^d, \qquad M = L_{2,0}(\Omega) := \left\{ v \in L_2(\Omega) : \int_\Omega v(x)\, dx = 0 \right\}, \tag{5.1.2}$$

and

$$a(u, v) := (\nabla u, \nabla v), \qquad b(v, q) := -\langle \nabla \cdot v, q\rangle. \tag{5.1.3}$$

It is well known that (2.1.1) holds in this case even with the stronger relation (2.4.1), so that (2.1.4) is true for (5.1.2). In view of the preceding discussion we have to address the following issues. First we identify a class of suitable wavelet bases which will be employed later in numerical experiments. Then we determine the compressibility range of the corresponding wavelet representations. Next, we discuss the regularity of the solution to (5.1.1) in a certain scale of Besov spaces. Although this information has no effect on the algorithmic realization it will allow us to determine under which principal circumstances the adaptive scheme offers even an asymptotically better work/accuracy balance than discretizations based on uniform mesh refinements. These results will guide the selection of our test examples.

5.2. Wavelet Representation. When Ω can be partitioned into regular parametric images $\Omega_l = \kappa_l(\square)$ of the unit d-cube $\square := (0,1)^d$, one can use the constructions from [6, 17] yielding conforming trial spaces for the velocities and pressure. We proceed now collecting the relevant properties of these bases in the present context.

We will reserve the notation Ψ_X for the wavelet basis for $X = H_0^1(\Omega)^d$, i.e., each wavelet $\psi_{X,\lambda}$ is a vector valued function with components $\psi_{\lambda,i}, \lambda \in \mathcal{J}_X$, $i = 1, \ldots, d$. A wavelet $\psi_{\lambda,i}$ which is supported in a single patch Ω_l is then constructed as a linear combination of tensor product B-splines of (coordinatewise) order m_X (which is for simplicity taken to be the same for each component i) composed with κ_l^{-1}. Wavelets whose support intersects several domains are obtained by suitably patching together such functions across interfaces, see [6, 17] for details. At this point a word on the nature of the indices λ is in order. Without going into details, λ encodes the spatial location of the wavelet $\psi_{X,\lambda}$ as well as its scale denoted by $|\lambda|$. We will only employ compactly supported wavelets whose supports then scale like $\operatorname{diam}(\operatorname{supp}\psi_\lambda) \sim 2^{-|\lambda|}$. The coarsest scale $|\lambda| = 0$ corresponds to finitely many functions, which roughly speaking span the polynomial part in an expansion. Thus for each component i the corresponding *multiresolution spaces* $S_{i,J} := \operatorname{span}\{\psi_{\lambda,i} : |\lambda| < J\}$ can be viewed as trial spaces on meshes of size 2^{-J}. To have a conforming discretization the $S_{i,J}$ are arranged to be contained in $H_0^1(\Omega)$. Being generated by m_X-th order B-splines they realize approximation order m_X in $H^{m_X}(\Omega) \cap H_0^1(\Omega)$. Such a basis can be realized for any order $m_X \in \mathbb{N}$. We will vary later this order keeping in mind that the restrictions to a patch Ω_l satisfy

$$\Psi_X|_{\Omega_l} \subset H^{m_X - 1/2}(\Omega_l)^d. \tag{5.2.1}$$

Moreover, recall that a wavelet basis consists of two disjoint collections of functions Ψ_X^+ and Ψ_X^- (and analogously for Ψ_M). As indicated above Ψ_X^+ is comprised of finitely many *scaling functions* of level $|\lambda| = 0$ whose preimages under the parametric mappings span all polynomials of order m_X on $\square$ (up to boundary conditions). The infinite collection Ψ_X^- contains the "true wavelets" in the following sense. In fact, the construction of Ψ_X involves a second important parameter $\tilde{m}_X$. Given any m_X one can take any $\tilde{m}_X \in \mathbb{N}$, $\tilde{m}_X \geq m_X$ such that $m_X + \tilde{m}_X$ is even, and arrange Ψ_X so that for any $\psi_{\lambda,i}$ supported in Ω_l the following m_X-th order *moment conditions* hold

$$(P, \psi_{\lambda,i})_{\Omega_l} = 0 \quad \text{for all } P \in \Pi_{\tilde{m}_X, \kappa_l}, \quad \psi_{\lambda,i} \in \Psi_i^-, \tag{5.2.2}$$

where $(\cdot,\cdot)_{\Omega_l}$ denotes the standard inner product on the subdomain Ω_l. Here $\Pi_{\tilde{m},\kappa_l} := \{P : P = g_l Q \circ \kappa_l^{-1}, Q \in \Pi_{\tilde{m}}\}$ where $g_l := |\det \partial \kappa_l^{-1}|$ and $\Pi_{\ m}$ denotes the space of all polynomials of degree $< \tilde{m}$. With a slight abuse of terminology we will refer to the elements of Π_{m,κ_l} simply as *polynomials*. In fact, since by assumption the g_l are smooth and bounded away from zero the local approximation properties of $\Pi_{\tilde{m},\kappa_l}$ are the same as those of $\Pi_{\tilde{m}}$ which is what matters for the compression properties.

The pressure functions will be expanded in a basis $\Psi_M = \{\psi_{M,\lambda} : \lambda \in \mathcal{J}_M\}$ which is also generated by B-splines of order m_M in the above sense. Likewise the order of moment conditions will be denoted by $\tilde{m}_M$, i.e.,

$$(P, \psi_{M,\lambda})_{\Omega_l} = 0 \quad \text{for all } P \in \Pi_{\tilde{m}_M, \kappa_l}, \ \psi_{M,\lambda} \in \Psi_M^-. \tag{5.2.3}$$

REMARK 5.1. *There are some important distinctions between Ψ_X and Ψ_M though (aside from the fact that Ψ_X is vector and Ψ_M is scalar valued). First, the $\psi_{M,\lambda}$ do not satisfy any boundary conditions. Moreover, the moment conditions hold everywhere in Ω since all wavelets are always fully supported in a single patch Ω_l. , i.e., the wavelets need not to be continuous across patch interfaces.*

Since by (5.2.3), the wavelets in Ψ_M^- have zero mean, an ab initio wavelet basis for $L_2(\Omega)$ can easily be transfromed into one for the constrained space $L_{2,0}(\Omega)$ by modifying only the finitely many elements in Ψ_M^+, a fact that will be important later in the numerical realization.

It has been shown in [6, 17] that bases Ψ_X and Ψ_M satisfy the norm equivalences (2.2.1) and (2.2.2) with scaling weights

$$(D_X)_\lambda := 2^{|\lambda|}, \quad (D_M)_\lambda := 1. \tag{5.2.4}$$

In fact, the alternative choice $(D_X)_\lambda := a(\psi_{X,\lambda}, \psi_{X,\lambda})^{1/2}$ typically gives rise to quantitatively better results but we will stick for simplicity with (5.2.4).

Hence, the resulting wavelet representations $\mathbf{A}$ and $\mathbf{B}^T$ are of the following form

$$\mathbf{A} = (a_{\lambda,\lambda'})_{\lambda,\lambda'\in\mathcal{J}_X}, \quad a_{\lambda,\lambda'} = \sum_{i,l=1}^{d} 2^{-(|\lambda|+|\lambda'|)} \int_\Omega \frac{\partial\psi_{\lambda,i}}{\partial x_l}\frac{\partial\psi_{\lambda',i}}{\partial x_l}dx, \tag{5.2.5}$$

$$\mathbf{B}^T = (b_{\lambda,\lambda'})_{\lambda\in\mathcal{J}_X,\lambda'\in\mathcal{J}_M}, \quad b_{\lambda,\lambda'} = -\sum_{i=1}^{d} 2^{-|\lambda|} \int_\Omega \psi_{M,\lambda'}(x)\frac{\partial\psi_{\lambda,i}}{\partial x_i}(x)dx. \tag{5.2.6}$$

5.3. Compression Properties. The matrices $\mathbf{A}$, $\mathbf{B}$, defined by (5.2.5) and (5.2.6), are known to be compressible in a range that depends on the regularity of the wavelets, see [8]. However, the special piecewise polynomial nature of the above bases allows us to establish a somewhat larger range of compressibility compared with the general estimates. In this subsection, we analyze the compression properties of matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{B}^T$ in detail. The analysis in this section is based on the following version of the *Schur lemma* (which is folklore).

LEMMA 5.2. *Let $\mathbf{T} = (T_{l,l'})_{l\in\mathcal{I},l'\in\mathcal{I}'}$ be a matrix and let $\mathcal{I},\mathcal{I}'$ be countable index sets. Suppose that there exist sequences $(\varpi_l)_{l\in\mathcal{I}}$ and $(\tilde{\varpi}_{l'})_{l'\in\mathcal{I}'}$ such that*

$$\sum_{l'\in\mathcal{I}'} |T_{l,l'}|\tilde{\varpi}_{l'} \le c\varpi_l \qquad \text{and} \qquad \sum_{l\in\mathcal{I}} |T_{l,l'}|\varpi_l \le c\tilde{\varpi}_{l'}, \qquad l\in\mathcal{I}, l'\in\mathcal{I}', \tag{5.3.1}$$

then $\|\mathbf{T}\| \le c$.

Our numerical examples refer to the L-shaped domain $\Omega = (-1,1)^2 \setminus (-1,0]^2$. Thus Ω can be decomposed e.g. into three subpatches $\Omega_l, l = 1,2,3$, each being a simple translate of the unit square $(0,1)^2$. The spaces Π_{m,κ_l} consist then of polynomials in the classical sense. The moment conditions (5.2.2) hold then on all of Ω also for those wavelets whose support overlaps more than one subdomain. In this case the truncation rule that produces the compressed matrices $\mathbf{A}_j$ from (4.1.9) reads as follows, see [8, 2]. In order to indicate the role of the spatial dimension we keep the general notation although the example refers to $d = 2$. Given j, set

$$\tilde{a}_{\lambda,\nu} := \begin{cases} a_{\lambda,\nu}, & \big||\lambda| - |\nu|\big| \le j/d, \\ 0, & \text{else.} \end{cases} \tag{5.3.2}$$

Unless otherwise stated, we shall henceforth use the abbreviation $m = m_X$, $\tilde{m} = \tilde{m}_X$.

THEOREM 5.3. *For the matrix $\mathbf{A}$ defined by (5.2.5) and any $\epsilon > 0$ the following compression estimate holds:*

$$\|\mathbf{A} - \mathbf{A}_J\| \lesssim 2^{-J(m-3/2-\epsilon)/d}, \quad \textit{i.e.,} \quad \mathbf{A}\in\mathcal{C}_s,\ s < s^* = (m-3/2)/d. \tag{5.3.3}$$

Proof: Eq. (5.3.3) can be established by using Lemma 5.2 with $\mathcal{I} = \mathcal{I}' = \mathcal{J}_X$ and $\varpi_\lambda = \tilde{\varpi}_\lambda = 2^{|\lambda|(1-d)}$ for all $\lambda\in\mathcal{J}_X$. The first step is to estimate a typical entry in the wavelet representation. Let $\Omega_{\lambda,i}$ denote the support of the i-component of ψ_λ. We recall that derivatives of wavelets are again wavelets with the order of vanishing moments increased by one, [23]. Exploiting this fact, we obtain for suitable

polynomials $P_{\lambda',i,l}$ on $\Omega_{\lambda',i}$ of degree at most m (recall $\tilde{m} \geq m$) and $|\lambda'| \geq |\lambda|$

$$|(\nabla\psi_{X,\lambda}, \nabla\psi_{X,\lambda'})| = \left|\sum_{i,l=1}^{d}\left(\frac{\partial\psi_{\lambda,i}}{\partial x_l}, \frac{\partial\psi_{\lambda',i}}{\partial x_l}\right)\right| = \left|\sum_{i,l=1}^{d}\left(\frac{\partial\psi_{\lambda,i}}{\partial x_l} - P_{\lambda',i,l}, \frac{\partial\psi_{\lambda',i}}{\partial x_l}\right)\right|$$
$$\lesssim \sum_{i,l=1}^{d}\left\|\frac{\partial\psi_{\lambda,i}}{\partial x_l} - P_{\lambda',i,l}\right\|_{L_2(\Omega_{\lambda',i})} 2^{|\lambda'|},$$

where we have applied (2.2.1) with the weights from (5.2.4) to estimate the term $\left\|\frac{\partial\psi_{\lambda',i}}{\partial x_l}\right\|_{L_2}$ by $2^{|\lambda'|}$. Setting $j := |\lambda|$, $j' := |\lambda'|$, since $\frac{\partial\psi_{\lambda,i}}{\partial x_l} \in H^s$, $s < m - 3/2$, a classical Whitney type estimate yields therefore

$$|\,(\nabla\psi_\lambda, \nabla\psi_{\lambda'})\,| \lesssim \sum_{i,l=1}^{d} 2^{j'}2^{-j'(m-3/2-\epsilon)}\left|\frac{\partial\psi_{\lambda,i}}{\partial x_l}\right|_{H^{m-3/2-\epsilon}}$$
$$\lesssim \sum_{i=1}^{d} 2^{j'}2^{-j'(m-3/2-\epsilon)}|\psi_{\lambda,i}|_{H^{m-1/2-\epsilon}}$$
$$\lesssim 2^{j'}2^{-j'(m-3/2-\epsilon)}2^{j(m-1/2-\epsilon)}$$
$$\lesssim 2^{(j-j')(m-3/2-\epsilon)}2^{j+j'},$$

so that, taking the scaling matrix $\mathbf{D}_X$ into account, we derive

$$|a_{\lambda,\lambda'}| \lesssim 2^{(j-j')(m-3/2-\epsilon)}, \qquad j' \geq j. \tag{5.3.4}$$

The case $j' < j$ can be treated analogously,

$$|a_{\lambda,\lambda'}| \lesssim 2^{(j'-j)(m-3/2-\epsilon)}, \qquad j' < j. \tag{5.3.5}$$

According to (5.3.3) and (5.3.1), we have to show that

$$\sum_{|j-j'|>J/d}\sum_{|\lambda'|=j'}|a_{\lambda,\lambda'}|2^{j'(1-d)} \lesssim 2^{-J(m-3/2-\epsilon)/d}\cdot 2^{j(1-d)}. \tag{5.3.6}$$

Let us again first consider the case $j' > j$. We start by observing that the crude estimate (5.3.4) does not tell the whole truth. If we combine the fact that Ψ_X^+ is spanned by cardinal B-splines with the vanishing moment property (5.2.2) of the wavelet basis, we see that for a fixed value of $|\lambda|$, many of the entries $|a_{\lambda,\lambda'}|$ are zero. Roughly speaking, the non-vanishing entries correspond only to the wavelets $\psi_{\lambda'}$ for which the support of one component $\psi_{\lambda',i}$ intersects the corresponding singular support $\mathcal{S}_i$ of $\psi_{\lambda,i}$. The set $\mathcal{S}_i$ can be viewed as a submanifold of dimension $d-1$ with measure of the order $2^{-j(d-1)}$. Consequently, for $j' > j$, there are at most a fixed constant multiple of $2^{(j'-j)(d-1)}$ many wavelets possessing a non-trivial intersection with $\mathcal{S}_i$. Therefore we get

$$\sum_{|\lambda'|=j'}|a_{\lambda,\lambda'}| \lesssim 2^{(j-j')(m-3/2-\epsilon)}2^{(j'-j)(d-1)} \lesssim 2^{(j-j')(m-3/2-\epsilon+1-d)}. \tag{5.3.7}$$

Hence we finally obtain

$$\sum_{j'-j>J/d}\sum_{|\lambda'|=j'}|a_{\lambda,\lambda'}|2^{j'(1-d)} \lesssim \sum_{j'=j+J/d}^{\infty}2^{(j-j')(m-3/2-\epsilon+1-d)}2^{j'(1-d)} \tag{5.3.8}$$
$$\lesssim 2^{j(m-3/2-\epsilon+1-d))}\sum_{j'=j+J/d}^{\infty}2^{-j'(m-3/2-\epsilon)}$$
$$\lesssim 2^{j(m-3/2-\epsilon+1-d)}2^{-(J/d+j)(m-3/2-\epsilon)}$$
$$\lesssim 2^{j(1-d)}2^{-J(m-3/2-\epsilon)/d}.$$

The case $j' \leq j$ can be treated analogously. The second condition in (5.3.1) can be checked in a similar fashion which confirms (5.3.3). □

REMARK 5.4. *By combining the results in* [8] *with the analysis in* [13]*, one derives the following bound for the range of compressibility of the wavelet representation of an elliptic differential operator of order* $2t$

$$s^* := \min\left\{\frac{\sigma}{d} - \frac{1}{2}, \frac{2t+2\tilde{m}}{d}\right\}.$$

Here the parameter σ *must satisfy* $t+\sigma < \gamma$*, where* γ *bounds the Sobolev regularity of the wavelets. In the present case one has* $t = 1$*,* $\gamma = m - 1/2$*, i.e.,* $\sigma = m - 3/2$*, and hence* $s^* = (m-3/2)/d - 1/2$*. Therefore* (5.3.3) *ensures in any spatial dimension a gain in the compression range by* $1/2$ *when compared with the usual estimate.*

For more general domains when the κ_l are no longer affine some constructions of wavelet bases guarantee the full order of vanishing moments (5.2.2) only for those wavelets that are supported in a single patch Ω_l. Those wavelets overlapping several subdomains still have at least first order moments and hence their gradients have second order moments. Of course, this occurs only along a $(d-1)$-dimensional manifold, and can be compensated by modifying the compression rule (5.3.2). Moreover, those entries $a(\psi_\lambda, \psi_{\lambda'})$, for which the supports overlap each other but their singular supports (cut regions of tensor product B-splines) do not intersect, are no longer zero. However, since one of the wavelets is arbitrarily smooth throughout the integration domain, the order of vanishing moments increases to $\tilde{m}_X + 1$ so that these entries are much smaller than the remaining ones which suffices as well. Alternatively, one can employ the construction from [18] where vanishing moments are not constrained through patch interfaces.

A similar result can also be established for the matrix $\mathbf{B}^T$ defined in (5.2.6).

THEOREM 5.5. *Suppose that the order* m_X *of the multiresolution spaces for the velocity space* X *and the order* $\tilde{m}_M$ *of the vanishing moments of the pressure wavelets defined in* (5.2.3) *satisfy* $\tilde{m}_M \geq m_X - 1$*. Then for the matrix* $\mathbf{B}^T$ *defined in* (5.2.6) *and any* $\epsilon > 0$ *the following compression estimate holds:*

$$\|\mathbf{B}^T - \mathbf{B}_J^T\| \lesssim 2^{-J(m-3/2-\epsilon)/d}, \quad \textit{i.e.,} \quad \mathbf{B}^T \in \mathcal{C}_s,\ s < s^* = (m-3/2)/d. \tag{5.3.9}$$

Proof: The proof follows the lines of the proof of Theorem 5.3, therefore we only sketch the arguments. We use Lemma 5.2 for the case $\mathcal{I} = \mathcal{J}_X$, $\mathcal{I}' = \mathcal{J}_M$, $\varpi_\lambda = 2^{|\lambda|(1-d)}$, $\lambda \in \mathcal{J}_X$, and $\tilde{\varpi}_{\lambda'} = 2^{|\lambda'|(1-d)}, \lambda' \in \mathcal{J}_M$. As before, for suitable polynomials $P_{\lambda,i}$ on $\Omega_{\lambda'}$ of degree $< m-1$ and $j' = |\lambda'| \geq j = |\lambda|$ we obtain

$$\begin{aligned}
|(\nabla \cdot \psi_{X,\lambda}, \psi_{M,\lambda'})| &\leq \left|\sum_{i=1}^d \left(\frac{\partial \psi_{\lambda,i}}{\partial x_l} - P_{\lambda,i}, \psi_{M,\lambda'}\right)\right| \lesssim \sum_{i=1}^d \left\|\frac{\partial \psi_{\lambda,i}}{\partial x_i} - P_{\lambda,i}\right\|_{L_2(\Omega_{\lambda'})} \|\psi_{M,\lambda'}\|_{L_2} \\
&\lesssim \sum_{i=1}^d 2^{-j'(m-3/2-\epsilon)} \left|\frac{\partial \psi_{\lambda,i}}{\partial x_i}\right|_{H^{m-3/2-\epsilon}} \lesssim \sum_{i=1}^d 2^{-j'(m-3/2-\epsilon)} |\psi_{\lambda,i}|_{H^{m-1/2-\epsilon}} \\
&\lesssim 2^{-j'(m-3/2-\epsilon)} 2^{j(m-1/2-\epsilon)},
\end{aligned}$$

so that

$$|b_{\lambda,\lambda'}| \lesssim 2^{(j-j')(m-3/2-\epsilon)}, \qquad j' \geq j. \tag{5.3.10}$$

The case $j' < j$ can again be treated analogously. According to (5.3.9) and (5.3.1), we have to show that

$$\sum_{|j-j'|>J/d} \sum_{|\lambda'|=j'} |b_{\lambda,\lambda'}| 2^{j'(1-d)} \lesssim 2^{-J(m-3/2-\epsilon)/d} \cdot 2^{j(1-d)}. \tag{5.3.11}$$

Let us again first consider the case $j' > j$. By using similar arguments as in the proof of Theorem 5.3, we get

$$\sum_{|\lambda'|=j'} |b_{\lambda,\lambda'}| \lesssim 2^{(j-j')(m-3/2-\epsilon)} 2^{(j'-j)(d-1)} \lesssim 2^{(j-j')(m-3/2-\epsilon+1-d)}, \tag{5.3.12}$$

hence

$$\sum_{j'-j>J/d}\sum_{|\lambda'|=j'}|b_{\lambda,\lambda'}|2^{j'(1-d)} \lesssim \sum_{j'=j+J/d}^{\infty}2^{(j-j')(m-3/2-\epsilon+1-d)}2^{j'(1-d)} \tag{5.3.13}$$
$$\lesssim 2^{j(m-3/2-\epsilon+1-d)}\sum_{j'=j+J/d}^{\infty}2^{-j'(m-3/2-\epsilon)}$$
$$\lesssim 2^{j(1-d)}2^{-J(m-3/2-\epsilon)/d}.$$

The case $j' \le j$ can again be treated analogously. The second condition in (5.3.1) can be verified by employing similar arguments. □

To determine finally the compressibility of the matrix $\mathbf{R}$ from (3.1.3) we can apply the same reasoning for $\partial\psi_{\lambda,i}/\partial x_l$ and $\psi_{M,\lambda'}$ replaced by $\tilde{\psi}_{M,\lambda}$. Since in this case no derivatives are involved and $\tilde{\Psi}_M$ just as Ψ_M is patchwise defined, the compressibility range is again determined by the order m_M of the primal basis Ψ_M (which limits the order of the polynomials that can be subtracted in the inner products) and the Sobolev regularity $\tilde{\gamma}_M$ of the dual basis $\tilde{\Psi}$ inside each patch Ω_l. The constructions in [6, 17] allow one to realize therefore any desired order s_R^* of compressibility for $\mathbf{R}$ provided m_M and $\tilde{\gamma}_M$ are chosen accordingly.

Theorems 5.3 and 5.5 tell us now in which range for a given choice of wavelet bases the general results Theorem 4.8 and Corollary 3.4 assert asymptotically optimal accuracy/work balance for the adaptive solution of the Stokes problem.

5.4. Regularity Theory for the Stokes Problem. So far we have presented some numerical tools to serve as input for an adaptive scheme that realizes asymptotically optimal convergence rates in (essentially) linear time within a certain range of error decay orders determined by the compressibility of the involved wavelet representations. A natural question is whether at all or under which circumstances the corresponding accuracy/work balance is better than for technically much simpler schemes based e.g. on uniformly refined meshes — in brief: when does adaptivity pay? It turns out that this question is inherently related to the *regularity* of the approximated solution. More precisely, while a given order of best approximation from trial spaces for preassigned uniform meshes (referred to as *linear schemes*) is characterized by the *Sobolev* regularity of the approximand, the order of *nonlinear* or *best N-term approximation* is (almost) characterized by the regularity in a certain *Besov scale* to be specified in a moment, see also [19]. To explain this let H^t denote a (closed subspace of a) Sobolev space such as $H_0^1(\Omega)$ respectively $H_0^1(\Omega)^d$ or $L_{2,0}(\Omega)$ for $t = 0$ and let Υ denote a wavelet basis in H^t satisfying a norm equivalence of the form (2.2.1) with suitable scaling matrix $\mathbf{D}^t$. In analogy to (4.1.1) let

$$\sigma_{N,H^t}(v) := \inf_{\mathbf{w},\#\mathbf{w}\le N}\|v - \mathbf{w}^T(\mathbf{D}^t)^{-1}\Upsilon\|_{H^t} \tag{5.4.1}$$

denote the error of best wavelet N-term approximation in H^t. The following fact has been shown in [12].

PROPOSITION 5.6. *Whenever $t \le s$ let*

$$\frac{1}{\alpha} = \frac{r-t}{d} + \frac{1}{2}. \tag{5.4.2}$$

Then (for a sufficiently regular basis Υ) one has

$$\sum_{N=1}^{\infty}\left(N^{(r-t)/d}\sigma_{N,H^t}(v)\right)^{\alpha} < \infty \quad \textit{iff} \quad v \in B_\alpha^r(L_\alpha(\Omega)). \tag{5.4.3}$$

Note that $B_\alpha^r(L_\alpha(\Omega))$ is the largest space of smoothnes r in L_α which is still embedded in H^t, since (5.4.2) marks the Sobolev embedding line. Clearly, (2.2.1) says that for $v = \mathbf{v}^T(\mathbf{D}^t)^{-1}\Upsilon$ one has

$$\sigma_{N,H^t}(v) \sim \sigma_{N,\ell_2}(\mathbf{v}). \tag{5.4.4}$$

Moreover, (5.4.3), (5.4.4) mean that when $v \in B_\alpha^r(L_\alpha(\Omega))$ then the best N-term approximation of its wavelet coefficients $\mathbf{v}$ decays at least like $\sigma_{N,\ell_2}(\mathbf{v}) \lesssim N^{-(r-t)/d}$. This is sharp in the sense that the

exponent $s = (r-t)/d$ is best possible. This subtle gap in the characterization of the Besov spaces is due to the small difference between the classical spaces ℓ_τ (characterizing wavelet coefficients for elements in the Besov space) and the weak type space ℓ_τ^w characterizing best N-term approximation of the wavelet coefficient sequences in ℓ_2, [19].

These facts suggest to ask for the regularity of the solution (u,p) of the Stokes problem (5.1.1) in the relevant Besov scales.

During the past years the Sobolev and Besov regularity theory for the Stokes problem has attracted the attention of several authors. We refer to [21, 24] for the Sobolev and to [10] for the Besov regularity theory. We consider here a planar polygonal domain $\Omega \subset \mathbb{R}^2$. This section can be viewed as both, a summary and a specific application of the results in [10, 21, 24] to the special case of (5.1.1) for the L-shaped domain. These results will be used later in Section 6 to select and properly interpret the numerical tests.

Some preparations are necessary. The smooth segments of $\partial\Omega$ are denoted by $\overline{\Gamma}_l$, Γ_l open, $l = 1,\ldots,N$, numbered in positive orientation. Furthermore, $\mathcal{V}_l$ denotes the endpoint of $\overline{\Gamma}_l$ and ω_l denotes the measure of the interior angle at $\mathcal{V}_l$. Moreover, we introduce polar coordinates (r_l,θ_l) in the vicinity of each vertex $\mathcal{V}_l$. By ζ_l we will always denote a suitable C^∞ truncation function. Finally, $z_{l,m}$ is a real solution of the transcendental equation

$$\sin^2(z\omega_l) = z^2\sin^2(\omega_l). \tag{5.4.5}$$

Unless otherwise stated, we shall always assume that

$$\tan(\omega_l) \neq \omega_l \qquad \text{for every} \quad l. \tag{5.4.6}$$

Let us first discuss the regularity of the velocity u. For $f \in L_2(\Omega)^2$, clearly the best we could expect is $u \in H^2(\Omega)^2$. However, it is well-known that even for smooth right-hand sides the Sobolev regularity of u may drop down due to certain singularity functions, see [21, 24]. In our case, a typical singular part u_S is of the form

$$u_S = \sum_l \sum_{0<z_{l,m}<1} c_{j,m}\zeta_l(r_l) r_l^{z_{l,m}} S_{l,z_{l,m}}, \tag{5.4.7}$$

where

$$S_{l,z} = (z\sin((z-2)\omega_l) + (2-z)\sin(z\omega_l))S_{l,z,1} - z(\cos((z-2)\omega_l) - \cos(z\omega_l))S_{l,z,2}, \tag{5.4.8}$$

with

$$\begin{aligned} S_{l,z,1} &= (z\sin((z-2)\theta_l) - (z+2)\sin(z\theta_l), z(\cos((z-2)\theta_l) - \cos(z\theta_l)))^T, \\ S_{l,z,2} &= (-z(\cos((z-2)\theta_l) - \cos(z\theta_l)), z\sin((z-2)\theta_l) - (z-2)\sin(z\theta_l))^T. \end{aligned} \tag{5.4.9}$$

We see that the singular part u_S describes the influence of the domain since it is independent of the given right-hand side f. Obviously, the Sobolev regularity of u_S decreases significantly as the angles ω_l increase. In contrast to this, the Besov regularity of u_S is almost independent of the shape of the domain in the following sense.

In fact, by following the lines in [10], the following result can be established.

THEOREM 5.7. *Any singular solution defined by* (5.4.7) *satisfies*

$$u_S \in B_\tau^r(L_\tau(\Omega))^2, \qquad \textit{for all} \quad r>0, \quad \frac{1}{\tau} = \frac{r-1}{2} + \frac{1}{2}. \tag{5.4.10}$$

This result can be proved by showing first that any singular solution defined by *(5.4.7)* satisfies

$$u_S \in B_\tau^r(L_\tau(\Omega))^2, \qquad \text{for all} \quad r>0, \quad \frac{1}{\tau} = \frac{r}{2} + \frac{1}{2}. \tag{5.4.11}$$

Note that the Besov scale in (5.4.11) corresponds to $H^t = L_2$ in Proposition 5.6 above and hence is related to nonlinear approximation in L_2. For the velocity components the best N-term approximation in H^1 is relevant though. To this end, obviously, u_S is contained in $H^1(\Omega)^2$. Therefore the result follows by interpolation between $H^1(\Omega)^2$ and $B^r_\tau(L_\tau(\Omega))^2$, $1/\tau = r/2+1/2$, see [11] for additional details. □

It remains to study the regularity of the pressure p. By writing

$$\nabla p = f + \Delta u, \tag{5.4.12}$$

inserting the singularity functions according to (5.4.7) and integrating (5.4.12), we see that for $f \in L_2(\Omega)^2$ the pressure can also possess a singular part p_S which in the vicinity of $\mathcal{V}_l$ can be written as

$$p_S = C_l(\theta_l) r_l^{z_{l,m}-1} \tag{5.4.13}$$

for some smooth function $C_l(\theta_l)$. Once again, by following the lines in [10], it can be shown that p_S has arbitrary high regularity in the nonlinear approximation scale $B^r_\tau(L_\tau(\Omega))$, $1/\tau = r/2 + 1/2$.

To construct the singular solutions according to (5.4.7), we have to determine the solutions of (5.4.5). This equation has been studied in detail in [21]. Let us briefly recall the results. We introduce the exceptional angle

$$\omega_0 = \tan(\omega_0). \tag{5.4.14}$$

Then the following Lemma holds.

LEMMA 5.8. *The equation* (5.4.5) *has no root in the strip* $0 < \Re(z) < 1$ *when* $\omega < \pi$. *It has only one single real root in that strip when* $\pi < \omega < \omega_0$ *and it has two distinct simple real roots in that strip when* $\omega_0 < \omega \leq 2\pi$.

Now we want to apply these results to the L-shaped domain. According to Lemma 5.8 and (5.4.7), we only have singular solutions corresponding to the reentrant corner. It can be checked that the exceptional angle ω_0 is given by $\omega_0 = 257.40$ degrees. Consequently, in our case (5.4.5) has two distinct roots and we have two singular solutions. The two roots can easily be computed numerically. We obtain

$$z_{1,1} = 0.9085291898461, \qquad z_{1,2} = 0.54448373678246.$$

Obviously, the second root is the 'worse' one. Consequently, to test the adaptive scheme, we designed the right-hand side f in such a way that the solution $\mathbf{u}$ is exactly the singularity function corresponding to $z_{1,2}$, i.e.,

$$u = \zeta(r) r^{0.54448373678246} S_{1,z_{1,2}},$$

see also Section 6. According to (5.4.12), we choose the pressure p as a solution to

$$\nabla p = \Delta(r^{z_{1,2}} S_{1,z_{1,2}}).$$

Finally, we normalize p in $L_{2,0}(\Omega)$. The suitable right-hand side can be obtained by another application of (5.4.12).

REMARK 5.9. *One can easily verify that*

$$u \in H^r(\Omega)^2, \quad r < r^*_X := 1.54448373678246 \qquad \textit{and} \quad p \in H^r(\Omega), \quad r < r^*_M := 0.54448373678246$$

(i.e., $u \notin H^{r^*_X}(\Omega)^2$, $p \notin H^{r^*_M}(\Omega)$*), which limits the convergence rate of uniform refinements. One the other hand,* u *and* p *both have* arbitrary *high Besov regularity. Hence, in principle, wavelet bases with high order regularity would give rise to corrispondingly high order adaptive approximation rates.*

6. Numerical Results. In this section, we present some numerical experiments for the Stokes-problem on the planar L-shaped domain $\Omega = (-1,1)^2 \setminus [-1,0]^2$. We employ different versions from the family of wavelet bases Ψ_X and Ψ_M from Section 5.2 for velocities and pressure, respectively.

Our objective is not to present a fully matured code but to gain additional quantitative insight that complements the preceding theoretical results of primarily asymptotic nature. This concerns the

quantitative effect of "violating" the LBB condition and the tradeoff between larger supports and better compressibility when using higher order wavelets as well as suggestions for further algorithmic variants and developments. For instance, the theoretical estimates, e.g. on the number K of iterations in **ADV**, are presumably overly conservative. So it would be interesting to see experimentally whether typically smaller numbers suffice or whether monitoring residuals pays to realize significantly earlier terminations. Furthermore, we wish to see how the scheme copes with highly singular cases suggested by the discussion in Section 5.4 compared with more regular solutions. More extensive tests of variants derived from first experiences will be presented elsewhere.

6.1. Discretization of the Pressure. Recall from (5.1.2) that $L_{2,0}(\Omega)$ is the appropriate pressure space. Hence the zero mean constraint requires special care. Here we exploit the fact that all wavelets in Ψ_M^- have, according to (5.2.3), vanishing moments of order $\tilde{m}_M \geq m_M \geq 1$, so that

$$\int_\Omega \psi_{M,\lambda}(x)\, dx = 0, \qquad \lambda \in \mathcal{J}_M, |\lambda| > j_0.$$

Hence for any $q = \mathbf{q}^T \Psi_M$ one has

$$\int_\Omega q(x)\, dx = \sum_{|\lambda|=j_0} q_\lambda \int_\Omega \psi_{M,\lambda}(x)\, dx =: \sum_{|\lambda|=j_0} q_\lambda \alpha_\lambda =: I_\Omega(q).$$

On the other hand, the scaling functions form a partition of unity, i.e.,

$$1 \equiv \sum_{|\lambda|=j_0} \tilde{\alpha}_\lambda\, \psi_{M,\lambda}(x), \qquad x \in \Omega, \qquad \tilde{\alpha}_\lambda := \int_\Omega \tilde{\psi}_{M,\lambda}(x)\, dx = (1, \tilde{\psi}_{M,\lambda}),$$

where $\{\tilde{\psi}_{M,\lambda} : |\lambda| = j_0\}$ is the (explicitly known) *dual basis* for the scaling functions in Ψ_M, i.e., $(\psi_{M,\lambda}, \tilde{\psi}_{M,\lambda'}) = \delta_{\lambda,\lambda'}$, see [6, 17]. Thus, denoting by $\mu(\Omega)$ the Lebesgue measure of Ω, we obtain a projection $P_0 : L_2(\Omega) \to L_{2,0}(\Omega)$ by

$$P_0(q) := \sum_{|\lambda|=j_0} \left(q_\lambda - \frac{I_\Omega(q)}{\mu(\Omega)} \tilde{\alpha}_\lambda \right) \psi_{M,\lambda} + \sum_{|\lambda|>j_0} q_\lambda \psi_{M,\lambda},$$

that factors out constants. Hence, realizing the zero mean constraint, requires modifications only on the *coarsest* level, whereas the wavelet coefficients remain unchanged. Since operators are only applied approximately, corresponding corrections are needed after applying **B** and also after coarsening. Since the projection P_0 depends on the particular primal wavelet basis for $L_2(\Omega)$ all arrays have to refer to the same basis so that the Riesz map $\mathbf{R} = (\tilde{\Psi}_M, \tilde{\Psi}_M)$ is needed in the second step (3.1.7)of the Uzawa iteration.

Note that the present way of factoring out constants is only a first convenient option. A drawback reflected by the experiments below is that due to the nature of P_0 always all coarse scale functions will be involved in the pressure approximations. In particular, for higher order trial functions this number grows, so that at least for the first few refinement steps the work/accuracy balance of the scheme is less favorable for the pressure component. Local coarse scale basis functions would remedy this effect.

A detailed description of the routines **APPLY** and **NCOARSE** can be found in [2, 8] combined with the above provisions with respect to the matrix **B**. As mentioned before, the routine **ELLSOLVE** is essentially the adaptive Poisson solver from [2]. This indicates the principal potential of recycling these basic routines for the treatment of problems with increasing complexity.

6.2. Description of the Test Cases. We wish to report below on two different test cases. Example (I) corresponds to *the most singular* solution described in Section 5.4. As can be seen in Figure 6.1, the pressure exhibits a strong singularity at the reentrant corner. In order to keep the effort for computing an exact reference solution as moderate as possible we have computed an approximation of the exact solution by truncating p. Of course, this limits the number of iterations of the adaptive algorithm for which meaningfull comparisions can be made.

Example (II) involves a pressure which is *localized* around the reentrant corner, has strong gradients but is smooth. More precisely, we have chosen an exact solution for the velocity which is very similar to the one above and a pressure solution which is constant around the reentrant corner and multiplied by a smooth cut off function. These functions are displayed in Figure 6.2.

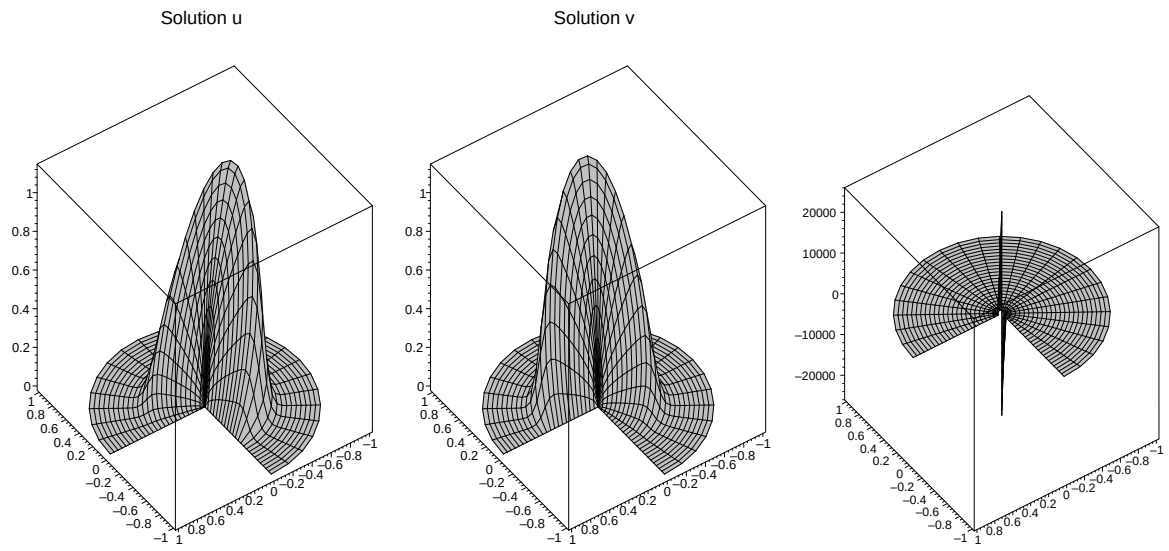

FIGURE 6.1. *Exact solutiopn for the first example. Velocity components (left and middle) and pressure (right). The pressure functions exhibits a strong singularity and is only shown up to* $r = 0.001$ *in polar coordinates.*

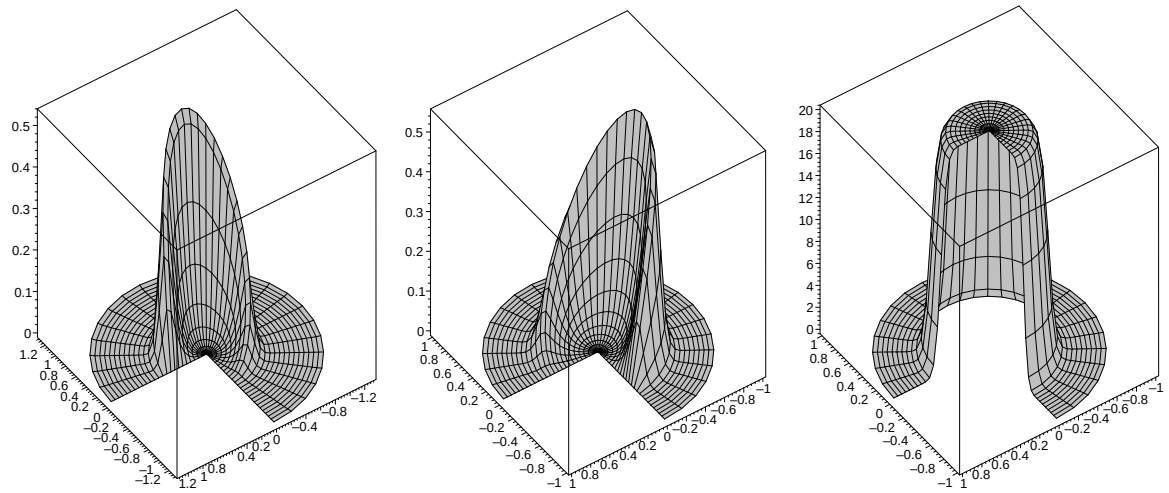

FIGURE 6.2. *Exact solutiopn for the second example. Velocity components (left and middle) and pressure (right).*

6.2.1. Choice of the Parameters. We expect that some of the constants resulting from the analysis are actually too pessimistic. For instance, deriving estimates for the constants in the norm equivalences, we have estimated K to be in the range of 15, which turned out to entail unnecessarily high accuracy in the treatment of the inner Poisson problems while the pressure approximation and hence the right-hand side for the Laplace problem are still poor. Several numerical experiments with different trial functions and for different test cases, indicate that $K = 3$ already seems to suffice and that the alternatives discussed in Section 3 are in these cases not necessary. All subsequent results are therefore based on this choice. Moreover, we have used $\rho = 0.6$ and $\omega = 1.3$ in all experiments.

6.3. Rate of Convergence. Table 6.1 displays the results for Example (I), employing piecewise linear trial functions for the velocity and piecewise constant functions for the pressure. We are interested in the relation between the error produced for a given number of degrees of freedom by the adaptive scheme and the error of best N-term approximation with respect to the underlying wavelet basis. To describe the results we denote by $\mathbf{u}^1, \mathbf{u}^2$ the wavelet coefficient arrays of the first and second velocity

component and define for $\mathbf{x} \in \{\mathbf{u^1}, \mathbf{u^2}, \mathbf{p}\}$ by

$$\rho_{\mathbf{x}} := \frac{\|\mathbf{x} - \mathbf{x}_{\mathbf{\Lambda}}\|_{\ell_2}}{\|\mathbf{x} - \mathbf{x}_{\#\mathbf{\Lambda}}\|_{\ell_2}}, \qquad r_{\mathbf{x}} := \frac{\|\mathbf{x} - \mathbf{x}_{\mathbf{\Lambda}}\|_{\ell_2}}{\|\mathbf{x}\|_{\ell_2}},$$

the ratio of the error of the adaptive approximation and the corresponding best N-term approximation, respectively the relative error. Recall from Corollary 3.4 that these quantities also reflect the error in the energy norms. We see that the velocity approximation is from the beginning very close to its best N-term approximation. For the reasons indicated above this is different for the pressure. The application of P_0 fills up the coarsest level which in this example has 768 degrees of freedom. To explain this in more detail assume that the adaptive method picks exactly one scaling function, so that the degree of freedom for the pressure would be 1. Since the integral of a scaling function is not zero, the pressure projection P_0 produces a non-zero constant whose expansion involves *all* scaling function coefficients. This is the reason why at the early stage of the refinement process the work accuracy balance for the pressure is less favorable. However, the last two iterates shown in the table indicate that the scheme catches up with the optimal rate. Local coarse scale bases would of course yield better results already from the beginning of the adaptive refinements.

It	δ	$\#\Lambda_{u^1}$	$\rho_{\mathbf{u^1}}$	$r_{\mathbf{u^1}}$	$\#\Lambda_{u^2}$	$\rho_{\mathbf{u^2}}$	$r_{\mathbf{u^2}}$	$\#\Lambda_p$	$\rho_{\mathbf{p}}$	$r_{\mathbf{p}}$
1	11.730947	33	1.04	0.6838	34	1.04	0.6744	768	130.35	1.0024
2	5.865474	84	1.26	0.3427	83	1.24	0.3447	768	130.40	1.0028
3	2.932737	193	1.32	0.1530	184	1.31	0.1541	768	15.37	0.5234
4	1.466368	446	1.29	0.0821	450	1.29	0.0897	929	4.15	0.2218
5	0.733184	1070	1.27	0.0434	1065	1.27	0.0456	1211	2.58	0.1034

TABLE 6.1

Results for the first example. Numbers of adaptively generated degrees of freedom, ratio to best N-term approximation and relative errors.

The results for Example (II) are shown in Table 6.2 and plots of the approximations are displayed in Figure 6.4. We see that the computed approximations differ only by a very moderate factor from the best N-term approximation. The results suggest the following directions for more systematic im-

It	δ	$\#\Lambda_{u^1}$	$\rho_{\mathbf{u^1}}$	$r_{\mathbf{u^1}}$	$\#\Lambda_{u^2}$	$\rho_{\mathbf{u^2}}$	$r_{\mathbf{u^2}}$	$\#\Lambda_p$	$\rho_{\mathbf{p}}$	$r_{\mathbf{p}}$
1	15.636636	278	28.20	1.2936	364	60.31	2.1867	768	6.96	0.3329
2	7.818318	261	8.30	0.4028	295	16.10	0.7003	768	3.76	0.1800
3	3.909159	234	3.72	0.1995	274	5.63	0.2617	768	1.80	0.0863
4	1.954580	180	1.25	0.0886	249	2.08	0.1056	810	1.22	0.0452
5	0.977290	233	1.14	0.0615	267	1.29	0.0615	980	1.07	0.0231
6	0.488645	298	1.11	0.0480	321	1.17	0.0470	1276	1.05	0.0117
7	0.244322	456	1.35	0.0398	505	1.43	0.0265	1551	1.09	0.0061
8	0.122161	704	1.36	0.0250	724	1.39	0.0177	1842	1.24	0.0035

TABLE 6.2

Results for the second example. Numbers of adaptively generated degrees of freedom, ratio to best N-term approximation and relative error.

plementations. The simple Richardson iteration should be replaced (possibly after a few initial steps)

by gradient or conjugate gradient steps. This should speed up convergence and avoid a necessarily pessimistic estimation of step size parameters. Since all algorithmic ingredients still require the same type of (approximate) matrix/vector multiplications one can employ the same routines. One should then include, however, monitoring residuals which, due to (2.3.1), should detect rapid convergence for a possible early termination of the iterations in **ADV (ii)**. Moreover, higher order wavelets should be tested to exploit larger compressibility ranges.

High order dircretizations.. Recall from Section 5.3 that the compressibility range of the wavelet representations grows with increasing regularity and hence order of the wavelet bases, see Theorems 5.3, 5.5. Moreover, the regularity results from Theorem 5.7 and Remark 5.9 indicate that the larger the compressibility range of the wavelet representations the more an adaptive scheme would gain at least asymptotically over uniform refinements. This suggest investigating the quantitative effect of employing higher order spline wavelets.

We compare now discretizations of various orders for the pressure in the second example. In Figure 6.3, we have shown the relative error versus the number of unknowns in a logarithmic scale. Comparing the slopes of the best N-term approximation, we obtain the expected asymptotic gain for increasing orders, again at the end with moderate values for the ratios $\rho_{\mathbf{x}}$. However, we also see that the fast decay

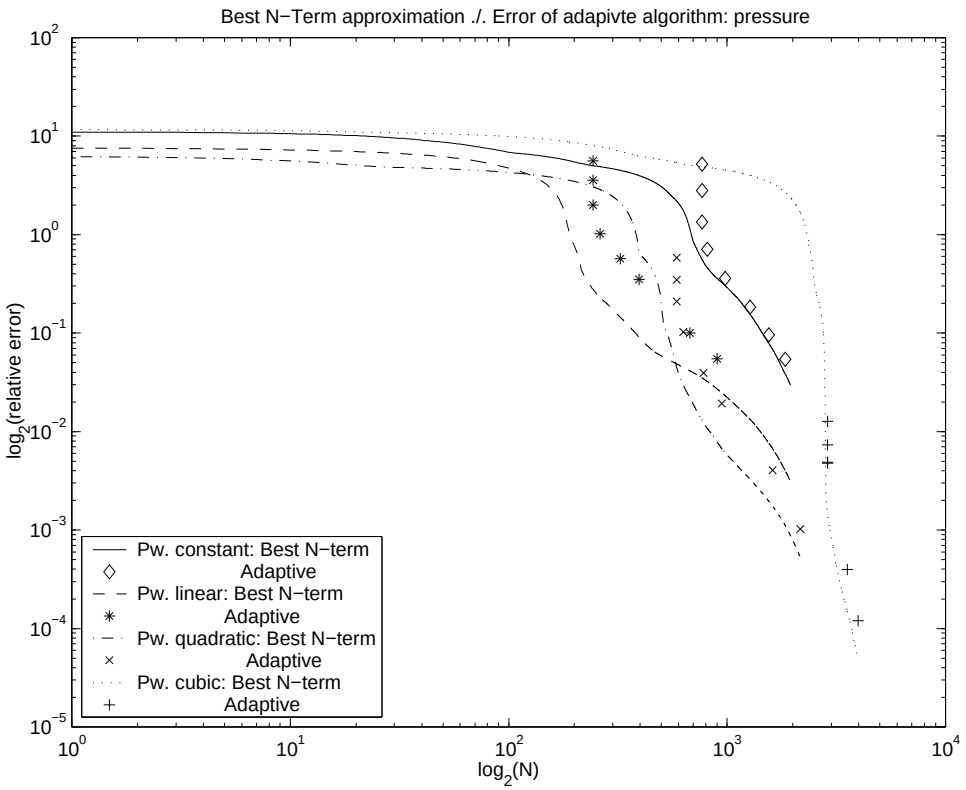

FIGURE 6.3. *Relative error versus number of unknowns for spline wavelets of different order for the discretization of the pressure in the second example.*

of the rate of the best N-term approximation is delayed more and more for an increasing order of trial functions. For instance, for piecewise cubic wavelets, we obtain an almost horizontal line until $N \approx 2000$. This is on one hand due to some technical restrictions of the particular patchwise tensor product wavelet bases used here requiring a certain coarsets level j_0 on each patch. The values for j_0 are shown in Table 6.3 for different orders. We see that j_0 increases with m (the case $m = 2$ is somewhat special due to the very local character of primal and dual functions). We display also the number of unknowns for the coarsest level, i.e., the number of scaling functions on level $j = j_0$. On the other hand, as pointed out before, the nature of P_0 keeps all coarse scale basis functions active. This explains why the slope of the best N-term approximation is almost horizontal until all scaling functions are used up. There are

several ways to alleviate this problem also for higher order discretizations. Aside from using local coarse scale basis functions with zero mean one can take a ficticous domain approach and append the boundary conditions by Lagrange multipliers. This allows one to use periodic wavelet bases on the fictitious domain where the minimal level can be always chosen as $j_0 = 0$ for all values of m and $\tilde{m}$. This issue will be addressed elsewhere.

$m, \tilde{m}$	1,3	2,2	3,3	4,4
j_0	4	3	4	5
N_Φ	705	242	587	2882

TABLE 6.3

Minimal level j_0 and number of scaling functions N_Φ on the minimal level for different order discretizations.

6.4. The LBB-Condition. At the first glance it is somewhat puzzling that in the analyis of the adaptive Uzawa method the LBB condition did not play any role. Roughly, speaking this is due to the fact that conceptually at every stage of the algorithm the full infinite dimensional operator is applied within a certain tolerance that has to be chosen tight enough to inherit the stability properties of the original infinite dimensional problem. This effect of adaptive schemes in connection with saddle point problems and also with more complex variational problems has been observed first in (a predecessor of) [9], see also [14] for saddle point problems. Hence it is interesting to study the quantitative influence of the choice of bases. Therefore, we have included a combination of bases for which pairs of fixed finite dimensional subspaces would violate the LBB-condition, namely piecewise linear trial functions for both velocity and pressure. The results are displayed in Table 6.4. We see that the rate of the best N-term approximation is still matched fairly well with ratios that are only slightly larger than in Table 6.2 for the piecewise linear/piecewise constant discretization. Note that the oscillations in the pressure approximation for

It	δ	$\#\Lambda_{u^1}$	$\rho_{\mathbf{u^1}}$	$r_{\mathbf{u^1}}$	$\#\Lambda_{u^2}$	$\rho_{\mathbf{u^2}}$	$r_{\mathbf{u^2}}$	$\#\Lambda_p$	$\rho_{\mathbf{p}}$	$r_{\mathbf{p}}$
1	16.743449	1	1.00	0.9293	1	1.00	0.9300	243	6.27552	0.3354
2	8.371724	1	1.00	0.9304	1	1.00	0.9292	243	3.98811	0.2131
3	4.185862	5	1.00	0.7586	5	1.00	0.7588	243	2.23810	0.1196
4	2.092931	20	1.13	0.4064	24	1.45	0.3979	262	2.08107	0.0612
5	1.046466	61	1.47	0.2107	77	1.79	0.2107	324	2.72102	0.0339
6	0.523233	178	1.33	0.1060	198	1.52	0.1306	396	2.81079	0.0209
7	0.261617	294	1.19	0.0533	286	1.46	0.0744	674	2.21371	0.0108
8	0.130808	478	1.25	0.0271	531	1.46	0.0362	899	1.83271	0.0071

TABLE 6.4

Results for the second example with piecewise linear trial functions for velocity and pressure. Note that in this case the number of degrees of freedom for the coarsest level is 243.

unstable elements shown by the experiments in [4] are not observed in the present context, see Figure 6.4. This seems to results from the different pressure update.

REFERENCES

[1] A. Barinka, *Fast Evaluation Tools for Adaptive Wavelet Schemes*, PhD Thesis, RWTH Aachen, in preparation.

[2] A. Barinka, T. Barsch, P. Charton, A. Cohen, S. Dahlke, W. Dahmen, and K. Urban, Adaptive wavelet schemes for elliptic problems – Implementation and numerical experiments, IGPM-Report # 173, RWTH Aachen, 1999, to appear in: *SIAM J. Scient. Comput.*

[3] A. Barinka, S. Dahlke, and W. Dahmen, Adaptive application of operators in standard wavelet representation, in preparation.
[4] E. Bänsch, P. Morin, and R.H. Nochetto, An adaptive Uzawa FEM for Stokes: Convergence without the inf-sup, preprint, 2001.
[5] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, 1991.
[6] A. Canuto, A. Tabacco, and K. Urban, The wavelet element method, part I: Construction and analysis, *Appl. Comp. Harm. Anal.*, **6** (1999), 1-52.
[7] A. Cohen, *Wavelet Methods in Numerical Analysis*, in: Handbook of Numerical Analysis, P.G. Ciarlet, J.L. Lions, eds., Elsevier, 2000, 417-711.
[8] A. Cohen, W. Dahmen, and R.A. DeVore, Adaptive wavelet methods for elliptic operator equations — Convergence rates, *Math. Comp.* **70** (2001), 27-75.
[9] A. Cohen, W. Dahmen, R.A. DeVore, Adaptive wavelet methods II - Beyond the elliptic case, IGPM Report # 199, RWTH Aachen, 2000.
[10] S. Dahlke, Besov regularity for the Stokes problem, in: Avances in Multivariate Approximation, W. Haußmann, K. Jetter, and M. Reimer, eds., Wiley VCH, Mathematical Research 107, Berlin, 1999, 129-138.
[11] S. Dahlke, Besov regularity for elliptic boundary value problems on polygonal domains, *Appl. Math. Lett.* **12(6)** (1999), 31-36.
[12] S. Dahlke, W. Dahmen, and R.A. DeVore, Nonlinear approximation and adaptive techniques for solving elliptic operator equations, in: Multiscale Wavelet Methods for PDEs, W. Dahmen, A. Kurdila, and P. Oswald, eds., Academic Press, San Diego, 1997, 237-284.
[13] S. Dahlke, W. Dahmen, R. Hochmuth, and R. Schneider, Stable multiscale bases and local error estimation for elliptic problems, *Appl. Numer. Math.* **23(1)** (1997), 21-48.
[14] S. Dahlke, R. Hochmuth, and K. Urban, Adaptive wavelet methods for saddle point problems, *Math. Model. Numer. Anal. (M2AN)* **34(5)** (2000), 1003-1022.
[15] W. Dahmen, Wavelet methods for PDEs — Some recent developments, J. Comp. Appl. Math. **128** (2001), 133-185.
[16] W. Dahmen, A. Kunoth, and K. Urban, Biorthogonal spline-wavelets on the interval — Stability and moment conditions, *Appl. Comp. Harm. Anal.* **6**, 132-196 (1999).
[17] W. Dahmen and R. Schneider, Composite wavelet bases for operator equations, *Math. Comp.*, **68** (1999), 1533-1567.
[18] W. Dahmen and R. Schneider, Wavelets on manifolds I. Construction and domain decomposition, *SIAM J. Math. Anal.*, **31** (1999), 184-230.
[19] R.A. DeVore, Nonlinear approximation, *Acta Numerica* **7** (1998), 51-150.
[20] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier-Stokes-Equations*, Springer-Verlag, 2nd edition, 1986.
[21] P. Grisvard, *Singularities in Boundary Value Problems*, Research Notes in Applied Mathematics **22**, Springer-Verlag, 1992.
[22] A. Kunoth, *Wavelet Methods for Minimization Problems Involving Elliptic Partial Differential Equations*, Teubner-Verlag, 2001
[23] P.G. Lemarié-Rieusset, *Analyses multi-résolutions non orthogonales, Commutation entre Projecteurs et Derivation et Ondelettes Vecteurs à divergence nulle* (in french), Rev. Mat. Iberoamericana **8** (1992), 221-236.
[24] J. Osborn, Regularity of solutions to the Stokes problem in a polygonal domain, in: Symposium on Numerical Solutions of Partial Differential Equations III, B. Hubbart, ed., Academic Press, 1975, 393-411.
[25] J. Vorloeper, *Multiskalenverfahren und Gebietszerlegungsmethoden* (in german), Master Thesis, RWTH Aachen, 1999.

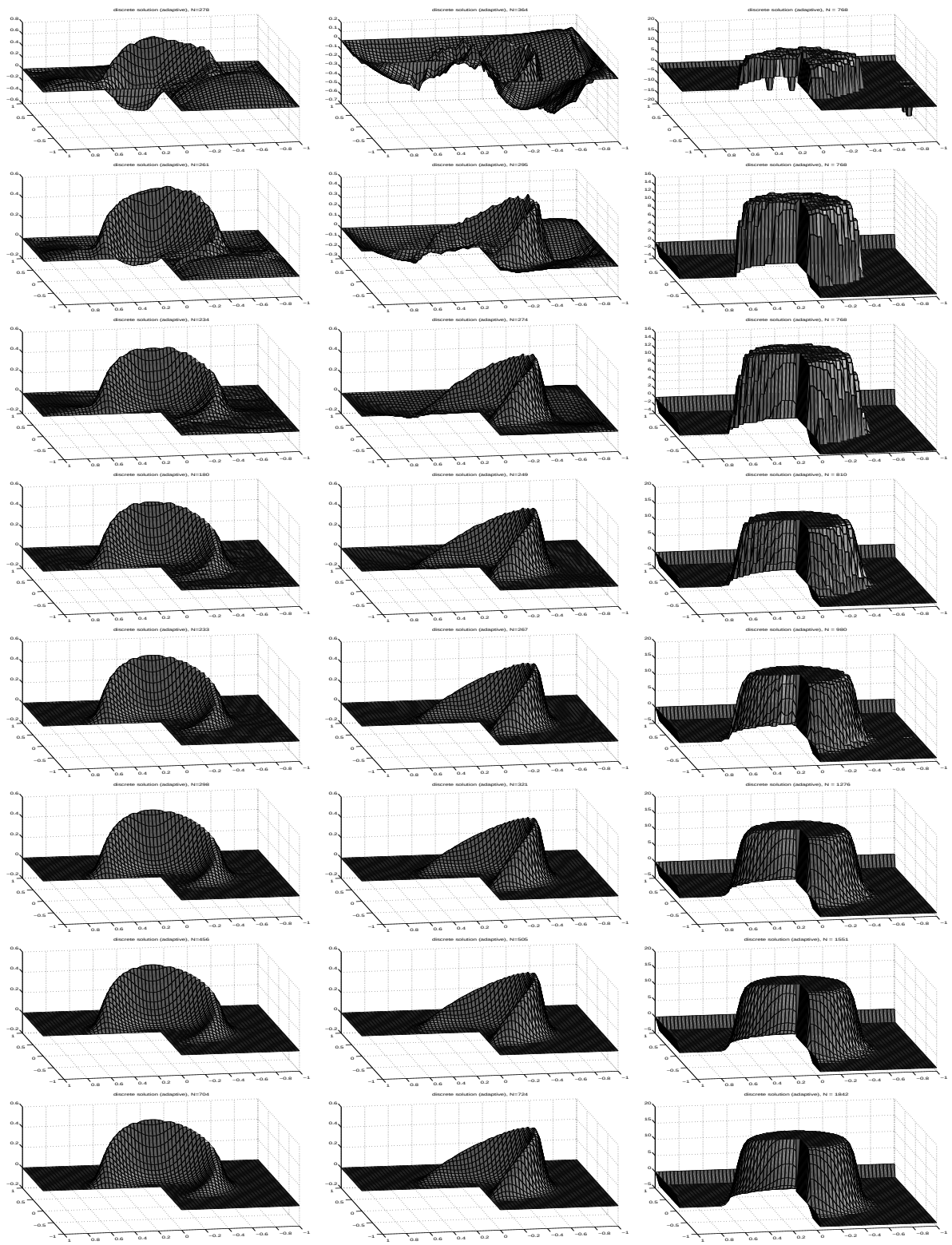

FIGURE 6.4. *Approximations for the second example. First and second velocity component (left and middle column) and pressure (right column).*

Berichte aus der Technomathematik ISSN 1435-7968

`http://www.math.uni-bremen.de/zetem/berichte.html`

— Vertrieb durch den Autor —

Reports **Stand: 17. Juli 2001**

98–01. Peter Benner, Heike Faßbender:
An Implicitly Restarted Symplectic Lanczos Method for the Symplectic Eigenvalue Problem, Juli 1998.

98–02. Heike Faßbender:
Sliding Window Schemes for Discrete Least-Squares Approximation by Trigonometric Polynomials, Juli 1998.

98–03. Peter Benner, Maribel Castillo, Enrique S. Quintana-Ortí:
Parallel Partial Stabilizing Algorithms for Large Linear Control Systems, Juli 1998.

98–04. Peter Benner:
Computational Methods for Linear–Quadratic Optimization, August 1998.

98–05. Peter Benner, Ralph Byers, Enrique S. Quintana-Ortí, Gregorio Quintana-Ortí:
Solving Algebraic Riccati Equations on Parallel Computers Using Newton's Method with Exact Line Search, August 1998.

98–06. Lars Grüne, Fabian Wirth:
On the rate of convergence of infinite horizon discounted optimal value functions, November 1998.

98–07. Peter Benner, Volker Mehrmann, Hongguo Xu:
A Note on the Numerical Solution of Complex Hamiltonian and Skew-Hamiltonian Eigenvalue Problems, November 1998.

98–08. Eberhard Bänsch, Burkhard Höhn:
Numerical simulation of a silicon floating zone with a free capillary surface, Dezember 1998.

99–01. Heike Faßbender:
The Parameterized SR Algorithm for Symplectic (Butterfly) Matrices, Februar 1999.

99–02. Heike Faßbender:
Error Analysis of the symplectic Lanczos Method for the symplectic Eigenvalue Problem, März 1999.

99–03. Eberhard Bänsch, Alfred Schmidt:
Simulation of dendritic crystal growth with thermal convection, März 1999.

99–04. Eberhard Bänsch:
Finite element discretization of the Navier-Stokes equations with a free capillary surface, März 1999.

99–05. Peter Benner:
Mathematik in der Berufspraxis, Juli 1999.

99–06. Andrew D.B. Paice, Fabian R. Wirth:
Robustness of nonlinear systems and their domains of attraction, August 1999.

99–07. Peter Benner, Enrique S. Quintana-Ortí, Gregorio Quintana-Ortí:
Balanced Truncation Model Reduction of Large-Scale Dense Systems on Parallel Computers, September 1999.

99–08. Ronald Stöver:
Collocation methods for solving linear differential-algebraic boundary value problems, September 1999.

99–09. Huseyin Akcay:
Modelling with Orthonormal Basis Functions, September 1999.

99–10. Heike Faßbender, D. Steven Mackey, Niloufer Mackey:
Hamilton and Jacobi come full circle: Jacobi algorithms for structured Hamiltonian eigenproblems, Oktober 1999.

99–11. Peter Benner, Vincente Hernández, Antonio Pastor:
On the Kleinman Iteration for Nonstabilizable System, Oktober 1999.

99–12. Peter Benner, Heike Faßbender:
A Hybrid Method for the Numerical Solution of Discrete-Time Algebraic Riccati Equations, November 1999.

99–13. Peter Benner, Enrique S. Quintana-Ortí, Gregorio Quintana-Ortí:
Numerical Solution of Schur Stable Linear Matrix Equations on Multicomputers, November 1999.

99–14. Eberhard Bänsch, Karol Mikula:
Adaptivity in 3D Image Processing, Dezember 1999.

00–01. Peter Benner, Volker Mehrmann, Hongguo Xu:
Perturbation Analysis for the Eigenvalue Problem of a Formal Product of Matrices, Januar 2000.

00–02. Ziping Huang:
Finite Element Method for Mixed Problems with Penalty, Januar 2000.

00–03. Gianfrancesco Martinico:
Recursive mesh refinement in 3D, Februar 2000.

00–04. Eberhard Bänsch, Christoph Egbers, Oliver Meincke, Nicoleta Scurtu:
Taylor-Couette System with Asymmetric Boundary Conditions, Februar 2000.

00–05. Peter Benner:
Symplectic Balancing of Hamiltonian Matrices, Februar 2000.

00–06. Fabio Camilli, Lars Grüne, Fabian Wirth:
A regularization of Zubov's equation for robust domains of attraction, März 2000.

00–07. Michael Wolff, Eberhard Bänsch, Michael Böhm, Dominic Davis:
Modellierung der Abkühlung von Stahlbrammen, März 2000.

00–08. Stephan Dahlke, Peter Maaß, Gerd Teschke:
Interpolating Scaling Functions with Duals, April 2000.

00–09. Jochen Behrens, Fabian Wirth:
A globalization procedure for locally stabilizing controllers, Mai 2000.

00–10. Peter Maaß, Gerd Teschke, Werner Willmann, Günter Wollmann:
Detection and Classification of Material Attributes – A Practical Application of Wavelet Analysis, Mai 2000.

00–11. Stefan Boschert, Alfred Schmidt, Kunibert G. Siebert, Eberhard Bänsch, Klaus-Werner Benz, Gerhard Dziuk, Thomas Kaiser:
Simulation of Industrial Crystal Growth by the Vertical Bridgman Method, Mai 2000.

00–12. Volker Lehmann, Gerd Teschke:
Wavelet Based Methods for Improved Wind Profiler Signal Processing, Mai 2000.

00–13. Stephan Dahlke, Peter Maass:
A Note on Interpolating Scaling Functions, August 2000.

00–14. Ronny Ramlau, Rolf Clackdoyle, Frédéric Noo, Girish Bal:
Accurate Attenuation Correction in SPECT Imaging using Optimization of Bilinear Functions and Assuming an Unknown Spatially-Varying Attenuation Distribution, September 2000.

00–15. Peter Kunkel, Ronald Stöver:
Symmetric collocation methods for linear differential-algebraic boundary value problems, September 2000.

00–16. Fabian Wirth:
The generalized spectral radius and extremal norms, Oktober 2000.

00–17. Frank Stenger, Ahmad Reza Naghsh-Nilchi, Jenny Niebsch, Ronny Ramlau:
A unified approach to the approximate solution of PDE, November 2000.

00–18. Peter Benner, Enrique S. Quintana-Ortí, Gregorio Quintana-Ortí:
Parallel algorithms for model reduction of discrete–time systems, Dezember 2000.

00–19. Ronny Ramlau:
A steepest descent algorithm for the global minimization of Tikhonov–Phillips functional, Dezember 2000.

01–01. Efficient methods in hyperthermia treatment planning:
Torsten Köhler, Peter Maass, Peter Wust, Martin Seebass, Januar 2001.

01–02. Parallel Algorithms for LQ Optimal Control of Discrete-Time Periodic Linear Systems:
Peter Benner, Ralph Byers, Rafael Mayo, Enrique S. Quintana-Ortí, Vicente Hernández, Februar 2001.

01–03. Peter Benner, Enrique S. Quintana-Ortí, Gregorio Quintana-Ortí:
Efficient Numerical Algorithms for Balanced Stochastic Truncation, März 2001.

01–04. Peter Benner, Maribel Castillo, Enrique S. Quintana-Ortí:
Partial Stabilization of Large-Scale Discrete-Time Linear Control Systems, März 2001.

01–05. Stephan Dahlke:
Besov Regularity for Edge Singularities in Polyhedral Domains, Mai 2001.

01–06. Fabian Wirth:
A linearization principle for robustness with respect to time-varying perturbations, Mai 2001.

01–07. Stephan Dahlke, Wolfgang Dahmen, Karsten Urban:
Adaptive Wavelet Methods for Saddle Point Problems - Optimal Convergence Rates, Juli 2001.