

AN ADAPTIVE FINITE ELEMENT METHOD IN L^2 -TV-BASED IMAGE DENOISING

MICHAEL HINTERMÜLLER AND M. MONSERRAT RINCON-CAMACHO

ABSTRACT. The first order optimality system of a total variation regularization based variational model with L^2 -data-fitting in image denoising (L^2 -TV problem) can be expressed as an elliptic variational inequality of the second kind. For a finite element discretization of the variational inequality problem, an a posteriori error residual based error estimator is derived and its reliability and (partial) efficiency are established. The results are applied to solve the L^2 -TV problem by means of the adaptive finite element method. The adaptive mesh refinement relies on the newly derived a posteriori error estimator and on an additional local variance estimator to cope with noisy data. The numerical solution of the discrete problem on each level of refinement is obtained by a superlinearly convergent algorithm based on Fenchel-duality and inexact semismooth Newton techniques and which is stable with respect to noise in the data. Numerical results justifying the advantage of adaptive finite elements solutions are presented.

1. INTRODUCTION

Image denoising is one of the fundamental tasks in mathematical image processing and aims at reconstructing an image from given noisy data. Often this is achieved by employing so-called variational methods which require to minimize an objective functional typically composed of a data fidelity and a regularization term; see, e.g., [44] for further details. Here we utilize an objective functional which was previously studied in [34] and several papers mentioned therein and is given by

$$(\mathcal{P}) \quad \text{minimize} \quad \frac{\mu}{2} \|\nabla v\|_{L^2(\Omega)}^2 + \frac{1}{2} \|v - z\|_{L^2(\Omega)}^2 + \alpha \int_{\Omega} |\nabla v|_{l^2} \, dx \quad \text{over } v \in H_0^1(\Omega),$$

where $\Omega \subset \mathbb{R}^2$ is the open, bounded and polygonal image domain, $z \in L^2(\Omega)$ denotes the given noisy image, and $\alpha > 0$ represents a regularization parameter. We always assume that $0 < \mu \ll \alpha$ such that this model is a very close approximation to the renowned total variation (TV) regularization model proposed by Rudin, Osher and Fatemi in their seminal work [35]:

$$(1.1) \quad \text{minimize} \quad \frac{1}{2} \|v - z\|_{L^2(\Omega)}^2 + \alpha \int_{\Omega} |Dv| \quad \text{over } v \in BV(\Omega).$$

Here, $BV(\Omega)$ denotes the space of functions of bounded variation and $\int_{\Omega} |Dv|$ the associated BV -seminorm; see [22] for their definitions and further properties. For $\mu \rightarrow 0$ in (\mathcal{P}) it may indeed be shown that the associated sequence of solutions (u_{μ}) of (\mathcal{P}) converges (weakly in $L^2(\Omega)$) to the solution u_0 of (1.1); see [34] and the references therein. Due to the importance of TV-regularization in edge preserving image restoration, many research efforts have been devoted to solving and extending the associated variational problem (1.1); see, e.g., [1, 12–14, 17, 33, 34, 37–40, 44] and the many references therein.

Due to given pixel frames, in order to solve the discrete version of (1.1) and/or its extensions or variations (like (\mathcal{P})) numerical solution schemes usually rely on finite difference (FD) approximations where every image pixel corresponds to a degree of freedom in the discretization; see the monograph

Date: March 20, 2012.

Key words and phrases. a posteriori error estimation, elliptic variational inequality of the second kind, adaptive finite elements, total variation, primal-dual method, semismooth Newton method.

Department of Mathematics, Humboldt-University of Berlin, Unter den Linden 6, 10099 Berlin, Germany, and START-Project “Interfaces and Free Boundaries” and SFB “Mathematical Optimization and Applications in Biomedical Science”, Institute of Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36, A-8010 Graz, Austria. (michal.hintermueller@uni-graz.at).

START-Project “Interfaces and Free Boundaries” and SFB “Mathematical Optimization and Applications in Biomedical Science”, Institute of Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36, A-8010 Graz, Austria. (maria.rincon-camacho@uni-graz.at).

[44] and the references therein or, for instance, [12] or [34] for so-called primal-dual methods. In contrast to these uniform mesh approaches, in the present work we proposed to solve the variational problem resulting from the first order optimality characterization of the solution to (\mathcal{P}) by means of an adaptive meshing technique.

Various discretization techniques including FD and finite element (FE) methods, respectively, admit adaptive techniques. Here we focus on adaptive finite element methods (AFEMs) due to their analytical advantages; see [3, 6, 43] for a general account of AFEM. Our interest in adaptive mesh refinement is motivated by the fact that it tries to achieve a given approximation quality of the numerical solution by ideally utilizing the smallest possible number of pixels. For instance, in our imaging context large homogeneous features, i.e. regions where u_μ is (almost) constant, may be recovered well already by locally discretizing by a small number of triangles in a FE discretization of u_μ . On the other hand, near edges and small scale features or texture the adaptive mechanism has to select a rather fine mesh automatically in order to guarantee a good approximation quality. In a posteriori residual-based AFEM this refinement process relies on local residual errors indicating triangles with large contribution to the overall error for refinement.

Returning to the above mentioned first order optimality conditions, setting

$$(1.2) \quad a(u, v) := \mu (\nabla u, \nabla v)_{L^2(\Omega)} + (u, v)_{L^2(\Omega)}, \quad l(u) := (z, u)_{L^2(\Omega)}, \quad j(u) := \alpha \int_\Omega |\nabla u|_{L^2} dx,$$

where $(\cdot, \cdot)_{L^2(\Omega)}$ denotes the usual $L^2(\Omega)$ -inner product such that $a : V \times V \rightarrow \mathbb{R}$, with $V := H_0^1(\Omega)$, becomes a V -coercive continuous symmetric bilinear form, $l \in V^*$, with $V^* = H^{-1}(\Omega)$ denoting the (topological) dual space of V , and $j : V \rightarrow \mathbb{R} := \mathbb{R} \cup \{\pm\infty\}$ is proper, convex and lower semi-continuous, the first order optimality characterization of the optimal solution u of (\mathcal{P}) (where we keep μ fixed and thus neglect it in the notation for the solution) reads

$$(1.3) \quad a(u, v - u) + j(v) - j(u) \geq \langle l, v - u \rangle_{V^*, V}, \quad \forall v \in V := H_0^1(\Omega).$$

Note that in terms of the above notation the objective of (\mathcal{P}) is given by

$$J_\mu(v) = J(v) + \frac{1}{2} \|z\|_{L^2(\Omega)}^2 := \frac{1}{2} a(v, v) + j(v) - \langle l, v \rangle_{V^*, V},$$

where $\langle \cdot, \cdot \rangle_{V^*, V}$ denotes the duality pairing between V^* and V . Since $\frac{1}{2} \|z\|_{L^2(\Omega)}^2$ is a constant only, it may be neglected in the variational context such that (\mathcal{P}) is equivalent to

$$(1.4) \quad \text{minimize } J(v) \quad \text{over } v \in H_0^1(\Omega).$$

Observe that (1.3) represents a so-called variational inequality of the second kind [23]. Previously, in [9] a posterior residual based error estimates for a specific (and with respect to j a somewhat simpler) class of variational inequalities of the second kind were derived. In view of (1.3), our aim is to extend these estimates to situations where j models the TV-regularization term in image denoising and where the variational inequality is yet formulated in a different primal-dual fashion (see the subsequent section) which is in particular suitable for numerical solution schemes.

We mention that under a regularity assumption on the data, in [20] a priori error estimates for the FE solution of a regularized gradient flow for the TV-model (1.1) are studied. The adaptive finite element method is used in tomographic imaging such as diffuse optical absorption imaging [26] or fluorescence diffuse optical tomography [5, 27] where either quadratic L^2 - or H^1 -regularization of Tikhonov-type (rather than the non-differentiable TV-regularization) is applied. Some adaptive finite element approaches are used to solve a modification of the Perona-Malik model for image processing in [8] and [36]. Image segmentation based on adaptive finite elements is studied in [21], and efficient solvers relying on OcTree-based adaptive mesh refinement for parametric as well as non-parametric approaches in image registration can be found in [28–30].

The outline of the rest of the paper is as follows. Fenchel duality results useful for solving the optimization problem (1.4) are stated in Section 2. An a posteriori error bound to the finite element approximation is derived in Section 3. The efficiency of this bound is studied in Section 4. An adaptive finite element algorithm that utilizes this a posteriori error bound is proposed in Section 5. Numerical results of the adaptive finite element algorithm are presented in Section 6. Finally some conclusions are drawn.

2. FENCHEL DUALITY RESULTS

The optimization problem (\mathcal{P}) can be solved by means of Fenchel duality; see [12, 19, 33, 34]. In [34] it was shown that the Fenchel-dual problem of (\mathcal{P}) is given by

$$(\mathcal{D}) \quad \text{maximize} \quad -\frac{1}{2} \|\operatorname{div} \vec{p} + z\|_{H^{-1}}^2 \quad \text{over } \mathbb{L}^2(\Omega) \quad \text{subject to } |\vec{p}|_{l^2} \leq \alpha \text{ a.e. in } \Omega,$$

where $\|\cdot\|_{H^{-1}} = \langle M_\mu w, w \rangle_{H_0^1(\Omega), H^{-1}(\Omega)}$ with $M_\mu = (I - \mu \Delta)^{-1}$ and $H^{-1}(\Omega)$ denotes the dual space of $H_0^1(\Omega)$. Here and below 'a.e.' stands for 'almost everywhere'.

Let \bar{u} be the solution of the problem (\mathcal{P}) and let \vec{p} be a solution of the dual problem (\mathcal{D}) . Then both quantities are related by the optimality system

$$(2.1) \quad -(\vec{p}, \nabla v)_{\mathbb{L}^2(\Omega)} = a(\bar{u}, v) - (z, v)_{L^2(\Omega)}, \quad \forall v \in H_0^1(\Omega),$$

$$(2.2) \quad (\vec{p}, \vec{t})_{\mathbb{L}^2(\Omega)} \leq \alpha \int_\Omega |\vec{t}|_{l^2} dx \quad \forall \vec{t} \in \mathbb{L}^2(\Omega), \quad (\vec{p}, \nabla \bar{u})_{\mathbb{L}^2(\Omega)} = \alpha \int_\Omega |\nabla \bar{u}|_{l^2} dx.$$

Note that the dual solution \vec{p} need not be unique in (2.1)-(2.2) which is due to the non-trivial kernel of the divergence operator occurring in weak form in (2.1). This fact may cause instabilities in primal-dual or dual numerical solution schemes. In [34] this non-uniqueness is overcome by considering the regularized version of (\mathcal{P}) given by

$$(\mathcal{P}_\gamma) \quad \text{minimize} \quad J_\gamma(u) := \frac{1}{2} a(u, u) - (z, u)_{L^2(\Omega)} + j_\gamma(u), \quad \text{over } u \in H_0^1(\Omega),$$

where, as before, $a(u, v) = \mu (\nabla u, \nabla v)_{\mathbb{L}^2(\Omega)} + (u, v)_{L^2(\Omega)}$ and $z \in L^2(\Omega)$, and $j(u)$ is replaced by the convex and Fréchet-differentiable function $j_\gamma(u)$, which is defined as

$$(2.3) \quad j_\gamma(u) := \alpha \int_\Omega \Phi_\gamma(\nabla u) dx, \quad \text{with} \quad \Phi_\gamma(\vec{y})(x) := \begin{cases} |\vec{y}(x)|_{l^2} - \frac{\gamma}{2} & \text{if } |\vec{y}(x)|_{l^2} \geq \gamma, \\ \frac{1}{2\gamma} |\vec{y}(x)|_{l^2}^2 & \text{if } |\vec{y}(x)|_{l^2} < \gamma, \end{cases}$$

with $\gamma > 0$ a (small) regularization parameter. In [34] the dual problem of (\mathcal{P}_γ) is shown to be

$$(\mathcal{D}_\gamma) \quad \begin{aligned} &\text{maximize} \quad \mathcal{H}_\gamma(\vec{p}) := -\frac{1}{2} \|\operatorname{div} \vec{p} + z\|_{H^{-1}}^2 - \frac{\gamma}{2} \|\vec{p}\|_{\mathbb{L}^2(\Omega)}^2 \quad \text{over } \vec{p} \in \mathbb{L}^2(\Omega), \\ &\text{subject to} \quad |\vec{p}|_{l^2} \leq \alpha, \quad \text{a.e. in } \Omega. \end{aligned}$$

The associated first order optimality condition relating the solutions of (\mathcal{P}_γ) and (\mathcal{D}_γ) , respectively, is given by

$$(2.4) \quad -(\vec{p}, \nabla v)_{\mathbb{L}^2(\Omega)} = a(\bar{u}, v) - (z, v)_{L^2(\Omega)}, \quad \forall v \in H_0^1(\Omega),$$

$$(2.5) \quad \max(\gamma, |\nabla \bar{u}(x)|_{l^2}) \vec{p}(x) = \alpha \nabla \bar{u}(x), \quad \text{for almost all } x \text{ in } \Omega;$$

see [34] and [24] for a related system in the context of Bingham fluids. We note that for small γ , $\frac{1}{\alpha} j_\gamma$ is close to the BV-seminorm. Further, observe that j_γ smoothes j locally around the kink of the norm-function involved in j . This local smoothing has a stabilizing effect on primal-dual numerical solution schemes; see [34] for a related observation. Of course, other types of smoothing are possible without altering our subsequent considerations. An example, which is often found in the literature is $\sqrt{\epsilon^2 + |\nabla \bar{u}|_{l^2}^2}$.

In the next section we derive residual-based a posteriori error estimators for a finite element approximation of the solution of the first order optimality system of either the original or regularized version of (\mathcal{P}) , respectively. By resorting to the formulation of the respective first order systems as a variational inequality of the second kind (in case of (\mathcal{P})) or a non-smooth variational system (in case of (\mathcal{P}_γ)), which can also be cast as a special variational inequality of second kind, both formulations will be treated within one framework. It, however, turns out that in case of (\mathcal{P}_γ) stronger results are obtained, which is due to (2.5) (with the latter replacing the set-inclusion (2.2) by a equality relation).

3. A POSTERIORI ERROR BOUND APPLIED TO FINITE ELEMENT APPROXIMATION

3.1. Finite element approach. We solve (2.1)-(2.2) or (2.4)-(2.5) by applying a finite element discretization and, in case of (2.4)-(2.5), by a subsequent semismooth Newton method.

For discretization purposes let $V_h \subset H_0^1(\Omega) =: V$ and $P_h \subset \mathbb{L}^2(\Omega) =: Y$ be two finite dimensional subspaces defined by

$$(3.1) \quad V_h := \{v_h \in C(\overline{\Omega}) : v_h|_T \in \Pi_k, \forall T \in \mathcal{T}_h\},$$

$$(3.2) \quad P_h := \{\vec{t}_h := (t_h^1, t_h^2) \in \mathbb{L}^2(\Omega) : t_h^1|_T, t_h^2|_T \in \Pi_{k-1}, \forall T \in \mathcal{T}_h\},$$

where Π_k denotes the set of polynomials with degree less than or equal to k defined on \mathbb{R}^2 and \mathcal{T}_h denotes a regular triangulation of Ω . For any open subset ω of Ω with Lipschitz boundary $\partial\omega$, we denote by $L^2(\omega)$, $\mathbb{L}^2(\omega)$, $H^1(\omega)$ and $L^2(\partial\omega)$ the standard Lebesgue or Sobolev spaces equipped with the usual norms [2]. For any $T \in \mathcal{T}_h$ let $\mathcal{E}(T)$ and $\mathcal{N}(T)$ denote the set of its edges and vertices, respectively, and let us define

$$(3.3) \quad \mathcal{E}_h := \bigcup_{T \in \mathcal{T}_h} \mathcal{E}(T), \quad \mathcal{N}_h := \bigcup_{T \in \mathcal{T}_h} \mathcal{N}(T), \quad \tilde{\omega}_T := \bigcup_{\mathcal{N}(T) \cap \mathcal{N}(T') \neq \emptyset} T', \quad \tilde{\omega}_E := \bigcup_{\mathcal{N}(E) \cap \mathcal{N}(T') \neq \emptyset} T'.$$

Let h_T be the diameter of $T \in \mathcal{T}_h$, and let h_E denote the length of $E \in \mathcal{E}_h$. We denote by $I_h : V \rightarrow V_h$ Clément's interpolation operator [16] satisfying the error estimates

$$(3.4) \quad \|\psi - I_h\psi\|_{L^2(T)} \leq c_1 h_T \|\psi\|_{H_0^1(\tilde{\omega}_T)}, \quad \|\psi - I_h\psi\|_{L^2(E)} \leq c_2 h_E^{1/2} \|\psi\|_{H_0^1(\tilde{\omega}_E)},$$

for all $\psi \in V$, where the positive constants c_1 and c_2 depend only on the smallest angle in the triangulation; see [15, 16, 43].

The finite element approximations $\bar{u}_h \in V_h$ and $\bar{p}_h \in P_h$ of the optimal solutions \bar{u} and \bar{p} to (P) and (D) solve

$$(3.5) \quad -(\bar{p}_h, \nabla v_h)_{\mathbb{L}^2(\Omega)} = a(\bar{u}_h, v_h) - (z, v_h)_{L^2(\Omega)}, \quad \forall v_h \in V_h,$$

$$(3.6) \quad (\bar{p}_h, \vec{t}_h)_{\mathbb{L}^2(\Omega)} \leq \alpha \int_{\Omega} |\vec{t}_h|^2 dx \quad \forall \vec{t}_h \in P_h, \quad (\bar{p}_h, \nabla \bar{u}_h)_{\mathbb{L}^2(\Omega)} = \alpha \int_{\Omega} |\nabla \bar{u}_h|^2 dx.$$

An analogous system is obtained when discretizing (2.4)-(2.5).

3.2. A posteriori error estimate. Let $e_h := \bar{u} - \bar{u}_h$ be the approximation error of the solution \bar{u}_h of (3.5)-(3.6) to the true solution \bar{u} of problem (P). The bilinear form $a : V \times V \rightarrow \mathbb{R}$ given by (1.2) is V -elliptic, i.e., there exist constants $c_{a_1} > 0$ and $c_{a_2} > 0$ such that

$$c_{a_1} \|v\|_V^2 \leq a(v, v) \leq c_{a_2} \|v\|_V^2, \quad \forall v \in V.$$

Thus, the approximation error e_h satisfies

$$(3.7) \quad \|e_h\|_V^2 = \|\bar{u} - \bar{u}_h\|_V^2 \leq \frac{1}{c_{a_1}} a(\bar{u} - \bar{u}_h, \bar{u} - \bar{u}_h).$$

Using the equivalence between the elliptic variational inequality of the second kind (1.3) and the optimization problem (P) and since $V_h \subset H_0^1(\Omega)$ we obtain

$$(3.8) \quad \frac{1}{2} a(\bar{u} - \bar{u}_h, \bar{u} - \bar{u}_h) = \frac{1}{2} a(\bar{u}_h, \bar{u}_h) - a(\bar{u}, \bar{u}_h) + \frac{1}{2} a(\bar{u}, \bar{u})$$

$$(3.9) \quad = \frac{1}{2} a(\bar{u}_h, \bar{u}_h) - a(\bar{u}, \bar{u}_h - \bar{u}) - \frac{1}{2} a(\bar{u}, \bar{u})$$

$$(3.10) \quad \leq \frac{1}{2} a(\bar{u}_h, \bar{u}_h) + j(\bar{u}_h) - j(\bar{u}) - l(\bar{u}_h - \bar{u}) - \frac{1}{2} a(\bar{u}, \bar{u})$$

$$(3.11) \quad = J(\bar{u}_h) - J(\bar{u})$$

$$(3.12) \quad =: (F(\bar{u}_h) + G(\nabla \bar{u}_h)) - (F(\bar{u}) + G(\nabla \bar{u})),$$

with $\nabla : H_0^1(\Omega) \rightarrow \mathbb{L}^2(\Omega)$, $F(u) := \frac{1}{2} a(u, u) - l(u)$ and $G(\vec{p}) = \alpha \int_{\Omega} |\vec{p}|^2 dx$ in case of (P) or $G(\vec{p}) := \alpha \int_{\Omega} \Phi_{\gamma}(\vec{p}) dx$ in case of (P_γ). The Fréchet-derivative $f(u) \in V^*$ of F at $u \in V$ satisfies

$$(3.13) \quad \langle f(u), v \rangle_{V^*, V} = a(u, v) - l(v) \quad \forall v \in V.$$

Moreover, since F is convex the inequality

$$(3.14) \quad F(\bar{u}_h) - F(v) \leq \langle f_{\bar{u}_h}, \bar{u}_h - v \rangle_{V^*, V} \quad \forall v \in V$$

holds true. Here we use $f_{\bar{u}_h} := f(\bar{u}_h)$. The convex function G need not be Fréchet differentiable, but (in both cases considered above) it admits a generalized derivative which is given by

$$(3.15) \quad \partial G(\vec{p}) := \{g \in Y^* : G(\vec{p}) - G(\vec{t}) \leq \langle g, \vec{p} - \vec{t} \rangle_{Y^*, Y} \quad \forall \vec{t} \in Y\}.$$

In the case of $G(\cdot) = \alpha \int_{\Omega} \Phi_{\gamma}(\cdot) dx$ becomes single-valued containing the Fréchet-derivative of j_{γ} . Note that upon identification we have $Y^* = Y$ and $\langle \cdot, \cdot \rangle_{Y^*, Y} = (\cdot, \cdot)_{\mathbb{L}^2(\Omega)}$. Below, in a slight misuse of notation, we use $g_{\bar{u}_h} := g(\nabla \bar{u}_h)$. These derivatives together with (3.12) yield

$$(3.16) \quad \frac{1}{2}a(\bar{u} - \bar{u}_h, \bar{u} - \bar{u}_h) \leq \langle \bar{u}_h - \bar{u}, f_{\bar{u}_h} \rangle_{V, V^*} + \langle \nabla \bar{u}_h - \nabla \bar{u}, g_{\bar{u}_h} \rangle_{Y, Y^*}$$

with equality holding if $G(\cdot) = \alpha \int_{\Omega} \Phi_{\gamma}(\cdot) dx$.

For our subsequent estimates we briefly recall the following relations.

Remark 1. If $\vec{p} \in \mathbb{L}^2(\Omega)$ satisfies $\operatorname{div} \vec{p} \in L^2(\Omega)$, then it admits a well defined normal trace $\vec{p} \cdot \vec{n}_{\Gamma}$ on $\Gamma = \partial\Omega$ which lies in $H^{-1/2}(\Gamma)$ [10, 11, 41]. Green's theorem implies

$$(\vec{p}, \nabla v)_{\mathbb{L}^2(\Omega)} + (\operatorname{div} \vec{p}, v)_{L^2(\Omega)} = \langle v, \vec{p} \cdot \vec{n}_{\Gamma} \rangle_{H^{1/2}(\Gamma), H^{-1/2}(\Gamma)} \quad \forall v \in H^1(\Omega).$$

For $\vec{p} \in \mathbb{L}^2(\Omega)$ and for all $v \in V$ we have that

$$(\nabla v, \vec{p})_{\mathbb{L}^2(\Omega)} = -\langle v, \operatorname{div} \vec{p} \rangle_{H_0^1(\Omega), H^{-1}(\Omega)}.$$

Using Green's first identity, the bilinear form a yields

$$(3.17) \quad a(\bar{u} - \bar{u}_h, \bar{u}_h) = \sum_{T \in \mathcal{T}_h} \left[\mu(\bar{u} - \bar{u}_h, -\Delta \bar{u}_h)_{L^2(T)} + \mu(\bar{u} - \bar{u}_h, \vec{n}_{\partial T} \cdot \nabla \bar{u}_h)_{L^2(\partial T)} + (\bar{u} - \bar{u}_h, \bar{u}_h)_{L^2(T)} \right],$$

where $\vec{n}_{\partial T}$ denotes the unit outward normal to ∂T .

Thus, from (3.16), Remark 1, (3.17) and the fact that \vec{p}_h is constant on T , we infer

$$(3.18) \quad \begin{aligned} \frac{1}{2}a(\bar{u} - \bar{u}_h, \bar{u} - \bar{u}_h) &\leq \sum_{T \in \mathcal{T}_h} \left[(\bar{u} - \bar{u}_h, z + \mu \Delta \bar{u}_h - \bar{u}_h)_{L^2(T)} - \mu(\bar{u} - \bar{u}_h, \vec{n}_{\partial T} \cdot \nabla \bar{u}_h)_{L^2(\partial T)} \right] \\ &\quad + \sum_{T \in \mathcal{T}_h} \left[-(\nabla \bar{u} - \nabla \bar{u}_h, \vec{p}_h)_{\mathbb{L}^2(T)} + (\nabla \bar{u} - \nabla \bar{u}_h, \vec{p}_h - g_{\bar{u}_h})_{\mathbb{L}^2(T)} \right] \\ &= \sum_{T \in \mathcal{T}_h} (\bar{u} - \bar{u}_h, \operatorname{div} \vec{p}_h + z + \mu \Delta \bar{u}_h - \bar{u}_h)_{L^2(T)} \\ &\quad - \sum_{T \in \mathcal{T}_h} \left[\mu(\bar{u} - \bar{u}_h, \vec{n}_{\partial T} \cdot \nabla \bar{u}_h)_{L^2(\partial T)} + (\bar{u} - \bar{u}_h, \vec{p}_h \cdot \vec{n}_{\partial T})_{L^2(\partial T)} \right] \\ &\quad - \sum_{T \in \mathcal{T}_h} \left[(I_h(\bar{u} - \bar{u}_h), \operatorname{div} \vec{p}_h + z + \mu \Delta \bar{u}_h - \bar{u}_h)_{L^2(T)} \right. \\ &\quad \quad \left. - (I_h(\bar{u} - \bar{u}_h), \vec{n}_{\partial T} \cdot \nabla \bar{u}_h)_{L^2(\partial T)} - (I_h(\bar{u} - \bar{u}_h), \vec{p}_h \cdot \vec{n}_{\partial T})_{L^2(\partial T)} \right] \\ &\quad + \sum_{T \in \mathcal{T}_h} (\nabla \bar{u} - \nabla \bar{u}_h, \vec{p}_h - g_{\bar{u}_h})_{\mathbb{L}^2(T)}. \end{aligned}$$

Rearranging terms in (3.18), using the Cauchy-Schwarz inequality and applying the estimates (3.4), the following inequality is obtained:

$$(3.19) \quad \begin{aligned} \frac{1}{2}a(\bar{u} - \bar{u}_h, \bar{u} - \bar{u}_h) &\leq \sum_{T \in \mathcal{T}_h} [c_1 h_T \|\operatorname{div} \vec{p}_h + z + \mu \Delta \bar{u}_h - \bar{u}_h\|_{L^2(T)} \|\bar{u} - \bar{u}_h\|_{L^2(\bar{\omega}_T)} \\ &\quad + \sum_{E \in \hat{\mathcal{E}}(T)} h_E^{1/2} c_2 (\mu \|\left[\frac{\partial \bar{u}_h}{\partial n} \right]_E\|_{L^2(E)} + \|\vec{p}_h \cdot \vec{n}\|_E) \|\bar{u} - \bar{u}_h\|_{H_0^1(\bar{\omega}_E)}] \\ &\quad + \sum_{T \in \mathcal{T}_h} \|\nabla \bar{u} - \nabla \bar{u}_h\|_{\mathbb{L}^2(T)} \|\vec{p}_h - g_{\bar{u}_h}\|_{\mathbb{L}^2(T)} \\ &\leq \left(\sum_{T \in \mathcal{T}_h} \max(c_1, c_2)^2 \|\bar{u} - \bar{u}_h\|_{H_0^1(\bar{\omega}_T)}^2 \right)^{1/2} \left(\sum_{T \in \mathcal{T}_h} h_T^2 \|\operatorname{div} \vec{p}_h + z + \mu \Delta \bar{u}_h - \bar{u}_h\|_{L^2(T)}^2 \right. \\ &\quad \left. + \sum_{E \in \hat{\mathcal{E}}_h} h_E (\mu \|\left[\frac{\partial \bar{u}_h}{\partial n} \right]_E\|_{L^2(E)} + \|\vec{p}_h \cdot \vec{n}\|_E)^2 \right)^{1/2} \\ &\quad + \|\nabla \bar{u} - \nabla \bar{u}_h\|_{\mathbb{L}^2(\Omega)} \left[\sum_{T \in \mathcal{T}_h} \|\vec{p}_h - g_{\bar{u}_h}\|_{\mathbb{L}^2(T)} \right] \\ &\leq 2 \max(c_1, c_2) \|\bar{u} - \bar{u}_h\|_{H_0^1(\Omega)} \left(\sum_{T \in \mathcal{T}_h} h_T^2 \|\operatorname{div} \vec{p}_h + z + \mu \Delta \bar{u}_h - \bar{u}_h\|_{L^2(T)}^2 \right. \\ &\quad \left. + \sum_{E \in \hat{\mathcal{E}}_h} h_E (\mu \|\left[\frac{\partial \bar{u}_h}{\partial n} \right]_E\|_{L^2(E)} + \|\vec{p}_h \cdot \vec{n}\|_E)^2 \right)^{1/2} \\ &\quad + \|\nabla \bar{u} - \nabla \bar{u}_h\|_{\mathbb{L}^2(\Omega)} \left[\sum_{T \in \mathcal{T}_h} \|\vec{p}_h - g_{\bar{u}_h}\|_{\mathbb{L}^2(T)} \right]. \end{aligned}$$

The quantities

$$\begin{aligned} \left[\frac{\partial \bar{u}_h}{\partial n} \right]_E &= \bar{n}_{\partial T} \cdot (\nabla \bar{u}_h)_T + \bar{n}_{\partial T'} \cdot (\nabla \bar{u}_h)_{T'}, \\ [\bar{\vec{p}}_h \cdot \bar{\vec{n}}]_E &= \bar{n}_{\partial T} \cdot (\bar{\vec{p}}_h)_T + \bar{n}_{\partial T'} \cdot (\bar{\vec{p}}_h)_{T'} \end{aligned}$$

are defined on an edge E shared by the elements T and T' and represent the jump discontinuities in the approximation to the normal flux and the dual variable on the interface E , respectively; see, e.g., [3]. Here $\hat{\mathcal{E}}(T)$ and $\hat{\mathcal{E}}_h$ denote the interior edges sets, i.e. $\hat{\mathcal{E}}(T) := \mathcal{E}(T) \setminus \partial\Omega$ and $\hat{\mathcal{E}}_h := \mathcal{E}_h \setminus \partial\Omega$.

Remark 2. The optimality condition (3.6) on $\bar{\vec{p}}_h$ requires that $\bar{\vec{p}}_h \in \partial G(\nabla \bar{u}_h)$. Since $\bar{u}_h \in V_h$ implies $\nabla \bar{u}_h \in P_h$ we have

$$(3.20) \quad \|\bar{\vec{p}}_h - g_{\bar{u}_h}\|_{L^2(T)} = 0.$$

The inequality (3.19) together with (3.7) and (3.20) prove the next theorem.

Theorem 3. Let $\bar{u} \in H_0^1(\Omega)$ be the solution of problem (P), let $\bar{u}_h \in V_h$ be the finite element approximation to this solution with $\bar{\vec{p}}_h$ a corresponding solution of the dual problem (D), both given by the system (3.5)-(3.6). Then we have the following a posteriori error bound

$$(3.21) \quad \begin{aligned} \|e_h\|_{H_0^1(\Omega)}^2 &= \|\bar{u} - \bar{u}_h\|_{H_0^1(\Omega)}^2 \leq C \left[\sum_{T \in \mathcal{T}_h} h_T^2 \|\operatorname{div} \bar{\vec{p}}_h + z + \mu \Delta \bar{u}_h - \bar{u}_h\|_{L^2(T)}^2 \right. \\ &\quad \left. + \sum_{E \in \hat{\mathcal{E}}_h} h_E \left(\mu \left\| \left[\frac{\partial \bar{u}_h}{\partial n} \right]_E \right\|_{L^2(E)} + \left\| [\bar{\vec{p}}_h \cdot \bar{\vec{n}}]_E \right\|_{L^2(E)} \right)^2 \right], \end{aligned}$$

with some constant $C > 0$ depending only on c_{a_1} and the shape regularity of the triangulation.

Based on our above findings we define the local error indicator

$$(3.22) \quad \varepsilon_T^{(k)} := \eta_T + \frac{1}{2}(\eta_{\partial T, u} + \eta_{\partial T, p}),$$

with

$$(3.23) \quad \begin{aligned} \eta_T &= h_T^2 \|\operatorname{div} \bar{\vec{p}}_h + z + \mu \Delta \bar{u}_h - \bar{u}_h\|_{L^2(T)}^2, \\ \eta_{\partial T, u} &= \sum_{E \in \hat{\mathcal{E}}(T)} \eta_{E, u}, \quad \eta_{E, u} = h_E \left\| \left[\frac{\partial \bar{u}_h}{\partial n} \right]_E \right\|_{L^2(E)}^2, \\ \eta_{\partial T, p} &= \sum_{E \in \hat{\mathcal{E}}(T)} \eta_{E, p}, \quad \eta_{E, p} = h_E \left\| [\bar{\vec{p}}_h \cdot \bar{\vec{n}}]_E \right\|_{L^2(E)}^2. \end{aligned}$$

We end this section by a closer inspection of our estimates in the case where $G(\cdot) = \int_{\Omega} \Phi_{\gamma}(\cdot) dx$. In this case, the difference $\|\bar{\vec{p}}_h - \bar{\vec{p}}\|_{L^2(\Omega)}$ is bounded above in terms of $\|\bar{u}_h - \bar{u}\|_{H_0^1(\Omega)}$ as the next result shows.

Lemma 4. Let $(\bar{u}, \bar{\vec{p}}) \in H_0^1(\Omega) \times \mathbb{L}^2(\Omega)$ and $(\bar{u}_h, \bar{\vec{p}}_h) \in V_h \times Y_h$ satisfy (2.4)-(2.5) and its finite element approximation, respectively. Then, for any $T \in \mathcal{T}_h(\Omega)$, we have

$$\|\bar{\vec{p}}_h - \bar{\vec{p}}\|_{L^2(T)} \leq C\alpha \min \left(2, \gamma^{-1} \|\bar{u}_h - \bar{u}\|_{H_0^1(T)} \right)$$

for some constant $C > 0$.

Proof. We prove the result by partitioning Ω into subset according to where the max is taken in $m(u) := \max(\gamma, |\nabla u|_{l_2})$ and $m(u_h) := \max(\gamma, |\nabla u_h|_{l_2})$.

(i) On $M_1 := \{x \in T : m(\bar{u})(x) = \gamma \wedge m(\bar{u}_h)(x) = \gamma\}$ we have

$$\|\bar{\vec{p}}_h - \bar{\vec{p}}\|_{L^2(M_1)}^2 = \frac{\alpha^2}{\gamma^2} \int_{M_1} |\nabla \bar{u} - \nabla \bar{u}_h|_{l_2}^2 dx \leq \frac{\alpha^2}{\gamma^2} \|\bar{u} - \bar{u}_h\|_{H_0^1(M_1)}^2.$$

(ii) Next consider $M_2 := \{x \in T : m(\bar{u})(x) = \gamma \wedge m(\bar{u}_h)(x) = |\nabla \bar{u}_h(x)|_{l_2}\}$. We obtain

$$\begin{aligned} \|\bar{\vec{p}}_h - \bar{\vec{p}}\|_{L^2(M_2)}^2 &= \alpha^2 \int_{M_2} \left| \frac{1}{\gamma} \nabla \bar{u} - \frac{1}{|\nabla \bar{u}_h|_{l_2}} \nabla \bar{u}_h \right|_{l_2}^2 dx \\ &= \frac{\alpha^2}{\gamma^2} \int_{M_2} \frac{1}{|\nabla \bar{u}_h|_{l_2}^2} |(\nabla \bar{u} - \nabla \bar{u}_h)|_{l_2}^2 + \nabla \bar{u}_h (|\nabla \bar{u}_h|_{l_2} - \gamma)|_{l_2}^2 dx \\ &\leq \frac{3\alpha^2}{\gamma^2} \int_{M_2} |\nabla \bar{u} - \nabla \bar{u}_h|_{l_2}^2 + |\nabla \bar{u}_h|_{l_2}^2 - \gamma|_{l_2}^2 dx \\ &\leq \frac{6\alpha^2}{\gamma^2} \int_{M_2} |\nabla \bar{u} - \nabla \bar{u}_h|_{l_2}^2 dx \leq \frac{6\alpha^2}{\gamma^2} \|\bar{u} - \bar{u}_h\|_{H_0^1(M_2)}^2 \end{aligned}$$

where we use $|\nabla \bar{u}(x)| \leq \gamma \leq |\nabla \bar{u}_h|_{l^2}$ on M_2 and the second triangle inequality for the last estimate.
 (iii) For $M_3 := \{x \in T : m(\bar{u}_h)(x) = \gamma \wedge m(\bar{u})(x) = |\nabla \bar{u}(x)|_{l^2}\}$ symmetric arguments to those in case (ii) yield the same estimate.
 (iv) Finally, consider $M_4 := \{x \in T : m(\bar{u}_h)(x) = |\nabla \bar{u}_h(x)|_{l^2} \wedge m(\bar{u})(x) = |\nabla \bar{u}(x)|_{l^2}\}$. We infer

$$\begin{aligned} \|\bar{p}_h - \bar{p}\|_{L^2(M_4)}^2 &= \int_{M_4} \frac{\alpha^2}{|\nabla \bar{u}_h|_{l^2}^2 |\nabla \bar{u}|_{l^2}^2} |\nabla \bar{u}| |\nabla \bar{u}_h|_{l^2} - |\nabla \bar{u}|_{l^2} |\nabla \bar{u}_h|_{l^2}^2 dx \\ &\leq 3\alpha^2 \int_{M_4} \left(\frac{1}{|\nabla \bar{u}|_{l^2}} + \frac{1}{|\nabla \bar{u}_h|_{l^2}} \right) |\nabla \bar{u} - \nabla \bar{u}_h|_{l^2}^2 dx \\ &\leq \frac{3\alpha^2}{\gamma^2} \|\bar{u} - \bar{u}_h\|_{H_0^1(M_4)}^2 \end{aligned}$$

where we used $\min(|\nabla \bar{u}(x)|_{l^2}, |\nabla \bar{u}_h(x)|_{l^2}) \geq \gamma$ on M_4 .

Since $T = M_1 \cup M_2 \cup M_3 \cup M_4$, we have

$$\|\bar{p}_h - \bar{p}\|_{L^2(T)}^2 \leq \frac{C\alpha^2}{\gamma^2} \|\bar{u} - \bar{u}_h\|_{H_0^1(T)}^2$$

for some $C > 0$.

Finally, recall that $|\bar{p}|_{l^2} \leq \alpha$ and $|\bar{p}_h|_{l^2} \leq \alpha$ a.e. on T , respectively. Hence,

$$\|\bar{p}_h - \bar{p}\|_{L^2(T)} \leq 2\alpha|T|,$$

where $|\cdot|$ denotes the two-dimensional Lebesgue measure. This concludes the proof. \square

From Theorem 3 and Lemma 4 we immediately conclude the following estimate.

Theorem 5. *Let $\bar{u} \in H_0^1(\Omega)$ be the solution of problem (\mathcal{P}_γ) , let $\bar{u}_h \in V_h$ be the finite element approximation to this solution with \bar{p}_h a corresponding solution of the dual problem (\mathcal{D}_γ) . Then we have the following a posteriori error bound*

$$\begin{aligned} \|e_h\|_{H_0^1(\Omega)}^2 + \gamma^2 \|\bar{p} - \bar{p}_h\|_{L^2(\Omega)}^2 &\leq C \left[\sum_{T \in \mathcal{T}_h} h_T^2 \|\operatorname{div} \bar{p}_h + z + \mu \Delta \bar{u}_h - \bar{u}_h\|_{L^2(T)}^2 \right. \\ &\quad \left. + \sum_{E \in \mathcal{E}_h} h_E \left(\mu \left\| \left[\frac{\partial \bar{u}_h}{\partial n} \right]_E \right\|_{L^2(E)} + \left\| [\bar{p}_h \cdot \vec{n}]_E \right\|_{L^2(E)} \right)^2 \right], \end{aligned}$$

with some constant $C > 0$ depending only on c_{a_1} and the shape regularity of the triangulation.

4. EFFICIENCY OF THE A POSTERIORI ERROR ESTIMATOR

The bound given in Theorem 3 is aimed to be used as a reference for an adaptive refinement process. In this section we show that the error bound is also reliable, i.e., up to data oscillations it also serves as a lower bound of the error.

We start by recalling a few results given in [43]. Let β_i^T , $i = 1, 2, 3$, be the barycentric coordinates of $T \in \mathcal{T}_h(\Omega)$ and $\mathbf{b}_T := 27 \prod_{i=1}^3 \beta_i^T$ its associated triangle-bubble function. Let β_i^E , $i = 1, 2$, be the barycentric coordinates of $E \in \mathcal{E}_h(\Omega)$ and $\mathbf{b}_E := 4 \prod_{i=1}^2 \beta_i^E$ its associated edge-bubble function. Then there exist constants c_3, c_4, c_5, c_6, c_7 that only depend on the shape regularity of the triangulation $\mathcal{T}_h(\Omega)$ such that for $\rho_T \in \Pi_k(T)$ and $\rho_E \in \Pi_k(E)$, $k \in \mathbb{N}_0$ we have the following inequalities:

$$(4.1) \quad \|\rho_T\|_{L^2(T)}^2 \leq c_3 (\rho_T^2, \mathbf{b}_T)_{L^2(T)}, \quad T \in \mathcal{T}_h(\Omega),$$

$$(4.2) \quad \|\rho_T \mathbf{b}_T\|_{L^2(T)} \leq \|\rho_T\|_{L^2(T)}, \quad T \in \mathcal{T}_h(\Omega),$$

$$(4.3) \quad \|\rho_T \mathbf{b}_T\|_{H_0^1(T)} \leq c_4 h_T^{-1} \|\rho_T\|_{L^2(T)}, \quad T \in \mathcal{T}_h(\Omega),$$

$$(4.4) \quad \|\rho_E\|_{L^2(E)}^2 \leq c_5 (\rho_E^2, \mathbf{b}_E)_{L^2(E)}, \quad E \in \mathcal{E}_h(\Omega),$$

$$(4.5) \quad \|\rho_E \mathbf{b}_E\|_{L^2(E)} \leq \|\rho_E\|_{L^2(E)}, \quad E \in \mathcal{E}_h(\Omega),$$

$$(4.6) \quad \|\tilde{\rho}_E \mathbf{b}_E\|_{L^2(\tilde{\omega}_E)} \leq c_6 h_E^{1/2} \|\rho_E\|_{L^2(E)}, \quad E \in \mathcal{E}_h(\Omega),$$

$$(4.7) \quad \|\tilde{\rho}_E \mathbf{b}_E\|_{H_0^1(\tilde{\omega}_E)} \leq c_7 h_E^{-1/2} \|\rho_E\|_{L^2(E)}, \quad E \in \mathcal{E}_h(\Omega),$$

where $\tilde{\rho}_E$ is the extension of ρ_E to the patch $\tilde{\omega}_E$. For fixed $E' \in \mathcal{E}_h(T) \setminus \{E\}$ and for $x \in T$ we define $\tilde{\rho}_E(x) := \rho_E(x_E)$, where $x_E \in E$ is such that $x - x_E$ is parallel to E' .

Let $z_h \in V_h$ be the finite element approximation to the given data z (see Remark 12) and let $\bar{u}_h \in V_h$ and $\bar{p}_h \in P_h$ be the finite element approximations to an optimal primal-dual pair (\bar{u}, \bar{p})

which satisfy the system of equations (3.5)-(3.6). We proceed by stating and proving some preliminary results.

Lemma 6. *There exists a positive constant C depending only on the shape regularity of the triangulation $\mathcal{T}_h(\Omega)$ such that*

$$(4.8) \quad \eta_T \leq C \left(h_T^2 \|z_h - z\|_{L^2(T)}^2 + \|e_h\|_{H_0^1(T)}^2 + \|\bar{p} - \bar{p}_h\|_{L^2(T)}^2 \right).$$

Proof. For any $v_h \in V_h$, according to the optimality conditions (2.4)-(2.5) and (3.5)-(3.6) we have

$$(4.9) \quad \begin{aligned} a(e_h, v_h) &= -(\bar{p}, \nabla v_h)_{L^2(\Omega)} + (z, v_h)_{L^2(\Omega)} - \mu(\nabla \bar{u}_h, \nabla v_h)_{L^2(\Omega)} - (\bar{u}_h, v_h)_{L^2(\Omega)} \\ &= \sum_{T \in \mathcal{T}_h(\Omega)} \left(-(\bar{p}, \nabla v_h)_{L^2(T)} + (z, v_h)_{L^2(T)} - \mu(\nabla \bar{u}_h, \nabla v_h)_{L^2(T)} - (\bar{u}_h, v_h)_{L^2(T)} \right) \\ &= \sum_{T \in \mathcal{T}_h(\Omega)} \left((\operatorname{div} \bar{p}_h + z_h + \mu \Delta \bar{u}_h - \bar{u}_h, v_h)_{L^2(T)} + (z - z_h, v_h)_{L^2(T)} + (\bar{p}_h - \bar{p}, \nabla v_h)_{L^2(T)} \right. \\ &\quad \left. - (\bar{p}_h \cdot \vec{n}_{\partial T}, v_h)_{L^2(\partial T)} \right) \end{aligned}$$

We set $f_h := (\operatorname{div} \bar{p}_h + z_h + \mu \Delta \bar{u}_h - \bar{u}_h)|_T \mathbf{b}_T$ and extend it by 0 to all of Ω (still keeping the notation f_h for the extended function). Note that $f_h \in V_h$. From $\mathbf{b}_T \in H_0^1(T)$ we infer

$$(4.10) \quad (\operatorname{div} \bar{p}_h + z_h + \mu \Delta \bar{u}_h - \bar{u}_h, f_h)_{L^2(T)} = (z_h - z, f_h)_{L^2(T)} + (\bar{p}_h - \bar{p}, \nabla f_h)_{L^2(T)} + a(e_h, f_h).$$

According to (4.2) and (4.3) we have the following inequalities:

$$(4.11) \quad \|f_h\|_{L^2(T)} \leq \|\operatorname{div} \bar{p}_h + z_h + \mu \Delta \bar{u}_h - \bar{u}_h\|_{L^2(T)},$$

$$(4.12) \quad \|f_h\|_{H_0^1(T)} \leq c_4 h_T^{-1} \|\operatorname{div} \bar{p}_h + z_h + \mu \Delta \bar{u}_h - \bar{u}_h\|_{L^2(T)}.$$

Using (4.10), (4.1) and the Cauchy-Schwarz inequality, we obtain the bound

$$\begin{aligned} h_T^2 \|\operatorname{div} \bar{p}_h + z_h + \mu \Delta \bar{u}_h - \bar{u}_h\|_{L^2(T)}^2 &\leq c_3 h_T^2 \|z_h - z\|_{L^2(T)} \|f_h\|_{L^2(T)} + c_3 h_T^2 \|\bar{p} - \bar{p}_h\|_{L^2(T)} \|f_h\|_{H_0^1(T)} \\ &\quad + c_3 c_{a_2} h_T^2 \|e_h\|_{H_0^1(T)} \|f_h\|_{H_0^1(T)} \\ &\leq c_3 h_T \|\operatorname{div} \bar{p}_h + z_h + \mu \Delta \bar{u}_h - \bar{u}_h\|_{L^2(T)} \left(h_T \|z_h - z\|_{L^2(\Omega)} \right. \\ &\quad \left. + c_4 \|\bar{p} - \bar{p}_h\|_{L^2(T)} + c_{a_2} c_4 \|e_h\|_{H_0^1(T)} \right) \end{aligned}$$

and consequently

$$(4.13) \quad \begin{aligned} h_T^2 \|\operatorname{div} \bar{p}_h + z_h + \mu \Delta \bar{u}_h - \bar{u}_h\|_{L^2(T)}^2 &\leq 2h_T^2 \|z_h - z\|_{L^2(T)}^2 + 2(c_{a_2} c_4)^2 \|e_h\|_{H_0^1(T)}^2 \\ &\quad + 2c_4^2 \|\bar{p} - \bar{p}_h\|_{L^2(T)}^2. \end{aligned}$$

Finally, (4.8) follows from (4.13) and

$$h_T^2 \|\operatorname{div} \bar{p}_h + z + \mu \Delta \bar{u}_h - \bar{u}_h\|_{L^2(T)}^2 \leq 2h_T^2 \|\operatorname{div} \bar{p}_h + z_h + \mu \Delta \bar{u}_h - \bar{u}_h\|_{L^2(T)}^2 + 2h_T^2 \|z - z_h\|_{L^2(T)}^2.$$

□

Lemma 7. *There exists a positive constant C depending only on the shape regularity of the triangulation $\mathcal{T}_h(\Omega)$ such that*

$$(4.14) \quad \eta_E \leq C \left(h_E^2 \|z_h - z\|_{L^2(\tilde{\omega}_E)}^2 + \sum_{t=1}^2 \|e_h\|_{H_0^1(T_t)}^2 + \|\bar{p} - \bar{p}_h\|_{L^2(\tilde{\omega}_E)}^2 \right),$$

where $T_1, T_2 \in \mathcal{T}_h(\Omega)$ and $E = T_1 \cap T_2 \in \mathring{\mathcal{E}}_h(\Omega)$.

Proof. We set $\rho_E := \mu \left[\frac{\partial \bar{u}_h}{\partial n} \right]_E + [\bar{p}_h \cdot \vec{n}]_E$ and $g_h := \bar{\rho}_E \mathbf{b}_E$, and we observe that $\mathbf{b}_E \in H_0^1(\tilde{\omega}_E)$, which, like g_h , can be extended by zero to a function in $V_h \subset H_0^1(\Omega)$ (again we keep the notation g_h and \mathbf{b}_E

for the extended functions). According to the optimality conditions (3.5)-(3.6) and Green's formula, we have

$$\begin{aligned}
 a(e_h, g_h) &= -(\bar{p}, \nabla g_h)_{\mathbb{L}^2(\tilde{\omega}_E)} + (z, g_h)_{L^2(\tilde{\omega}_E)} - \mu(\nabla \bar{u}_h, \nabla g_h)_{\mathbb{L}^2(\tilde{\omega}_E)} - (\bar{u}_h, g_h)_{L^2(\tilde{\omega}_E)} \\
 &= -(\bar{p} - \bar{p}_h, \nabla g_h)_{\mathbb{L}^2(\tilde{\omega}_E)} - (\bar{p}_h, \nabla g_h)_{\mathbb{L}^2(\tilde{\omega}_E)} + (z, g_h)_{L^2(\tilde{\omega}_E)} \\
 &\quad + \mu(\Delta \bar{u}_h, g_h)_{L^2(\tilde{\omega}_E)} - \mu \sum_{t=1}^2 (\vec{n}_{\partial T_t} \cdot \nabla \bar{u}_h, g_h)_{L^2(\partial T_t)} - (\bar{u}_h, g_h)_{L^2(\tilde{\omega}_E)} \\
 &= (\operatorname{div} \bar{p}_h + z_h + \mu \Delta \bar{u}_h - \bar{u}_h, g_h)_{L^2(\tilde{\omega}_E)} + (z - z_h, g_h)_{L^2(\tilde{\omega}_E)} \\
 &\quad + (\bar{p}_h - \bar{p}, \nabla g_h)_{\mathbb{L}^2(\tilde{\omega}_E)} - \left(\mu \left[\frac{\partial \bar{u}_h}{\partial n} \right]_E + [\bar{p}_h \cdot \vec{n}]_E, \rho_E \mathbf{b}_E \right)_{L^2(E)}
 \end{aligned} \tag{4.15}$$

From (4.15), (4.4), (4.6) and (4.7) we infer

$$\begin{aligned}
 h_E \left(\mu \left\| \left[\frac{\partial \bar{u}_h}{\partial n} \right]_E \right\|_{L^2(E)} + \left\| [\bar{p}_h \cdot \vec{n}]_E \right\|_{L^2(E)} \right)^2 &\leq c_5 h_E \left(\mu \left[\frac{\partial \bar{u}_h}{\partial n} \right]_E + [\bar{p}_h \cdot \vec{n}]_E, \rho_E \mathbf{b}_E \right)_{L^2(E)} \\
 &= c_5 h_E \left[(\operatorname{div} \bar{p}_h + z_h + \mu \Delta \bar{u}_h - \bar{u}_h, g_h)_{L^2(\tilde{\omega}_E)} + (z - z_h, g_h)_{L^2(\tilde{\omega}_E)} \right. \\
 &\quad \left. + (\bar{p}_h - \bar{p}, \nabla g_h)_{L^2(\tilde{\omega}_E)} - a(e_h, g_h) \right] \\
 &\leq c_5 h_E \left[\|\operatorname{div} \bar{p}_h + z_h + \mu \Delta \bar{u}_h - \bar{u}_h\|_{L^2(\tilde{\omega}_E)} \|g_h\|_{L^2(\tilde{\omega}_E)} + \|z - z_h\|_{L^2(\tilde{\omega}_E)} \|g_h\|_{L^2(\tilde{\omega}_E)} \right. \\
 &\quad \left. + \|\bar{p}_h - \bar{p}\|_{L^2(\tilde{\omega}_E)} \|g_h\|_{H_0^1(\tilde{\omega}_E)} + c_{a_2} \|e_h\|_{H_0^1(\tilde{\omega}_E)} \|g_h\|_{H_0^1(\tilde{\omega}_E)} \right] \\
 &\leq c_5 h_E^{1/2} \|\rho_E\|_{L^2(E)} \left[c_6 h_E \|\operatorname{div} \bar{p}_h + z_h + \mu \Delta \bar{u}_h - \bar{u}_h\|_{L^2(\tilde{\omega}_E)} \right. \\
 &\quad \left. + c_6 h_E \|z - z_h\|_{L^2(\tilde{\omega}_E)} + c_7 \|\bar{p}_h - \bar{p}\|_{L^2(\tilde{\omega}_E)} + c_{a_2} c_7 \|e_h\|_{H_0^1(\tilde{\omega}_E)} \right]
 \end{aligned}$$

Since $\|\rho_E\|_{L^2(E)} \leq \mu \left\| \left[\frac{\partial \bar{u}_h}{\partial n} \right]_E \right\|_{L^2(E)} + \left\| [\bar{p}_h \cdot \vec{n}]_E \right\|_{L^2(E)}$, we obtain

$$\begin{aligned}
 h_E \left\| \left[\frac{\partial \bar{u}_h}{\partial n} \right]_E \right\|_{L^2(E)}^2 &\leq 2c_5^2 \left(c_6^2 h_E^2 \|\operatorname{div} \bar{p}_h + z_h + \mu \Delta \bar{u}_h - \bar{u}_h\|_{L^2(\tilde{\omega}_E)}^2 + c_6^2 h_E^2 \|z - z_h\|_{L^2(\tilde{\omega}_E)}^2 \right. \\
 &\quad \left. + c_{a_2}^2 c_7^2 \|e_h\|_{H_0^1(\tilde{\omega}_E)}^2 + c_7^2 \|\bar{p}_h - \bar{p}\|_{L^2(\tilde{\omega}_E)}^2 \right).
 \end{aligned} \tag{4.16}$$

Due to the regularity of the triangulation \mathcal{T}_h , the ratio h_T/h_E is uniformly bounded from below and above. Hence, (4.8) can be used in (4.16) to obtain (4.14). \square

Combining the previous auxiliary results, we conclude with the following theorem.

Theorem 8. *Let $\bar{u} \in H_0^1(\Omega)$ be the solution of the problem (\mathcal{P}) , let $\bar{u}_h \in V_h$ the finite element approximation to this solution, and let \bar{p}_h be an approximation to a solution of the dual problem (\mathcal{D}) , given by the system of equations (3.5)-(3.6). Further let ε_T be given as in (3.23). Then there exists a positive constant C that depends only on the shape regularity of $\mathcal{T}_h(\Omega)$ such that*

$$\begin{aligned}
 \sum_{T \in \mathcal{T}_h(\Omega)} \varepsilon_T &\leq C \left(\|e_h\|_{H_0^1(\Omega)}^2 + \sum_{T \in \mathcal{T}_h(\Omega)} h_T^2 \|z - z_h\|_{L^2(T)}^2 + \sum_{T \in \mathcal{T}_h(\Omega)} \|\bar{p}_h - \bar{p}\|_{L^2(T)}^2 \right) \\
 &= C \left(\|e_h\|_{H_0^1(\Omega)}^2 + \|\bar{p}_h - \bar{p}\|_{L^2(\Omega)}^2 + \sum_{T \in \mathcal{T}_h(\Omega)} h_T^2 \|z - z_h\|_{L^2(T)}^2 \right).
 \end{aligned}$$

Observe that in the case where $G(\cdot) = \alpha \int_{\Omega} \Phi_{\gamma}(\cdot)$ Theorem 5 and Theorem 8 establish reliability and efficiency of the residual-based a posteriori error estimator for $\|\bar{u} - \bar{u}_h\|_{H_0^1(\Omega)}^2 + \gamma^2 \|\bar{p} - \bar{p}_h\|_{L^2(\Omega)}^2$ for any $\gamma > 0$ (and, thus, for $\|\bar{u} - \bar{u}_h\|_{H_0^1(\Omega)}^2 + \|\bar{p} - \bar{p}_h\|_{L^2(\Omega)}^2$ with a constant depending on γ^{-1} for estimating the difference in the dual variables). In the case of $G(\bar{p}) = \alpha \int_{\Omega} |\bar{p}|_{L^2} dx$, however, only an estimator for $\|\bar{u} - \bar{u}_h\|_{H_0^1(\Omega)}$ is available where the efficiency bound depends on the difference of the dual variables \bar{p} and \bar{p}_h . This is due to the non-uniqueness of \bar{p} in (2.1)-(2.2) and similarly for \bar{p}_h in the discrete analogue of this system. As a consequence, for decreasing h the quantity \bar{p}_h need not get closer to \bar{p} in the \mathbb{L}^2 -norm. One only has that both vector-fields are bounded by α in the pointwise almost everywhere sense and, hence, $\|\bar{p} - \bar{p}_h\|_{L^2(\Omega)} \leq 2\alpha|\Omega|$.

5. ADAPTIVE FINITE ELEMENT ALGORITHM

In the rest of the present work we consider the regularized problem (\mathcal{P}_γ) , where \bar{u} and \bar{p} satisfy the optimality system (2.4)-(2.5). The finite dimension spaces $V_h \subset H_0^1(\Omega)$ and $P_h \subset \mathbb{L}^2(\Omega)$, where the finite element solutions \bar{u}_h and \bar{p}_h lie, are as follows

$$(5.1) \quad V_h := \{v_h \in C(\bar{\Omega}) : v_h|_T \in \Pi_1, \forall T \in \mathcal{T}_h\},$$

$$(5.2) \quad P_h := \{\vec{t}_h := (t_h^1, t_h^2) \in \mathbb{L}^2(\Omega) : t_h^1|_T, t_h^2|_T \in \Pi_0, \forall T \in \mathcal{T}_h\}.$$

These finite dimension spaces have the following bases

$$(5.3) \quad \mathcal{B}_V^h := \{\phi_1, \dots, \phi_n\}, \quad \mathcal{B}_P^h := \{\vec{\eta}_1, \dots, \vec{\eta}_{2m}\}.$$

Thus, the finite element solutions have the form

$$(5.4) \quad \bar{u}_h = \sum_{i=1}^n u_i \phi_i \in V_h \quad \text{and} \quad \bar{p}_h = \sum_{j=1}^{2m} p_j \vec{\eta}_j \in P_h.$$

The coefficients u_i , $i = 1, \dots, n$, and p_j , $j = 1, \dots, 2m$, are found by solving the system

$$(5.5) \quad -(\bar{p}_h, \nabla \phi_i)_{\mathbb{L}^2(\Omega)} = a(\bar{u}_h, \phi_i) - (z, \phi_i)_{L^2(\Omega)}, \quad \forall \phi_i \in \mathcal{B}_V^h,$$

$$(5.6) \quad \max(\gamma, |\nabla \bar{u}_h(x)|_{l^2}) \bar{p}_h(x) - \nabla \bar{u}_h(x) = 0.$$

Note that (5.6) is non-smooth due to the presence of the max-term. For this reason we choose a generalized Newton method for solving (5.5)-(5.6). In fact, it turns out that the system (5.5)-(5.6) is semismooth and the generalized Newton iteration converges locally at a superlinear rate; see [34] for further details in the imaging context and [32] for a more general discussion of semismooth Newton methods.

5.1. Adaptive finite element algorithm for image reconstruction. Throughout the remainder of this paper, the domain is given by $\Omega := (0, 1)^2$ and the initial triangulation in our numerics is a uniform one consisting of 16 triangles.

Given a regular triangulation $\mathcal{T}_{h^{(k)}}$ and its corresponding finite element spaces $V_{h^{(k)}}$, $P_{h^{(k)}}$, we compute the finite element solutions $\bar{u}_{h^{(k)}} \in V_{h^{(k)}}$ and $\bar{p}_{h^{(k)}} \in P_{h^{(k)}}$ by solving (5.5)-(5.6). We refer to [18, 34] for the numerical solution algorithm which is of semismooth Newton type.

We observe that for such finite element spaces we obtain that $\Delta \bar{u}_{h^{(k)}} = 0$ and $\text{div } \bar{p}_{h^{(k)}} = 0$, on each triangle. For each element $T \in \mathcal{T}_{h^{(k)}}$ the local error indicator is given by

$$(5.7) \quad \varepsilon_T^{(k)} := \eta_T + \frac{1}{2}(\eta_{\partial T, u} + \eta_{\partial T, p}),$$

where

$$(5.8) \quad \begin{aligned} \eta_T &= h_T^2 \|z - \bar{u}_{h^{(k)}}\|_{L^2(T)}^2, \\ \eta_{\partial T, u} &= \sum_{E \in \hat{\mathcal{E}}(T)} \eta_{E, u}, & \eta_{E, u} &= h_E \left\| \left[\frac{\partial \bar{u}_{h^{(k)}}}{\partial n} \right]_E \right\|_{L^2(E)}, \\ \eta_{\partial T, p} &= \sum_{E \in \hat{\mathcal{E}}(T)} \eta_{E, p}, & \eta_{E, p} &= h_E \left\| [\bar{p}_{h^{(k)}} \cdot \vec{n}]_E \right\|_{L^2(E)}. \end{aligned}$$

Assumption 9. *Based on the maximal resolution of the data image z , we assume that there exists a limit for the refinement, i.e. there exists $h_{\min} > 0$ such that for any $T \in \mathcal{T}_{h^{(k)}}$ and all iterations k , we have $h_T > h_{\min}$.*

5.2. Bulk criterion. For the local error for a triangle $T \in \mathcal{T}_{h^{(k)}}$ given by $\varepsilon_T^{(k)} := \eta_T + \frac{1}{2}\eta_{\partial T}$ we observe that the orders of the element error η_T and the jump discontinuity $\eta_{\partial T}$ are very different, respectively. The order of the element error is 10^{-7} , whereas the order of the jump discontinuity in the approximation to the normal flux is 10 and the one for the jump discontinuity in the approximation to the dual variable on the interface ∂T is 10^{-2} . Therefore we propose to use the bulk criterion for refinement; see [31]. At each iteration an element $T \in \mathcal{T}_h$ is marked to be divided if

- $T \in \mathcal{M}_h^T$ or
- at least two edges $E \in \hat{\mathcal{E}}_h(T)$ belong to \mathcal{M}_h^E ,

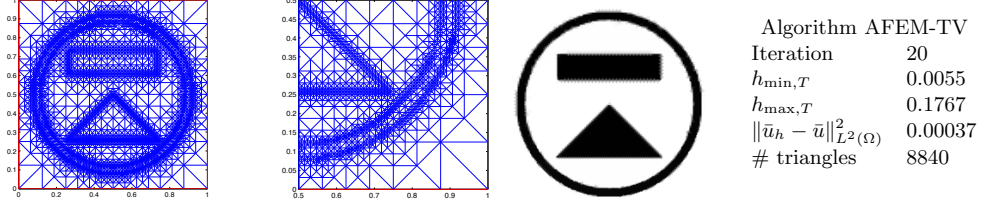


FIGURE 1. Restoration of a noise free image.

where, given the universal constants $\theta_1 > 0$, $\theta_2 > 0$ and $\theta_3 > 0$, with $\theta_i \in (0, 1)$, $i = 1, 2, 3$, the sets \mathcal{M}_h^T and \mathcal{M}_h^E are such that

$$(5.9) \quad \begin{aligned} \theta_1(\sum_{E \in \mathcal{E}_h} \eta_{E,u}) &\leq \sum_{E \in \mathcal{M}_h^{E,u}} \eta_{E,u}, \\ \theta_2(\sum_{E \in \mathcal{E}_h} \eta_{E,p}) &\leq \sum_{E \in \mathcal{M}_h^{E,p}} \eta_{E,p}, \\ \theta_3(\sum_{T \in \mathcal{T}_h} \eta_T) &\leq \sum_{T \in \mathcal{M}_h^T} \eta_T, \end{aligned}$$

$\mathcal{M}_h^E = \mathcal{M}_h^{E,u} \cup \mathcal{M}_h^{E,p}$. At each iteration, the resulting set of triangles $T \in \mathcal{T}_{h^{(k)}}$ marked for refinement is denoted by $\mathcal{D}_{h^{(k)}}^T$.

Remark 10. As it is well-known, the subdivision of an element T might lead to the division of one of its adjacent triangles as well in order to preserve the regularity of the triangulation. We refer to [42] for a description of possible mesh refinement techniques.

5.3. Noise level consideration. The image data is given by $z = \hat{u} + \eta$ where η represents white Gaussian noise with mean zero and variance σ^2 . The solution \bar{u} for an appropriate regularization parameter α satisfies the condition

$$(5.10) \quad \|z - \bar{u}\|_{L^2(\Omega)}^2 \leq \sigma^2,$$

see [13]. Thus, if the variance σ^2 is known, an over-refinement in homogeneous image regions can be avoided by imposing the following constraint:

$$(C_\sigma) \text{ An element } T \in \mathcal{T}_{h^{(k)}} \text{ may belong to the set } \mathcal{M}_{h^{(k)}}^T \text{ only if } \|z - \bar{u}_{h^{(k)}}\|_{L^2(T)}^2 > \tilde{\sigma}_{(\sigma^2, T, h_T^{(k)}, h_{\min})}^2.$$

Remark 11. The bound $\tilde{\sigma}_{(\sigma^2, T, h_T^{(k)}, h_{\min})}^2$ depends on the noise level σ^2 , on the size of the element T and on the resolution of the image given by h_{\min} . Numerically it needs to be computed in a similar way as it was done in Section 3.2 of [18]. For more details, we refer to Appendix A.

5.4. Coarsening. As a coarsening method we consider the strategy presented in [7]. Four adjacent triangles of the same level of division (the same size) as the ones shown in Figure 3 (a) can be assembled into 2 triangles R_1 and R_2 as the ones shown in Figure 3 (b) by eliminating the edges $[4, 5]$ and $[5, 2]$. Since $u_{h^{(k)}} \in H_0^1(\Omega)$, the edges $[4, 5]$ and $[5, 2]$ can be eliminated if

$$(5.11) \quad \|\nabla u_{h^{(k)}}\|_{L^2(T_i)} \leq \overline{\text{TOL}}, \quad i = 1, 2, 3, 4.$$

5.5. Final algorithm. Our proposed image restoration algorithm is as follows.

Algorithm AFEM-TV

- (1) Choose an initial regular mesh $\mathcal{T}_{h^{(0)}}$.
- (2) Compute the finite element approximations $\bar{u}_{h^{(k)}} \in V_{h^{(k)}}$, $\bar{p}_{h^{(k)}} \in P_{h^{(k)}}$ by solving (5.5)-(5.6).
- (3) Compute, for each triangle $T \in \mathcal{T}_{h^{(k)}}$, the values η_T and η_E for all $E \in \mathcal{E}_h(T)$.
- (4) Following the bulk criterion presented in section 5.2, determine the set $\mathcal{D}_{h^{(k)}}$ of elements to be divided.

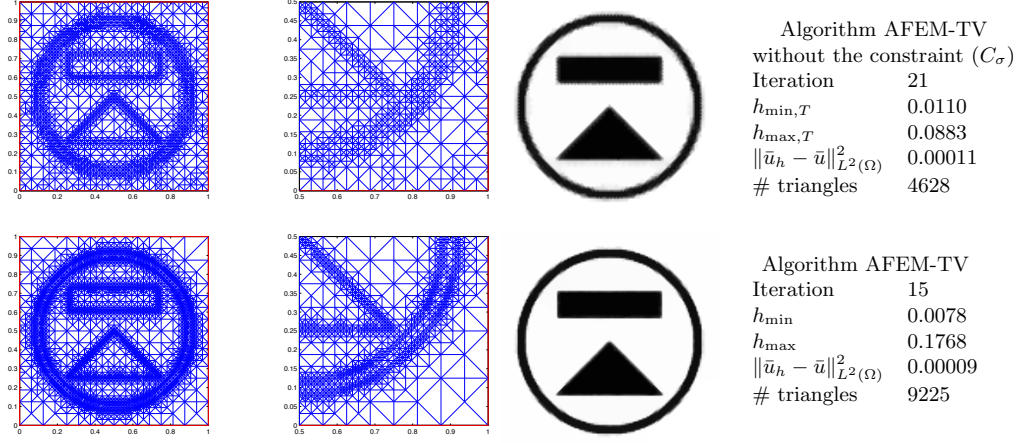
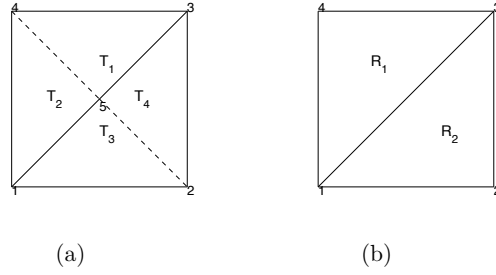
FIGURE 2. Restoration of an image corrupted with Gaussian ($\sigma = 0.1$).

FIGURE 3. Coarsening.

- (5) For every set of triangles $\{T_1, T_2, T_3, T_4\} \subset \mathcal{T}_{h^{(k)}} \setminus \mathcal{D}_{h^{(k)}}$ as the ones presented in Figure 3 (a), eliminate the corresponding edges $[4, 5]$ and $[5, 2]$, whenever

$$\|\nabla u_{h^{(k)}}\|_{L^2(T_i)} \leq \overline{\text{TOL}}, \quad i = 1, 2, 3, 4.$$

- (6) Refine the elements $T \in \mathcal{D}_{h^{(k)}}$. Set $k := k + 1$ and return to (2).

Remark 12. The approximation of the data z given by z_h is crucial for the results obtained by Algorithm AFEM-TV. In the case of noise free images, we use the following approximation

$$(5.12) \quad (z_h)_i = \int_T z \phi_i d\bar{x} \approx \frac{1}{3} |T| z(x_c, y_c),$$

where $|T|$ is the area of the triangle T and (x_c, y_c) are the coordinates of the center of gravity of T , as proposed in [4]. However in the case of noisy images, the information given by $z(x_c, y_c)$ is corrupted by noise and can no longer be used. Thus, instead of only considering $z(x_c, y_c)$ we consider $\bar{z}_T = \frac{1}{|T|} \int z d\bar{x}$ and we obtain

$$(5.13) \quad \int_T z \phi_i d\bar{x} \approx \frac{1}{3} |T| \bar{z}_T.$$

Note that in our case we approximate \bar{z}_T by taking the average of the discrete data values of z contained in T .

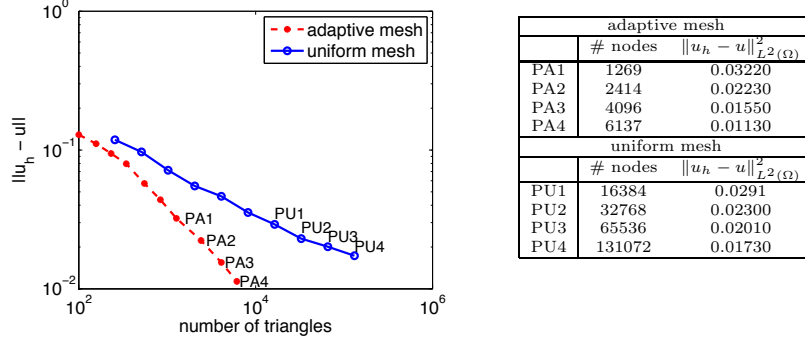


FIGURE 4. Performance of the algorithm, two dimensional case.

5.5.1. *Numerical examples.* The reconstruction of a noise-free benchmark image by the algorithm AFEM-TV is presented in Figure 1. The left plot shows the final adaptive mesh, the next to the right plot provides a zoomed region of the mesh showing the local refinement and the third plot depicts the reconstruction result. Finally, the number of refinement steps (20), the minimal and maximal length of the hypotenuses of triangles is denoted by $h_{\min,T}$ and $h_{\max,T}$ and ‘# triangles’ denotes the final number of triangles. The numerical result of this algorithm for the restoration of the same image but now corrupted by white Gaussian noise with $\sigma = 0.1$ is shown in the second row of Figure 2. In the first row of Figure 2 we show the results obtained when not imposing the constraint (C_σ), i.e. there might be triangles $T \in \mathcal{M}_{h(k)}^T$ satisfying

$$(5.14) \quad \|z - \bar{u}_{h(k)}\|_{L^2(T)}^2 \leq \tilde{\sigma}_{(\sigma^2, T, h_T^{(k)}, h_{\min})}^2,$$

which are subject to refinement. This increases the number of triangles marked for refinement in homogeneous regions. On the other hand, the result shown in the second row of Figure 2 relies on the condition (C_σ). It can be clearly observed that, although more triangles are used, the triangulation of the homogeneous region in the lower right corner remains coarser than in the first row.

5.5.2. *Numerical specifications.* In step (2) of algorithm AFEM-TV, the finite element solutions $\bar{u}_{h(k)} \in V_{h(k)}$ and $\bar{p}_{h(k)} \in P_{h(k)}$ are computed by means of the semismooth Newton algorithm described in [34]. The choice of the parameters for the semismooth Newton method is as follows: $\mu = 10^{-8}$, $\gamma = 10^{-6}$, $\alpha = 6 \cdot 10^{-4}$ in the case of a noise free image and $\alpha = 10^{-3}$ in the case of an image corrupted with white Gaussian noise with $\sigma = 0.1$.

In step (5) of Algorithm AFEM-TV we use $\overline{\text{TOL}} = 10^{-5}$. The parameters in the bulk criterion are $\theta_1 = 0.6$, $\theta_2 = 0.2$ and $\theta_3 = 0.6$. The maximal resolution of the image is given $h_{\min} = \sqrt{2}/256$.

6. NUMERICAL RESULTS

6.1. **Efficiency of the algorithm.** We consider the noisy image related to the result in Figure 2. The associated efficiency graph is presented in Figure 4. We observe that the algorithm using the adaptive mesh strategy outperforms the one using a uniform mesh. In order to reach an error of $\|\bar{u} - \bar{u}_h\|_{L^2(\Omega)}^2 \approx 0.01$, the adaptive mesh needs around 6000 triangles and the uniform mesh needs more than 130000; compare points PA4 and PU4. The numerical specifications are the ones given in section 5.5.2. The performance of this algorithm for more complex images is evaluated in the rest of the paper.

6.2. **General remarks.** The test images used in this section are the ones shown in Figure 5. The original images are the so called “Clock” in (a) and the “Cameraman” in (c), both of size 512×512 pixels. Thus, the maximal resolution of the image is $h_{\min} = \sqrt{2}/512$. These images are corrupted by white Gaussian noise with standard deviation $\sigma = 0.1$; see (b) and (d). Unless otherwise indicated, the general numerical specifications in the following tests are as described in section 5.5.2. We present

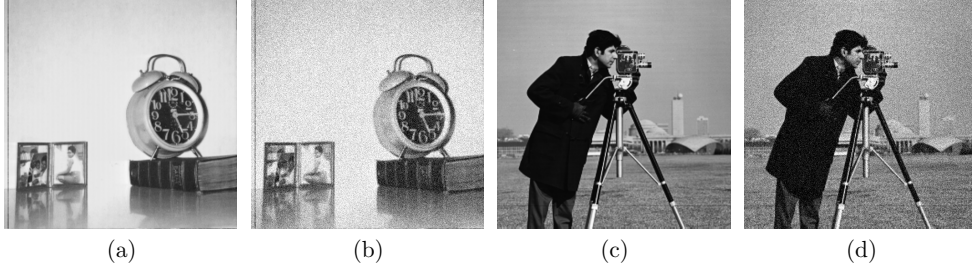


FIGURE 5. **Test images.** Clock: (a) noise free, (b) with white Gaussian noise $\sigma = 0.1$. Cameraman: (c) noise free, (d) with white Gaussian noise $\sigma = 0.1$.

the final adapted mesh, the reconstructed image and, in some cases, a zoom of this mesh. Further, we compare the results with the reconstruction of the image by finite elements with a very fine uniform mesh given by $h_{\min,T} = h_{\max,T} = 1/256$. In Figures 7 and 9 we present the evolution of the peak signal-to-noise ratio (PSNR) according to the number of triangles.

6.3. Test of algorithm AFEM-TV. We start by reconstructing the original images of the Clock and the Cameraman presented in Figure 5 (a) and (c), respectively. The reconstruction of the Clock with Algorithm AFEM-TV is shown in the first row of Figure 6 (a)-(b) and the reconstruction of the Cameraman is in the second row (d)-(e). In both cases we observe that homogeneous regions like the background of the Clock or the sky in the Cameraman require only a coarse triangulation for their successful reconstruction. In fact, we observe that not only the sky of the Cameraman remains homogeneous but also that the shaded background of the Clock is well preserved. Detail regions like the portrait and the clock in the Clock or the camera in the Cameraman require more refinement. The edges also need more refinement, of course. In the resulting adapted meshes we can clearly observe the contours of the portrait, the clock and the book in the Clock and the contour of the man and the camera in the Cameraman. Small details like the numbers in the Clock are well recovered although the white buildings in the back of the Cameraman are not as well reconstructed. This is possibly due to the low contrast between the buildings and the sky. In Figure 6 (c), (f) we present the reconstructions given by the fine uniform mesh. We observe that although details like the buildings in the cameraman are better reconstructed with the uniform fine mesh, the cost of this reconstruction is extremely high when compared to the results of Algorithm AFEM-TV. In Figure 7 we show that a higher PSNR is obtained when using an adaptive mesh rather than using a uniform mesh with a similar number of degrees of freedom. Note here that AFEM is allowed to refine down to the scale given by the underlying pixel frame of the image.

In the following experiment we restore the noisy images of the Clock and the Cameraman presented in Figure 5 (b) and (d). The results are depicted in Figure 8 and the performance of the algorithm is shown in Figure 9. The homogeneous regions like the background of the Clock or the sky in the Cameraman require a coarser mesh than the regions with smaller features, like the clock or the portrait in the Clock or the camera in the Cameraman. The edges are also well detected by the mesh which shows a stronger refinement at the contour of the book, the clock or the portrait in the Clock or the contour of the man in the Cameraman. Although the quality of the final reconstruction is not as good as the one given by the fine uniform mesh with $h_{\min,T} = h_{\max,T} = 1/256$, the reconstruction is still acceptable (compare, for instance, the PSNR-values shown in Figure 9). Similar to the reconstruction of noise free images, we observe in Figure 9 that a higher PSNR is obtained when using an adaptive mesh rather than when using a uniform mesh, both with approximately the same number of degrees of freedom.

6.4. Discussion on the convergence of the dual variable \bar{p}_h . Our numerics utilize the case where $G = j_\gamma$ in Section 3, i.e., the finite elements solutions \bar{u}_h and \bar{p} satisfy equation (5.6). Hence, according to Theorem 5, we observe that the a posteriori error bound also estimates the error in the dual variable given by $\|\bar{p} - \bar{p}_h\|_{L^2(\Omega)}$ since Lemma 4 shows that this error is bounded by the error of

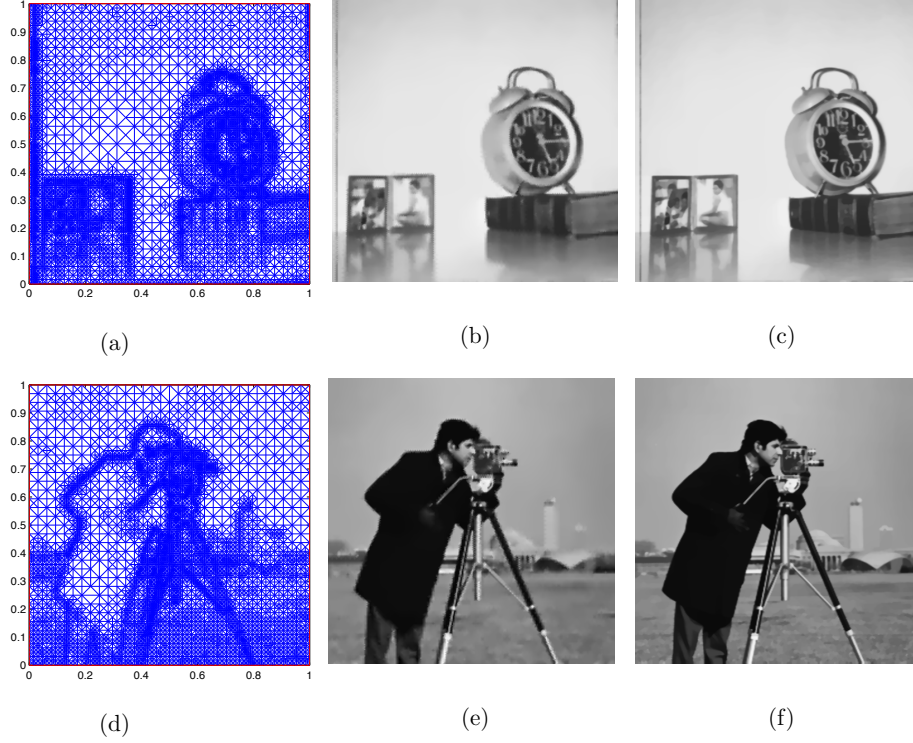


FIGURE 6. **Restoration of noise free images** $\alpha = 1e - 004$. (a),(d): Adapted meshes given by Algorithm AFEM-TV. (b), (e): Restored images given by Algorithm AFEM-TV. (c), (f): Restored images given with the fine uniform mesh.

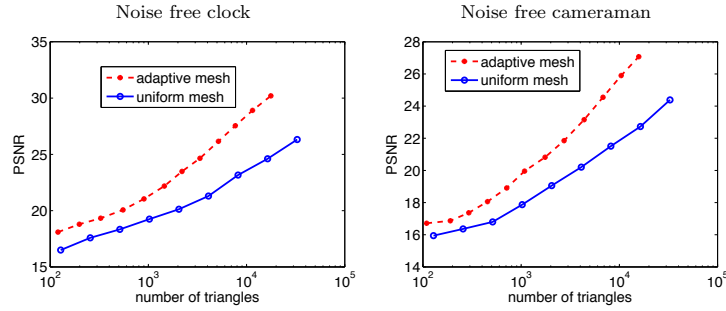


FIGURE 7. Restoration of noise free images

the approximation \bar{u}_h given by $\|\bar{u}_h - \bar{u}\|_{H_0^1(\Omega)}$. Numerically it can indeed be shown that this error decreases as the number of triangles increases. As an example, in Figure 10 we show the convergence of the error $\|\bar{p} - \bar{p}_h\|_{L^2(\Omega)}$ for the reconstruction of the noise free and the noisy image of the Clock, respectively.

7. CONCLUSIONS

The results of this paper show that the adaptive finite element method (AFEM) may be used as an efficient discretization tool in total variation based image denoising. Based on a suitably chosen

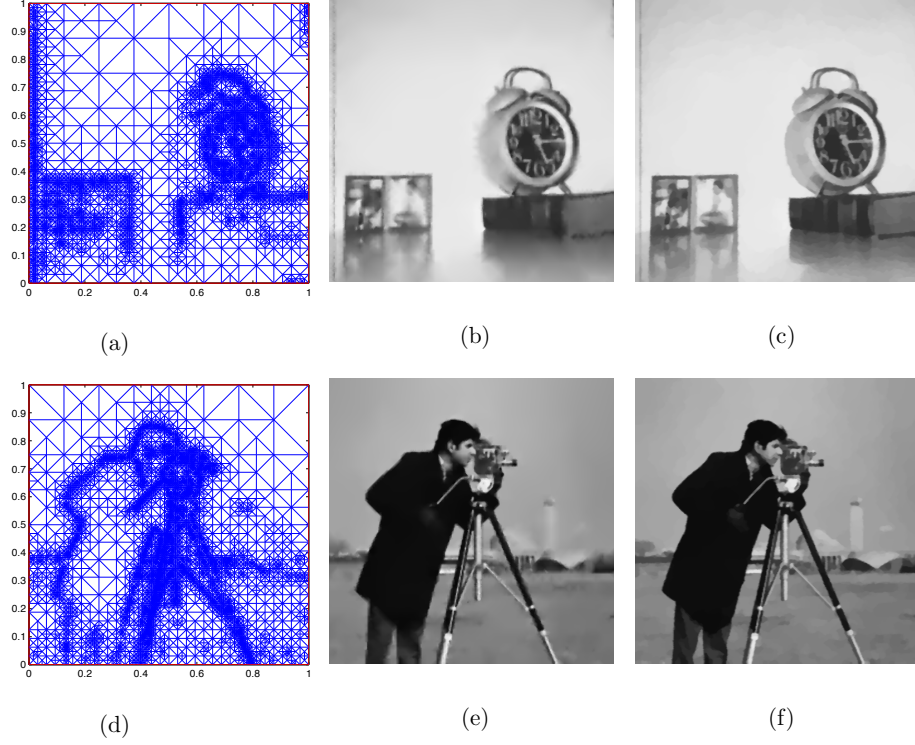


FIGURE 8. **Restoration of noisy images** $\alpha = 1e - 004$. (a),(d): Adapted meshes given by Algorithm AFEM-TV. (b), (e): Restored images given by Algorithm AFEM-TV. (c), (f): Restored images given with the fine uniform mesh.

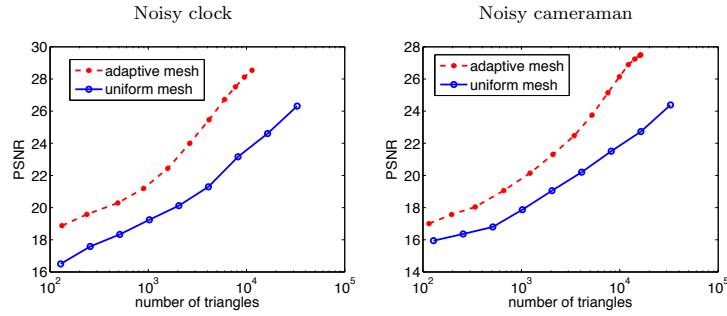
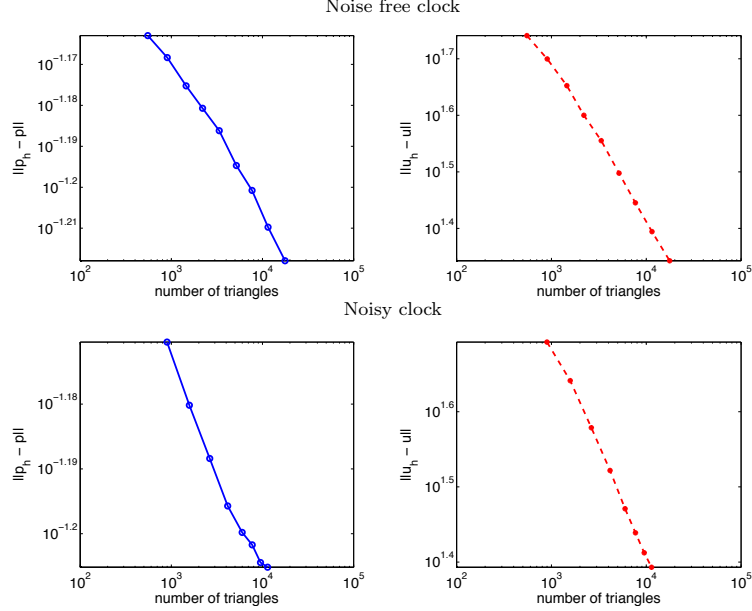


FIGURE 9. Restoration of noisy images

error estimator, it automatically adapts in image regions with details and edges while homogeneous regions are kept on a coarse discretization level. The proposed a posteriori error bound for adapting the mesh and finding a finite element approximation to the image is highly accurate for signals or images without noise. However, when noise is present, AFEM needs to be equipped with additional features to avoid over-refinement. For this purpose, condition (C_σ) of this paper relies on an estimate of the noise variance which takes the number of pixels (sample points) in each triangle into account and prevents an excessive refinement. A similar technique was used by the authors in earlier work for locally adapted regularization. Concerning the chosen finite element spaces we note that when $\mu = 0$, then one has to resort to Raviart-Thomas FE spaces [10] as the pre-dual problem of (1.1) is

FIGURE 10. Convergence of the error on the dual variable $\|\vec{p} - \vec{p}_h\|_{\mathbb{L}^2(\Omega)}$.

then posed in $H_0(\text{div})$ [33], the space of vector-fields \vec{p} in $\mathbb{L}^2(\Omega)$ with $\text{div } \vec{p} \in L^2(\Omega)$ and $\vec{p} \cdot \vec{n} = 0$ on $\Gamma = \partial\Omega$. Still, the refinement techniques and the tools for dealing with noise in the data etc. which are developed in the present paper will be important in this case.

Acknowledgement. The authors acknowledge financial support by the Austrian Science Fund FWF under START-program Y305 “Interfaces and Free Boundaries” and the SFB F32 “Mathematical Optimization and Applications in Biomedical Sciences”. Further, the authors would like to thank S. Gonzalez (Escuela Politécnica Nacional Quito-Ecuador) for scientific discussions.

APPENDIX A. BOUND FOR THE VARIANCE

In this section we utilize the theory presented in [18] for finding local variance estimators.

We assume that the data z is a matrix of $N \times N$ pixels. Thus, in the discrete setting the noise η in $z = \hat{u} + \eta$ can be regarded as a sample of size $N \times N$ of a normally distributed random variable with zero mean and variance σ^2 . Each element $T \in \mathcal{T}_{h^{(k)}}$ contains $\omega = \lceil |T|/(h_{\min}^2) \rceil$ pixels. Then the random variable

$$(A.1) \quad W_T^\omega = \frac{1}{\sigma^2} \sum_{(i,j) \in T} (\eta_{i,j})^2$$

has the χ^2 -distribution with ω degrees of freedom, i.e. $W_T^\omega \sim \chi_\omega^2$. Here, the notation $(i,j) \in T$ refers to the associated point of Ω lying in T . If $u = \hat{u}$ satisfies $\eta = z - \hat{u}$, then

$$(A.2) \quad S_T^\omega = \sum_{(i,j) \in T} (z_{i,j} - \hat{u}_{i,j})^2 = \sum_{(i,j) \in T} (\eta_{i,j})^2 = \sigma^2 W_T^\omega.$$

If, on the other hand, u is not the true signal and $z - u$ contains image details rather than noise only, then we expect that

$$(A.3) \quad S_T^\omega = \sum_{(i,j) \in T} (z_{i,j} - u_{i,j})^2 > \sum_{(i,j) \in T} (\eta_{i,j})^2 = \sigma^2 W_T^\omega.$$

Therefore, we search for a bound B such that $S_T^\omega > B$ implies that the element T contains details. Given $N \times N$ the total number of pixels in the image, we consider the expected maximum of a sample of size $l = \lceil N \times N / \omega \rceil$ coming from a χ^2 -distribution with ω degrees of freedom. We write $B^{\omega,l} = B(\omega, l)$ with

$$(A.4) \quad B^{\omega,l} := \sigma^2 \mathfrak{E}(\max_{i=1,\dots,l} W_i^\omega),$$

i.e. the expected value. This bound can be computed as follows: Let $\kappa = 0.577215$ be the Euler-Mascheroni constant, \mathfrak{f} be the χ^2 -distribution with ω degrees of freedom and \mathfrak{F} denote its cumulative distribution function, i.e.

$$(A.5) \quad \mathfrak{F}(W) = \int_{-\infty}^W \mathfrak{f}(w) dw.$$

Now, let \tilde{W} be such that

$$(A.6) \quad \mathfrak{F}(\tilde{W}) = 1 - \frac{1}{l},$$

and let $\beta = l\mathfrak{f}(\tilde{W})$. Thus

$$(A.7) \quad \mathfrak{E}(\max_{i=1,\dots,l} W_i^\omega) = \tilde{W} + \frac{\kappa}{\beta}.$$

For a detailed description see [18] and [25].

Thus the bound $\tilde{\sigma}_{(\sigma \cdot h_i^{(k)}, h_{\min})}^2$ is

$$(A.8) \quad \tilde{\sigma}_{(\sigma \cdot h_i^{(k)}, h_{\min})}^2 = \sigma^2 \left(\tilde{W} + \frac{\kappa}{\beta} \right).$$

REFERENCES

- [1] R. Acar and C.R. Vogel. Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Problems*, 1994.
- [2] R.A. Adams and J.J. Fournier. *Sobolev Spaces*, volume 140. Academic Press, 2nd edition, 2003.
- [3] M. Ainsworth and J.T. Oden. *A Posteriori Error Estimation in Finite Element Analysis*. Wiley, New York, 2000.
- [4] J. Albery, C. Carstensen, and S. Funken. Remarks around 50 lines of matlab: short finite element implementation. *Numerical Algorithms*, 20:239–263, 1999.
- [5] W. Bangerth and A. Joshi. Adaptive finite element methods for the solution of inverse problems in optical tomography. *Inverse Problems*, 24:22p., 2008.
- [6] W. Bangerth and R. Rannacher. *Adaptive Finite Element Methods for Differential Equations*. Birkhäuser, 2003.
- [7] E. Bänsch and K. Mikula. A coarsening finite element strategy in images selective smoothing. *Computing and Visualization in Science*, 1:53–61, 1997.
- [8] C. Bazan and P. Blomgren. Adaptive finite element method for image processing. In *Proceedings of the COMSOL Multiphysics Conference 2005 Boston*, 2005.
- [9] V. Bostan, W. Han, and B.D. Reddy. A posteriori error estimation and adaptive solution of elliptic variational inequalities of the second kind. *Applied Numerical Mathematics*, 52:13–38, 2005.
- [10] D. Braess and R. Verfürth. A posteriori error estimators for the Raviart-Thomas element. *SIAM J. Numer. Anal.*, 33(6):2431–2444, 1996.
- [11] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods*. Springer-Verlag, New York, 1991.
- [12] A. Chambolle. An algorithm for total variation minimization and application. *Journal of Mathematical Imaging and Vision*, 20:89–97, 2004.
- [13] A. Chambolle and P.-L. Lions. Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76:167–188, 1997.
- [14] Q. Chang and I.-L. Chern. Acceleration methods for total variation-based image denoising. *SIAM J. Applied Mathematics*, 25:982–994, 2003.
- [15] P.G. Ciarlet. *The Finite Element Method for Elliptic Problems*. North Holland, 1978.
- [16] P. Clément. Approximation by finite element functions using local regularization. *RAIRO Analyse Numérique*, 2:77–84, 1975.
- [17] D. C. Dobson and C. R. Vogel. Convergence of an iterative method for total variation denoising. *SIAM J. Numer. Anal.*, 34:1779–1791, 1997.
- [18] Y. Dong, M. Hintermüller, and M.M. Rincon-Camacho. Automated Regularization Parameter Selection in Multi-Scale Total Variation Models for Image Restoration. *Journal of Mathematical Imaging and Vision*, 40(1):82, 2011.
- [19] I. Ekeland and R. Témam. *Convex Analysis and Variational Problems*. Classics Appl. Math. 28, SIAM, 1999.
- [20] X. Feng and A. Prohl. Analysis of total variation flow and its finite element approximations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 37(3):533–556, 2003.
- [21] M. Fried. Multichannel image segmentation using adaptive finite elements. *Computing and Visualization in Science*, 12(3):125–135, 2008.
- [22] E. Giusti. *Minimal Surfaces and Functions of Bounded Variation*. Birkhäuser, 1984.
- [23] R. Glowinski, J.L. Lions, and R. Trémolieres. *Analyse Numérique des inéquations variationnelles*. Dunod, Paris, France, 1976.
- [24] S. González. *Semismooth Newton and Path-Following Methods for the Numerical Simulation of Bingham Fluids*. PhD thesis, Escuela Politécnica Nacional, Ecuador, 2008.
- [25] E.J. Gumbel. Les valeurs extrêmes des distributions statistiques. *Annales de l’institut Henri Poincaré*, 5(2):115–158, 1935.

- [26] M. Guven et al. Effect of discretization error analysis and adaptive mesh generation in diffuse optical absorption imaging: I. *Inverse Problems*, 23:1115–1133, 2007.
- [27] M. Guven et al. Discretization error analysis and adaptive meshing algorithms for fluorescence diffuse optical tomography: part I. *IEEE Trans. Med. Imag.*, 29:217–229, 2010.
- [28] E. Haber, S. Heldmann, and U. Ascher. Adaptive finite volume method for distributed non-smooth parameter identification. *Inverse Problems*, 23(4):1659–1676, 2007.
- [29] E. Haber, S. Heldmann, and J. Modersitzki. An octree method for parametric image registration. *SIAM J. on Scientific Computing*, 29:2008–2023, 2007.
- [30] E. Haber, S. Heldmann, and J. Modersitzki. Adaptive mesh refinement for non parametric image registration. *SIAM J. on Scientific Computing*, 30:3012–3027, 2008.
- [31] M. Hintermüller, R.H.W. Hoppe, Y. Iliash, and M. Kieweg. An a posteriori error analysis of adaptive finite element methods for distributed elliptic control problems with control constraints. *ESAIM: COCV*, 14(3):540 – 560, 2008.
- [32] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semismooth newton method. *SIAM Journal on Optimization*, 13(3):865–888, 2002.
- [33] M. Hintermüller and K. Kunisch. Total bounded variation regularization as bilaterally constrained optimization problem. *SIAM J. Appl. Math.*, 64:1311–1333, 2004.
- [34] M. Hintermüller and G. Stadler. An infeasible primal-dual algorithm for total bounded variation-based inf-convolution-type image restoration. *SIAM Journal on Scientific Computing*, 28(1):1–23, 2006.
- [35] L. I. Rudin, S. Osher and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.
- [36] T. Preusser and M. Rumpf. An adaptive finite element method for large scale image processing. In *Proceedings of the Second International Conference on Scale-Space Theories in Computer Vision*, 223–234.
- [37] L. Rudin. MTV-multiscale total variation principle for a pde-based solution to nonsmooth ill-posed problem. Technical report, Cognitech, Inc. Talk presented at the Workshop on Mathematical Methods in Computer Vision, University of Minnesota, 1995.
- [38] S. Osher, M. Burger, D. Goldfarb, J. Xu and W. Yin. An iterative regularization method for total variation-based image restoration. *SIAM Multiscale Model. and Simu.*, 4:460–489, 2005.
- [39] D. Strong and T. Chan. Spatially and scale adaptive total variation based regularization and anisotropic diffusion in image processing. Technical report, UCLA, 1996.
- [40] D. Strong and T. Chan. Edge-preserving and scale-dependent properties of total variation regularization. *Inverse Problems*, 19:165–187, 2003.
- [41] R. Temam. *Navier-Stokes Equations*. North-Holland, Amsterdam, 1977.
- [42] R. Verfürth. *A review of a posteriori error estimation and adaptive mesh-refinement techniques*. Teubner, Stuttgart, 1994.
- [43] R. Verfürth. *A Review of A Posteriori Error Estimation and Adaptive Mesh Refinement Techniques*. Wiley Teubner, 1996.
- [44] C.R. Vogel. *Computational Methods for Inverse Problems*, volume 23 of *Frontiers Appl. Math.* SIAM–Society of Industrial and Applied Mathematics, 2002.