Convergence of Adaptive Wavelet Methods for Goal-Oriented Error Estimation

Wolfgang Dahmen, Angela Kunoth, Jürgen Vorloeper

no. 286

Diese Arbeit erscheint in:
ENUMATH 2005 Proceedings

Sie ist mit Unterstützung des von der Deutschen Forschungsgemeinschaft getragenen Sonderforschungsbereiches 611 an der Universität Bonn entstanden und als Manuskript vervielfältigt worden.

Bonn, September 2006

Convergence of Adaptive Wavelet Methods for Goal–Oriented Error Estimation *

Wolfgang Dahmen[1], Angela Kunoth[2], and Jürgen Vorloeper[1]

[1] Institut für Geometrie und Praktische Mathematik
RWTH Aachen, Templergraben 55, 52056 Aachen, Germany
{dahmen,jvor}@igpm.rwth-aachen.de
[2] Institut für Angewandte Mathematik und Institut für Numerische Simulation
Universität Bonn, Wegelerstr. 6, 53115 Bonn, Germany
kunoth@iam.uni-bonn.de

Summary. We investigate adaptive wavelet methods which are *goal–oriented* in the sense that a *functional* of the solution of a linear elliptic PDE is computed up to arbitrary accuracy at possibly low computational cost measured in terms of degrees of freedom. In particular, we propose a scheme that can be shown to exhibit convergence to the target value without insisting on energy norm convergence of the primal solution. The theoretical findings are complemented by first numerical experiments.

1 Introduction

The importance of *adaptive* solution concepts for large scale computational tasks arising in Numerical Simulation based on PDEs or integral equations is nowadays well accepted. The evidence provided by numerical experience is, however, nor quite in par with the theoretical foundation of such schemes. A thorough analytical understanding, in turn, has recently proven to lead to new algorithmic paradigms in connection with wavelet based schemes. Rigorous complexity and convergence estimates were obtained for adaptive wavelet methods for a wide class of linear and nonlinear variational problems, see, e.g., [8, 9, 12, 14]. These estimates relate for the first time the computational work and the adaptively generated number of degrees of freedom to the *target accuracy* of the approximate solution. This accuracy refers to the approximation in some (energy) norm, i.e., the whole unknown solution is recovered. These

* This research was supported in part by the EEC Human Potential Programme under contract HPRN-CT-2002-00286, "Breaking Complexity", the SFB 401, "Flow Modulation and Fluid-Structure Interaction at Airplane Wings", and SFB 611, "Singular Phenomena and Scaling in Mathematical Models", funded by the German Research Foundation.

developments have meanwhile spilled over to the Finite Element setting where analogous results could be obtained for a much more restricted problem class, though, see, e.g., [3, 24].

However, in many applications one is only interested in some *functional* of the solution which, in particular, might be local such as point values or integrals on some lower dimensional manifold. In such a case one might expect to obtain the desired information at a much lower expense than computing the whole solution. This is exactly the objective of *goal-oriented* error estimation which gives rise to the so called *dual weighted residual method* (DWR), see, e.g., [4] and the references cited therein.

Many striking examples indicate that one may indeed reach the goal with the aid of this paradigm at the expense of much less computational work in comparison with schemes driven by *norm approximation*. On the other hand, a rigorous analysis of the DWR faces a number of severe obstructions related, in particular, to the fact that the central error representation involves the (unknown) solution to the dual problem. Thus, the dual solution has to be estimated along the way. Although this problem arises, in principle, already when dealing with linear problems, it becomes more delicate in the nonlinear case since the dual solution depends then on the primal one. It is fair to say that the mutual intertwinement of the accuracies of dual and primal solutions, especially with regard to the spatial distribution of degrees of freedom, is far from a rigorous understanding. It is not even clear in the linear case that adaptive refinements based on the practiced versions of the DWR paradigm actually converge in the sense that the searched value is actually approached better and better by the computed one as the refinement goes on. It is this issue that will be the primary concern of this paper.

To appreciate this issue, it is helpful to keep a few principal facts in mind. Approximability of a function in some norm can always be understood in terms of the *regularity* of that function (with respect to some nonclassical regularity measure). In a typical application of the DWR, adaptivity is not driven by the regularity of the searched for object, but primarily by the *locality* of the targeted information, conveyed by the dual solution which is often termed *generalized Green's function*, see, e.g., [19]. This generalized Green's function indicates the influence of parts of the primal solution away from the spatial location of the target functional. Thus, the experience gained with adaptive wavelet schemes for energy norm approximation is not immediately seen to be helpful in the context of the DWR.

Nevertheless, the primary goal of this paper is to contribute to the understanding of the DWR by looking at this paradigm from a wavelet point of view. Here is a rough indication why this might indeed be a promising perspective: The key to the above mentioned results from [8, 9] is to formulate an iteration (e.g., a gradient or a Newton scheme) for the full infinite dimensional problem formulated in wavelet coordinates. This idealized iteration is then mimicked by the adaptive evaluation of the involved operators within any desired error

tolerance. Staying in that sense controllably close to the infinite dimensional problem may therefore be expected to help also in the context of the DWR.

In this note we wish to explore this aspect for an admittedly simple class of model problems, namely, linear elliptic boundary value problems. Moreover, we shall consider only linear evaluation functionals that belong to the dual of the energy space. Further linearization and/or regularization can be, of course, performed as explained in many foregoing investigations. The main point is to identify the key mechanisms so as to draw also conclusions for more complex problems.

We shall occasionally use the following convention for estimates containing generic constants. The relation $a \sim b$ always stands for $a \lesssim b$ and $a \gtrsim b$, i.e., a can be estimated from above and below by a constant multiple of b independent of all parameters on which a or b may depend.

2 Goal–Oriented Error Estimation

2.1 Problem Formulation

Let V denote a Hilbert space living on some bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$ and let V' be its topological dual. Its associated dual form will be denoted as $\langle \cdot, \cdot \rangle_{V \times V'}$, or shortly as $\langle \cdot, \cdot \rangle$.

Moreover, let $a(\cdot, \cdot)$ be a symmetric bilinear form which will here always supposed to be continuous and elliptic on V, i.e., there exist constants c_A, C_A such that

$$\sqrt{c_A}\|v\|_V \leq a(v,v)^{1/2} \leq \sqrt{C_A}\|v\|_V, \qquad v \in V. \tag{2.1}$$

In this case the variational problem: given any $f \in V'$, find $u \in V$ such that

$$a(v,u) = \langle v, f \rangle, \qquad v \in V, \tag{2.2}$$

is well posed. It will be convenient to introduce the induced operator $A : V \to V'$ given by $\langle v, Aw \rangle := a(v,w)$ for all $v, w \in V$.

Instead of approximating the whole solution u we are interested in evaluating only a *functional* of the unknown solution. Specifically, we consider the following problem: Given a fixed linear functional $J \in V'$, compute

$$J(u) := \langle u, J \rangle, \tag{2.3}$$

where u is the solution of (2.2). $J(u)$ may be a very local quantity, such as the point evaluation of u at some point $x_* \in \Omega$, if the Dirac functional is in V' (as in the case of Plateau's equation on an interval), or a local quantity like the mean of u over some small domain $\Omega_\delta \subset \Omega$, i.e., $J(u) = |\Omega_\delta|^{-1} \int_{\Omega_\delta} u(x)\, dx$, or a weighted integral of u over some lower dimensional manifold in Ω. We shall exclude first more general situations such as *nonlinear* functionals J which would require an additional linearization process as shown in [4], as well as

functionals that are not contained in V' but require additional regularity of the solution.

Of course, one might approximate the quantity $J(u)$ by determining first some approximation u_Λ to u sitting in some finite dimensional trial space indicated by the subscript Λ, and take then $J(u_\Lambda)$ as an approximation to the desired value $J(u)$. Moreover, in the above framework it is natural to take u_Λ as a *Galerkin solution* with respect to some subspace $V_\Lambda \subset V$, i.e.,

$$a(v, u_\Lambda) = \langle v, f\rangle, \qquad v \in V_\Lambda. \tag{2.4}$$

Under the circumstances (2.1), (2.2), u_Λ is uniquely determined for any $V_\Lambda \subset V$. (For conceptual reasons that will become clear later, we deliberately do not even insist at this point on V_Λ being finite dimensional.) We shall frequently use the shorthand notation

$$e_\Lambda := u - u_\Lambda.$$

Our goal now is to determine u_Λ such that for a given *target accuracy* $\varepsilon > 0$

$$|J(u) - J(u_\Lambda)| = |J(u - u_\Lambda)| = |J(e_\Lambda)| \le \varepsilon, \tag{2.5}$$

while the computational cost needed to determine u_Λ is to be kept as low as possible. Since, by assumption, $J \in V'$, we have

$$|J(e_\Lambda)| \le \|J\|_{V\to\mathbb{R}}\, \|e_\Lambda\|_V, \tag{2.6}$$

where, as usual, $\|J\|_{V\to\mathbb{R}} := \sup_{v\in V, \|v\|_V\le 1}\langle v, J\rangle$.

Remark 1. When $J \notin V'$ but $J \in (V^+)'$ where $V^+ \hookrightarrow V$ and $u, u_\Lambda \in V^+$, we obtain an analogous estimate of the form $|J(e_\Lambda)| \le \|J\|_{V^+\to\mathbb{R}}\, \|e_\Lambda\|_{V^+}$.

Staying with the simpler former situation, a principal gain is that the target accuracy ε can be achieved by solving two problems, namely, the primal (2.2) and the dual one (2.8) with accuracies of the order $\sqrt{\varepsilon}$. Thus, choosing some subspace V_Λ, based on some a-priori estimates, such that the Galerkin error satisfies

$$\|u - u_\Lambda\|_V < \varepsilon/\|J\|_{V\to\mathbb{R}}, \tag{2.7}$$

this, together with (2.6), would yield (2.5). In general, such an a-priori choice would require a too large V_Λ. In any case, an adaptive choice of V_Λ with respect to the energy norm may lead to an overestimation since such a norm approximation does not take the locality of J into account.

2.2 The Dual Weighted Residual Method: Error Representation

It is the very purpose of the *dual weighted residual method* (DWR) to take the locality of J into account when refining a given discretization so as to improve on the accuracy of the approximate value, possibly without approximating

the whole solution everywhere in the domain with a comparable accuracy. In order to motivate the subsequent development, we briefly review some basic facts concerning this methodology from [4, 19]. The key is to obtain an *error representation* comprised of local quantities that reflect residual terms which *can* be evaluated. The derivation of such representations relies on *duality arguments* to be explained next.

Let $z \in V$ be the solution of the *dual problem*

$$a(z, w) = \langle w, J \rangle, \qquad w \in V, \tag{2.8}$$

with $J \in V'$ serving as right hand side. Inserting $w = u - u_\Lambda = e_\Lambda$ yields the error representation

$$J(e_\Lambda) = \langle e_\Lambda, J \rangle = a(z, e_\Lambda) = a(z - y_\Lambda, e_\Lambda), \qquad \text{for any } y_\Lambda \in V_\Lambda, \tag{2.9}$$

where we have used Galerkin orthogonality in the last step. This suggests several options for bounding these residuals. First, we obtain the estimate

$$|J(u - u_\Lambda)| = |a(z - y_\Lambda, u - u_\Lambda)| \lesssim \|u - u_\Lambda\|_V \inf_{y_\Lambda \in V_\Lambda} \|z - y_\Lambda\|_V. \tag{2.10}$$

Thus, if the computational work (measured in terms of problem size expressed as the number of degrees of freedom N) needed to compute such approximations for the primal and dual solution with accuracy ε scales like $N(\varepsilon) = \varepsilon^{-\alpha}$ for some $\alpha > 0$, the error in (2.10) can be bounded by ε^2. So the computational work needed to determine the value $J(u)$ within a tolerance ε scales like $2\varepsilon^{-\alpha/2}$. This is asymptotically better than just computing the primal solution with tolerance ε in the energy norm (2.7).

This still does not exploit the locality of the functional J of interest. In the framework of Finite Element discretizations, one usually treats this latter objective by bounding the error representation $a(z - y_\Lambda, u - u_\Lambda)$ by a sum of local computable quantities. To specify this, let Λ denote then a current triangulation of the domain Ω. Such estimates have then the form

$$|a(z - y_\Lambda, u - u_\Lambda)| \lesssim \sum_{T \in \Lambda} w_T(y_\Lambda)\, r_T(u_\Lambda), \tag{2.11}$$

where the $r_T(u_\Lambda)$ are *local residuals* of the approximate solution u_Λ and the $w_T(y_\Lambda)$ are *weights* computed in terms of the dual solution. For the simple case $a(v, w) = \int_\Omega (\nabla y)^T \nabla w dx$, they look like

$$r_T(u_\Lambda) = \|f + \Delta u_\Lambda\|_{L_2(T)} + \frac{1}{2} h_T^{-1/2} \left\| \left[\frac{\partial u_\Lambda}{\partial n} \right] \right\|_{L_2(\partial T)}. \tag{2.12}$$

The weights or *stability factors* are of the form

$$w_T(y_\Lambda) = \|z - y_\Lambda\|_{L_2(T)} + h_T^{1/2} \|z - y_\Lambda\|_{L_2(\partial T)}, \tag{2.13}$$

see, e.g., [4, 19].

Note that, while the $r_T(u_\Lambda)$ are computable, the weights $w_T(y_\Lambda)$ depend on the unknown dual solution z. One can argue that, in practical applications it suffices to know only the "trend" of these weights to see the influence of the local residual $r_T(u_\Lambda)$ and, consequently, of the local error caused by u_Λ. There are several ways of obtaining approximations to these weights:

(i) One can compute an approximate solution $\bar{z}$ of z on some finer mesh than the one used for the primal solution and substitute $\bar{z}$ for z.
(ii) One can compute a *higher order Galerkin approximation* as a substitute for z in (2.13).
(iii) Instead of computing the difference $z - y_\Lambda$, one determines a higher order Galerkin approximation $\bar{z}$ to z, computes its second order derivatives and replaces $w_T(y_\Lambda)$ by some constant multiple of $h_T^2 \|\bar{z}\|_{H^2(T)}$.
(iv) A lower order Galerkin approximation is postprocessed to provide second order approximations that can then be used as in (iii).

In simple cases, all these strategies are expected to work fine. Nevertheless, even in the simple linear model case, none of them give rigorous bounds for the actual error resulting from any refinement strategy and from corresponding decisions on how accurately the dual solution needs to be approximated. The amount of confidence one can put in either of them may vary considerably: Neither is it clear that any fixed mesh refinement or a higher order approximation is sufficiently closer to the true solution to provide a reliable trend (in particular, near singularities), nor is it clear that the second order derivatives behave as those of the true dual solution (again, especially, when singularities interfere).

Thus, already at a rather basic level, one faces the essential question as to how accurately should the dual solution be computed and how localized the distribution of degrees of freedom can be chosen without loosing essential information.

The subsequent discussion attempts to shed some further light on these issues exploiting some concepts that have been developed in connection with adaptive wavelet schemes, see, e.g., [7, 8, 9].

2.3 Wavelet Coordinates

Let $\Psi := \{\psi_\lambda : \lambda \in \mathcal{I}\} \subset V$ be a wavelet basis for V. By this we mean that every $v \in V$ has a unique expansion $v = \sum_{\lambda \in \mathcal{I}} v_\lambda \psi_\lambda$ with coefficient array $\mathbf{v} = (v_\lambda)_{\lambda \in \mathcal{I}}$ such that for fixed constants c_Ψ, C_Ψ one has

$$c_\Psi \|\mathbf{v}\| \le \|v\|_V \le C_\Psi \|\mathbf{v}\|, \tag{2.14}$$

where $\|\mathbf{v}\|^2 := \sum_{\lambda \in \mathcal{I}} |v_\lambda|^2 = \mathbf{v}^T\mathbf{v}$ denotes the ℓ_2-norm. Only when the ℓ_2-norm with respect to a specific subset $\Lambda \subset \mathcal{I}$ is meant we write for clarity $\|\mathbf{v}\|^2_{\ell_2(\Lambda)} := \sum_{\lambda \in \Lambda} |v_\lambda|^2$. Recall that, by a simple duality argument (see, e.g., [13]), one has

$$C_\Psi^{-1}\|\langle\psi_\lambda, w\rangle\| \le \|w\|_{V'} \le c_\Psi^{-1}\|\langle\psi_\lambda, w\rangle\|, \quad w \in V'. \tag{2.15}$$

For typical constructions of wavelet bases that are suitable, e.g., for $V = H_0^1(\Omega)$, we refer to [5, 6, 15, 16, 11, 17]. Here it suffices to add a few remarks on the structure of the index set $\mathcal{I}$. Each index λ comprises information on the scale, denoted by $|\lambda|$, and on the spatial location of the associated basis function $k(\lambda)$. There is usually a finite number of "scaling function type" basis functions on some coarsest level of resolution j_0. This subset will be denoted by $\mathcal{I}_\phi$. All remaining indices refer to "true" wavelets gathered in $\mathcal{I}_\psi$. These wavelets are always of compact support whose diameter scale like $2^{-|\lambda|}$. Moreover, these true wavelets have cancellation properties of some specified order $\tilde{m}$ usually derived from a corresponding order of vanishing moments $\langle\psi_\lambda, P\rangle = 0$ for all $\lambda \in \mathcal{I}_\psi$ and any polynomial P of total order at most $\tilde{m}$. Furthermore, it follows from (2.14) that the wavelets are normalized such that $\|\psi_\lambda\|_V \sim 1$.

Testing (2.2) by $v = \psi_\lambda$, $\lambda \in \mathcal{I}$, we obtain an equivalent formulation in wavelet coordinates

$$\mathbf{A}\mathbf{u} = \boldsymbol{f}, \tag{2.16}$$

where

$$\mathbf{A} = \big(a(\psi_\lambda, \psi_\nu)\big)_{\lambda,\nu\in\mathcal{I}} \tag{2.17}$$

is the wavelet representation of the operator $A : V \to V'$ induced by $a(v, w) = \langle v, Aw\rangle$ for all $v, w \in V$. Likewise the dual problem (2.8) is equivalent to

$$\mathbf{A}^T\mathbf{z} = \mathbf{J}, \tag{2.18}$$

where $\mathbf{J} := \big(\langle\psi_\lambda, E\rangle\big)_{\lambda\in\mathcal{I}}$. Combining (2.14), (2.15) with (2.1) yields

$$c_\Psi^2 c_A\|\mathbf{v}\| \le \|\mathbf{A}\mathbf{v}\| \le C_\Psi^2 C_A\|\mathbf{v}\|, \quad \mathbf{v} \in \ell_2, \tag{2.19}$$

i.e., the wavelet representation is well conditioned in the Euclidean metric ℓ_2, see e.g. [9].

For any subset $\Lambda \subset \mathcal{I}$ we let $\Psi_\Lambda := \{\psi_\lambda : \lambda \in \mathcal{I}\} \subset V$ be the corresponding subset of wavelets and denote by V_Λ the closure in V of the linear span of Ψ_Λ. We continue denoting by u_Λ the Galerkin solution, now with respect to the subspace V_Λ, and by $\mathbf{u}_\Lambda$ the corresponding array of wavelet coefficients supported in Λ.

Note that for any $w = \sum_{\lambda\in\mathcal{I}} w_\lambda\psi_\lambda =: \mathbf{w}^T\Psi$

$$J(w) = \sum_{\lambda\in\mathcal{I}} w_\lambda J(\psi_\lambda) = \mathbf{J}^T\mathbf{w}. \tag{2.20}$$

Thus, abbreviating $\mathbf{e}_\Lambda := \mathbf{u} - \mathbf{u}_\Lambda$, $e_\Lambda := (\mathbf{u} - \mathbf{u}_\Lambda)^T\Psi$, the representation (2.9) then takes on the form

$$J(u) - J(u_\Lambda) = \mathbf{J}^T\mathbf{e}_\Lambda = (\mathbf{z} - \mathbf{y}_\Lambda)^T(\boldsymbol{f} - \mathbf{A}\mathbf{u}_\Lambda) = (\mathbf{A}^T(\mathbf{z} - \mathbf{y}_\Lambda))^T(\mathbf{u} - \mathbf{u}_\Lambda), \tag{2.21}$$

where $\mathbf{y}_\Lambda$ is any vector supported in Λ and the primal residual is given by

$$\mathbf{r}_\Lambda(\mathbf{u}) := \boldsymbol{f} - \mathbf{A}\mathbf{u}_\Lambda = \mathbf{A}\mathbf{e}_\Lambda. \tag{2.22}$$

It is important to note here that (2.22) is the *true* residual for the infinite dimensional operator $\mathbf{A}$.

We shall frequently exploit that, by definition, one has

$$\mathbf{r}_\Lambda(\mathbf{u})|_\Lambda = \mathbf{0}. \tag{2.23}$$

Moreover, it immediately follows from (2.19) that

$$c_A c_\Psi^2 \|\mathbf{u} - \mathbf{u}_\Lambda\| \le \|\mathbf{r}_\Lambda(\mathbf{u})\| \le C_A C_\Psi^2 \|\mathbf{u} - \mathbf{u}_\Lambda\|. \tag{2.24}$$

Hence, approximations in V and V' on the function side reduce to approximation in ℓ_2 for the primal and dual wavelet coefficient arrays.

Of course, the problem that the representation (2.21) involves the unknown dual solution remains the same as in conventional discretization settings. However, while the terms in (2.11) reflect primarily spatial localization, the summands in (2.21) convey spatial and frequency information in terms of (dual) wavelet coefficients (of the residual) and of the error. We shall explore next whether this can be exploited for a reliable error estimation.

3 Adaptive Error Estimation

Our objective is to develop a-posteriori refinement strategies that aim at computing $J(u)$ within some error tolerance at possibly low computational cost. This amounts to a DWR method in wavelet coordinates. (2.20) suggests to take (the computable quantity)

$$J(u_\Lambda) = \mathbf{J}(\mathbf{u}_\Lambda) = \sum_{\lambda\in\Lambda} \mathbf{J}^T\mathbf{u}_\Lambda \tag{3.25}$$

as an approximate value of the target functional, where Λ is a suitable finite index set. Concerning the incurred error, since, by (2.23), one has $\mathbf{r}_\Lambda(\mathbf{u})|_\Lambda = 0$, we infer from (2.21)

$$\mathbf{J}^T\mathbf{e}_\Lambda = \sum_{\lambda\in\mathcal{I}\setminus\Lambda} z_\lambda(\mathbf{r}_\Lambda(\mathbf{u}))_\lambda. \tag{3.26}$$

As a natural heuristics this suggests an analog to option (i) in the Finite Element context, namely, to select some larger index set $\hat{\Lambda} \supset \Lambda$ and replace $\mathbf{z}$ in (3.26) by the Galerkin solution $\mathbf{z}_{\hat{\Lambda}}$ in $V_{\hat{\Lambda}}$. But again the question remains, how large has $\hat{\Lambda}$ to be chosen in order to provide a reliable estimate. The following simple observations suggest how to deal with this question. By (2.21) we have

$$|\mathbf{J}^T(\mathbf{u}-\mathbf{u}_\Lambda)| \le \Big|\sum_{\lambda\in\Lambda_\delta\setminus\Lambda} z_{\hat\Lambda,\lambda}\, r_{\Lambda,\lambda}(\mathbf{u})\Big| + \sum_{\lambda\in\mathbb{I}\setminus\Lambda} |(z_\lambda - z_{\hat\Lambda,\lambda})\, r_{\Lambda,\lambda}(\mathbf{u})|. \quad (3.27)$$

The first part is a finite sum that is computable through the primal residual on a finite set and the computed $\mathbf{z}_{\hat\Lambda}$. The second part can be estimated as follows:

$$|\mathbf{J}^T(\mathbf{u}-\mathbf{u}_\Lambda)| \le \Big|\sum_{\lambda\in\hat\Lambda\setminus\Lambda} z_{\hat\Lambda,\lambda}\, r_{\Lambda,\lambda}(\mathbf{u})\Big| + \inf_{1\le p,p'\le\infty \frac{1}{p}+\frac{1}{p'}=1} \|\mathbf{z}-\mathbf{z}_{\hat\Lambda}\|_{\ell_p}\|\mathbf{r}_\Lambda(\mathbf{u})\|_{\ell_{p'}}. \quad (3.28)$$

Specifically, $p = p' = 1/2$ yields

$$|\mathbf{J}^T(\mathbf{u}-\mathbf{u}_\Lambda)| \le \Big|\sum_{\lambda\in\hat\Lambda\setminus\Lambda} z_{\hat\Lambda,\lambda}\, r_{\Lambda,\lambda}(\mathbf{u})\Big| + \|\mathbf{z}-\mathbf{z}_{\hat\Lambda}\|\,\|\mathbf{r}_\Lambda(\mathbf{u})\|. \quad (3.29)$$

Thus, due to the norm equivalences (2.24), (2.15), (2.14) the second term on the right hand side is the product of the primal and dual energy norm error. Thus, whenever the dual solution is approximated in the energy norm and the growth of Λ depends on the energy norm approximation of $\mathbf{z}$ the target value is approximated with increasing accuracy even though the global primal residual does not tend to zero at all in ℓ_2. It may tend to zero in some weaker norm which, according to (3.28), could give a better estimate.

Led by the above considerations, we formulate now in precise terms an algorithm which, for any given target accuracy ε, computes $J(u_\Lambda) = \mathbf{J}^T(\mathbf{u}_\Lambda)$ such that $|J(e_\Lambda)| = |\mathbf{J}^T(\mathbf{e}_\Lambda)| \le \varepsilon$. A central ingredient is the adaptive wavelet scheme from [9] that will be formulated next. The resulting well-posedness in ℓ_2 (2.19) allows one to contrive an (idealized) iteration

$$\mathbf{u}^{n+1} = \mathbf{u}^n - \mathbf{B}(\mathbf{A}\mathbf{u}^n - \boldsymbol{f}), \quad n = 0,1,2,\dots, \quad (3.30)$$

where $\mathbf{B}$ is (a possibly stage dependent) preconditioner, such that for some $\rho < 1$

$$\|\mathbf{u}-\mathbf{u}^{n+1}\| \le \rho\|\mathbf{u}-\mathbf{u}^n\|, \quad n \in \mathbb{N}_0, \quad (3.31)$$

see [8, 9] for various examples covering also noncoercive problems.

The idea is now to mimic (3.30) numerically by evaluating the weighted residual $\mathbf{B}(\mathbf{A}\mathbf{u}^n - \boldsymbol{f})$ within a stage dependent dynamical accuracy tolerance. This, in turn, hinges on the *adaptive* evaluation of the involved (at this stage still infinite dimensional) operators when applied to a finitely supported array. We refer to [9, 10, 2] for the precise description of such evaluation schemes for a range of (linear and nonlinear) operators. Therefore we may assume at this point to have a routine of the following form at hand:

RES$[\eta, \mathbf{B}, \mathbf{A}, \boldsymbol{f}, \mathbf{v}] \to \mathbf{r}_\eta$ COMPUTES FOR ANY FINITELY SUPPORTED INPUT $\mathbf{v}$ AND ANY POSITIVE TOLERANCE η AN APPROXIMATE FINITELY SUPPORTED RESIDUAL $\mathbf{r}_\eta$ SUCH THAT

$$\|\mathbf{B}(\mathbf{A}\mathbf{v} - \boldsymbol{f}) - \mathbf{r}_\eta\| \le \eta. \quad (3.32)$$

We further need the routine
COARSE$[\eta, \mathbf{v}] \to \mathbf{w}_\eta$ DETERMINES FOR ANY FINITELY SUPPORTED INPUT $\mathbf{v}$ AN OUTPUT $\mathbf{w}_\eta$ WITH POSSIBLY SMALL SUPPORT SUCH THAT STILL

$$\|\mathbf{v} - \mathbf{w}_\eta\| \le \eta. \tag{3.33}$$

Following [9] the announced adaptive solution scheme can now be described as follows.
SOLVE $[\varepsilon, \mathbf{A}, \boldsymbol{f}, \overline{\mathbf{u}}^0] \to (\overline{\mathbf{u}}_\varepsilon, \Lambda_\varepsilon)$ COMPUTES FOR ANY GIVEN TARGET ACCURACY $\varepsilon > 0$ AND ANY INITIAL GUESS $\overline{\mathbf{u}}^0$, SATISFYING $\|\mathbf{u} - \overline{\mathbf{u}}^0\| \le \delta$, AN APPROXIMATION $\overline{\mathbf{u}}_\varepsilon$ TO (2.2), SUPPORTED IN SOME FINITE (TREE LIKE) INDEX SET Λ_ε, SUCH THAT

$$\|\mathbf{u} - \overline{\mathbf{u}}_\varepsilon\| \le \varepsilon, \tag{3.34}$$

ACCORDING TO THE FOLLOWING STEPS:

(I) CHOOSE SOME $C^* > 1$, $\bar\rho \in (0,1)$. SET $\varepsilon_0 := \delta$ ACCORDING TO THE ABOVE INITIALIZATION, AND $j = 0$;
(II) IF $\varepsilon_j \le \varepsilon$ STOP AND OUTPUT $\overline{\mathbf{u}}_\varepsilon := \bar{\mathbf{u}}^j$; ELSE SET $\mathbf{v}^0 := \bar{\mathbf{u}}^j$ AND $k = 0$
(II.1) SET $\eta_k := \omega_k \bar\rho^k \varepsilon_j$ AND COMPUTE

$$\mathbf{r}^k = \mathbf{RES}\,[\eta_k, \mathbf{B}, \mathbf{A}, \boldsymbol{f}, \mathbf{v}^k], \qquad \mathbf{v}^{k+1} = \mathbf{v}^k - \mathbf{r}^k.$$

(II.2) IF

$$\beta\big(\eta_k + \|\mathbf{r}^k\|\big) \le \varepsilon_j/(2(1 + C^*)), \tag{3.35}$$

SET $\tilde{\mathbf{v}} := \mathbf{v}^k$ AND GO TO (III). ELSE SET $k+1 \to k$ AND GO TO (II.1).
(III) COARSE$[\frac{C^*\varepsilon_j}{2(1+C^*)}, \tilde{\mathbf{v}}] \to \bar{\mathbf{u}}^{j+1}$, $\varepsilon_{j+1} = \varepsilon_j/2$, $j+1 \to j$, GO TO (II).

Step (ii) is a block of perturbed iterations of the form (3.30). As soon as the approximate residual is small enough, the iteration is interrupted by a coarsening step. The constant β in step (ii.2) depends on the constants in (2.19). It can be shown that the number of perturbed iterations between two coarsening steps remains uniformly bounded. Things are arranged such that after an iteration block and a coarsening step the error in the energy norm is at least halved. Thus, under the above conditions the scheme SOLVE terminates always after finitely many steps. Moreover, its computational complexity is in some sense asymptotically optimal in that the number of adaptively generated degrees of freedom and the respective computational work grow at the rate of the best N-term approximation, see [9]. For more general problem classes, the coarsening step ensures optimal complexity rates. It has recently been shown in [20], however, that coarsening can be avoided for the current class of problems.

We shall use (variants of) this algorithm as ingredients in the present weighted dual residual scheme. The routine RES is based on the following ingredients. Suppose for simplicity that $\boldsymbol{f}$ is a finitely supported array, possibly as a result of a preprocessing step. In addition, one needs an approximate application of $\mathbf{A}$:

APPLY$[\eta, \mathbf{A}, \mathbf{v}] \to \mathbf{w}$ COMPUTES FOR ANY FINITELY SUPPORTED INPUT $\mathbf{v}$ AND ANY TOLERANCE $\eta > 0$ A FINITELY SUPPORTED OUTPUT $\mathbf{w}$ SUCH THAT

$$\|\mathbf{A}\mathbf{v} - \mathbf{w}\| \leq \eta. \tag{3.36}$$

Realizations of such a routine satisfying all requirements that render SOLVE having optimal complexity can be found in [1]. For the current type of elliptic problems we can, in principle, choose the preconditioner $\mathbf{B} = \alpha\mathbf{I}$ as a stage independent damped identity which gives rise to a Richardson iteration. In this case the residual approximation scheme takes the form

$$\text{RES}\,[\eta, \mathbf{A}, \boldsymbol{f}, \mathbf{v}] := \alpha\,(\text{APPLY}\,[\eta/2\alpha, \mathbf{A}, \mathbf{v}] - \text{COARSE}\,[\eta/2\alpha, \boldsymbol{f}])\,. \tag{3.37}$$

The quantitative performance of this choice is usually rather poor and we refer to [18] for more efficient versions that are actually used in our experiments here as well.

Since SOLVE produces energy norm approximants, a few preparatory comments on its use in the present context are in order. Let again $\Lambda \subset \mathcal{I}$ be any (possibly infinite) subset of $\mathcal{I}$. For any two such subsets Λ, Λ' let

$$\mathbf{A}_{\Lambda,\Lambda'} := \big(a(\psi_\lambda, \psi_\nu)\big)_{\lambda\in\Lambda,\nu\in\Lambda'}$$

be the section of $\mathbf{A}$ determined by Λ and Λ'. For simplicity we set $\mathbf{A}_\Lambda := \mathbf{A}_{\Lambda,\Lambda}$. Clearly, (2.4) is then equivalent to

$$\mathbf{A}_\Lambda \mathbf{u}_\Lambda = \boldsymbol{f}_\Lambda := \boldsymbol{f}|_\Lambda. \tag{3.38}$$

Of course, (2.19) remains valid when replacing ℓ_2 by $\ell_2(\Lambda)$ and $\mathbf{A}$ by $\mathbf{A}_\Lambda$ uniformly in Λ. Solving the original problem in V_Λ can therefore be done by running the scheme SOLVE while restricting all arrays to Λ. An adaptive application of the operator $\mathbf{A}$ in this constrained setting can be thought of for the moment as employing the usual (unconstrained) scheme to the constrained input and cutting the result back to Λ. (There may be even better ways taking the special circumstances into account but this satisfies all the properties needed in [9] to establish corresponding error and complexity estimates for the restricted case.) We identify this version of SOLVE by writing $\text{SOLVE}_\Lambda[\eta, \mathbf{A}, \boldsymbol{f}, \overline{\mathbf{u}}^0]$ (and accordingly $\text{RES}_\Lambda[\eta, \mathbf{A}, \boldsymbol{f}, \mathbf{v}]$). As before, the subscript Λ is omitted when $\Lambda = \mathcal{I}$. All arrays generated by this scheme are then by definition supported in Λ.

It will be important to distinguish between the residual $\alpha(\mathbf{A}_\Lambda\mathbf{v} - \boldsymbol{f}_\Lambda)$ in $\ell_2(\Lambda)$ which is approximated by $\text{RES}_\Lambda[\eta, \mathbf{A}, \boldsymbol{f}, \mathbf{v}]$ and the *full* residual $\mathbf{A}\mathbf{v} - \boldsymbol{f}$ which appears in (2.21). The latter one reflects the global deviation of $\mathbf{v}$ from the exact solution $\mathbf{u}$. In fact, for the *exact* solution $\mathbf{u}_\Lambda$ of the restricted problem (3.38) one has $\mathbf{A}\mathbf{u}_\Lambda = \mathbf{A}_{\mathcal{I},\Lambda}\mathbf{u}_\Lambda$ and therefore

$$\mathbf{r}_\Lambda(\mathbf{u}) = \mathbf{A}_{\mathcal{I},\Lambda}\mathbf{u}_\Lambda - \boldsymbol{f} = \begin{pmatrix} \mathbf{A}_\Lambda\mathbf{u}_\Lambda - \boldsymbol{f}_\Lambda \\ \mathbf{A}_{\mathcal{I}\setminus\Lambda,\Lambda}\mathbf{u}_\Lambda - \boldsymbol{f}_{\mathcal{I}\setminus\Lambda} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{A}_{\mathcal{I}\setminus\Lambda,\Lambda}\mathbf{u}_\Lambda - \boldsymbol{f}_{\mathcal{I}\setminus\Lambda} \end{pmatrix}, \tag{3.39}$$

reflecting the pollution caused by the restricted wavelet coordinate domain. A more careful analysis of this aspect will be given in a forthcoming paper. We have collected now the main ingredients for the following scheme:

ALGORITHM I$[\varepsilon, \mathbf{A}, \mathbf{J}, \boldsymbol{f}] \to \bar{J}$ COMPUTES FOR ANY TARGET ACCURACY $\varepsilon > 0$ A VALUE $\bar{J}$ SUCH THAT

$$|\bar{J} - J(u)| \le \varepsilon, \tag{3.40}$$

WHERE u IS THE SOLUTION TO (2.2), AS FOLLOWS:

(I) FIX PARAMETERS $c_u, c_z, c_r \in (0,1)$, $m_0 \ge 2$ AND SET $j = 0$, $\delta_u := c_A^{-1}\|\boldsymbol{f}\|$, $\delta_z := c_A^{-1}\|\mathbf{J}\|$ AND CHOOSE $\varepsilon_0 := \min\{\delta_u/2, \delta_z/2\}$.
APPLY SOLVE $[\varepsilon_0, \mathbf{A}, \boldsymbol{f}, \mathbf{0}] \to (\overline{\mathbf{u}}^0, \hat{\Lambda}_0)$;
APPLY SOLVE $[\varepsilon_0, \mathbf{A}^T, \mathbf{J}, \mathbf{0}] \to (\overline{\mathbf{z}}^0, \hat{\Upsilon}_0)$;
SET $\Lambda_0 := \hat{\Lambda}_0 \cup \hat{\Upsilon}_0$.

(II) APPLY SOLVE$[c_z\varepsilon_j, \mathbf{A}^T, \mathbf{J}, \overline{\mathbf{z}}^j] \to (\hat{\mathbf{z}}_j, \hat{\Lambda}_j)$;
APPLY SOLVE$_{\Lambda_j}[c_u\varepsilon_j, \mathbf{A}, \boldsymbol{f}, \overline{\mathbf{u}}^j] \to \overline{\mathbf{u}}_{\Lambda_j}$;
APPLY RES$[c_r\varepsilon_j, \mathbf{A}, \boldsymbol{f}, \overline{\mathbf{u}}_{\Lambda_j}]|_{\mathcal{I}\setminus\Lambda_j} \to \mathbf{r}$;
SET $\tilde{\mathbf{w}} := \hat{\mathbf{z}}_j|_{\hat{\Lambda}_j\setminus\Lambda_j}$ AND COMPUTE

$$e_j := \Big|\sum_{\lambda\in\hat{\Lambda}_j\setminus\Lambda_j} \tilde{w}_\lambda r_\lambda\Big|. \tag{3.41}$$

IF

$$e_j + \varepsilon_j\Big\{(C_A c_u + c_r)(\|\tilde{\mathbf{w}}|_{\hat{\Lambda}_j\setminus\Lambda_j}\| + c_z\varepsilon_j) + c_z\|\mathbf{r}\|\Big\} \le \varepsilon \tag{3.42}$$

STOP AND ACCEPT

$$\bar{J} = \mathbf{J}^T\overline{\mathbf{u}}_{\Lambda_j} := \sum_{\lambda\in\Lambda_j} \bar{u}_{\Lambda_j,\lambda} J_\lambda \tag{3.43}$$

AS TARGET VALUE.
OTHERWISE

(III) SET

$$\bar{\mathbf{u}}^{j+1} := \bar{\mathbf{u}}_{\Lambda_j},\ \bar{\mathbf{z}}^{j+1} := \hat{\mathbf{z}}_j, \quad \Lambda_{j+1} := \Lambda_j \cup \hat{\Lambda}_j, \quad \varepsilon_{j+1} = \varepsilon_j/m_0, \quad j+1 \to j, \tag{3.44}$$

AND GO TO (II).

A few comments on this scheme are in order. Step (i) should be viewed as an initialization where ε_0 is a crude initial tolerance whose square is typically still larger than the target accuracy ε. The initial approximate solutions for the primal and dual problem are energy norm approximations. Because of the crude target accuracy, one expects that the degrees of freedom generated in Λ_0 are necessary anyway.

Note that the approximations $\bar{\mathbf{u}}_{\Lambda_j}$ are then generated through the *restricted scheme* SOLVE_{Λ_j} while the corresponding residual approximations are unrestricted. Moreover, the application of SOLVE for the dual problem in step (ii) is unconstrained. We have explained the rationale of this step above. It essentially enforces the approximation of $\mathbf{z}$ in the norm but is expected to draw in only the relevant degrees of freedom concentrated near the support of J. It presumably requires only a few iterations with the initial guess $\overline{\mathbf{z}}_{\Lambda_j}$ which already is a good norm approximation for a somewhat larger tolerance.

In summary, in the above version the primal problem is *always* solved in a *constrained* subspace determined by the *norm approximation* of the dual solution.

Theorem 1. *For any target accuracy $\varepsilon > 0$ the above scheme terminates after a finite number of steps and outputs a result $\overline{J}$ satisfying $|J(u) - \overline{J}| \le \varepsilon$.*

Proof: First note that at the jth stage we have, according to (3.26),

$$\begin{aligned} J(e_{\Lambda_j}) = \mathbf{z}^T\mathbf{r}_{\Lambda_j}(\mathbf{u}) &= \sum_{\lambda\in\hat{\Lambda}_j\setminus\Lambda_j} \tilde{w}_\lambda r_\lambda + \sum_{\lambda\in\hat{\Lambda}_j\setminus\Lambda_j} \tilde{w}_\lambda(r_{\Lambda_j,\lambda}(\mathbf{u}) - r_\lambda) \\ &\quad + \sum_{\lambda\in I\!I\setminus\Lambda_j} (z_\lambda - \tilde{w}_\lambda) r_\lambda + \sum_{\lambda\in I\!I\setminus\Lambda_j} (z_\lambda - \tilde{w}_\lambda)(r_{\Lambda_j,\lambda}(\mathbf{u}) - r_\lambda) \\ &= \left(\tilde{\mathbf{w}}|_{\hat{\Lambda}_j\setminus\Lambda_j}\right)^T \mathbf{r} + \left(\tilde{\mathbf{w}}|_{\hat{\Lambda}_j\setminus\Lambda_j}\right)^T (\mathbf{r}_{\Lambda_j}(\mathbf{u}) - \mathbf{r}) \\ &\quad + \left((\mathbf{z}-\tilde{\mathbf{w}})|_{I\!I\setminus\Lambda_j}\right)^T \mathbf{r} + \left((\mathbf{z}-\tilde{\mathbf{w}})|_{I\!I\setminus\Lambda_j}\right)^T (\mathbf{r}_{\Lambda_j}(\mathbf{u}) - \mathbf{r}), \end{aligned}$$

so that

$$\begin{aligned} |J(e_{\Lambda_j})| \le e_j &+ \|\tilde{\mathbf{w}}|_{\hat{\Lambda}_j\setminus\Lambda_j}\|\,\|\mathbf{r}_{\Lambda_j}(\mathbf{u}) - \mathbf{r}\| + \|(\mathbf{z}-\tilde{\mathbf{w}})|_{I\!I\setminus\Lambda_j}\|\,\|\mathbf{r}\| \\ &+ \|(\mathbf{z}-\tilde{\mathbf{w}})|_{I\!I\setminus\Lambda_j}\|\,\|\mathbf{r} - \mathbf{r}_{\Lambda_j}(\mathbf{u})\|. \end{aligned} \tag{3.45}$$

We collect now several auxiliary estimates for the various terms in (3.45). By definition of $\tilde{\mathbf{w}}$ we have

$$\|(\mathbf{z}-\tilde{\mathbf{w}})|_{I\!I\setminus\Lambda_j}\| \le \|\mathbf{z}-\tilde{\mathbf{w}}\| \le c_z\varepsilon_j. \tag{3.46}$$

As for the exact residual of the exact Galerkin solution $\mathbf{u}_{\Lambda_j}$, we have, on account of (3.38), the very rough estimate

$$\|\mathbf{r}_\Lambda(\mathbf{u})\| \le \|\boldsymbol{f}\| + \|\mathbf{A}\mathbf{u}_\Lambda\| = \|\boldsymbol{f}\| + \|\mathbf{A}\mathbf{A}_\Lambda^{-1}\boldsymbol{f}_\Lambda\|. \tag{3.47}$$

Alternatively, because the exact Galerkin solution $\mathbf{u}_\Lambda$ is a best approximation to $\mathbf{u}$ from $\ell_2(\Lambda)$ in the norm $|||\mathbf{v}|||^2 := \mathbf{v}^T\mathbf{A}\mathbf{v}$, one could argue that

$$\|\mathbf{r}_\Lambda(\mathbf{u})\| \le C_A^{1/2}\|\mathbf{A}^{1/2}(\mathbf{u}-\mathbf{u}_\Lambda)\| \le C_A^{1/2}\|\mathbf{A}^{1/2}(\mathbf{u}-\bar{\mathbf{u}}^0)\| \le C_A\varepsilon_0, \tag{3.48}$$

which would allow us to use the initial norm approximation to $\mathbf{u}$ in step (I) of ALGORITHM I to influence the constant.

Moreover, the approximate residual $\mathbf{r}$ deviates from the exact one for the exact Galerkin solution $\mathbf{u}_{\Lambda_j}$ by

$$\begin{aligned}\|\mathbf{r}_{\Lambda_j}(\mathbf{u}) - \mathbf{r}\| &\le \|\mathbf{A}(\mathbf{u}_{\Lambda_j} - \overline{\mathbf{u}}_{\Lambda_j})\| + \|\mathbf{A}\overline{\mathbf{u}}_{\Lambda_j} - \boldsymbol{f} - \mathbf{r}\| \\ &\le \|\mathbf{A}(\mathbf{u}_{\Lambda_j} - \overline{\mathbf{u}}_{\Lambda_j})\| + c_r\varepsilon_j \le (C_A c_u + c_r)\varepsilon_j. \end{aligned} \tag{3.49}$$

Inserting (3.46) and (3.49) into (3.45), yields

$$|J(e_{\Lambda_j})| \le e_j + \|\tilde{\mathbf{w}}|_{\hat{\Lambda}_j\setminus\Lambda_j}\|(C_A c_u + c_r)\varepsilon_j + c_z\varepsilon_j\Big(\|\mathbf{r}\| + (C_A c_u + c_r)\varepsilon_j\Big), \tag{3.50}$$

which is the computable error bound (3.42). Thus the termination criterion ensures that the asserted target tolerance is met.

In order to prove convergence it remains to estimate the terms $\|\tilde{\mathbf{w}}|_{\hat{\Lambda}_j\setminus\Lambda_j}\|$, $\|\mathbf{r}\|$ and e_j. Clearly

$$\begin{aligned}\|\tilde{\mathbf{w}}|_{\hat{\Lambda}_j\setminus\Lambda_j}\| &\le \|\big(\mathbf{z} - \tilde{\mathbf{w}}\big)|_{\mathcal{I}\setminus\Lambda_j}\| + \|\mathbf{z}|_{\hat{\Lambda}_j\setminus\Lambda_j}\| \\ &\le c_z\varepsilon_j + \|\mathbf{z} - \hat{\mathbf{z}}_{j-1}\| \le c_z(\varepsilon_j + \varepsilon_{j-1}) \\ &= c_z(1 + m_0)\varepsilon_j. \end{aligned} \tag{3.51}$$

Furthermore, by (3.47) and (3.49),

$$\|\mathbf{r}\| \le \|\mathbf{r} - \mathbf{r}_{\Lambda_j}(\mathbf{u})\| + \|\mathbf{r}_{\Lambda_j}(\mathbf{u})\| \le (C_A c_u + c_r)\varepsilon_j + C_A\varepsilon_0. \tag{3.52}$$

Finally, by (3.51) and (3.52), we obtain

$$e_j \le \|\tilde{\mathbf{w}}|_{\hat{\Lambda}_j\setminus\Lambda_j}\|\,\|\mathbf{r}\| \le c_z(1 + m_0)\varepsilon_j\big((C_A c_u + c_r)\varepsilon_j + C_A\varepsilon_0\big), \tag{3.53}$$

which also tends to zero as j grows. This finishes the proof. □

To prepare for the numerical experiments in the subsequent section, we address next several further issues concerning the scheme ALGORITHM I.

We have not specified yet the choice of the parameters c_u, c_z, c_r. Of course, the smaller these parameters are chosen, the more will the computed error terms e_j dominate the true error. It is also clear that one should take $c_z < c_u$. The numerical experiments in the subsequent section will shed some more light on the quantitative behavior of ALGORITHM I regarding this point.

Concerning the progressive improvement of accuracy, let

$$\bar{e}_j(\tilde{\mathbf{w}}, \mathbf{r}) := e_j + \varepsilon_j\Big\{(C_A c_u + c_r)(\|\tilde{\mathbf{w}}|_{\hat{\Lambda}_j\setminus\Lambda_j}\| + c_z\varepsilon_j) + c_z\|\mathbf{r}\|\Big\}, \tag{3.54}$$

see step (ii) in ALGORITHM I. An alternative choice of the tolerances ε_j might be

$$\varepsilon_{j+1} := \frac{1}{m_0}\min\{\varepsilon_j, \bar{e}_j(\tilde{\mathbf{w}}, \mathbf{r})\}, \tag{3.55}$$

in order to exploit the fact that the error decay is superlinear. In fact, in view of (3.50) and (3.51), the estimate (3.42) says that

$$|J(e_{\Lambda_j})| \leq c\,\varepsilon_j(\|\mathbf{r}\| + \varepsilon_j).$$

Thus, up to the approximate residual $\|\mathbf{r}\|$, the error decay is quadratic in the refinement tolerances ε_j. If instead of using the constraint scheme $\textsc{Solve}_{\Lambda_j}$ for the primal problem in step (ii) of Algorithm I, one applies the unconstraint Solve also to the primal problem, the term $\|\mathbf{r}\|$ would decay like ε_j as well. In this case, an overall quadratic error decay would result which is the point of view taken in [22]. In fact, during the final stage of this work, we became aware of recent results by M. S. Mommer and R. P. Stevenson [22] who derive convergence *rates* for a goal oriented scheme in the Finite Element framework. There, however, they combine adaptive energy norm approximations to the primal *and* dual solution to arrive at concrete rates. Of course, this may increase the number of degrees of freedom required for the primal solution even in regions where they may only weakly contribute to the accuracy of the target functional. We shall address this issue in the experiments in the subsequent section.

Even though in the present scheme the primal problem is solved only in a constrained way, one expects that the third term on the right hand side of (3.45) is too crude an estimate. In fact, as shown in later experiments, $\|\mathbf{r}\|$ may not tend to zero at all but $\mathbf{r}$ may be "locally" small where $\mathbf{z}$ has its most significant terms and large contributions may be damped by negligible components of $\mathbf{z}$. Therefore, the Cauchy Schwarz inequality produces a significant overestimation. Better estimates would require some a-priori knowledge about the decay of the coefficients in the dual solution $\mathbf{z}$ which will be discussed in a forthcoming paper.

As another practical variant, one could tame the increase of degrees of freedom by modifying step (ii) in Algorithm I as follows. When (3.42) is not satisfied, for $g_\lambda := |\tilde{w}_\lambda r_\lambda|$, $\lambda \in \hat{\Lambda}_j \setminus \Lambda_j$, let $\mathbf{g} := (g_\lambda)_{\lambda\in\hat{\Lambda}_j\setminus\Lambda_j}$ and determine the smallest subset $\Gamma \subset \hat{\Lambda}_j \setminus \Lambda_j$ such that

$$\|\mathbf{g}|_\Gamma\|_{\ell_1(\Gamma)} \geq \frac{1}{2}\|\mathbf{g}\|_{\ell_1(\hat{\Lambda}_j\setminus\Lambda_j)}. \tag{3.56}$$

In the subsequent step (iii), one would then set

$$\bar{\mathbf{u}}^{j+1} := \bar{\mathbf{u}}_{\Lambda_j},\ \bar{\mathbf{z}}^{j+1} := \hat{\mathbf{z}}_j, \quad \Lambda_{j+1} := \Lambda_j \cup \Gamma, \quad \varepsilon_{j+1} = \varepsilon_j/m_0, \quad j+1 \to j, \tag{3.57}$$

and go to (ii). This may be viewed as a coarsening based on the error representation. To ensure convergence, one could add in (3.57), in addition, the support of a norm approximation to $\mathbf{z}$ with respect to the coarser tolerance $c_z'\varepsilon_j$, $c_z' > c_z$. The reasoning remains then the same while the constants change somewhat.

As for the computational complexity of any of these versions, most of the applications of Solve are actually just tightenings of already good initial

guesses where the current accuracy is improved only by a constant factor. So the corresponding computational work remains, in principle, proportional to the current number of degrees of freedom.

4 Numerical Experiments

We complement next the above findings by some first numerical experiments that are to shed some light on the quantitative behavior of the various error components.

Our test case is the Poisson equation on the L–shaped domain $\Omega = (-1,1)^2\backslash\big((-1,0]\times[0,1)\big)$ so that

$$a(u,v) = \int_\Omega (\nabla u)^T \nabla v \, dx \tag{4.58}$$

and $V = H_0^1(\Omega)$ in (2.2). This problem is interesting since the solution may exhibit a singularity caused by the shape of the domain even for smooth right hand sides, see, e.g., [21]. Thus, we can monitor the quantitative influence of such a singularity on the growth of the sets Λ_j. For the discretization, we use a globally continuous and piecewise linear wavelet basis.

The linear functional in our experiments is given by

$$J(u) = \frac{1}{|\Omega_{v,\delta}|}\int_{\Omega_{v,\delta}} u(x)dx \tag{4.59}$$

with

$$\Omega_{v,\delta} := \{x \in \mathbb{R}^2 : \|v - x\|_\infty \le \delta\} \subset \Omega.$$

We choose $v = (0.5, 0.5)^T$ and $\delta = 0.1$. The right hand side is scaled such that $J(u) \approx 1$. Hence $J(e_\Lambda)$ is close to the relative error $|J(e_\Lambda)|/|J(u)|$. Using approximations to u of very high accuracy, we use the resulting value of J for the validation of the results.

In the experiments below, e_j is defined as before by (3.41) while the second summand on the right hand side of (3.42) is denoted by f_j, so that $e_j + f_j$ is the computed error bound at the jth stage of Algorithm I.

4.1 Example 1: Smooth Right Hand Side

In the first example, we choose $f := 10$ so that the solution u of (2.2) exhibits only a singularity at the reentrant corner.

Table 1 shows that the "true" error $J(e_\Lambda)$ decays at least as fast as the parameter ε_j. The component e_j is much smaller than the true error and the computed error bound $e_j + f_j$ exceeds the true error only by a factor around 2. This is illustrated in Figure 3 which displays the computed dual error and the computed primal residual. While the dual energy norm error is halved within

j	ε_j	$e_j + f_j$	e_j	f_j	$J(e_\Lambda)$
1	2.07e+00	8.11e-01	3.10e-01	5.00e-01	1.02e+00
2	1.03e+00	8.91e-01	5.77e-01	3.14e-01	7.47e-01
3	5.17e-01	3.82e-01	2.20e-01	1.61e-01	2.55e-01
4	2.58e-01	1.21e-01	3.98e-02	8.08e-02	1.32e-01
5	1.29e-01	3.45e-02	3.72e-03	3.07e-02	4.21e-02
6	6.46e-02	2.03e-02	5.05e-03	1.53e-02	2.35e-02
7	3.23e-02	9.03e-03	1.70e-03	7.34e-03	7.30e-03
8	1.61e-02	4.24e-03	6.84e-04	3.56e-03	3.63e-03
9	8.07e-03	1.93e-03	2.24e-04	1.71e-03	8.77e-04

Table 1. Convergence history of ALGORITHM I in Example 1.

each iteration, the primal residual shows very poor convergence in accordance with the spirit of the scheme. As mentioned earlier, the slight overestimation is probably due to the crude estimate in the third term of the right hand side of (3.45). This is substantiated by Figure 1 which depicts the computed primal and dual solution $\mathbf{u}_{\Lambda_j}$ and $\mathbf{z}_{\Lambda_j}$ for $j = 1, \ldots, 5$. The strong concentration of the generalized Green's function around the support of J indicates that the primal residual, being large far away from the support of J, would hardly influence accuracy.

Moreover, the actual behavior of the primal approximate solutions is illustrated in Figures 2 and 4. With each wavelet ψ_λ, we associate a reference point $\kappa_\lambda \in \mathbb{R}^2$ which is located in the 'center' of its support. Locations where wavelets on many scales overlap therefore appear darker. Therefore, plotting the reference points $(\kappa_\lambda)_{\lambda\in\Lambda}$ gives an impression of the distribution of active indices in $u = \sum_{\lambda\in\Lambda} \bar{u}_\lambda$. Specifically, in Figure 2 the distribution of the elements of Λ_9 is displayed. As expected, most wavelets are located near the support of J and near the reentrant corner.

To see where the largest coefficients of the primal residual $\mathbf{r}$ are located, we plot the reference points of the largest (in modulus) 5% of the coefficients r_λ. The result is displayed in Figure 4. It can be seen that, near the support of J, the residual is small, reflecting a 'local' (in the wavelet coordinate domain) convergence behavior of $\bar{\mathbf{u}}_{\Lambda_j}$.

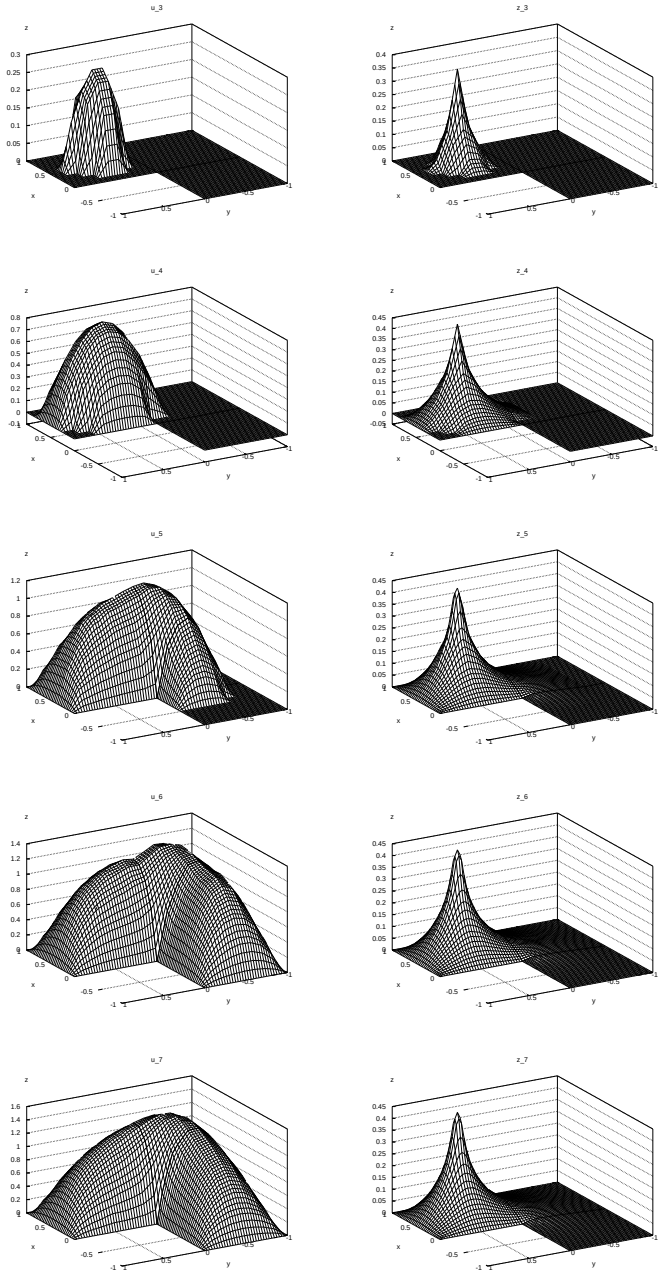

Fig. 1. Computed primal and dual solution in Example 1.

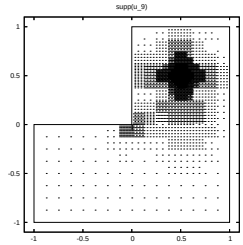

Fig. 2. Set of active coefficients Λ_9 used to evaluate $J(u_\Lambda)$ in Example 1.

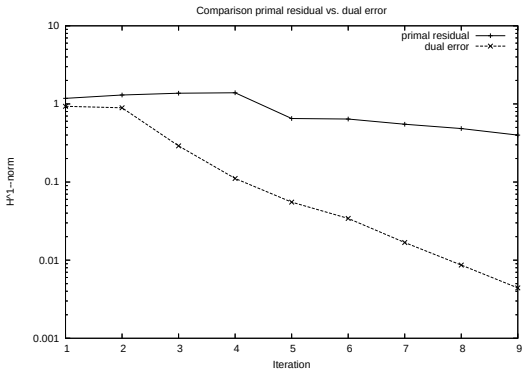

Fig. 3. Convergence of primal residual and dual solution in Example 1.

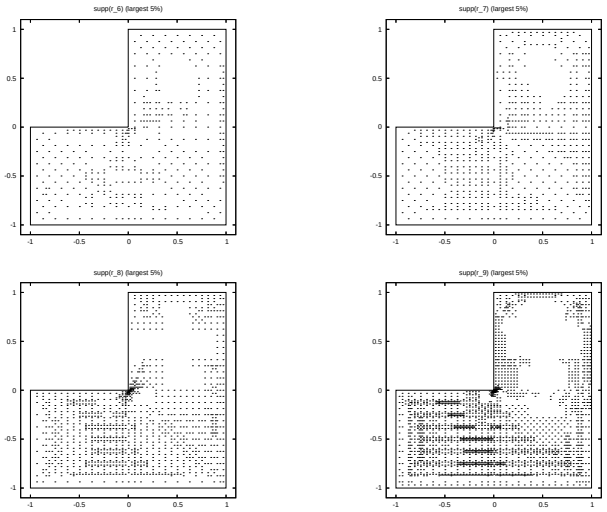

Fig. 4. Largest (in modulus) 5% of coefficients appearing in the primal residual vector in Example 1.

j	ε_j	$e_j + f_j$	e_j	f_j	$J(e_\Lambda)$	$\#\Lambda_j$
1	1.03e+00	1.31e+00	5.77e-01	7.33e-01	7.5092e-01	16
2	5.17e-01	5.85e-01	2.20e-01	3.65e-01	2.5913e-01	53
3	2.58e-01	2.27e-01	4.47e-02	1.82e-01	1.3628e-01	139
4	1.29e-01	9.21e-02	3.72e-03	8.84e-02	5.7297e-02	279
5	6.46e-02	4.90e-02	4.87e-03	4.41e-02	2.7194e-02	570
6	3.23e-02	2.37e-02	1.68e-03	2.20e-02	6.8861e-03	1752
7	1.61e-02	1.17e-02	6.95e-04	1.10e-02	2.7267e-03	5726

Table 2. Convergence history of ALGORITHM I in Example 2.

4.2 Example 2: Singular Right Hand Side

Next we wish to test the influence of a strong singularity of the primal solution u located far away from the support of J. This is realized by constructing a corresponding right hand side as follows. All (dual) wavelet coefficients of f are set equal to zero except the ones that overlap a fixed given point in the domain. These coefficients are chosen as $\langle \psi_\lambda, f\rangle := 1/(|\lambda| + 1)$. Since on each dyadic level only a uniformly bounded finite number of indices contributes and since the sequence $(\langle \psi_\lambda, f\rangle)_{\lambda\in I\!I}$ therefore belongs to ℓ_2, the resulting functional f is not contained in $L_2(\Omega)$, but certainly in $H^{-1}(\Omega)$. We finally add to f the constant function from Example 1. We expect that the singularity of the right hand side causes a strong concentration of relevant coefficients in the solution u that are spatially close to the singularity of f and comprise a wide range of relevant scales.

As we see from Table 2, the overestimation of the true error is slightly stronger than in Example 1. The reason is that, according to Figure 5, the primal residual is in this case larger (away from the support of J) due to the unresolved singularity caused by the right hand side f, so that the third term on the right hand side of (3.45) is overly pessimistic.

Table 3 sheds some more light on the local behavior of the primal residual. It shows that in the lower left patch where the singularity of f is located it does not converge to zero at all which, however, does not appear to affect the accuracy in a strong way.

The complexity of the scheme is indicated in Figure 6 which shows that the true error actually decays like N^{-1}, where N is the size of the index set needed to compute the approximate target value. Note that the rate for the energy norm error would be $N^{-1/2}$ at best.

j	$\|\tilde{\mathbf{w}}\|$	$\|\mathbf{r}\|$	$\|\mathbf{r}\|_{P1}\|$	$\|\mathbf{r}\|_{P3}\|$
1	6.92e-01	5.35e+00	8.60e-01	5.23e+00
2	2.23e-01	5.37e+00	6.80e-01	5.23e+00
3	1.02e-01	5.40e+00	3.22e-01	5.29e+00
4	5.37e-02	5.20e+00	3.23e-01	5.19e+00
5	3.30e-02	5.20e+00	2.85e-01	5.18e+00
6	1.63e-02	5.19e+00	1.94e-01	5.18e+00
7	8.46e-03	5.19e+00	1.13e-01	5.18e+00
8	4.34e-03	5.18e+00	5.66e-02	5.17e+00

Table 3. Convergence of dual error, primal residual, primal residual restricted to upper right patch $P1$ and lower left patch $P3$ in Example 2.

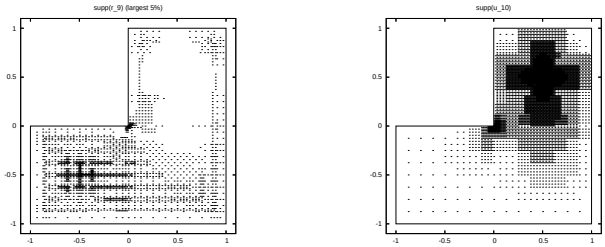

Fig. 5. Largest (in modulus) 5% of coefficients appearing in the primal residual vector and index set Λ_{10} generated in Example 2.

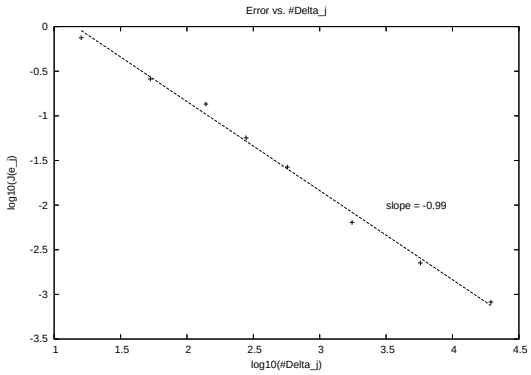

Fig. 6. Error $J(e_\Lambda)$ vs. number of degrees of freedom in Example 2.

References

1. A. Barinka, T. Barsch, Ph. Charton, A. Cohen, S. Dahlke, W. Dahmen, K. Urban, Adaptive wavelet schemes for elliptic problems — Implementation and numerical experiments, SIAM J. Sci. Comp., 23 (2001), 910–939.
2. A. Barinka, W. Dahmen, R. Schneider, Fast Computation of Adaptive Wavelet Expansions, IGPM Report 244, RWTH Aachen, August 2004.
3. P. Binev, W. Dahmen, R. DeVore, Adaptive Finite Element Methods with Convergence Rates, Numer. Math., 97(2004), 219–268.
4. R. Becker, R. Rannacher, An optimal error control approach to a–posteriori error estimation, Acta Numerica 2001, 1–102.
5. C. Canuto, A. Tabacco, K. Urban, The wavelet element method, part I: Construction and analysis, Appl. Comput. Harm. Anal., 6 (1999), 1–52.
6. C. Canuto, A. Tabacco, K. Urban, The Wavelet Element Method, Part II: Realization and additional features in 2D and 3D, Appl. Comp. Harm. Anal. **8** (2000), 123–165.
7. A. Cohen, W. Dahmen, R. DeVore, Adaptive wavelet methods for elliptic operator equations – Convergence rates, Math. Comp. 70, 2001, 27–75.
8. A. Cohen, W. Dahmen, R. DeVore, Adaptive wavelet methods II – Beyond the elliptic case, Found. Computat. Math. 2, 2002, 203–245.
9. A. Cohen, W. Dahmen, R. DeVore, Adaptive wavelet scheme for nonlinear variational problems, SIAM J. Numer. Anal. 41 (5), 2003, 1785–1823.
10. A. Cohen, W. Dahmen, R. DeVore, Sparse evaluation of compositions of functions using multiscale expansions, SIAM J. Math. Anal., 35 (2003), 279–303.
11. A. Cohen, R. Masson, Wavelet adaptive methods for second order elliptic problems, boundary conditions and domain decomposition, Numer. Math., 86 (2000), 193–238.
12. S. Dahlke, W. Dahmen, K. Urban, Adaptive wavelet methods for saddle point problems – Convergence rates, SIAM J. Numer. Anal., 40 (No. 4) (2002), 1230–1262.
13. W. Dahmen, Multiscale and Wavelet Methods for Operator Equations, C.I.M.E. Lecture Notes in Mathematics, *Multiscale Problems and Methods in Numerical Simulation*, Springer Lecture Notes in Mathematics, Vol. 1825, Springer-Verlag, Heidelberg, 2003, 31–96.
14. W. Dahmen, A. Kunoth, Adaptive wavelet methods for linear–quadratic elliptic control problems: Convergence rates, SIAM J. Contr. Optim. 43(5)(2005), 1640–1675.
15. W. Dahmen, R. Schneider, Composite wavelet bases for operator equations, Math. Comp., 68 (1999), 1533–1567.
16. W. Dahmen, R. Schneider, Wavelets on manifolds I: Construction and domain decomposition, SIAM J. Math. Anal., 31 (1999), 184–230.
17. W. Dahmen, R. P. Stevenson, Element-by-element construction of wavelets – stability and moment conditions, SIAM J. Numer. Anal., 37 (1999), 319–325.
18. W. Dahmen, J. Vorloeper, Adaptive Application of Operators and Newton's Method, in preparation.
19. K. Eriksson, D. Estep, P. Hansbo, C. Johnson, Introduction to adaptive methods for differential equations, Acta Numerica 1995, 105–158.
20. T. Gantumur, H. Harbrecht, R. P. Stevenson, An optimal adaptive wavelet method without coarsening of the iterands, Preprint No. 1325, Department of

Mathematics, Utrecht University, March 2005, submitted, revised version, June 2005.
21. P. Grisvard, Elliptic Problems in Nonsmooth Domains, Pitman, Boston, 1985.
22. M. S. Mommer, R. P. Stevenson, A goal oriented adaptive finite element method with a guaranteed convergence rate, December 21, 2005, lecture at the "Utrecht Workshop on Fast Numerical Solutions of PDEs", (December 20–22, 2005, Utrecht, The Netherlands).
23. R. P. Stevenson, On the compressibility of operators in wavelet coordinates, SIAM J. Math. Anal. 35(5), 1110–1132 (2004).
24. R. P. Stevenson, Optimality of a standard adaptive finite element method, Preprint No. 1329, Department of Mathematics, Utrecht University, May 2005, submitted, revised version, January 2006.

Bestellungen nimmt entgegen:

Institut für Angewandte Mathematik
der Universität Bonn
Sonderforschungsbereich 611
Wegelerstr. 6
D - 53115 Bonn

Telefon: 0228/73 3411
Telefax: 0228/73 7864
E-mail: anke@iam.uni-bonn.de Homepage: http://www.iam.uni-bonn.de/sfb611/

Verzeichnis der erschienenen Preprints ab No. 255

275. Otto, Felix; Reznikoff, Maria G.: Slow Motion of Gradient Flows

276. Albeverio, Sergio; Baranovskyi, Oleksandr; Pratsiovytyi, Mykola; Torbin, Grygoriy: The Ostrogradsky Series and Related Probability Measures

277. Albeverio, Sergio; Koroliuk, Volodymyr; Samoilenko, Igor: Asymptotic Expansion of Semi-Markov Random Evolutions

278. Frehse, Jens; Kassmann, Moritz: Nonlinear Partial Differential Equations of Fourth Order under Mixed Boundary Conditions

279. Juillet, Nicolas: Geometric Inequalities and Generalized Ricci Bounds in the Heisenberg Group

280. DeSimone, Antonio; Grunewald, Natalie; Otto, Felix: A New Model for Contact Angle Hysteresis

281. Griebel, Michael; Oeltz, Daniel: A Sparse Grid Space-Time Discretization Scheme for Parabolic Problems

282. Fattler, Torben; Grothaus, Martin: Strong Feller Properties for Distorted Brownian Motion with Reflecting Boundary Condition and an Application to Continuous N-Particle Systems with Singular Interactions

283. Giacomelli, Lorenzo; Knüpfer, Hans: Flat-Data Solutions to the Thin-Film Equation Do Not Rupture

284. Barbu, Viorel; Marinelli, Carlo: Variational Inequalities in Hilbert Spaces with Measures and Optimal Stopping

285. Philipowski, Robert: Microscopic Derivation of the Three-Dimensional Navier-Stokes Equation from a Stochastic Interacting Particle System

286. Dahmen, Wolfgang; Kunoth, Angela; Vorloeper, Jürgen: Convergence of Adaptive Wavelet Methods for Goal-Oriented Error Estimation; erscheint in: ENUMATH 2005 Proceedings