ADAPTIVE WAVELET METHODS FOR ELLIPTIC VARIATIONAL INEQUALITIES I: ANALYSIS

MARIO ROMETSCH AND KARSTEN URBAN

Abstract. We introduce an adaptive wavelet method for solving elliptic variational inequalities. We show convergence and optimality of the scheme as compared to the best N-term approximation of the solution. The scheme is based upon a projected Richardson method w.r.t. the (infinite) sequence space of wavelet coefficients. By combining well-known approximate operator applications and a new adaptive projection, we obtain a fully computable algorithm. We describe results of numerical experiments in 1D that show the quantitative performance of our scheme. Further numerical investigations are devoted to Part II.

1. Introduction

Variational inequalities appear in several problems in various fields of applications. Let us just mention elastoplastic hardening or the valuation of American type options in finance. There is a huge literature both on theory and numerical methods for solving variational inequalities, in particular of elliptic type (EVI). However, the construction and numerical analysis of adaptive numerical methods is still a quite active area of research, e.g. [4, 13, 16, 17, 20, 25]. In particular, questions like convergence and optimality of adaptive schemes are partly open.

In this paper, we introduce an adaptive wavelet scheme, prove its convergence and optimality as compared to the rate of the best N-term approximation. We combine ideas that are around already for some time. On one hand, already in [19] a convergent Wavelet-Rothe method for a variational inequality from elastoplasticity was introduced. However, there was no complexity analysis for this scheme and also numerical tests were not optimal from the quantitative point of view. On the other hand, significant progress has been made for adaptive wavelet methods for operator equations in the last years. Convergence and optimality have been proven for a wide range of operator equations, e.g. [8, 15, 26].

The idea is to reformulate an EVI equivalently as a discrete but still infinite problem for the wavelet coefficients of the desired solution. For this discrete problem, we consider a projected Richardson scheme which is known to converge. In order to obtain a fully computable version, we introduce known approximate operator applications and a new adaptive projection in order to respect the convex constraint set. It turns out that it is often useful to further correct the current

Date: September 26, 2010.

2000 *Mathematics Subject Classification.* 65T60,35Q68.

Key words and phrases. Wavelets, adaptive numerical methods, variational inequalities.

This work has been supported by the Deutsche Forschungsgemeinschaft within the Research Training Group (Graduiertenkolleg) GrK1100 *Modellierung, Analyse und Simulation in der Wirtschaftsmathematik* at the University of Ulm.

iteration by solving the discretized EVI on the set of active wavelet coefficients. Because of this and also for comparison reasons with uniform methods we start by introducing a corresponding method.

We have performed several numerical experiments in order to investigate the quantitative performance of our scheme. In order to concentrate on the properties of the scheme only, we have restricted ourselves to 1D examples, of course clearly having in mind that this is not the ultimate interest. We comment on this in the conclusions in Section 7 and also refer to Part II of this paper, [24].

The remainder of this paper is organized as follows. In Section 2 we recall the basic facts on (elliptic) variational inequalities and in Section 3 we introduce the wavelet formulation of EVI as well as the necessary background on wavelets itself. Since we perform various numerical tests, we collect the data for these tests in Section 4. We also determine the corresponding best N-term approximation as a benchmark for our adaptive schemes. In Section 5 we introduce a wavelet method for a given (but possibly adaptively generated) index set Λ. This is done for two reasons. First of all, we show the deficiency of standard multiresolution Galerkin methods, where a multiresolution space is used as trial and test space. Secondly, as already mentioned, we will use this solver later in Section 6 within the adaptive method for a fixed index set Λ. In the sequel, we use the abbreviation $A \lesssim B$ to indicate the existence of a constant $c > 0$ such that $A \leq c\,B$. By $A \sim B$ we denote $A \lesssim B$ and $B \lesssim A$ (in general with different constants, of course).

2. Variational Inequalities

In this section, we recall the main facts on elliptic variational inequalities. Let V be a Hilbert space and $\mathcal{K} \subseteq V$ a closed, convex and non-empty subset. Then we define the projection operator $P_{\mathcal{K}}$ onto $\mathcal{K}$ as the best approximation of $f \in V$ by some element $P_{\mathcal{K}}(f) \in \mathcal{K}$, i.e., $\|f - P_{\mathcal{K}}(f)\|_V = \min_{g\in\mathcal{K}} \|f - g\|_V$. It is well-known, that $P_{\mathcal{K}}(f)$ is uniquely defined, [16, Lemma 2.1]. Moreover, the following properties are known ([16, Theorem 2.3])

$$(P_{\mathcal{K}}(f), g - P_{\mathcal{K}}(f))_V \geq (f, g - P_{\mathcal{K}}(f))_V, \qquad \forall g \in \mathcal{K}, \tag{2.1}$$

and the projection is non-expansive, namely

$$\|P_{\mathcal{K}}(u) - P_{\mathcal{K}}(v)\|_V \leq \|u - v\|_V, \qquad \forall u, v \in V. \tag{2.2}$$

For the special case $V = L_2(\Omega)$ and $\mathcal{K} = \{v \in L_2(\Omega) \mid v \geq g \text{ a.e. in } \Omega\}$ with some function $g \in L_2(\Omega)$, the projection $P_{\mathcal{K}}(u)$ is given by $P_{\mathcal{K}}(u) = \max\{u, g\}$, i.e., $(P_{\mathcal{K}}(u))(x) = \max\{u(x), g(x)\}$, see [16, Section II.3].

In order to introduce elliptic variational inequalities, let $V \hookrightarrow H \cong H^* \hookrightarrow V^*$ be a Gelfand triple and again let $\mathcal{K} \subseteq V$ be non-empty, closed and convex. Moreover, let $A : V \to V^*$ be an elliptic operator and $f \in V^*$ be given. Then, the following problem is called an *Elliptic Variational Inequality (EVI)*: Find $u \in V$ such that

$$\langle Au, v - u\rangle_{V^*\times V} \geq \langle f, v - u\rangle_{V^*\times V}, \qquad \forall v \in \mathcal{K}, \tag{2.3}$$

where $\langle\cdot,\cdot\rangle_{V^*\times V}$ denotes the duality pairing of V^* and V using H as pivot space. If we define (as usual) the bilinear form $a(u, v) = \langle Au, v\rangle_{V^*\times V}$, then we can state (2.3) also in the form $a(u, v - u) \geq \langle f, v - u\rangle_{V^*\times V}$, for all $v \in \mathcal{K}$. From [17] it is known that the EVI (2.3) admits a unique solution $u \in V$ for every $f \in V^*$.

3. Wavelet Formulation of Variational Inequalities

3.1. **Wavelets.** In this section, we briefly recall the main facts on wavelets. Since we have EVIs in mind, we concentrate on wavelet bases for elliptic differential equations having in mind that this can easily be extended also to other classes of operators. For more details on wavelets, we refer e.g. to [26].

Let $\Psi = \{\psi_\lambda : \lambda \in \mathcal{J}\} \subset H_0^1(\Omega)$ be a wavelet basis, where as usual the index $\lambda = (j,k)$ encodes the scale $|\lambda| := j$ and the location $k(\lambda) := k$ of a wavelet. We will pose some assumptions on Ψ as follows:

(B.1) The wavelets are local, i.e., with $\sigma_\lambda := \operatorname{supp}\ (\psi_\lambda)$ we have $\operatorname{diam}(\sigma_\lambda) \sim 2^{-|\lambda|}$.

(B.2) Let $\Psi = \{\psi_\lambda : \lambda \in \mathcal{J}_\phi\} \cup \{\psi_\lambda : \lambda \in \mathcal{J}_\psi = \mathcal{J} \setminus \mathcal{J}_\phi\}$ with $\mathcal{J}_\phi$ the index set encoding scaling functions. Then $(\psi_\lambda, q)_{L_2(\Omega)} = 0$ for all $\lambda \in \mathcal{J}_\psi$ and $q \in \mathcal{P}_m$, where $\mathcal{P}_m$ denote the polynomials of maximal order $m \in \mathbb{N}$. This means that ψ_λ, $\lambda \in \mathcal{J}_\psi$, have vanishing moments of order m.

(B.3) The wavelet system $\Psi = \{\psi_\lambda : \lambda \in \mathcal{J}\}$ forms a Riesz basis for $L_2(\Omega)$ and gives rise to norm equivalences at least for the Sobolev spaces $H^s(\Omega)$, $s \in [-1,1]$ in the following sense: There exist constants $0 < c_\Psi \le C_\Psi < \infty$ such that

$$c_\Psi \|\mathbf{D}^s \mathbf{d}\|_{\ell_2(\mathcal{J})} \le \|\mathbf{d}^T \Psi\|_{H^s(\Omega)} \le C_\Psi \|\mathbf{D}^s \mathbf{d}\|_{\ell_2(\mathcal{J})}, \quad s \in [-1,1],$$

where $\mathbf{d} = (d_\lambda)_{\lambda\in\mathcal{J}}$, $\mathbf{D} = \{2^{|\lambda|}\delta_{\lambda,\mu}\}_{\lambda,\mu\in\mathcal{J}}$, $\mathbf{d}^T\Psi := \sum_{\lambda\in\mathcal{J}} d_\lambda \psi_\lambda$. We use the normalization $\|\psi_\lambda\|_{L_2(\Omega)} = 1$, $\lambda \in \mathcal{J}$.

Given any operator $R : H_0^1(\Omega) \to H^{-1}(\Omega)$, we can define its wavelet representation as $\mathbf{R}(\mathbf{u}) := (R(\mathbf{u}^T\Psi), \Psi)$. In particular, if $A : H_0^1(\Omega) \to H^{-1}(\Omega)$ is given in terms of a bilinear form $a(\cdot,\cdot)$, we get $\mathbf{A} := a(\Psi,\Psi)$. In this case, there are constants $0 < c_A \le C_A < \infty$ such that

$$c_A\|\mathbf{v}\| \le \|\mathbf{A}\mathbf{v}\| \le C_A\|\mathbf{v}\|. \tag{3.4}$$

Note that these constants may be estimated from the infinite matrix $\mathbf{A}$.

Often, a wavelet basis is generated by a Multiresolution Analysis (MRA), formed by the span of all scaling functions on a certain level, i.e., $S_j := \operatorname{span}\{\phi_{j,k} : k \in \mathcal{I}_j\}$. Here, for $\lambda \in \mathcal{J}_\phi$, we have $\psi_\lambda = \phi_{0,k}$, $\lambda = (0,k)$, $k \in \mathcal{I}_j$. If we use S_j as trial and test space for a Galerkin method, this is called *Multiresolution Galerkin method (MGM)*. Wavelet spaces $W_j := \operatorname{span}\{\psi_\lambda : |\lambda| = j\}$ then satisfy $S_{j+1} = S_j \oplus W_j$. This means that any $v_{j+1} \in S_{j+1}$ has two representations

$$v_{j+1} = \sum_{k\in\mathcal{I}_{j+1}} c_{j+1,k}\phi_{j+1,k} = \sum_{k\in\mathcal{I}_0} c_{0,k}\phi_{0,k} + \sum_{\ell=0}^{j}\sum_{k\in\mathcal{J}_\ell} d_{\ell,k}\psi_{\ell,k},$$

the left one called *single scale* and the right one *multiscale representation*. The change of basis $\mathbf{G}_j : \mathbf{c}_{j+1} \mapsto (\mathbf{c}_0, \mathbf{d}_0, \dots, \mathbf{d}_j)^T$ (with $\mathbf{c}_\ell = (c_{\ell,k})_{k\in\mathcal{I}_\ell}$, $\mathbf{d} = (d_{\ell,k})_{k\in\mathcal{J}_\ell}$) is called *wavelet decomposition* and its inverse $\mathbf{M}_j = \mathbf{G}_j^{-1}$ *wavelet reconstruction*.

3.2. **An equivalent ℓ_2-problem.** Now, we use the wavelet basis Ψ as trial and test functions, which results in an EVI for the coefficients $\mathbf{u} \in \ell_2(\mathcal{J})$. First, we define the convex set in wavelet space as $\mathbf{K} := \{\mathbf{d} \in \ell_2(\mathcal{J}) : v = \mathbf{d}^T\Psi \in \mathcal{K}\}$. Then, the problem reads: Find $\mathbf{u} \in \mathbf{K}$ such that

$$(\mathbf{A}\mathbf{u}, \mathbf{v}-\mathbf{u}) \ge (\mathbf{f}, \mathbf{v}-\mathbf{u}), \qquad \forall \mathbf{v} \in \mathbf{K}. \tag{3.5}$$

It is readily seen that (3.5) is equivalent to (2.3). In fact, for $u = \mathbf{u}^T\Psi$, $v = \mathbf{v}^T\Psi \in \mathcal{K}$, we have $\mathbf{u}, \mathbf{v} \in \mathbf{K}$. Then, $\langle Au, v-u\rangle_{V^*\times V} = a(u, v-u) = \mathbf{u}^T\, a(\Psi,\Psi)\,(\mathbf{v}-\mathbf{u}) =$

$(\mathbf{Au}, \mathbf{v}-\mathbf{u})$ and for $f = \mathbf{f}^T\tilde{\Psi} \in V^*$ we get by biorthogonality $\langle f, v-u\rangle_{V^*\times V} = \mathbf{f}^T(\tilde{\Psi}, \Psi)(\mathbf{v}-\mathbf{u}) = (\mathbf{f}, \mathbf{v}-\mathbf{u})$.

4. Numerical Examples

In this section, we collect the data of all our numerical examples that we consider in the remainder of this paper. As a typical example of an elliptic inequality we consider the following obstacle problem for the Helmholtz operator on $\Omega = (0,1)$ with some $\beta \geq 0$ and an obstacle $g \in H_0^1(\Omega)$: Find $u \in H_0^1(\Omega)$ such that

$$-\Delta u + \beta u \geq f, \quad u \geq g, \quad (u-g)(-\Delta u + \beta u - f) = 0 \text{ in } \Omega. \tag{4.6}$$

Note that for (4.6) the convex set is given by $\mathcal{K} = \left\{v \in H_0^1(\Omega) \;\middle|\; v \geq g\right\}$. We will sometimes (at least for numerical purposes) reduce ourselves to the case $g = 0$ which is by no means a restriction as can be seen by considering $\hat{u} := u - g$ and $\hat{f} := f + \Delta g - \beta\, g$ which reduces the general case to the specific obstacle $g = 0$.

Remark 4.1. *As in the case of elliptic equations, we would like to estimate the error in the energy norm for a finite approximation u_Λ, i.e., we are interested in $\|u-u_\Lambda\|_a$, which is equivalent to $\|u-u_\Lambda\|_{H^1(\Omega)}$. Unfortunately, we cannot apply the standard argument $\|u-u_\Lambda\|_a^2 = \|u\|_a^2 - \|u_\Lambda\|_a^2$ here, as for the inequality (2.3) there is no Galerkin orthogonality. We will therefore follow a different approach that was also used in* [3] *for the estimation of the energy error. We know that (in case that $a(\cdot,\cdot)$ is symmetric) the solution u of (2.3) minimizes the energy functional $\Pi(v) = \frac{1}{2}a(v,v) - \langle f, v\rangle$ over $\mathcal{K} \subseteq V$. Thus by* [3, (16)]*, we have*

$$\frac{1}{2}\|u-u_\Lambda\|_a \leq \sqrt{|\Pi(u) - \Pi(u_\Lambda)|} \tag{4.7}$$

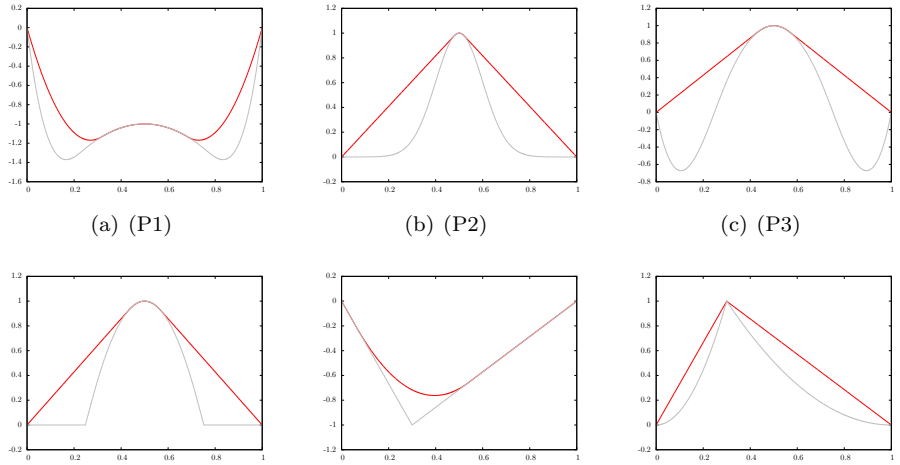
and we will use the right-hand side of (4.7) for estimating the convergence rate w.r.t. $\|\cdot\|_{H^1(\Omega)}$. By doing so, we are left with computing

$$\Pi(u_\Lambda) = \frac{1}{2}\mathbf{u}_\Lambda^T\mathbf{A}_{\Lambda\times\Lambda}\mathbf{u}_\Lambda - \mathbf{f}_\Lambda^T\mathbf{u}_\Lambda, \qquad u_\Lambda = \mathbf{u}_\Lambda^T\Psi_\Lambda,$$

as long as we know $\Pi(u)$ for the solution u to (2.3). For our test examples (P1)-(P6), as described in [22, Remark 5.1 and Subsection 5.1.2]*, this is the case.* □

4.1. **Examples for the Helmholtz inequality.** The data for our six examples is displayed in Table 4.4 (page 22). These have been chosen in order to consider different possible difficulties like lacking regularity, or as in (P6) a contact zone being only a single point. The corresponding plots are shown in Figure 4.1.

4.2. **Best N-term approximation.** In order to setup a benchmark for the adaptive schemes, we determine numerically the rate of the best N-term approximation for our test examples. The results are shown in Table 4.1. Already for the examples with smooth data (P1) to (P3) the inherent lack of regularity, that is due to the operator which governs the EVI (2.3), is responsible for a limited rate of $s = 1.5$. We investigate, if a higher convergence rate is possible by using the best N-term approximation and observe a slight gain in convergence speed $s \approx 2.1$ for $d = 3$. So even for smooth data, a uniform scheme has to be expected to have a deficient rate as compared to the best N-term approximation for wavelets of sufficiently high order. Next, we consider the singular examples (P4) to (P6). Here the possible rate $s \approx 1.0$ for $d = 2$ is quadratically higher than for the uniform schemes. This is due to the Besov regularity of the solution u which is $u \in B^s_{\tau,\tau}, \forall s > 0$ (see Lemma 6.1),



(a) (P1) (b) (P2) (c) (P3)

(d) (P4) (e) (P5) (f) (P6)

FIGURE 4.1. Plots of the examples for the Helmholtz inequality problem. Solutions in red, obstacles in gray.

resulting in a nonlinear approximation rate of $s_{22} = 1$ for $d = 2$ and $s_{33} = 2$ for $d = 3$, that is limited only by the smoothness of the wavelets.

	(P1)	(P2)	(P3)		(P4)	(P5)	(P6)
N-term, $d = 3$	2.16	2.14	2.19	N-term, $d = 2$	1.11	1.05	1.04
uniform, $d = 3$	1.50*	1.50*	1.50*	uniform, $d = 2$	0.50*	0.50*	0.50*

TABLE 4.1. Decay rates of the best N-term approximation for u_i as well as the maximal rates for the uniform approximation determined by the Sobolev regularity of u_i. The star values are the predicted ones.

5. Methods for given fixed index sets

We start by considering uniform schemes for the solution of the variational inequalities introduced above. As already mentioned above, we have two goals, namely to show the deficiency of (uniform) Multiresolution Galerkin schemes and to setup a framework for solving an EVI for adaptively generated index sets Λ. For the EVI, the discretized problem for fixed trial and test spaces V_Λ reads: Find $u_\Lambda \in \mathcal{K} \cap V_\Lambda = \mathcal{K}_\Lambda$ such that

$$a(u_\Lambda, v_\Lambda - u_\Lambda) \geq (f, v_\Lambda - u_\Lambda), \qquad \forall v_\Lambda \in \mathcal{K}_\Lambda. \tag{5.8}$$

This leads to a problem for the coefficients of the function u: Find $\mathbf{u}_\Lambda \in \mathbf{K}_\Lambda := \{\mathbf{v} \in \ell_2(\Lambda) : \mathbf{d}^T\Psi \in \mathcal{K}\} = \mathbf{K} \cap \ell_2(\Lambda)$ such that

$$(\mathbf{v}_\Lambda - \mathbf{u}_\Lambda)^T \mathbf{A}_{\Lambda\times\Lambda}\mathbf{u}_\Lambda \geq (\mathbf{v}_\Lambda - \mathbf{u}_\Lambda)^T \mathbf{f}_\Lambda, \qquad \forall \mathbf{v}_\Lambda \in \mathbf{K}_\Lambda, \tag{5.9}$$

with the matrix $\mathbf{A}_{\Lambda\times\Lambda} = (a(\psi_\mu, \psi_\lambda))_{\lambda,\mu\in\Lambda}$ and the right-hand side $\mathbf{f}_\Lambda = (\langle f, \psi_\lambda\rangle)_{\lambda\in\Lambda}$.

5.1. **Convergence Analysis.** The error analysis of (5.8) resp. (5.9) follows standard lines ([13, Theorem 1]) and results in the estimate

$$\|u-u_\Lambda\|_V^2 \le \frac{C_1^2}{C_2^2}\|u-w_\Lambda\|_V^2 + \frac{2}{C_2}\|f-Au\|_{V^*}(\|u-w_\Lambda\|_V + \|u_\Lambda - w\|_V),$$

for all $w \in \mathcal{K}$ and $w_\Lambda \in \mathcal{K}_\Lambda$. For our conforming method we can choose $w = u_\Lambda$ and get

$$\|u-u_\Lambda\|_V^2 \lesssim \inf_{w_\Lambda \in \mathcal{K}_\Lambda} \left(\|u-w_\Lambda\|_V^2 + \|f-Au\|_{V^*}\|u-w_\Lambda\|_V\right). \tag{5.10}$$

As shown in [13, Theorem 2], for the special case of the Dirichlet obstacle problem ($A=-\Delta$, $V=H_0^1(\Omega)$ and $V^*=H^{-1}(\Omega)$), if $f \in L_2(\Omega)$ we get for $\Lambda = \mathcal{I}_J$, $V_\Lambda = S_J$ the higher rate

$$\|u-u_J\|_{H^1(\Omega)} \lesssim 2^{-J}\|u\|_{H^2(\Omega)}, \qquad J \ge j_0.$$

However, if we take a closer look at (5.10), for a *general* EVI driven by an operator $A : \tilde{H}^s(\Omega) \to H^{-s}(\Omega)$ and right-hand side $f \in H^{-s}(\Omega)$, we see that (neglecting the first term, which is of higher order) one only gets

$$\|u-u_\Lambda\|_V^2 \lesssim \inf_{v_\Lambda \in \mathcal{K}\cap V_\Lambda} \|f-Au\|_{V^*}\|u-v_\Lambda\|_V$$

and so we cannot expect a convergence rate that is higher than $\|u-u_J\|_{H^t(\Omega)} \lesssim 2^{-\frac{s}{2}J}\|u\|_{H^{t+s}(\Omega)}$, which is half of the Dirichlet obstacle problem case.

5.2. **Projected Richardson Iteration.** One possible algorithm for solving (5.9) for fixed Λ is the projected Richardson iteration:

$$\mathbf{u}^{n+1} = \mathbf{P}_{\mathbf{K}_\Lambda}(\mathbf{u}^n + \alpha_\Psi(\mathbf{f} - \mathbf{A}\mathbf{u}^n))$$

for a suitable relaxation parameter α_Ψ and $\mathbf{P}_{\mathbf{K}_\Lambda}$ denotes the projection onto $\mathbf{K}_\Lambda$. The discretized iteration may also be written as

$$(\mathbf{u}_\Lambda^{n+1}, \mathbf{v}_\Lambda - \mathbf{u}_\Lambda^{n+1}) \ge (\mathbf{u}_\Lambda^n + \alpha_\Psi(\mathbf{f}_\Lambda - \mathbf{A}_{\Lambda\times\Lambda}\mathbf{u}_\Lambda^n), \mathbf{v}_\Lambda - \mathbf{u}_\Lambda^{n+1}), \quad \forall \mathbf{v}_\Lambda \in \mathbf{K}_\Lambda.$$

Recall that, the projection is also characterized by the minimal distance, i.e., $\|\mathbf{f}_\Lambda - \mathbf{P}_{\mathbf{K}_\Lambda}(\mathbf{f}_\Lambda)\| = \min_{\mathbf{g}_\Lambda \in \mathbf{K}_\Lambda}\|\mathbf{f}_\Lambda - \mathbf{g}_\Lambda\|$. Hence, we can compute the projection as the solution of the quadratic program

$$\mathbf{P}_{\mathbf{K}_\Lambda}(\mathbf{f}_\Lambda) = \operatorname*{arg\,min}_{\mathbf{g}_\Lambda \in \mathbf{K}_\Lambda}\left\{\frac{1}{2}\mathbf{g}_\Lambda^T\mathbf{g}_\Lambda - 2\mathbf{f}_\Lambda^T\mathbf{g}_\Lambda\right\}.$$

If we are able to characterize $\mathbf{u}_\Lambda \in \mathbf{K}_\Lambda$ equivalently as $\mathbf{C}_\Lambda\mathbf{u}_\Lambda \ge \mathbf{d}_\Lambda$ with some matrix $\mathbf{C}_\Lambda$ and a vector $\mathbf{d}_\Lambda$, we can compute $\mathbf{P}_{\mathbf{K}_\Lambda}$ as a constraint approximation on all $\mathbb{R}^{|\Lambda|}$. This is realized in Algorithm 5.1. We will comment on the choice of $\mathbf{C}_\Lambda$ later. The particular choice of the stopping criterion in line 6 will be discussed in Section 5.5 below.

5.3. **Preconditioning.** As pointed out in [18], the use of a properly scaled wavelet basis leads to a preconditioned variant of Algorithm 5.1 which then in turn converges independently of the discretization level.

Algorithm 5.1 PRICHARDSON: $[\mathbf{A}, \mathbf{f}, \Lambda, \delta_{\text{fp}}] \to \mathbf{v}_\Lambda$

1: choose $0 < \alpha_\Psi < 2\frac{c_A}{C_A^2}$
2: set $\mathbf{u}_\Lambda^0 = \mathbf{0} \in \ell_2(\Lambda)$, $n = 0$
3: **repeat**
4: $\quad \hat{\mathbf{u}}_\Lambda^{n+1} = \mathbf{u}_\Lambda^n + \alpha_\Psi \left(\mathbf{f}_\Lambda - \mathbf{A}_{\Lambda\times\Lambda}\mathbf{u}_\Lambda^n\right)$
5: $\quad$ Calculate $\mathbf{u}_\Lambda^{n+1}$ as the solution of the quadratic program,

$$\mathbf{u}_\Lambda^{n+1} = \operatorname*{arg\,min}_{\mathbf{u}_\Lambda \in \mathbb{R}^{|\Lambda|}} \{\mathbf{u}_\Lambda^T \mathbf{u}_\Lambda - 2\left(\hat{\mathbf{u}}_\Lambda^{n+1}\right)^T \mathbf{u}_\Lambda\} \qquad \text{subject to} \qquad \mathbf{C}_\Lambda \mathbf{u}_\Lambda \geq \mathbf{d}_\Lambda$$

6: **until** $\|\mathbf{u}_\Lambda^{n+1} - \mathbf{u}_\Lambda^n\| < \delta_{\text{fp}}$
7: **return** $\mathbf{v}_\Lambda := \mathbf{u}_\Lambda^{n+1}$

5.4. **Choice of the constraint matrix.** In the case of a Multiresolution Galerkin method ($V_\Lambda = \mathcal{S}_J$) and the homogeneous obstacle problem, one possible choice for $\mathbf{C}_\Lambda$ is the reconstruction matrix $\mathbf{C}_\Lambda = \mathbf{M}_{j_0} \cdots \mathbf{M}_J$ and for $\mathbf{d}_\Lambda$ the vector $\mathbf{d}_\Lambda = 0$. This choice results in $\mathbf{C}_\Lambda : \ell_2(\Lambda) \to \ell_2(\Lambda)$. If we think, however, of an adaptively generated index set $\Lambda \subset \mathcal{J}$, such a choice for the constraint matrix would be prohibitive as we would obtain $\mathbf{C}_\Lambda : \ell_2(\Lambda) \to \ell_2(\mathcal{I}_J)$, with J being the maximal level of *all* indices contained in Λ. In nD, the number of rows (and therefore the number of constraints) scales like 2^{nJ} – totally destroying the complexity of the algorithm. Anticipating our adaptive projection algorithm we note that we use a *local reconstruction*, see [22, Remark 4.15], for which the number of constraints stays proportional to the number of unknowns and the constraint matrix is sparse. Finally and as mentioned above, the presented method extends to general EVIs if we are able to characterize the constraint set. For example, the constraint set $\mathcal{K} = \left\{v \in H_0^1(\Omega) \;\middle|\; v' \geq g\right\}$ in case of spline wavelets may easily be characterized as $\mathbf{C}_\Lambda = \mathbf{M}_{j_0} \cdots \mathbf{M}_J \mathbf{E}$, where $\mathbf{E}$ is the differentiation matrix for the splines.

5.5. **Numerical results.** We describe some numerical experiments concerning the projected Richardson iteration for a fixed index set Λ.

5.5.1. *Residual estimation for obstacle problems.* We first consider the question whether the stopping criterion which is used in line 6 of Algorithm 5.1 is reasonable. In Galerkin solvers for equations on some set Λ one usually requires that the residual $\|\mathbf{f}_\Lambda - \mathbf{A}_{\Lambda\times\Lambda}\mathbf{v}_\Lambda\| = \|\mathbf{r}\| \leq \eta \|\mathbf{r}_0\|$ has to be sufficiently small, because this implies $\|\mathbf{u}_\Lambda - \mathbf{v}_\Lambda\| \leq \frac{\eta}{c_A}\|\mathbf{r}_0\|$. For an inequality, we would like to use a similar stopping criterion. As for the exact solution $\mathbf{u}$ of the homogeneous variant of (4.6) one has $(\mathbf{u}, \mathbf{A}\mathbf{u} - \mathbf{f}) = 0$, a natural error measure would be

$$r := |(\mathbf{u}^n, \mathbf{A}\mathbf{u}^n - \mathbf{f})| \in \mathbb{R}_+. \tag{5.11}$$

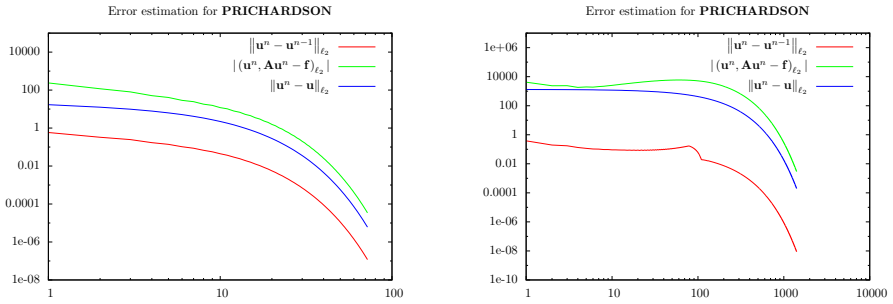
Because of $(\mathbf{u}, \mathbf{A}\mathbf{u} - \mathbf{f}) = 0$ we get

$$\begin{aligned} r \leq & |(\mathbf{u}^n - \mathbf{u}, \mathbf{A}\mathbf{u}^n - \mathbf{f})| + |(\mathbf{u}, \mathbf{A}\mathbf{u}^n - \mathbf{A}\mathbf{u})| + |(\mathbf{u}, \mathbf{A}\mathbf{u} - \mathbf{f})| \\ \leq & \|\mathbf{u}^n - \mathbf{u}\| \|\mathbf{A}\mathbf{u}^n - \mathbf{f}\| + \|\mathbf{u}\| \|\mathbf{A}(\mathbf{u}^n - \mathbf{u})\| \\ \leq & \|\mathbf{u}^n - \mathbf{u}\| (\|\mathbf{A}\mathbf{u}^n - \mathbf{f}\| + C_A \|\mathbf{u}\|). \end{aligned}$$

From this we see that $r \to 0$ for $\mathbf{u}^n \to \mathbf{u}$. Unfortunately the reverse (and even more important) implication $\mathbf{u}^n \to \mathbf{u}$ for $r \to 0$ is not clear and we were not able to

prove this. This is the reason why we want to conduct some experiments whether r in (5.11) is a reliable error estimator.

To this end, we choose the operator to be $A = -\Delta$, then generate a random right-hand side $\mathbf{f}_\Lambda$ with $f_\lambda \sim \mathcal{U}_{[-10,10]}$ and solve the resulting system (5.9) by a standard method to obtain an approximation $\mathbf{u}^{\mathrm{app}}$ for the exact solution. This is then compared with the iterates generated by **PRICHARDSON**, i.e. we check $\|\mathbf{u}^{\mathrm{app}} - \mathbf{u}^n\|$ against r. We choose the biorthogonal basis from [12] and $J = 9$.



(a) $d = \tilde{d} = 2$ and $j_0 = 3$ (b) $d = \tilde{d} = 3$ and $j_0 = 4$

FIGURE 5.2. Error estimate for the projected Richardson scheme, Dirichlet inequality, random right-hand side.

Let us comment on the results that are shown in Figure 5.2. In all our experiments we saw that our error estimator r is reliable in the sense that it is an upper bound for the error. We can also observe that it is efficient – the error estimator decreases as our iterate approaches the solution of our discretized variational inequality (5.9). Finally, we got these results both for linear and quadratic basis functions, therefore our stopping criterion seems appropriate.

5.5.2. *Uniform convergence rates.* Now we examine the convergence order of the uniform method, i.e., for $\Lambda = \mathcal{I}_J$. For solving the finite-dimensional linear complementary problems, we chose the projected Richardson iteration, where the constraint matrix corresponds to the reconstruction operator.

d	(P1)	(P2)	(P3)	(P4)	(P5)	(P6)
2	0.97	0.99	0.90	0.59	0.49	0.51
3	1.23	1.27	1.30	0.50	0.48	0.52

TABLE 5.2. Convergence rates in H^1-norm of the uniform scheme.

The results are shown in Table 5.2. For the first three examples (P1) to (P3), which all have $H^{5/2-\varepsilon}$-regularity, the uniform scheme with linear ansatz functions ($d = \tilde{d} = 2$) almost attains the optimal convergence order of $s = 1$. For quadratic ansatz functions ($d = \tilde{d} = 3$), we obtain a rate which is close to the optimal rate of $s = 1.5$ for the uniform approximation of u_i. If we would consider an operator *equation* with the same data we would get $u \in H^s(\Omega)$ for all $s > 0$ and therefore a convergence rate of $s = 2$ limited by the regularity of the ansatz functions. Hence, the inequality (2.3) limits the regularity, regardless that the obstacle and right-hand

side of the problem are smooth. Already at this point, we thus have a reasoning for designing a more sophisticated algorithm than the uniform scheme for solving the operator inequality. In the singular examples (P4) to (P6), there seems to be no hope that the uniform schemes are well-suited to solve the obstacle problem. Here the regularity drops to $u \in H^{3/2-\varepsilon}$ resulting in a very poor convergence rate of $s = \frac{1}{2}$. Overall, the rates were all within the range of the theoretical ones predicted in Section 5 and Table 4.1.

6. Adaptive Methods for EVIs

Now, we introduce our fully adaptive wavelet projected Richardson solver for EVI. In order to set a benchmark, we start by investigating the Besov regularity of solutions of EVI and hence the rate of their best N-term approximation. Then, we introduce, analyze and realize a fully computable adaptive solver called **EVISOLVE**. In addition to this, we also show a simplified version which performs numerically well but for which we do not have theoretical justification.

6.1. Besov regularity for EVIs.

Our first goal is to show that the solution of the EVI (2.3) admits a higher Besov regularity than in the Sobolev scale.

Lemma 6.1. *Let $u \in V := H^1_0(\Omega)$, $\Omega := (0,1)$, solve (2.3) for the special case of $A = -\Delta$. Then we have $u \in B^s_\tau(L_\tau(\Omega))$ for all $s > 0$ with $\tau = (s + 1/2)^{-1}$.*

Proof. Suppose that $u \in V$ solves (2.3) for $A = -\Delta$. Then, $f^{\mathrm{Eq}} := -\Delta u \in V^*$ and the equation $-\Delta u^{\mathrm{Eq}} = f^{\mathrm{Eq}}$ has a unique solution $u^{\mathrm{Eq}} \in V$ with $u^{\mathrm{Eq}} \in B^s_\tau(L_\tau(\Omega))$ for all $s > 0$, as known from the Besov regularity of the Dirichlet problem in 1D, see [14, Section 2.6.1], [15, Section 3] and [11, 10]. By $u = u^{\mathrm{Eq}}$ we obtain that $u \in B^s_\tau(L_\tau(\Omega))$ for all $s > 0$. □

Let us briefly comment on this result. It shows that the solution of the 1D Helmholtz inequality has arbitrary high Besov regularity. Hence, the rate of convergence of an adaptive scheme is only limited by the smoothness (order) of the wavelets, not by the operator. The above proof also shows that the Besov regularity of the variational inequality is at least the one of the variational *equality*. Since the Sobolev smoothness of the inequality is lower then for the equation,[1] this shows that the potential gain of adaptivity is higher for inequalities, adaptivity pays off!

6.2. EVISOLVE.

Now we develop the announced adaptive wavelet method for the solution of (2.3).

6.2.1. *Abstract adaptive method.* We start from the operator inequality (3.5). As for the uniform scheme, we consider the following projected Richardson iteration in wavelet coordinates for $\ell_2(\mathcal{J})$. For any starting point $\mathbf{u}^0 \in \ell_2(\mathcal{J})$ we define the operator $\mathbf{S} : \ell_2(\mathcal{J}) \to \ell_2(\mathcal{J})$ by

$$\mathbf{u}^{n+1} = \mathbf{S}\mathbf{u}^n := \mathbf{P}_{\mathbf{K}}\left(\mathbf{u}^n + \alpha_\Psi\left(\mathbf{f} - \mathbf{A}\mathbf{u}^n\right)\right). \tag{6.12}$$

For the last equation, we define the projection $\mathbf{P}_{\mathbf{K}}\hat{\mathbf{u}}$ by

$$\mathbf{P}_{\mathbf{K}}\hat{\mathbf{u}} := (P_{\mathcal{K}}(\hat{\mathbf{u}})^T\Psi, \tilde{\Psi})_{L_2(\Omega)}. \tag{6.13}$$

[1] We mention, for example, that according to [5, Theorem 2] the solution u of the Dirichlet obstacle problem in any dimension has no more regularity than $u \in C^1$.

It is readily seen that the operator $\mathbf{S}$ is contractive for all $0 < \alpha_\Psi < 2\frac{c_A}{C_A^2}$ with the optimal choice for $\alpha_\Psi = \frac{c_A}{C_A^2}$, which guarantees a fixed error reduction of $\|\mathbf{u}^{n+1} - \mathbf{u}\| \leq \rho_\Psi \|\mathbf{u}^n - \mathbf{u}\|$, where

$$\rho_\Psi = \|\mathbf{I} - \alpha_\Psi \mathbf{A}\| = \left(1 - c_A^2/C_A^2\right)^{\frac{1}{2}}, \tag{6.14}$$

in particular $\rho_\Psi < 1$. Then, we have:

Lemma 6.2. *Let $\mathbf{u} \in \ell_2(\mathcal{J})$ be a fixed point of $\mathbf{S}$. Then, $\mathbf{u}$ solves (3.5).*

Proof. First note that by definition of $\mathbf{S}$ we have $\mathbf{u} \in \mathbf{K}$. Furthermore, since $\mathbf{u}$ is the projection of $\hat{\mathbf{u}} = \mathbf{u} + \alpha_\Psi(\mathbf{f} - \mathbf{A}\mathbf{u})$ onto $\mathbf{K}$, we have by (2.1) that for all $\mathbf{v} \in \mathbf{K}$

$$(\mathbf{u}, \mathbf{v} - \mathbf{u}) \geq (\hat{\mathbf{u}}, \mathbf{v} - \mathbf{u}) = (\mathbf{u}, \mathbf{v} - \mathbf{u}) + \alpha_\Psi(\mathbf{f} - \mathbf{A}\mathbf{u}, \mathbf{v} - \mathbf{u}),$$

which is equivalent to $\alpha_\Psi(\mathbf{A}\mathbf{u}, \mathbf{v} - \mathbf{u}) \geq \alpha_\Psi(\mathbf{f}, \mathbf{v} - \mathbf{u})$, for all $\mathbf{v} \in \mathbf{K}$ and since $\alpha_\Psi > 0$ we obtain our claim. □

6.2.2. *Adaptive projection.* For a computable version of (6.12) we need the routines $\mathbf{RHS}[\eta, \mathbf{f}] \to \mathbf{f}_\eta$, $\mathbf{APPLY}[\eta, \mathbf{A}, \mathbf{u}] \to \mathbf{w}_\eta$ and $\mathbf{COARSE}[\eta, \mathbf{v}] \to \mathbf{v}_\eta$ which are well-known from adaptive wavelet methods for elliptic problems, [6].

Now, we develop a computable version of the projection in (6.13). This is the major ingredient for the realization of a computable version of the adaptive iteration defined in (6.12). The idea is to view the projection $P_\mathcal{K}$ as a nonlinear map and to construct a particular version of **RECOVERY**, [2, 8, 9, 26]. Thus, we are going to verify the assumptions posed in [9], namely

Assumption 1. $\mathbf{F} : \ell_2 \to \ell_2$ is a Lipschitz map, i.e. we assume that there is some function $\mathbf{L_F} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that

$$\|\mathbf{F}\mathbf{u} - \mathbf{F}\mathbf{v}\|_{\ell_2} \leq \mathbf{L_F}(\max\{\|\mathbf{u}\|_{\ell_2}, \|\mathbf{v}\|_{\ell_2}\}) \|\mathbf{u} - \mathbf{v}\|_{\ell_2}. \tag{6.15}$$

Assumption 2. If $\mathbf{w} = \mathbf{F}\mathbf{u}$ for some finite $\mathbf{u} \in \ell_2$, then we have the estimate

$$|w_\lambda| \leq \mathbf{L_F}(\|\mathbf{u}\|_{\ell_2}) \sup_{\mu:\, \mathrm{supp}\ (\psi_\lambda) \cap \mathrm{supp}\ (\psi_\mu) \neq \emptyset} |u_\mu|^{-\gamma(|\lambda| - |\mu|)}. \tag{6.16}$$

For the first assumption (6.15), already from (2.2) we have that we can set $\mathbf{L_F} \equiv 1$. Next we are going to prove the second assumption

Lemma 6.3. *The $H^t(\Omega)$-projection onto the closed, convex and non-empty set $\mathcal{K}$ satisfies the assumption of the* **RECOVERY** *scheme from* [9] *for all $\gamma \leq 2t + 1/2$. This in turn ensures s^*-compressibility with*

$$s^* = \gamma - 1/2 = \min\{d - 2 + t, 2t\}. \tag{6.17}$$

Proof. The estimate (6.16) reads in this case

$$|u_\lambda| \leq \sup_{\mu:\, \mathrm{supp}\ (\psi_\lambda) \cap \mathrm{supp}\ (\psi_\mu) \neq \emptyset} |\hat{u}_\mu|^{-\gamma(|\lambda| - |\mu|)},$$

for $\gamma = d - 2 + t + \frac{1}{2} = r + t + \frac{1}{2}$. Because the biorthogonal spline wavelets have $\tilde{d} \geq d$ vanishing moments we start with

$$\begin{aligned} |u_\lambda| = |(u, \psi_\lambda)| = \inf_{P \in \mathcal{P}_r} |(u - P, \psi_\lambda)| &\lesssim |u|_{W^r(L_\infty(\mathrm{supp}\ (\psi_\lambda)))}\, 2^{-(r+t+\frac{1}{2})|\lambda|} \\ &= |P_\mathcal{K}\hat{u}|_{W^r(L_\infty(\mathrm{supp}\ (\psi_\lambda)))}\, 2^{-(r+t+\frac{1}{2})|\lambda|} \end{aligned}$$

and then follow the lines of the proof of [9, Theorem 4.1] in order to show

$$|P_{\mathcal{K}}\hat{u}|_{W^r(L_\infty(\text{supp }(\psi_\lambda)))} \lesssim |\hat{u}|_{W^r(L_\infty(\text{supp }(\psi_\lambda)))} \tag{6.18}$$

for $r = d - 2$. The case $\hat{u} \equiv 0$ is trivial since then we get $P_{\mathcal{K}}\hat{u} \equiv 0$ and therefore $u_\lambda = 0, \lambda \in \mathcal{J}$. We will prove (6.18) by contradiction. Therefore, we assume that there exists $\hat{u} \in H^t$ and $\lambda \in \mathcal{J}$ such that for every $c > 0$

$$|u|_{W^{r,\infty}(\text{supp }(\psi_\lambda))} > c\,|\hat{u}|_{W^r(L_\infty(\text{supp }(\psi_\lambda)))}\,.$$

This means that $\operatorname{ess\,sup}_{\text{supp }(\psi_\lambda)} \partial^r u = \infty$ or that $\partial^r u$ is not essentially bounded on supp (ψ_λ). This, however, conflicts with (2.2), namely

$$\|u\|_{W^t(L_2(\Omega))} = \|u\|_{t;\Omega} \leq \|\hat{u}\|_{t;\Omega} = \|\hat{u}\|_{W^t(L_2(\Omega))}\,,$$

so (6.18) holds at least for $r \leq t$. □

Note that the latter result means for $t = 1$ we get $r \leq 1$ or $d \leq 3$, hence we achieve the optimal rates for linear and quadratic wavelets.

Ultimately, we want to get an algorithm $\mathbf{PROJECT}[\eta, \hat{\mathbf{v}}] \to \mathbf{w}_\eta$ that determines a finite approximation $\mathbf{w}_\eta \in \mathbf{K}$ of the projection $\mathbf{P_K}\hat{\mathbf{v}}$ of a *finite* input $\hat{\mathbf{v}}$ within some pre-specified error bound η, i.e., $\|\mathbf{P_K}\hat{\mathbf{v}} - \mathbf{w}_\eta\| \leq \eta$. From the definition of the projection we know that $\|\hat{\mathbf{v}} - \mathbf{P_K}\hat{\mathbf{v}}\| = \min_{\mathbf{g}\in\mathbf{K}} \|\hat{\mathbf{v}} - \mathbf{g}\|$, so we can use a quadratic program as in the finite-dimensional case. Hence, the projection $\mathbf{P_K}\hat{\mathbf{v}}$ is given by

$$\mathbf{P_K}\hat{\mathbf{v}} = \arg\min_{\mathbf{g}\in\mathbf{K}}\{\mathbf{g}^T\mathbf{g} - 2\hat{\mathbf{v}}^T\mathbf{g}\}. \tag{6.19}$$

For simplicity, we restrict the description of our adaptive projection to the homogeneous obstacle problem, $\mathcal{K} = \left\{v \in V \;\middle|\; v \geq 0\right\}$. For more general convex sets we refer to Section 5. As we use preconditioned sequences $\mathbf{u} \in \ell_2$, we modify $\mathbf{K}$ to be

$$\mathbf{K} := \{\mathbf{v} \in \ell_2(\mathcal{J}) :\ (\mathbf{D}^{-1}\mathbf{v})^T\Psi(x) \geq 0\}.$$

Since $\mathbf{K}$ is infinite-dimensional, we cannot solve the corresponding quadratic problem. The aim is to replace $\mathbf{K}$ by some finite $\mathbf{K}_\eta$ so that the solution $\mathbf{g}_\eta$ of the finite quadratic program still satisfies $||\mathbf{g}_\eta - \mathbf{P_K}\hat{\mathbf{v}}\| \leq \eta$. To this end, we use the scheme **PREDICTION**, [8] that results in a finite index set $\hat{\Lambda} = \hat{\Lambda}(\mathbf{P_K}\hat{\mathbf{v}}, \eta)$ containing all significant indices of $\mathbf{P_K}\hat{\mathbf{v}}$, i.e., $\|(\mathbf{P_K}\hat{\mathbf{v}})\big|_{\hat{\Lambda}} - \mathbf{P_K}\hat{\mathbf{v}}\| \leq \eta$. Then, we set

$$\mathbf{K}_\eta := \{\mathbf{v} \in \ell_2(\hat{\Lambda}) :\ (\mathbf{D}^{-1}\mathbf{v})^T\Psi(x) \geq 0\},$$

which is finite-dimensional. With these settings, we obtain Algorithm 6.1.

Algorithm 6.1 $\mathbf{PROJECT}[\eta, \hat{\mathbf{v}}] \to \mathbf{w}_\eta$

1: Predict the set $\hat{\Lambda}$ as $\mathbf{PREDICTION}[P_{\mathcal{K}}(\hat{v}), \eta]$
2: Solve the quadratic program on $\hat{\Lambda}$

$$\mathbf{w}_\eta := \arg\min_{\mathbf{u}\in\mathbf{K}_\eta}\{\mathbf{u}^T\mathbf{u} - 2\hat{\mathbf{v}}^T\mathbf{u}\} \tag{6.20}$$

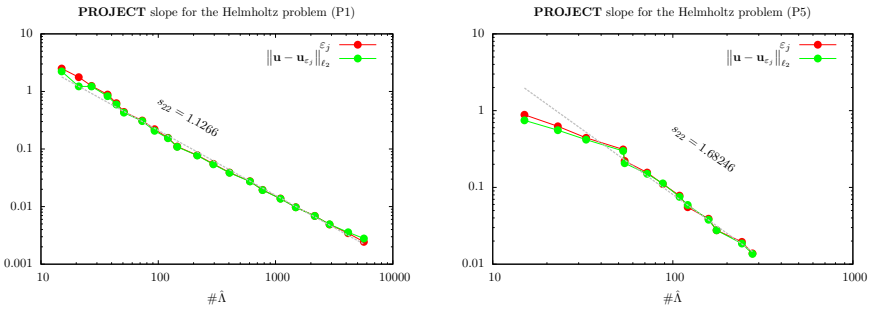
The non-negativity of $\mathbf{K}_\eta$ may again be controlled by the *local reconstruction*, see [22, Subsection 4.5.3].

Lemma 6.4. *The adaptive projection of Algorithm 6.1 computes an approximation* $\mathbf{w}_\eta$ *to* $\mathbf{P}_\mathbf{K}\hat{\mathbf{v}}$ *that satisfies* $\|\mathbf{P}_\mathbf{K}\hat{\mathbf{v}} - \mathbf{w}_\eta\| \leq \eta$ *with* $\#\hat{\Lambda} \lesssim \|\mathbf{u}\|_{\mathcal{A}^s}\eta^{-1/s}$, *with* $s \leq s^*$ *from (6.17). Moreover, we have* $\mathbf{P}_\mathbf{K}\hat{\mathbf{v}} \in \mathcal{A}^s$ *and* $\|\mathbf{P}_\mathbf{K}\hat{\mathbf{v}}\|_{\mathcal{A}^s} \lesssim 1 + \|\hat{\mathbf{v}}\|_{\mathcal{A}^s}$. *The computational complexity is* $\mathcal{O}(N^{1.75})$.

Proof. The error bound for the approximation on the tree $\mathbf{K}_\eta$ follows from [9, Theorem 3.2] as indicated in the analysis of **RECOVERY**. □

Remark 6.5. *Obviously, the rate* $N^{1.75}$ *is not optimal. This comes* only *from the complexity of solving the quadratic program. Thus, if we can replace this by a direct pointwise evaluation of the projection, we obtain optimal complexity. This can e.g. easily be done if the projection reduces to* $\max\{u, 0\}$. *This remark applies to all subsequent remarks concerning complexity.*

6.2.3. *Numerical tests for* **PROJECT**. Now, we test the approximation properties of Algorithm 6.1. We measure the true error and compare it with the tolerances. Our test is inspired by the way the projection appears in our adaptive algorithm (6.12). Thus, first we sample the solutions of the Helmholtz inequality problem from Section 4 at a maximal level of $J = 18$ resulting in some $\mathbf{u}$. Then, for some given tolerances $\varepsilon_j = 2^{-j/2}$ we first threshold $\mathbf{u}$ with **COARSE** to get $\mathbf{v}_{\varepsilon_j}$. With $\mathbf{v}_{\varepsilon_j}$ we compute with **RHS** and **APPLY** the approximations $\mathbf{w}_{\varepsilon_j}$ and $\mathbf{f}_{\varepsilon_j}$ resulting in an approximation $\mathbf{r}_{\varepsilon_j}$ of the residual $\mathbf{r} = (\mathbf{f} - \mathbf{A}\mathbf{u})$. Then, the projected Richardson iteration takes the form $\hat{\mathbf{u}}_{\varepsilon_j} = \mathbf{v}_{\varepsilon_j} + \alpha_\Psi \mathbf{r}_{\varepsilon_j}$ with the relaxation parameter $\alpha_\Psi = \frac{c_A}{C_A^2}$ and c_A, C_A from (3.4). Finally, we take $\hat{\mathbf{u}}_{\varepsilon_j}$ as input of Algorithm 6.1 with the tolerance ε_j to get $\mathbf{u}_{\varepsilon_j}$. Since $\mathbf{u}$ is a fixed point of $\mathbf{S}$, we can measure the ℓ_2-error of the application of $\mathbf{S}$ by $\|\mathbf{u} - \mathbf{u}_{\varepsilon_j}\|$.



(a) (P1), $d = \tilde{d} = 2$, error (b) (P5), $d = \tilde{d} = 2$, error

FIGURE 6.3. **PROJECT**, (P1) and (P5) for $d = \tilde{d} = 2$.

In Figure 6.3, we have shown the results for (P1) and (P5). As we see, the tolerances are met very good. The other examples give similar results. Convergence rates and complexity for all examples are shown in Table 6.3. We obtain rates of about $s \approx 1.2$ for the smooth examples (P1) to (P3) and also for (P4). For the examples (P5) and (P6), we get better rates. In all cases, the expected rate of $s^* = 1$ is excelled. Moreover, the complexity of solving the quadratic program (6.20) is in the range $\mathcal{O}(N^{1.37})$– $\mathcal{O}(N^{1.57})$ and therefore a bit superior than our analysis.

	(P1)	(P2)	(P3)	(P4)	(P5)	(P6)
s	1.13	1.26	1.08	1.26	1.68	1.68
complexity $N^\bullet$	1.38	1.44	1.39	1.56	1.55	1.57

TABLE 6.3. Convergence rates and complexity of **PROJECT**.

6.2.4. *Computable adaptive method.* With the adaptive projection at hand, we now have all ingredients for a completely computable variant of (6.12). The design of the method draws heavily on the adaptive wavelet algorithm developed in [7]. Fixing the constant $K := \min\left\{\ell \in \mathbb{N} : \ ((1+2\alpha)\ell + \rho)\rho^{\ell-1} \leq \frac{1}{10}\right\}$, we arrive at **EVISOLVE** in Algorithm 6.2.

Algorithm 6.2 EVISOLVE$[\varepsilon, \mathbf{A}, \mathbf{f}] \to \mathbf{u}(\varepsilon)$

Initialization $\mathbf{u}_\Lambda^0 = \mathbf{0}$, $\varepsilon_0 = c_A^{-1}\|\mathbf{f}\|$, $j = 0$
while $\varepsilon_j > \varepsilon$ **do**
  $\mathbf{v}_0 = \mathbf{u}_\Lambda^j$
  for $\ell = 0, \ldots, K-1$ **do**
    RHS$[\rho^\ell \varepsilon_j, \mathbf{f}] \to \mathbf{f}_\ell$
    APPLY$[\rho^\ell \varepsilon_j, \mathbf{A}, \mathbf{v}^\ell] \to \mathbf{w}_\ell$
    $\hat{\mathbf{v}}^{\ell+1} = \mathbf{v}^\ell + \alpha\,(\mathbf{f}_\ell - \mathbf{w}_\ell)$
    PROJECT$[\rho^\ell \varepsilon_j, \hat{\mathbf{v}}^{\ell+1}] \to \mathbf{v}^{\ell+1}$
  end for
  COARSE$[\frac{4}{10}\varepsilon_j, \mathbf{v}^K] \to \mathbf{u}_\Lambda^{j+1}$
  $\varepsilon_{j+1} = \varepsilon_j/2$; $j \to j+1$
end while
return $\mathbf{u}(\varepsilon) := \mathbf{u}_\Lambda^j$

We obtain the following convergence result.

Proposition 6.6. *The iterates $\mathbf{u}_\Lambda^j$ of Algorithm 6.2 satisfy $\|\mathbf{u} - \mathbf{u}_\Lambda^j\| \leq \varepsilon_j$, where $\varepsilon_j = 2^{-j}\varepsilon_0$.*

Proof. From (3.4) and a stability estimate[2] we have $\|\mathbf{u} - \mathbf{u}_\Lambda^0\| = \|\mathbf{u}\| \leq \|\mathbf{A}^{-1}\|\|\mathbf{f}\| \leq c_A^{-1}\|\mathbf{f}\| = \varepsilon_0$. We proceed by induction and assume that for some $j \geq 0$ we have $\|\mathbf{u} - \mathbf{u}_\Lambda^j\| \leq \varepsilon_j$. If we denote by $\mathbf{u}^j(\mathbf{v})$ the exact iterates of (6.12) with initial guess $\mathbf{v}$, we get

$$\begin{aligned}\|\mathbf{v}^{\ell+1} - \mathbf{u}^{\ell+1}(\mathbf{u}_\Lambda^j)\| &\leq \|\mathbf{v}^{\ell+1} - \mathbf{P}_\mathbf{K}(\mathbf{v}^\ell + \alpha(\mathbf{f}_\ell - \mathbf{w}_\ell))\| \\ &\quad + \|\mathbf{P}_\mathbf{K}(\mathbf{v}^\ell + \alpha(\mathbf{f}_\ell - \mathbf{w}_\ell)) - \mathbf{P}_\mathbf{K}(\mathbf{u}^\ell(\mathbf{u}_\Lambda^j) + \alpha(\mathbf{f} - \mathbf{A}\mathbf{u}^\ell(\mathbf{u}_\Lambda^j)))\|\end{aligned}$$

[2]The stability estimate $\|\mathbf{u}\| \leq \|\mathbf{A}^{-1}\|\|\mathbf{f}\|$ is only valid if $\mathbf{0} \in \mathbf{K}$. If this is not the case, another estimate depending on the type of inequality has to be used.

Next, by using the properties of **PROJECT** and the non-expansiveness of $\mathbf{P}_\mathbf{K}$

$$\begin{aligned}\|\mathbf{v}^{\ell+1} - \mathbf{u}^{\ell+1}(\mathbf{u}_\Lambda^j)\| &\leq \eta_\ell + \|\mathbf{P}_\mathbf{K}(\mathbf{v}^\ell + \alpha(\mathbf{f}_\ell - \mathbf{w}_\ell)) - \mathbf{P}_\mathbf{K}(\mathbf{u}^\ell(\mathbf{u}_\Lambda^j) + \alpha(\mathbf{f} - \mathbf{A}\mathbf{u}^\ell(\mathbf{u}_\Lambda^j)))\| \\ &\leq \eta_\ell + \|(\mathbf{v}^\ell + \alpha(\mathbf{f}_\ell - \mathbf{w}_\ell)) - (\mathbf{u}^\ell(\mathbf{u}_\Lambda^j) + \alpha(\mathbf{f} - \mathbf{A}\mathbf{u}^\ell(\mathbf{u}_\Lambda^j)))\| \\ &= \eta_\ell + \|(\mathbf{I} - \alpha\mathbf{A})(\mathbf{v}^\ell - \mathbf{u}^\ell(\mathbf{u}_\Lambda^j)) + \alpha(\mathbf{f}_\ell - \mathbf{f}) + \alpha(\mathbf{A}\mathbf{v}^\ell - \mathbf{w}_\ell)\| \\ &\leq \rho\|\mathbf{v}^\ell - \mathbf{u}^\ell(\mathbf{u}_\Lambda^j)\| + \eta_\ell(1+2\alpha) \\ &\leq \rho^\ell\|\mathbf{v}^0 - \mathbf{u}^0(\mathbf{u}_\Lambda^j)\| + (1+2\alpha)\sum_{k=0}^{\ell}\eta_k\rho^{\ell-k} = (1+2\alpha)\sum_{k=0}^{\ell}\eta_k\rho^{\ell-k},\end{aligned}$$

with $\eta_\ell = \rho^\ell\varepsilon_j$. Furthermore, with the definition of η_ℓ we conclude

$$\begin{aligned}\|\mathbf{v}^K - \mathbf{u}^K(\mathbf{u}_\Lambda^j)\| &\leq (1+2\alpha)\sum_{k=0}^{K-1}\eta_k\rho^{K-1-k} = (1+2\alpha)\sum_{k=0}^{K-1}\rho^k\varepsilon_j\rho^{K-1-k} \\ &= (1+2\alpha)K\rho^{K-1}\varepsilon_j.\end{aligned}$$

Therefore we obtain by the definition of K

$$\begin{aligned}\|\mathbf{v}^K - \mathbf{u}\| &\leq \|\mathbf{v}^K - \mathbf{u}^K(\mathbf{u}_\Lambda^j)\| + \|\mathbf{u}^K(\mathbf{u}_\Lambda^j) - \mathbf{u}\| \\ &\leq \|\mathbf{v}^K - \mathbf{u}^K(\mathbf{u}_\Lambda^j)\| + \rho^K\|\mathbf{u}_\Lambda^j - \mathbf{u}\| \\ &\leq \|\mathbf{v}^K - \mathbf{u}^K(\mathbf{u}_\Lambda^j)\| + \rho^K\varepsilon_j \leq ((1+2\alpha)K + \rho)\rho^{K-1}\varepsilon_j \leq \frac{1}{10}\varepsilon_j\end{aligned}$$

and finally with the properties of **COARSE** we arrive at

$$\|\mathbf{u} - \mathbf{u}_\Lambda^{j+1}\| \leq \|\mathbf{u} - \mathbf{v}^K\| + \|\mathbf{v}^K - \mathbf{u}_\Lambda^{j+1}\| \leq \frac{1}{10}\varepsilon_j + \frac{4}{10}\varepsilon_j = \frac{\varepsilon_j}{2},$$

which concludes the proof. □

Remark 6.7. *Similar to* [7]*, the constant K in the inner loop of Algorithm 6.2 could be quite large. Moreover, in the case of the projected Richardson iteration, it is not possible to use the approximate residuals $\mathbf{w}_\ell - \mathbf{f}_\ell$ as an indicator for $\|\mathbf{u} - \mathbf{v}_\ell\|$. Of course we could exit the inner loop earlier if we could guarantee that $\|\mathbf{u} - \mathbf{v}^\ell\| \leq \frac{1}{10}\varepsilon_j$. We will consider this issue later.*

Remark 6.8. *Note that the new iterate $\mathbf{v}^{\ell+1}$ in the inner loop may not be the best approximation in the linear space $\Lambda^{\ell+1} = \operatorname{supp}\left(\mathbf{v}^{\ell+1}\right)$. Therefore, a few iterations of Algorithm 5.1 should decrease the number of inner loops substantially. This is one of the reasons why we introduced the scheme for a given index set Λ. It is also important to apply Algorithm 5.1 to the iterate $\mathbf{u}_\Lambda^{j+1}$ after the coarsening step in order to ensure $\mathbf{u}_\Lambda^{j+1} \in \mathbf{K}$.*

We continue with the discussion of the convergence rate of Algorithm 6.2. For $0 < \tau < 2$ we denote by $\mathbf{v}^*$ the non-decreasing rearrangement, set

$$|\mathbf{v}|_{\ell_\tau^{\mathrm{w}}(\mathcal{J})} := \sup_{n\in\mathbb{N}} n^{1/\tau}|v_n^*|$$

and $\|\mathbf{v}\|_{\ell_\tau^{\mathrm{w}}(\mathcal{J})} := \|\mathbf{v}\|_{\ell^2(\mathcal{J})} + |\mathbf{v}|_{\ell_\tau^{\mathrm{w}}(\mathcal{J})}$. We recall some well-known facts from [6].

Proposition 6.9. *Let $\tau = (s + \frac{1}{2})^{-1}$, then $\mathbf{v}$ belongs to $\ell_\tau^{\mathrm{w}}(\mathcal{J})$ if and only if $\|\mathbf{v} - \mathbf{v}_N\| = \mathcal{O}(N^{-s})$ and we have $\|\mathbf{v} - \mathbf{v}_N\|_{\ell_\tau^{\mathrm{w}}(\mathcal{J})} \lesssim N^{-s}\|\mathbf{v}\|_{\ell_\tau^{\mathrm{w}}(\mathcal{J})}$.* □

Proposition 6.10. *For any threshold $\eta > 0$ the output $\mathbf{w} = \mathbf{COARSE}[\eta, \mathbf{v}]$ satisfies $\|\mathbf{w}\|_{\ell_\tau^{\mathrm{w}}(\mathcal{J})} \leq \|\mathbf{v}\|_{\ell_\tau^{\mathrm{w}}(\mathcal{J})}$. Moreover, assume that for a sequence $\mathbf{u}$ and a tolerance $0 < \eta < \|\mathbf{u}\|_{\ell_2(\mathcal{J})}$ an approximation $\mathbf{v}$ of $\mathbf{u}$ is given satisfying $\|\mathbf{u} - \mathbf{v}\|_{\ell_2(\mathcal{J})} \leq \frac{\eta}{5}$. If in addition $\mathbf{u} \in \ell_\tau^{\mathrm{w}}(\mathcal{J})$ for $\tau = (s + \frac{1}{2})^{-1}$ and some $s > 0$, then we get for the output $\mathbf{w}$ of $\mathbf{COARSE}[4\eta/5, \mathbf{v}]$ that $\|\mathbf{u} - \mathbf{w}\|_{\ell_2(\mathcal{J})} \leq \eta$ and:*

(1) *The cardinality of $\Lambda = \operatorname{supp}(\mathbf{w})$ is bounded by $\#\Lambda \lesssim \|\mathbf{u}\|^{1/s}_{\ell_\tau^{\mathrm{w}}(\mathcal{J})} \eta^{-1/s}$,*
(2) *it holds $\|\mathbf{u} - \mathbf{w}\|_{\ell_2(\mathcal{J})} \lesssim \|\mathbf{u}\|_{\ell_\tau^{\mathrm{w}}(\mathcal{J})} (\#\Lambda)^{-s}$, and*
(3) $\|\mathbf{w}\|_{\ell_\tau^{\mathrm{w}}(\mathcal{J})} \lesssim \|\mathbf{u}\|_{\ell_\tau^{\mathrm{w}}(\mathcal{J})}$. □

We use these known results in order to investigate under which conditions Algorithm 6.2 exhibits asymptotically optimal complexity. The complexity of the algorithms **APPLY** and **RHS** has been shown in [6] to be of the order $\mathcal{O}(\eta^{-1/s})$. The algorithm **PROJECT** was shown also to have this complexity. Finally, the cardinality of the output of **PROJECT** is of the order $\mathcal{O}(\eta^{-1/s})$ w.r.t. the input which in turns is the output of **APPLY** and **RHS**, so that the overall cardinality is $\mathcal{O}(\eta^{-1/s})$. However, the arithmetic operations needed to solve the QP is of the order $\mathcal{O}(N^{1.75})$. We summarize our findings concerning Algorithm 6.2.

Theorem 6.11. *Assume that (6.14) holds. Under the assumptions on* **APPLY**, **RHS**, **PROJECT** *and* **COARSE** *for any target accuracy ε the scheme* **EVISOLVE** *determines an approximate solution $\mathbf{u}(\varepsilon)$ of (3.5) after a finite number of steps such that $\|\mathbf{u} - \mathbf{u}(\varepsilon)\| \leq \varepsilon$ Moreover, if the routines* **APPLY**, **RHS** *and* **PROJECT** *are s^*-admissible for some $s^* > 0$ and if the solution $\mathbf{u}$ of (3.5) has a convergence rate $s < s^*$ of its best N-term approximation, then the following statements hold:*

(1) *The support of $\mathbf{u}(\varepsilon)$ is bounded by $\#\operatorname{supp} \mathbf{u}(\varepsilon) = \#\Lambda(\varepsilon) \lesssim \|\mathbf{u}\|^{1/s}_{\ell_\tau^{\mathrm{w}}(\mathcal{J})} \varepsilon^{-1/s}$.*
(2) *We have $\|\mathbf{u}(\varepsilon)\|_{\ell_\tau^{\mathrm{w}}(\mathcal{J})} \lesssim \|\mathbf{u}\|_{\ell_\tau^{\mathrm{w}}(\mathcal{J})}$.*
(3) *The number of arithmetic and sort operations is $\mathcal{O}(\varepsilon^{-1.75/s})$.*

Proof. The first statement follows from Proposition 6.6. Now we assume that the error of the best N-term approximation of $\mathbf{u}$ is smaller than CN^{-s} for some constant $C > 0$ and some $0 < s < s^*$. Then, by Proposition 6.9 we have that $\mathbf{u} \in \ell_\tau^{\mathrm{w}}(\mathcal{J})$ and also $\mathbf{f} \in \ell_\tau^{\mathrm{w}}(\mathcal{J})$ as well as $\|\mathbf{f}\|_{\ell_\tau^{\mathrm{w}}(\mathcal{J})} \lesssim \|\mathbf{u}\|_{\ell_\tau^{\mathrm{w}}(\mathcal{J})}$. Therefore, we are only left to show that for $j \in \mathbb{N}$ the iterates $\mathbf{u}_\Lambda^j$ satisfy (1)–(3) with $\varepsilon > 0$, $\mathbf{u}(\varepsilon)$, $\Lambda(\varepsilon)$ replaced by ε_j, $\mathbf{u}_\Lambda^j$ and Λ_j. Of course this holds for $j = 0$. Now suppose that (1)–(3) are true for some $j \geq 0$. Then, we get that the application of **RHS** in Algorithm 6.2 requires

$$N_j^{\mathbf{RHS}} \lesssim (\sum_{\ell=0}^{K} \rho^{-\ell/s}) \|\mathbf{f}\|^{1/s}_{\ell_\tau^{\mathrm{w}}(\mathcal{J})} \varepsilon^{-1/s}$$

floating point and sorting operations. A similar statement holds true for the approximate matrix-vector multiplication, namely

$$N_j^{\mathbf{APPLY}} \lesssim \varepsilon_j^{-1/s} \sum_{\ell=0}^{K} \rho^{-\ell/s} (\|\mathbf{v}^\ell\|^{1/s}_{\ell_\tau^{\mathrm{w}}(\mathcal{J})} + \#\operatorname{supp} \mathbf{v}^\ell).$$

Finally, the number $N_j^{\mathbf{PROJECT}}$ of flops for **PROJECT** is bounded by $\mathcal{O}(N^{1.75})$ and govern the complexity of the whole scheme. Because of for $\mathbf{w} = \mathbf{PROJECT}[\eta, \mathbf{v}]$ we have that $\|\mathbf{w}\|_{\ell_\tau^{\mathrm{w}}(\mathcal{J})} \lesssim \|\mathbf{v}\|_{\ell_\tau^{\mathrm{w}}(\mathcal{J})}$ and in view of the inner loop in **EVISOLVE**

it follows that

$$\begin{aligned}\|\mathbf{v}^{\ell+1}\|_{\ell^{\mathrm{w}}_\tau(\mathcal{J})} &\lesssim \|\mathbf{v}^\ell + \alpha(\mathbf{w}_\ell - \mathbf{f}_\ell)\|_{\ell^{\mathrm{w}}_\tau(\mathcal{J})} \leq \|\mathbf{v}^\ell\|_{\ell^{\mathrm{w}}_\tau(\mathcal{J})} + \alpha(\|\mathbf{w}_\ell\|_{\ell^{\mathrm{w}}_\tau(\mathcal{J})} + \|\mathbf{f}_\ell\|_{\ell^{\mathrm{w}}_\tau(\mathcal{J})}) \\ &\lesssim \|\mathbf{f}\|_{\ell^{\mathrm{w}}_\tau(\mathcal{J})} + \|\mathbf{v}^\ell\|_{\ell^{\mathrm{w}}_\tau(\mathcal{J})}.\end{aligned}$$

Since K is fixed and bounded we conclude that

$$\max_{\ell \leq K} \|\mathbf{v}^\ell\|_{\ell^{\mathrm{w}}_\tau(\mathcal{J})} \lesssim \|\mathbf{f}\|_{\ell^{\mathrm{w}}_\tau(\mathcal{J})} + \|\mathbf{u}^j_\Lambda\|_{\ell^{\mathrm{w}}_\tau(\mathcal{J})} \lesssim \|\mathbf{f}\|_{\ell^{\mathrm{w}}_\tau(\mathcal{J})} + \|\mathbf{u}\|_{\ell^{\mathrm{w}}_\tau(\mathcal{J})}$$

where we have used (2) as the induction assumption. In summary, we obtain

$$\begin{aligned}N_j^{\mathbf{RHS}}, N_j^{\mathbf{APPLY}} &\lesssim \left(\|\mathbf{f}\|_{\ell^{\mathrm{w}}_\tau(\mathcal{J})} + \|\mathbf{u}\|_{\ell^{\mathrm{w}}_\tau(\mathcal{J})}\right) \varepsilon_j^{-1/s}, \\ N_j^{\mathbf{PROJECT}} &\lesssim \left(\|\mathbf{f}\|_{\ell^{\mathrm{w}}_\tau(\mathcal{J})} + \|\mathbf{u}\|_{\ell^{\mathrm{w}}_\tau(\mathcal{J})}\right) \varepsilon_j^{-1.75/s}.\end{aligned}$$

We conclude that the support of the inner iterates stay controlled according to

$$\#\mathrm{supp}\ \mathbf{v}^\ell \lesssim (\|\mathbf{f}\|_{\ell^{\mathrm{w}}_\tau(\mathcal{J})} + \|\mathbf{u}\|_{\ell^{\mathrm{w}}_\tau(\mathcal{J})})\varepsilon_j^{-1/s}, \qquad \ell \leq K.$$

Recall that after at most K inner iterates of **EVISOLVE** we apply **COARSE** on $\mathbf{v}^K$ to get $\mathbf{u}_\Lambda^{j+1}$. Hence, we have $\|\mathbf{u}_\Lambda^{j+1}\|_{\ell^{\mathrm{w}}_\tau(\mathcal{J})} \lesssim \|\mathbf{u}\|_{\ell^{\mathrm{w}}_\tau(\mathcal{J})}$ as well as $\#\Lambda_{j+1} \lesssim \|\mathbf{u}\|^{1/s}\varepsilon_{j+1}^{-1/s}$. Since the coarsening is of constant order, this concludes the proof. □

6.2.5. *Numerical Results for* **EVISOLVE**. Now we want to conduct tests for **EVISOLVE**. In particular, we want to perform tests how to leave the inner loop (lines 4-9) since the estimated number K may be by far too large. In order that our adaptive wavelet solver applied to the smooth examples (P1) to (P3) should reach the optimal rate of $s = -2$, we would have to use a biorthogonal wavelet basis with at least $d = \tilde{d} = 3$. However, using wavelets from [12] would yield a minimal level for $d = \tilde{d} = 3$ of $j_0 = 4$ and $C_A = 2.35$, $c_A = 0.0003$, $\kappa = 7704$, $\alpha_\Psi = 0.000052$, $\rho_\Psi \approx 1$ so that we would need $K \approx 3.000.000.000$ inner iterations in **EVISOLVE**. Of course, this situation may change using other bases.

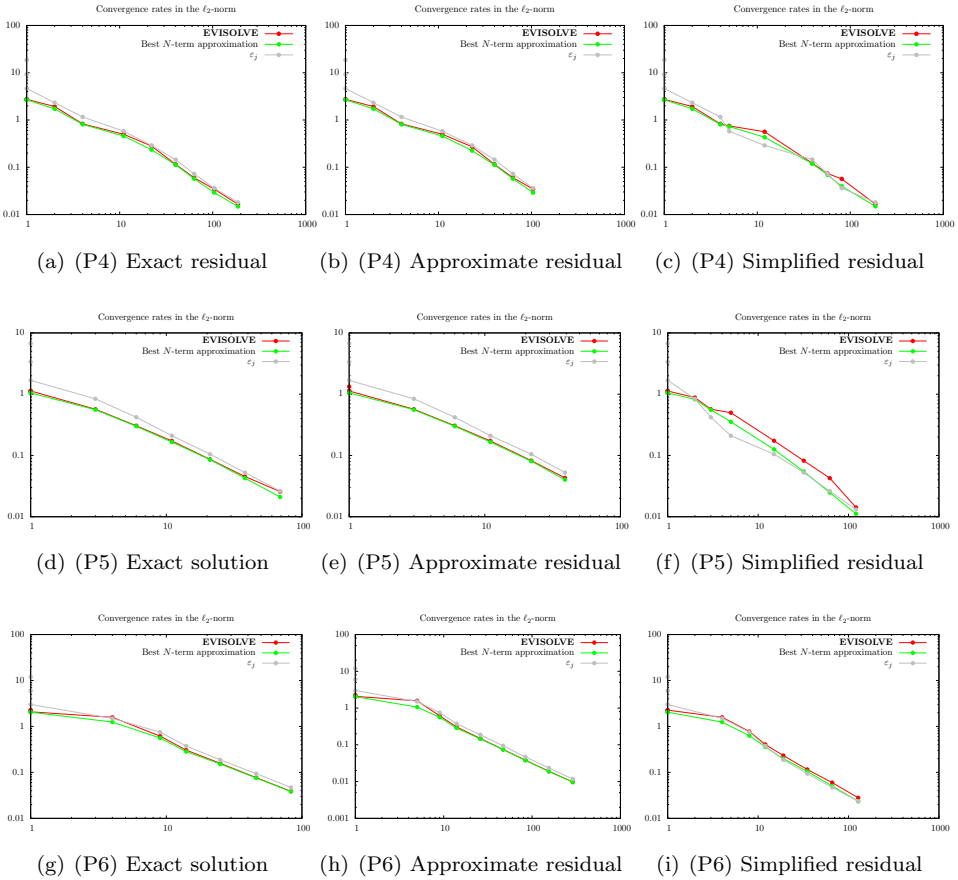
Because of this observation we test our adaptive wavelet solver only on the singular examples (P4), (P5) and (P6), where it suffices to chose the parameters of the underlying wavelet basis as $d = \tilde{d} = 2$. To this end, we test three variants, namely:

(1) Since in our experiments we know the analytic solution, we can compute also the exact error for the inner iteration. This is of course only used for comparison as a benchmark and dubbed exact residual.
(2) We take the approximate residual based upon the following observation. For $r^j := |(\mathbf{v}^j, \mathbf{A}\mathbf{v}^j - \mathbf{f})|$ we get

$$\begin{aligned}r^j &\leq |(\mathbf{v}^\ell, \mathbf{A}\mathbf{v}^\ell - \mathbf{w}_\ell)| + |(\mathbf{v}^\ell, \mathbf{f}_\ell - \mathbf{f})| + |(\mathbf{v}^\ell, \mathbf{w}_\ell - \mathbf{f}_\ell)| \\ &\leq 2\|\mathbf{v}^\ell\|\eta + |(\mathbf{v}^\ell, \mathbf{w}_\ell - \mathbf{f}_\ell)|, \qquad (6.21)\end{aligned}$$

whose components are in fact computed in each inner iteration.
(3) We take only $|(\mathbf{v}^\ell, \mathbf{w}_\ell - \mathbf{f}_\ell)|$ as a simplified variant of (6.21).

The comparison of the convergence history for the examples (P4) to (P6) are shown in Figure 6.4. We see not much differences in the slope using the variants of the residual from the use of the exact solution. This shows indeed that K as above is heavily overestimated and that our residual estimators seem appropriate for building a stopping criterion upon it. Moreover, the bound ε_j from Proposition 6.6 is in fact valid for the exact and the approximate residual.



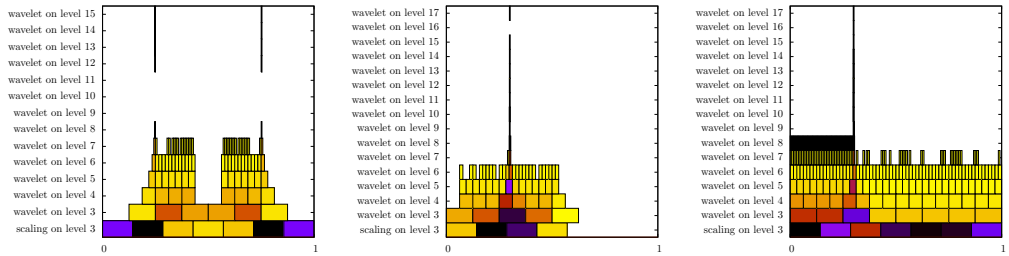
(a) (P4) Exact residual (b) (P4) Approximate residual (c) (P4) Simplified residual

(d) (P5) Exact solution (e) (P5) Approximate residual (f) (P5) Simplified residual

(g) (P6) Exact solution (h) (P6) Approximate residual (i) (P6) Simplified residual

FIGURE 6.4. Leaving the inner loop of **EVISOLVE** with different stopping criteria.

The plots in Figure 6.5 demonstrate that the wavelet coefficients clearly reflect the singularities also when using the approximate residual.



(a) (P4) $N = 104$ (b) (P5) $N = 70$ (c) (P6) $N = 289$

FIGURE 6.5. Wavelet coefficients of the iterates.

6.3. **S-ADWAV-EVISOLVE for EVIs.** Inspired by [1], we also propose a simplified variant of our adaptive solver for the elliptic inequality (3.5), which also works with a *security zone* $\mathcal{C}$. The fundamental difference to [1, Algorithm 3.10] is that we cannot use the residual $\mathbf{r}_{\partial\Lambda_j} = \mathbf{f}_{\partial\Lambda_j} - \mathbf{A}_{\partial\Lambda_j\times\Lambda_j}\mathbf{u}^j$ on some set $\partial\Lambda_j$ for the estimation of the error $\|\mathbf{u}-\mathbf{u}^j\|_{\ell_2}$ simply because for the elliptic inequality this term does not converge to zero as $\mathbf{u}^j \overset{j\to\infty}{\longrightarrow} \mathbf{u}$. Therefore we take a different heuristic approach for this issue.

Remark 6.12 (Error estimation for **S-ADWAV-EVISOLVE**)**.** *From Remark 4.1 we know, at least for elliptic inequalities that are driven by a self-adjoint operator A, that the error bound $\|u-u_\Lambda\|_a \le 2\sqrt{|\Pi(u)-\Pi(u_\Lambda)|}$ is valid. Furthermore, as the energy norm is equivalent to the $\|\cdot\|_{\ell_2}$-norm, we have also*

$$\frac{1}{2}\|\mathbf{u}-\mathbf{u}_\Lambda\|_{\ell_2} \lesssim \sqrt{|\Pi(\mathbf{u})-\Pi(\mathbf{u}_\Lambda)|} = \sqrt{\left|\Pi(\mathbf{u}) - \left(\frac{1}{2}\mathbf{u}_\Lambda^T\mathbf{A}_{\Lambda\times\Lambda}\mathbf{u}_\Lambda - \mathbf{f}_\Lambda^T\mathbf{u}_\Lambda\right)\right|}.$$

That means, given the optimal value of the energy functional $\Pi(\mathbf{u})$, we have an error estimator. Unfortunately, for a general $\mathbf{u}$ this value is not known. Our approach is now to estimate this value from the values of the approximations $\mathbf{u}_\Lambda$ and use an extrapolation for the calculation of $\hat{\Pi}(\mathbf{u}_\Lambda) \approx \Pi(\mathbf{u})$. If it holds then that $\hat{\Pi}(\mathbf{u}_\Lambda)$ converges faster to $\Pi(\mathbf{u})$ than $\mathbf{u}_\Lambda$ converges to $\mathbf{u}$, the quantity $2\sqrt{\left|\hat{\Pi}(\mathbf{u}_\Lambda)-\Pi(\mathbf{u}_\Lambda)\right|}$ will be a reliable error estimator.

For the example calculations in Subsection 6.3.1 below we use a linear polynomial extrapolation (denoted by **EXTRAPOLATE**) as described in [21, Section 3.2].

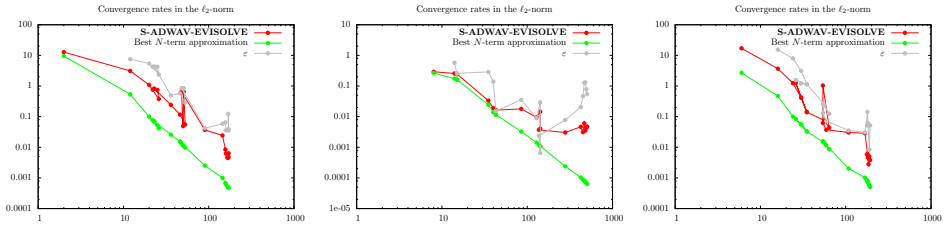
Algorithm 6.3 S-ADWAV-EVISOLVE$[\varepsilon, \mathbf{A}, \mathbf{f}] \to \mathbf{u}(\varepsilon)$

1: Initialization $j=1$, $\mathbf{u}^0 = \mathbf{0}$, $\Lambda_0 = \mathcal{I}_{j_0}$
2: **repeat**
3: $\quad\mathbf{v} =$**PRICHARDSON**$[\Lambda_j, \mathbf{A}_{\Lambda_j\times\Lambda_j}, \mathbf{u}^{j-1}, \mathbf{f}_{\Lambda_j}]$
4: $\quad\mathbf{w} =$**COARSE**$[\mathrm{tol}_{\mathrm{off}}, \mathbf{v}]$
5: $\quad\Lambda_j = \mathrm{supp}\ (\mathbf{w})$
6: $\quad\mathbf{u}^j =$ **PRICHARDSON**$[\Lambda_j, \mathbf{A}_{\Lambda_j\times\Lambda_j}, \mathbf{w}, \mathbf{f}_{\Lambda_j}]$
7: $\quad\hat{\Pi}(\mathbf{u}^j) =$**EXTRAPOLATE**$[\mathbf{u}^j]$
8: $\quad\Pi(\mathbf{u}^j) = \frac{1}{2}\left(\mathbf{u}^j, \mathbf{A}_{\Lambda_j\times\Lambda_j}\mathbf{u}^j\right)_{\ell_2} - \left(\mathbf{f}_{\Lambda_j}, \mathbf{u}^j\right)_{\ell_2}$
9: $\quad\varepsilon_j = 2\sqrt{\left|\hat{\Pi}(\mathbf{u}^j)-\Pi(\mathbf{u}^j)\right|}$
10: $\quad\partial\Lambda_j = \mathcal{C}(\Lambda_j, c)$
11: $\quad\mathbf{r}_{\partial\Lambda_j} = \mathbf{f}_{\partial\Lambda_j} - \mathbf{A}_{\partial\Lambda_j\times\Lambda_j}\mathbf{u}^j$
12: $\quad\Lambda_{j+1} = \Lambda_j \cup \mathrm{supp}\left(\mathbf{COARSE}\ [\mathrm{tol}_{\mathrm{on}}, \mathbf{r}_{\partial\Lambda_j}]\right)$
13: **until** $\varepsilon_j < \varepsilon$

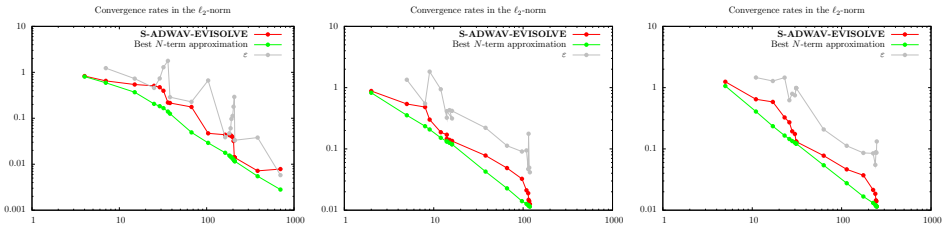
A few comments on Algorithm 6.3 are in order. First of all, we have to make an additional call to the Galerkin solver in line 6 because there is no guarantee that the thresholded vector $\mathbf{w}$ is an element of the convex set $\mathbf{K}$. The next part of the algorithm consists of the error estimation as described Remark 6.12. Finally, we have to decide how new coefficients are added to the set Λ. Here we resort to the approach from [1], although there will be large coefficients in $\mathbf{r}_{\partial\Lambda_j}$ that do

not correspond to significant coefficients of the solution $\mathbf{u}$. For example, in the homogeneous elliptic obstacle problem this is the case for regions in Ω where $u \equiv 0$. However, these coefficients are erased after the next call to the Galerkin solver. In this way, our simplified adaptive algorithm iterates until the estimated error is smaller than the target accuracy ε. Likewise than in the case for the simplified adaptive solver for the elliptic equation, it is not obvious how a rigorous convergence proof for this algorithm may be established. Algorithm 6.3 might be optimized even further if the accuracy of **PRICHARDSON** is adapted to the current error ε_j.

6.3.1. *Numerical Results for* **S-ADWAV-EVISOLVE**. We describe some results of numerical experiments for the simplified adaptive wavelet method. In order to achieve a convergence rate of $s = -2$ for (P1)-(P3) as expected from the best N-term approximation in Section 4, we have to take a wavelet basis with higher regularity than $d = 2$ and so we take the biorthogonal wavelet basis with parameters $d = \tilde{d} = 3$. The results are shown in Figure 6.6, (a)-(c).



(a) (P1): Convergence. (b) (P2): Convergence. (c) (P3): Convergence.



(d) (P4): Convergence. (e) (P5): Convergence. (f) (P6): Convergence.

FIGURE 6.6. **S-ADWAV-EVISOLVE** for (P1)-(P6).

For (P1), we see that our algorithm is close to the best N-term approximation. For (P2), the simplified adaptive scheme does not perform quite well. While at the beginning the algorithm tracks the best N-term approximation it seems then difficult to equilibrate the indices in order to keep the convergence rate. Again we see that after some time the algorithm tracks well the slope of the best N-term approximation of the third example (P3).

For the singular examples (P4)-(P6), we choose the biorthogonal basis corresponding to $d = \tilde{d} = 2$ and show the results in Figure 6.6, (d)-(f). We readily confirm that in case of a singular solution the algorithm performs very well in distinguishing the relevant coefficients and captures the sharp transitions of the target function u. For (P5) the picture is somewhat similar. Lastly for (P6) it again takes

the adaptive algorithm some time to figure out where the singularity is but then it nearly drops to the line of the corresponding best N-term approximation.

7. Conclusions and Outlook

In this paper, we have introduced a fully adaptive projected Richardson scheme for solving EVIs using wavelets. We have proven its convergence and its optimality as compared to the best N-term approximation (with almost optimal complexity). We have performed several numerical tests (in 1D only) also for a simplified version of the algorithm.

We have seen that the quantitative properties of the wavelet bases (here mainly their condition number) has a severe influence on the behavior of the scheme. In particular, the number of necessary inner loops in the projected Richardson iteration turns out to be a problem. If computed exactly, this number becomes unrealistically large. To this end, we have introduced two variants of an estimate for the residual.

Note that the software Lawa, that was used to conduct all experiments, may be freely obtained from [23].

All our numerical experiments are in 1D. The reason for this is twofold. First of all, the principle behavior of our scheme can already be shown in 1D, in particular its convergence rate, its complexity and also the comparison with the best N-term approximation. Of course, we are aware that EVIs in higher dimension offer more challenging problems such as a contact zone which is a possibly non-smooth manifold in $d-1$ dimensions. Of course, our algorithm will also work in this situation but it is interesting to see the quantitative performance in representing the contact zone. This, however, is merely a question which kind of wavelet functions (e.g. isotropic, anisotropic etc.) is used. We consider this as well as EVIs arising from singular integral operators in the second part of this paper, [24].

References

[1] S. Berrone and T. Kozubek, *An Adaptive WEM Algorithm for Solving Elliptic Boundary Value Problems in Fairly General Domains*, SIAM Journal on Scientific Computing, 28 (2006), pp. 2114–2138.

[2] K. Bittner and K. Urban, *Adaptive Wavelet Methods using Semiorthogonal Spline Wavelets: Sparse Evaluation of Nonlinear Functions*, Applied and Computational Harmonic Analysis, 24 (2008), pp. 94–119.

[3] D. Braess, C. Carstensen, and R. H. Hoppe, *Convergence Analysis of a Conforming Adaptive Finite Element Method for an Obstacle Problem*, Numerische Mathematik, 107 (2007), pp. 455–471.

[4] F. Brezzi, W. W. Hager, and P. Raviart, *Error Estimates for the Finite Element Solution of Variational Inequalities – Part I. Primal Theory*, Numerische Mathematik, 28 (1977), pp. 431–443.

[5] L. A. Caffarelli, *The Obstacle Problem Revisited*, Journal of Fourier Analysis and Applications, 4 (1998), pp. 383–402.

[6] A. Cohen, W. Dahmen, and R. A. DeVore, *Adaptive Wavelet Methods for Elliptic Operator Equations: Convergence Rates*, Mathematics of Computation, 70 (2001), pp. 27–76.

[7] ———, *Adaptive Wavelet Methods II – Beyond the Elliptic Case*, Foundations of Computational Mathematics, 2 (2002), pp. 203–245.

[8] ———, *Adaptive Wavelet Schemes for Nonlinear Variational Problems*, SIAM Journal on Numerical Analysis, 41 (2003), pp. 1785–1823.

[9] ———, *Sparse Evaluation of Compositions of Functions using Multiscale Expansions*, SIAM Journal on Mathematical Analysis, 35 (2004), pp. 279–303.

[10] S. Dahlke, *Besov Regularity for Elliptic Boundary Value Problems in Polygonal Domains*, Applied Mathematics Letters, 12 (1999), pp. 31–36.
[11] S. Dahlke and R. A. DeVore, *Besov Regularity for Elliptic Boundary Value Problems*, Communications in Partial Differential Equations, 22 (1997), p. 16.
[12] W. Dahmen, A. Kunoth, and K. Urban, *Biorthogonal Spline Wavelets on the Interval–Stability and Moment Conditions*, Applied and Computational Harmonic Analysis, 6 (1999), pp. 132–196.
[13] R. S. Falk, *Error Estimates for the Approximation of a Class of Variational Inequalities*, Mathematics of Computation, 28 (1974), pp. 963–971.
[14] M. Fornasier, *Compressive Algorithms–Adaptive Solutions of PDEs and Variational Problems*, Habilitationsschrift, Faculty of Mathematics of the University of Vienna, `http://www.ricam.oeaw.ac.at/people/page/fornasier/Habil.pdf`, January 2008.
[15] T. Gantumur, *An Optimal Adaptive Wavelet Method for Nonsymmetric and Indefinite Elliptic Problems*, Journal of Computational and Applied Mathematics, 211 (2008), pp. 90–102.
[16] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and Their Applications*, Academic Press, 1980.
[17] J.-L. Lions and G. Stampacchia, *Variational Inequalities*, Communications on Pure and Applied Mathematics, 20 (1967).
[18] A.-M. Matache, P. Nitsche, and C. Schwab, *Wavelet Galerkin Pricing of American Options on Lévy Driven Assets*, Quantitative Finance, 5 (2005), pp. 403–424.
[19] G. Naldi, K. Urban, and P. Venini, *A convergent adaptive wavelet-rothe method for elastoplastic hardening*, Arabian Journal for Science and Engineering, 29 (2004), pp. 17–32.
[20] R. H. Nochetto, T. von Petersdorff, and C.-S. Zhang, *A Posteriori Error Analysis for a Class of Integro-differential Operators*. Submitted to *Numerische Mathematik*, 2008.
[21] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C++*, Cambridge University Press, 2007.
[22] M. Rometsch, *A Wavelet Tour of Option Pricing*, Dissertation, Institut für Numerische Mathematik an der Fakultät für Mathematik und Wirtschaftswissenschaften, Universität Ulm, 2010.
[23] M. Rometsch and A. Stippler, Lawa – *Library for Adaptive Wavelet Applications*. Website, 2010. Available online at `http://lawa.sourceforge.net`.
[24] M. Rometsch and K. Urban, *Adaptive Wavelet Methods for Elliptic Variational Inequalities II: Further Numerical Investigations*, preprint, Ulm University, 2010.
[25] G. Savaré, *Weak Solutions and Maximal Regularity for Abstract Evolution Inequalities*, Advances in Mathematical Ssciences and Applications, 6 (1996), pp. 377–418.
[26] K. Urban, *Wavelet Methods for Elliptic Partial Differential Equations*, Oxford University Press, 2009.

Mario Rometsch, University of Ulm, Institute for Numerical Mathematics, Helmholtzstrasse 18, D-89069 Ulm, Germany
E-mail address: `roman.rometsch@uni-ulm.de`

Karsten Urban, University of Ulm, Institute for Numerical Mathematics, Helmholtzstrasse 18, D-89069 Ulm, Germany
E-mail address: `karsten.urban@uni-ulm.de`

Ex.	g, f, u	Reg.
(P1)	$g(x) = \begin{cases} -128x^3+92x^2-20x & 0 \le x \le \frac{1}{4}, \\ -4x^2+4x-2 & \frac{1}{4} \le x \le \frac{3}{4}, \\ -128(1-x)^3+92(1-x)^2-20(1-x) & \frac{3}{4} \le x \le 1 \end{cases}$ $f(x) \equiv -32$ $u(x) = \begin{cases} 16x^2+4(1-\sqrt{10})x & 0 \le x \le \frac{\sqrt{10}}{10}, \\ g(x) & \frac{\sqrt{10}}{10} < x \le 1-\frac{\sqrt{10}}{10}, \\ 16x^2+4(\sqrt{10}-9)x+4(5-\sqrt{10}) & 1-\frac{\sqrt{10}}{10} < x \le 1. \end{cases}$	$\frac{5}{2}-\varepsilon$
(P2)	$g(x) = \exp\{-50(x-\frac{1}{2})^2\}$ $f(x) \equiv 0$, and by setting $\hat{x} := \frac{1}{4}+\sqrt{21}/20$, $\alpha := \frac{g(\hat{x})}{\hat{x}}$ $u(x) = \begin{cases} \alpha x & 0 \le x < \hat{x} \\ g(x) & \hat{x} \le x < 1-\hat{x}\ , \\ \alpha(1-x) & x \le 1 \end{cases}$	$\frac{5}{2}-\varepsilon$
(P3)	$g(x) = \begin{cases} 1-16(x-\frac{1}{2})^2 & \lvert x-\frac{1}{2}\rvert \le 0.2 \\ \frac{1000}{9}\lvert x-\frac{1}{2}\rvert^3 - \frac{248}{3}\lvert x-\frac{1}{2}\rvert^2 + \frac{40}{3}\lvert x-\frac{1}{2}\rvert + \frac{1}{9} & \lvert x-\frac{1}{2}\rvert > 0.2 \end{cases}$ $f(x) \equiv 0$, setting $\hat{x} := \frac{\sqrt{3}}{4}$ and $\alpha := \frac{g(\hat{x})}{\hat{x}}$. $u(x) = \begin{cases} \alpha x & 0 \le x \le \hat{x}, \\ g(x) & \hat{x} < x \le 1-\hat{x}, \\ \alpha(1-x) & 1-\hat{x} < x \le 1 \end{cases}$	$\frac{5}{2}-\varepsilon$
(P4)	$g(x) = \begin{cases} 1-16(x-\frac{1}{2})^2 & \lvert x-\frac{1}{2}\rvert \le 0.2 \\ \frac{1000}{9}\lvert x-\frac{1}{2}\rvert^3 - \frac{248}{3}\lvert x-\frac{1}{2}\rvert^2 + \frac{40}{3}\lvert x-\frac{1}{2}\rvert + \frac{1}{9} & 0.2 < \lvert x-\frac{1}{2}\rvert \le x_1 \\ 0 & \lvert x-\frac{1}{2}\rvert > x_1, \end{cases}$ $x_1 := \frac{189-3\sqrt{469}}{500}$, $f(x) \equiv 0$, setting $\hat{x} := \frac{\sqrt{3}}{4}$ and $\alpha := \frac{g(\hat{x})}{\hat{x}}$. $u(x) = \begin{cases} \alpha x & 0 \le x \le \hat{x}, \\ g(x) & \hat{x} < x \le 1-\hat{x}, \\ \alpha(1-x) & 1-\hat{x} < x \le 1 \end{cases}$	$\frac{3}{2}-\varepsilon$
(P5)	$g(x) = \begin{cases} -\frac{10}{3}x & 0 \le x \le \frac{3}{10}, \\ \frac{10}{7}(x-1) & \frac{3}{10} < x \le 1. \end{cases}$ $f(x) \equiv -10$, setting $x_1 := \frac{13}{210}$, $x_2 := \frac{113}{210}$, $a := 5$, $b := -\frac{83}{21}$ and $c := \frac{169}{8820}$ $u(x) = \begin{cases} g(x) & 0 \le x \le x_1, \\ ax^2+bx+c & x_1 < x \le x_2 \\ g(x) & x_2 < x \le 1, \end{cases}$	$\frac{3}{2}-\varepsilon$
(P6)	$g(x) = \begin{cases} \frac{100}{9}x^2 & 0 \le x \le \frac{3}{10}, \\ \frac{100}{49}(1-x)^2 & \frac{3}{10} < x \le 1, \end{cases}$ $f(x) \equiv 0.$ $u(x) = \begin{cases} \frac{10}{3}x & 0 \le x \le \frac{3}{10}, \\ \frac{10}{7}(1-x) & \frac{3}{10} < x \le 1, \end{cases}$	$\frac{3}{2}-\varepsilon$

TABLE 4.4. Data for our examples (P1) to (P6). The last column indicates the maximal Sobolev regularity of u.