Comput. Methods Appl. Mech. Engrg. xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

Comput. Methods Appl. Mech. Engrg.

journal homepage: www.elsevier.com/locate/cma

Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems ☆

Linda El Alaoui [a], Alexandre Ern [b], Martin Vohralík [c,d,*]

[a] LAGA, Université Paris 13, 99 avenue Jean-Baptiste Clément, 93430 Villetaneuse, France
[b] Université Paris-Est, CERMICS, Ecole des Ponts, 6 & 8 av. B. Pascal, 77455 Marne-la-Vallée, France
[c] UPMC Univ. Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, 75005 Paris, France
[d] CNRS, UMR 7598, Laboratoire Jacques-Louis Lions, 75005 Paris, France

ARTICLE INFO

Article history:
Received 31 July 2009
Received in revised form 4 February 2010
Accepted 23 March 2010
Available online xxxx

Keywords:
A posteriori error estimate
Monotone nonlinear problem
Linearization
Stopping criterion
Guaranteed upper bound
Robustness

ABSTRACT

We derive a posteriori error estimates for a class of second-order monotone quasi-linear diffusion-type problems approximated by piecewise affine, continuous finite elements. Our estimates yield a guaranteed and fully computable upper bound on the error measured by the dual norm of the residual, as well as a global error lower bound, up to a generic constant independent of the nonlinear operator. They are thus fully robust with respect to the nonlinearity, thanks to the choice of the error measure. They are also locally efficient, albeit in a different norm, and hence suitable for adaptive mesh refinement. Moreover, they allow to distinguish, estimate separately, and compare the discretization and linearization errors. Hence, the iterative (Newton–Raphson, fixed point) linearization can be stopped whenever the linearization error drops to the level at which it does not affect significantly the overall error. This can lead to important computational savings, as performing an excessive number of unnecessary linearization iterations can be avoided. A strategy combining the linearization stopping criterion and adaptive mesh refinement is proposed and numerically tested for the p-Laplacian.

1. Introduction

Let Ω be an open polyhedron of $\mathbb{R}^d$, $d \geqslant 2$. We consider the nonlinear problem in conservative form

$$-\nabla\cdot\boldsymbol{\sigma}(\nabla u) = f \quad \text{in } \Omega, \tag{1.1a}$$
$$u = 0 \quad \text{on } \partial\Omega. \tag{1.1b}$$

The scalar-valued unknown function u is termed the *potential*, and the $\mathbb{R}^d$-valued function $-\boldsymbol{\sigma}(\nabla u)$ is termed the *flux*. We assume that the flux function $\boldsymbol{\sigma}:\mathbb{R}^d\to\mathbb{R}^d$ takes the following quasi-linear form

$$\forall\boldsymbol{\xi}\in\mathbb{R}^d,\quad \boldsymbol{\sigma}(\boldsymbol{\xi}) = a(|\boldsymbol{\xi}|)\boldsymbol{\xi}, \tag{1.2}$$

where $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^d$ and where $a:\mathbb{R}_+\to\mathbb{R}$ is a given function. The function a is assumed below to satisfy a growth condition of the form $a(x)\sim x^{p-2}$ as $x\to+\infty$ for some real number $p\in(1,+\infty)$, so that the natural energy space V for the above model problem is the Sobolev space $W_0^{1,p}(\Omega)$. The data f is taken in $L^q(\Omega)$ where $q := \frac{p}{p-1}$ so that $\frac{1}{p}+\frac{1}{q}=1$. Hence, the model problem in weak form amounts to finding $u\in V$ such that

$$(\boldsymbol{\sigma}(\nabla u),\nabla v) = (f,v)\quad \forall v\in V, \tag{1.3}$$

where $(\cdot,\cdot)$ denotes the integral over Ω of the (scalar) product of the two arguments. The function a satisfies monotonicity and continuity conditions stated in Section 2 and ensuring that the problem (1.3) is well-posed.

The prototypical example for the present model problem is the so-called p-Laplacian, for which $a(x)=x^{p-2}$. The a priori error analysis for approximating the p-Laplacian by piecewise affine, continuous finite elements has been started by Glowinski and Marrocco [23,24]; see also Ciarlet [15, p. 312]. One well-known difficulty when working with the natural energy norm is that the derived error estimates are not sharp. This drawback has been circumvented by Barrett and Liu [6] upon introducing a so-called quasi-norm, thereby achieving optimal approximation results. The quasi-norm of the error between the exact solution u and the approximate solution, say u_h, is a weighted L^2-norm of the gradient $\nabla(u-u_h)$, where the weight depends on ∇u and ∇u_h.

The a posteriori error analysis of finite element approximations to a large class of nonlinear problems, including the present model problems, has been started by Verfürth; see [33] and [34, p. 47]. The main result is a two-sided bound of the energy error by the dual norm of the residual multiplied by suitable norms of the

☆ This work was partially supported by the GNR MOMAS (PACEN/CNRS, ANDRA, BRGM, CEA, EDF, IRSN, France).

* Corresponding author at: UPMC Univ. Paris 06 UMR 7598, Laboratoire Jacques-Louis Lions, 75005 Paris, France.
E-mail addresses: elalaoui@math.univ-paris13.fr (L. El Alaoui), ern@cermics.enpc.fr (A. Ern), vohralik@ann.jussieu.fr (M. Vohralík).

0045-7825/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.cma.2010.03.024

linearized operator at the exact solution, under the assumption that this latter operator is invertible and locally Lipschitz-continuous and that the approximate solution is sufficiently close to the exact solution. This yields in particular residual-based estimators in the energy norm. These estimators have been exploited, in particular, by Veeser [32] to prove the convergence of an adaptive finite element method for the p-Laplacian. Alternatively, quasi-norm error estimates for the p-Laplacian have been analyzed by Liu and Yan [28–30], leading to weighted residual-based estimators. Quasi-norm residual-based estimators have been further explored by Carstensen and Klose [9] with a focus on evaluating the constants in the estimates and under the assumption that the gradient norm of the approximate solution is positive everywhere in the domain. Moreover, gradient recovery techniques have been analyzed by Carstensen et al. [10] to estimate the quasi-norm of the error. Quite recently, Diening and Kreuzer [19] have obtained two-sided bounds for an appropriate measure of the error and proven the linear convergence of a suitable adaptive finite element method. The error measure is the L^2-norm of the difference $\mathbf{F}(\nabla u) - \mathbf{F}(\nabla u_h)$, where the auxiliary vector field $\mathbf{F}$ is such that $\mathbf{F}(\xi) = |\boldsymbol{\sigma}(\xi)|^{\frac{1}{2}}|\xi|^{-\frac{1}{2}}\xi$. This error measure turns out to be equivalent to the quasi-norm of the error, with constants depending on the nonlinearity (that is, the properties of the function a in (1.2)).

We observe that, whatever the error measure, the above bounds on the error involve constants depending on the function a. In the case of the p-Laplacian, they depend on the Lebesgue exponent p. Moreover, with a few exceptions, e.g., [9], the error upper bounds involve unknown generic constants. Therefore, the first objective of this work is to derive *guaranteed* bounds on the error, that is, error upper bounds without undetermined constants, and at the same time ensure *robustness*, that is, two-sided error bounds whose ratio is independent of the nonlinearity. To this purpose, we use as error measure a residual flux-based dual norm, namely

$$\mathcal{J}_u(u_h) = \sup_{v\in V\setminus\{0\}} \frac{(\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}(\nabla u_h), \nabla v)}{\|v\|_V}.$$

Working with residual flux-based quantities to measure the error is somewhat natural since fluxes satisfy basic conservation properties that are at the heart of approximation methods, even using continuous finite elements. Furthermore, the idea of using a dual norm is inspired by the work of Verfürth where dual norms have been considered, e.g., in the context of parabolic [38] and convection-dominated stationary convection–diffusion equations [40]. Dual residual norms have also been considered for nonlinear problems in [33], and the present dual norm has been considered in [11,12]. More recently, it has been observed in [46] that residual flux-based error measures are also natural in the context of diffusion problems with heterogeneous coefficients. Furthermore, we remark that although our error upper bounds are fully computable, the actual error measure is not, even if the exact solution is known; we will discuss below how the error measure can be approximated in numerical experiments with synthetic exact solutions so as to compute effectivity indices. Note, however, that in practical computations, the exact solution is never known and hence the error is never computable. We also point out that achieving robust error estimates does not mean necessarily that the error bounds can be extended to the limit cases $p = 1$ or $p = +\infty$, similar to the vanishing-diffusion limit in convection–diffusion equations.

Our a posteriori error estimates are formulated in terms of a $\mathbf{H}$(div)-conforming flux reconstruction. For conforming finite element methods, related earlier work in the linear case includes [1] (here the flux is not explicitly reconstructed) and [7,17,27,31]. In the spirit of Luce and Wohlmuth [31], guaranteed a posteriori estimates of the present type were proposed in [45] for the Laplace equation. They have been shown robust for inhomogeneous and anisotropic diffusion in [46] and for the reaction–diffusion case in [13]. We also refer to [22] for a unified setting encompassing various discretization methods in the context of the heat equation. Recently, Verfürth [41] derived another estimate based on flux reconstruction for singularly perturbed diffusion problems and, similar to [13,22,45,46], proved (see [41, Proposition 2.2]) that this estimate is a lower bound for the classical residual one of [34]. In the nonlinear case, the only work deriving a posteriori estimates based on flux reconstruction we are aware of is [26]. Therein, quasi-linear diffusion problems similar to (1.1a)–(1.1b) are discretized by various nonconforming locally conservative methods.

In the present paper, the a posteriori error analysis based on $\mathbf{H}$(div)-conforming flux reconstruction proceeds as follows. The error upper bound hinges on a local conservation property of the reconstructed flux, say $\mathbf{t}_h$; see Assumption 3.4. The error lower bound hinges instead on an approximation property of $\mathbf{t}_h$; see Assumption 4.1. This approximation property enables us to prove that our estimates are lower bounds for the classical residual ones. We provide two examples for reconstructing the flux $\mathbf{t}_h$ satisfying Assumptions 3.4 and 4.1 in the context of piecewise affine, continuous finite elements. Higher-order methods are not considered herein. This is motivated, in part, by the fact that in many cases the exact solution u may not have much additional regularity beyond that of the natural energy space V; see [15, p. 324] for a similar remark concerning the p-Laplacian.

The discrete problem amounts to a system of nonlinear equations, and, in practice, is solved using an iterative method involving some kind of linearization. Given an approximate solution, say $u_{L,h}$, at a given stage of the iterative process and on a given mesh, there are actually two sources of error, namely linearization and discretization. Balancing these two sources of error can be of paramount importance in practice, since it can avoid performing an excessive number of nonlinear solver iterations if the discretization error dominates. Therefore, the second objective of this work is to design a posteriori error estimates distinguishing linearization and discretization errors in the context of an adaptive procedure. This type of analysis has been started by Chaillou and Suri [11,12] for a certain class of nonlinear problems similar to the present one and in the context of iterative solution of linear algebraic systems in [25]; we also refer to [16] for adaptive numerical approximation of nonlinear problems in the wavelets context. Chaillou and Suri only considered a fixed stage of the linearization process, while we take here the analysis one step further in the context of an iterative loop. Furthermore, they only considered a specific form for the linearization, namely of fixed point-type, while we allow for a wider choice, including Newton–Raphson methods. We consider an adaptive loop in which at each step, a fixed mesh is considered and the nonlinear solver is iterated until the linearization error estimate is brought below the discretization error estimate; then, the mesh is adaptively refined and the loop is advanced. In this work, we will not tackle the delicate issue of proving the convergence of the above adaptive algorithm. We will also assume that at each iterate of the nonlinear solver, a well-posed problem is obtained. This property is by no means granted in general; it amounts, for the p-Laplacian, to assume, as mentioned before in [9], that the gradient norm of the approximate solution is positive everywhere in the domain. We mention that in our numerical experiments, all the discrete problems were indeed found to be well-posed.

This paper is organized as follows. Section 2 describes the setting for the nonlinear problem together with its discretization and linearization. Section 3 is devoted to the derivation of the guaranteed error upper bounds, while Section 4 is concerned with the efficiency of the estimators. Section 5 presents two possible approaches to reconstruct the flux $\mathbf{t}_h$ in the context of piecewise affine, continuous finite elements. Section 6 contains the numerical results. Finally, Appendix A collects various auxiliary lemmas.

Please cite this article in press as: L. El Alaoui et al., Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems, Comput. Methods Appl. Mech. Engrg. (2010), doi:10.1016/j.cma.2010.03.024

2. The setting

We describe here the considered nonlinear problem together with its discretization and linearization.

2.1. The continuous problem

Henceforth, for a real number $r \in (1,+\infty)$ and a subset $\omega \subset \Omega$, $\|\cdot\|_{r,\omega}$ denotes the canonical norm in $L^r(\omega)$ or $[L^r(\omega)]^d$ with appropriate Lebesgue measure; the subscript ω is omitted whenever $\omega = \Omega$ while the subscript r is omitted whenever $r = 2$. Moreover, the space V is equipped with the norm $\|\cdot\|_V := \|\nabla(\cdot)\|_p$. It is convenient to introduce the operator $A : V \to V'$ such that for all $u, v \in V$,

$$\langle Au, v\rangle_{V',V} := (\boldsymbol{\sigma}(\nabla u), \nabla v). \tag{2.1}$$

The function a in the definition (1.2) of the flux function $\boldsymbol{\sigma}$ is assumed to satisfy the following *growth condition*: There are positive constants R_1, c_1, and c_2 such that for all $x \geqslant R_1$,

$$c_1 x^{p-2} \leqslant a(x) \leqslant c_2 x^{p-2}. \tag{2.2}$$

Owing to the upper bound in (2.2) and using the Hölder inequality, it is readily seen that for all $u \in V$, $\boldsymbol{\sigma}(\nabla u) \in [L^q(\Omega)]^d$ so that Au is indeed in V'. The lower bound in (2.2) is classically used to infer an a priori estimate on the solution to (1.3).

Concerning the operator A, we make the following assumptions:

(i) *Monotonicity*. For all $v, w \in V$, there holds

$$\langle Av - Aw, v - w\rangle_{V',V} \geqslant 0. \tag{2.3}$$

Moreover, there exist a function $\gamma : \mathbb{R}_+ \to \mathbb{R}_+^*$ (taking positive values) and a strictly increasing function $\chi : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\chi(0) = 0$ and $\lim_{t\to+\infty}\chi(t) = +\infty$ such that for all $R \geqslant 0$ and for all $v, w \in B_V(0,R)$, the open ball in V centered at the origin and of radius R, there holds

$$\langle Av - Aw, v - w\rangle_{V',V} \geqslant \gamma(R)\chi(\|v - w\|_V)\|v - w\|_V. \tag{2.4}$$

(ii) *Hölder continuity*. There exist a function $\Gamma : \mathbb{R}_+ \to \mathbb{R}_+$ and a real number $\alpha \in (0,1]$ such that for all $R \geqslant 0$ and for all $v, w \in B_V(0,R)$,

$$\|Av - Aw\|_{V'} \leqslant \Gamma(R)\|v - w\|_V^{\alpha}. \tag{2.5}$$

Under these assumptions, the model problem (1.3) is well-posed; see, e.g., [15, p. 321].

Remark 2.1 (*Link with a minimization problem*). Introducing the energy density $\varphi(x) := \int_0^x ya(y)\mathrm{d}y$ for $x \in \mathbb{R}_+$ and the functional

$$J : V \ni v \mapsto J(v) := (\varphi(|\nabla v|), 1) - (f, v),$$

it is readily seen that J is differentiable in V with $J'(v) = Av - f$ so that J is convex owing to (2.3). Hence, solving (1.3) amounts to seeking the unique minimizer of the functional J over V.

Example 2.2 (*p-Laplacian*). In the case of the p-Laplacian, for $p \geqslant 2$, the operator A is strongly monotone over the whole space V with $\chi(t) = t^{p-1}$ (the function γ in (2.4) is then taken to be constant), while it is Lipschitz-continuous for bounded arguments, that is $\alpha = 1$ in (2.5) with $\Gamma(R) \sim (2R)^{p-2}$. For $p \leqslant 2$, the operator A is strongly monotone for bounded arguments with $\chi(t) = t$ and $\gamma(R) \sim (2R)^{p-2}$, while it is Hölder-continuous with parameter $\alpha = p - 1$ over the whole space V. Furthermore, we mention that the setting for the p-Laplacian fits the more general framework of the so-called φ-Laplacian; see, e.g., [18].

2.2. Discretization

Let $(\mathcal{T}_h)_h$ be a shape-regular family of affine meshes of Ω consisting of simplices. We assume that the meshes cover Ω exactly. We also suppose that each mesh $\mathcal{T}_h$ is matching in the sense that it contains no "hanging nodes". Let $\mathbb{P}_k(\mathcal{T}_h)$, $k \geqslant 0$, be spanned by piecewise polynomials of total degree $\leqslant k$ on the mesh $\mathcal{T}_h$ and let

$$V_h := \mathbb{P}_1(\mathcal{T}_h) \cap V,$$

be the usual first-order, continuous finite element space on the mesh $\mathcal{T}_h$. The discrete nonlinear problem takes the following form:

$$u_h \in V_h, \quad (\boldsymbol{\sigma}(\nabla u_h), \nabla v_h) = (f, v_h) \quad \forall v_h \in V_h. \tag{2.6}$$

The discrete nonlinear problem (2.6) is well-posed and its solution u_h satisfies an a priori estimate similar to that satisfied by the exact solution u. The properties of the operator A also imply that u_h converges to u in V.

In addition to the partition of Ω induced by the mesh $\mathcal{T}_h$, we will consider two other partitions. Firstly, let $\mathcal{D}_h$ be the dual mesh formed around the vertices of $\mathcal{T}_h$ using element and face barycenters; see the left part of Fig. 1 for an illustration in dimension $d = 2$. The set $\mathcal{D}_h$ is partitioned into $\mathcal{D}_h = \mathcal{D}_h^{\mathrm{int}} \cup \mathcal{D}_h^{\mathrm{ext}}$, where $\mathcal{D}_h^{\mathrm{int}}$ collects the dual volumes associated with interior vertices and $\mathcal{D}_h^{\mathrm{ext}}$ those associated with boundary vertices. For $D \in \mathcal{D}_h$, h_D denotes its diameter. Secondly, $\mathcal{S}_h$ is the simplicial mesh which is simultaneously a submesh (refinement) of both $\mathcal{T}_h$ and $\mathcal{D}_h$ formed by the meshes $\mathcal{S}_D$; see the right part of Fig. 1. The meshes $(\mathcal{S}_h)_h$ are assumed to be shape regular with parameter denoted by κ. The mesh $\mathcal{D}_h$ will be used in the error upper bound; the error estimators will also be localized on that mesh. The mesh $\mathcal{S}_h$ will be used in the construction of the equilibrated flux $\mathbf{t}_h$ and in the error lower bound.

The faces of the mesh $\mathcal{S}_h$ are collected into the set $\mathcal{G}_h$ which is partitioned into $\mathcal{G}_h = \mathcal{G}_h^{\mathrm{int}} \cup \mathcal{G}_h^{\mathrm{ext}}$, where $\mathcal{G}_h^{\mathrm{int}}$ collects the faces inside Ω and $\mathcal{G}_h^{\mathrm{ext}}$ those located on the boundary $\partial\Omega$. The set $\mathcal{G}_h^{\mathrm{int}}$ is further partitioned into $\mathcal{G}_h^{\mathrm{int}} = \mathcal{G}_h^{\mathcal{D}} \cup \mathcal{G}_h^{\mathcal{T}} \cup \mathcal{G}_h^{\mathcal{DT}}$, where

$$\mathcal{G}_h^{\mathcal{D}} = \{F \in \mathcal{G}_h^{\mathrm{int}}, F \subset \partial D \text{ for some } D \in \mathcal{D}_h\},$$
$$\mathcal{G}_h^{\mathcal{T}} = \{F \in \mathcal{G}_h^{\mathrm{int}}, F \subset \partial T \text{ for some } T \in \mathcal{T}_h\}.$$

The set $\mathcal{G}_h^{\mathcal{D}}$ collects the interior faces located on the boundary of the dual volumes (indicated by a dashed line in the right part of Fig. 1), $\mathcal{G}_h^{\mathcal{T}}$ those located on the boundary of the original elements (indicated by a solid line), and $\mathcal{G}_h^{\mathcal{DT}}$ the remaining ones (indicated by a dotted line).

2.3. Linearization

The continuous nonlinear problem (1.3) can be linearized at a given function $u_0 \in V$. To this purpose, a linear or affine flux function $\boldsymbol{\sigma}_{\mathrm{L}} : \mathbb{R}^d \to \mathbb{R}^d$ is introduced (the dependence of $\boldsymbol{\sigma}_{\mathrm{L}}$ on u_0 is left implicit to alleviate the notation), leading to the continuous linearized problem

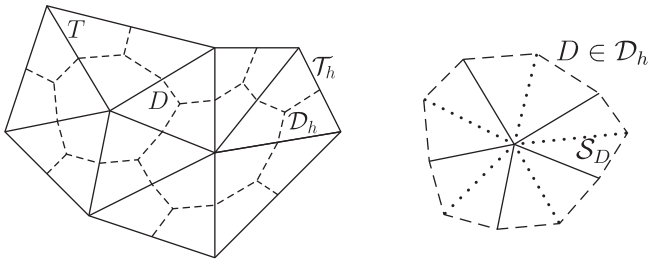

Fig. 1. Simplicial mesh $\mathcal{T}_h$ and the associated vertex-centered dual mesh $\mathcal{D}_h$ (left) and the fine simplicial mesh $\mathcal{S}_D$ of $D \in \mathcal{D}_h$ (right).

Please cite this article in press as: L. El Alaoui et al., Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems, Comput. Methods Appl. Mech. Engrg. (2010), doi:10.1016/j.cma.2010.03.024

$$u_L \in V, \quad (\boldsymbol{\sigma}_L(\nabla u_L), \nabla v) = (f, v) \quad \forall v \in V. \tag{2.7}$$

Similarly, for the discrete nonlinear problem (2.6), the discrete linearized problem is

$$u_{L,h} \in V_h, \quad (\boldsymbol{\sigma}_L(\nabla u_{L,h}), \nabla v_h) = (f, v_h) \quad \forall v_h \in V_h. \tag{2.8}$$

An important property is that this yields the following commuting diagram

$$\begin{array}{ccc} \boxed{\text{continuous nonlinear (1.3)}} & \xrightarrow{\text{discretization}} & \boxed{\text{discrete nonlinear (2.6)}} \\ \downarrow \text{linearization} & & \downarrow \text{linearization} \\ \boxed{\text{continuous linearized (2.7)}} & \xrightarrow{\text{discretization}} & \boxed{\text{discrete linearized (2.8)}} \end{array}.$$

It is convenient to introduce the linearized operator $A_L : V \to V'$ such that for all $u, v \in V$,

$$\langle A_L u, v\rangle_{V',V} := (\boldsymbol{\sigma}_L(\nabla u), \nabla v). \tag{2.9}$$

In practice, starting from an initial guess $u^0_{L,h} \in V_h$, a sequence of discrete solutions $\left\{u^i_{L,h}\right\}_{i\geqslant 1}$ is generated through the following iterative algorithm: For $i \geqslant 1$,

(1) Linearize the flux function at $u^{i-1}_{L,h}$;
(2) Solve the discrete linearized problem (2.8) for $u^i_{L,h}$;
(3) If the desired precision is reached, then stop, else set $i \leftarrow (i+1)$ and go to step (1).

One of the central points of the present paper is the choice of the *stopping criterion*. We propose it in Section 3.3 with the purpose to balance discretization and linearization errors.

There are two basic choices for linearizing the flux function $\boldsymbol{\sigma}_L$ at a given u_0. The first one is,

$$\forall \boldsymbol{\xi} \in \mathbb{R}^d, \quad \boldsymbol{\sigma}_L(\boldsymbol{\xi}) := a(|\nabla u_0|)\boldsymbol{\xi}. \tag{2.10}$$

This choice has been considered by Chaillou and Suri [11,12]. The second one, assuming that a is differentiable on $\mathbb{R}_+$ and that $xa'(x)$ is bounded by x^{p-2} for large x, is

$$\forall \boldsymbol{\xi} \in \mathbb{R}^d, \quad \boldsymbol{\sigma}_L(\boldsymbol{\xi}) := a(|\nabla u_0|)\boldsymbol{\xi} + a'(|\nabla u_0|)\frac{1}{|\nabla u_0|}(\nabla u_0 \otimes \nabla u_0)(\boldsymbol{\xi} - \nabla u_0). \tag{2.11}$$

In the context of the above iterative loop, the choice (2.10) corresponds to a fixed point iteration for solving the discrete nonlinear problem (2.6), while (2.11) corresponds to the full Newton–Raphson iteration. In the case (2.11), we slightly abuse the terminology since the operator A_L is actually affine.

Proving the well-posedness of the linearized problems (2.7) and (2.8) and the convergence of the above iterative loop goes beyond the scope of the present paper. Henceforth, we make the assumption that these properties indeed hold. Write generally $\boldsymbol{\sigma}_L(\boldsymbol{\xi}) = \mathbb{A}(\nabla u_0)\boldsymbol{\xi} + \mathbf{b}(\nabla u_0)$, where $\mathbb{A}(\nabla u_0)$ takes values in $\mathbb{R}^{d\times d}$ and $\mathbf{b}(\nabla u_0)$ takes values in $\mathbb{R}^d$. Then the linearized problems (2.7) and (2.8) are well-posed when $\mathbf{b}(\nabla u_0) \in [L^q(\Omega)]^d$ and when the linear operator $\widetilde{A} : V \to V'$, $\langle \widetilde{A}u, v\rangle_{V',V} := (\mathbb{A}(\nabla u_0)\nabla u, \nabla v)$, $u, v \in V$, satisfies the hypotheses of Section 2.1. Incidentally, we observe that if the iterative loop converges, then necessarily $u^i_{L,h} \to u_h$, the unique solution to (2.6), as $i \to +\infty$. We also remark that in the context of the p-Laplacian, we are thus led to assume that for all $i \geqslant 0$, the gradient norm of $u^i_{L,h}$ is positive everywhere in the domain. In our numerical experiments, all the discrete problems were indeed found to be well-posed, and the iterative loop did converge. Actually, the choice (2.11) associated with the Newton–Raphson iteration led to much faster convergence rates than the choice (2.10) associated with the fixed point iteration.

Remark 2.3 (*Anisotropic diffusion for Newton–Raphson linearization*). We observe that the choice (2.11) for the linearized flux function amounts to adding anisotropic (rank-one) diffusion to the isotropic diffusion $a(|\nabla u_0|)$ obtained using the linearization (2.10).

3. A posteriori error estimates

We state and prove here our a posteriori error estimates and give our stopping criterion for iterative linearizations.

Using the definition (2.1) of the nonlinear operator A, the goal of this section is to derive guaranteed upper bounds for the quantity

$$\mathcal{J}_u(u_{L,h}) := \|Au - Au_{L,h}\|_{V'} = \sup_{v\in V\setminus\{0\}} \frac{(\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}(\nabla u_{L,h}), \nabla v)}{\|\nabla v\|_p}. \tag{3.1}$$

Although the error measure $\mathcal{J}_u(u_{L,h})$ is not a norm for the difference $u - u_{L,h}$ because of the nonlinearity, we observe that under the assumption that $u_{L,h}$ is uniformly bounded in V, the monotonicity and Hölder continuity properties of the operator A readily imply that $\mathcal{J}_u(u_{L,h}) \to 0$ if and only if $\|u - u_{L,h}\|_V \to 0$. We also notice that $\mathcal{J}_u(u_{L,h})$ is not equivalent to the energy error $\|u - u_{L,h}\|_V$ in the sense that, in general, there do not exist positive $\boldsymbol{\sigma}$-independent constants c_1 and c_2 such that $c_1\|u - w\|_V \leqslant \mathcal{J}_u(w) \leqslant c_2\|u - w\|_V$ uniformly in $w \in V$.

3.1. Abstract a posteriori error estimate distinguishing the discretization and linearization errors

Following Chaillou and Suri [12], we now distinguish the discretization and linearization errors:

Proposition 3.1 (Abstract a posteriori error estimate distinguishing the discretization and linearization errors). *Let $u \in V$ be the solution of (1.3) and let $u_{L,h} \in V$ be arbitrary. Then,*

$$\mathcal{J}_u(u_{L,h}) \leqslant \|f - A_L u_{L,h}\|_{V'} + \|A_L u_{L,h} - Au_{L,h}\|_{V'}. \tag{3.2}$$

Proof. Since $Au = f$ in V', we infer that

$$\|Au - Au_{L,h}\|_{V'} = \|f - Au_{L,h}\|_{V'} = \|f - A_L u_{L,h} + A_L u_{L,h} - Au_{L,h}\|_{V'},$$

and we conclude using the triangle inequality. □

Remark 3.2 (*Discretization and linearization errors*). Following Chaillou and Suri [12], we call the first term in the right-hand side of (3.2) the *discretization error* and the second one the *linearization error*. Since $A_L u_L = f$ in V', where u_L is the solution to the continuous linearized problem (2.7), the discretization error can also be written as $\|A_L u_L - A_L u_{L,h}\|_{V'}$.

Remark 3.3 (*Arbitrary $u_{L,h}$*). In the statement of Proposition 3.1, the discrete function $u_{L,h}$ need not be the solution of the discrete linearized problem (2.8), but can instead be arbitrary in V. The same remark holds for the statement of Theorem 3.5.

3.2. Guaranteed and fully computable a posteriori error estimate

To define our a posteriori error estimate, we suppose the following:

Assumption 3.4 (*Local conservation*). There exists a vector field $\mathbf{t}_h \in \mathbf{W}^q(\mathrm{div}, \Omega) := \{\mathbf{v} \in \mathbf{L}^q(\Omega)\,;\, \nabla\cdot\mathbf{v} \in \mathbf{L}^q(\Omega)\}$ such that

$$(\nabla\cdot\mathbf{t}_h, 1)_D = (f, 1)_D \quad \forall D \in \mathcal{D}_h^{\mathrm{int}}.$$

Please cite this article in press as: L. El Alaoui et al., Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems, Comput. Methods Appl. Mech. Engrg. (2010), doi:10.1016/j.cma.2010.03.024

We will verify Assumption 3.4 in Section 5.

For each $D \in \mathcal{D}_h$, let the *residual estimator* $\eta_{R,D}$, the *diffusive flux estimator* $\eta_{DF,D}$, and the *linearization estimator* $\eta_{L,D}$ be defined as

$$\eta_{R,D} := m_D \|f - \nabla \cdot \mathbf{t}_h\|_{q,D}, \tag{3.3a}$$

$$\eta_{DF,D} := \|\boldsymbol{\sigma}_L(\nabla u_{L,h}) + \mathbf{t}_h\|_{q,D}, \tag{3.3b}$$

$$\eta_{L,D} := \|\boldsymbol{\sigma}(\nabla u_{L,h}) - \boldsymbol{\sigma}_L(\nabla u_{L,h})\|_{q,D}. \tag{3.3c}$$

Here, $m_D = C_{P,p,D} h_D$ if $D \in \mathcal{D}_h^{int}$ and $m_D = C_{F,p,D,\partial\Omega} h_D$ if $D \in \mathcal{D}_h^{ext}$ and $C_{P,p,D}$ is the constant from the generalized Poincaré inequality

$$\|\varphi - \varphi_D\|_{p,D} \leqslant C_{P,p,D} h_D \|\nabla \varphi\|_{p,D} \quad \forall \varphi \in W^{1,p}(D), \tag{3.4}$$

$D \in \mathcal{D}_h^{int}$ (here $\varphi_D = (\varphi, 1)_D / |D|$) and $C_{F,p,D,\partial\Omega}$ is the constant from the generalized Friedrichs inequality

$$\|\varphi\|_{p,D} \leqslant C_{F,p,D,\partial\Omega} h_D \|\nabla \varphi\|_{p,D} \quad \forall \varphi \in W^{1,p}(D) \text{ such that } \varphi = 0 \text{ on } \partial\Omega \cap \partial D, \tag{3.5}$$

$D \in \mathcal{D}_h^{ext}$. If $p = 2$, $C_{P,p,D} = 1/\pi$ if D is convex and $C_{F,p,D,\partial\Omega} = 1$ in general, cf. [46] and the references therein. For $p \geqslant 2$, $C_{P,p,D} = \pi^{-\frac{2}{p}} d^{1-\frac{1}{p}}$ if D is convex, see [36], and for all $p \in (1, +\infty)$, $C_{P,p,D} = p^{\frac{1}{p}} 2^{\frac{(p-1)}{p}}$ if D is convex, see [14]. Note, however, that using the construction of $\mathbf{t}_h$ by (5.4a) and (5.4b) from Section 5, the values of these constants are actually not needed whenever f is piecewise constant. We can now state the main result of this section:

Theorem 3.5 (A posteriori error estimate). *Let $u \in V$ be the solution of (1.3) and let $u_{L,h} \in V$ be arbitrary. Let the error estimators $\eta_{R,D}$, $\eta_{DF,D}$, and $\eta_{L,D}$ be given by (3.3a)–(3.3c). Then, under Assumption 3.4, there holds*

$$\mathcal{J}_u(u_{L,h}) \leqslant \eta := \left\{ \sum_{D \in \mathcal{D}_h} (\eta_{R,D} + \eta_{DF,D})^q \right\}^{\frac{1}{q}} + \left\{ \sum_{D \in \mathcal{D}_h} \eta_{L,D}^q \right\}^{\frac{1}{q}}.$$

Proof. We estimate the two terms in the right-hand side of (3.2).

(i) Estimate on the discretization error. Observe that for all $v \in V$ with $\|v\|_V = 1$, there holds

$$\langle f - A_L u_{L,h}, v \rangle_{V',V} = \langle f, v \rangle_{V',V} - (\nabla \cdot \mathbf{t}_h, v) + (\nabla \cdot \mathbf{t}_h, v) - \langle A_L u_{L,h}, v \rangle_{V',V}.$$

We first bound

$$\begin{aligned}
\langle f, v \rangle_{V',V} - (\nabla \cdot \mathbf{t}_h, v) &= (f - \nabla \cdot \mathbf{t}_h, v) \\
&= \sum_{D \in \mathcal{D}_h^{int}} (f - \nabla \cdot \mathbf{t}_h, v)_D \\
&\quad + \sum_{D \in \mathcal{D}_h^{ext}} (f - \nabla \cdot \mathbf{t}_h, v)_D \\
&= \sum_{D \in \mathcal{D}_h^{int}} (f - \nabla \cdot \mathbf{t}_h, v - v_D)_D \\
&\quad + \sum_{D \in \mathcal{D}_h^{ext}} (f - \nabla \cdot \mathbf{t}_h, v)_D \\
&\leqslant \sum_{D \in \mathcal{D}_h^{int}} \|f - \nabla \cdot \mathbf{t}_h\|_{q,D} \|v - v_D\|_{p,D} \\
&\quad + \sum_{D \in \mathcal{D}_h^{ext}} \|f - \nabla \cdot \mathbf{t}_h\|_{q,D} \|v\|_{p,D} \\
&\leqslant \sum_{D \in \mathcal{D}_h} m_D \|f - \nabla \cdot \mathbf{t}_h\|_{q,D} \|\nabla v\|_{p,D} \\
&= \sum_{D \in \mathcal{D}_h} \eta_{R,D} \|\nabla v\|_{p,D},
\end{aligned}$$

where we have used the Hölder inequality together with (3.4) and (3.5). Furthermore, using the Green theorem, definition (2.9) of the linearized operator, and the Hölder inequality yields

$$\begin{aligned}
(\nabla \cdot \mathbf{t}_h, v) - \langle A_L u_{L,h}, v \rangle_{V',V} &= -(\mathbf{t}_h + \boldsymbol{\sigma}_L(\nabla u_{L,h}), \nabla v) \\
&\leqslant \sum_{D \in \mathcal{D}_h} \eta_{DF,D} \|\nabla v\|_{p,D}.
\end{aligned}$$

Collecting the two above bounds leads to

$$\langle f - A_L u_{L,h}, v \rangle_{V',V} \leqslant \sum_{D \in \mathcal{D}_h} (\eta_{R,D} + \eta_{DF,D}) \|\nabla v\|_{p,D},$$

whence the Hölder inequality yields

$$\begin{aligned}
\|f - A_L u_{L,h}\|_{V'} &= \sup_{v \in V,\ \|v\|_V = 1} \langle f - A_L u_{L,h}, v \rangle_{V',V} \\
&\leqslant \left\{ \sum_{D \in \mathcal{D}_h} (\eta_{R,D} + \eta_{DF,D})^q \right\}^{\frac{1}{q}}.
\end{aligned}$$

(ii) Estimate on the linearization error. For all $v \in V$ with $\|v\|_V = 1$, using definitions (2.1) and (2.9) along with the Hölder inequality yields

$$\begin{aligned}
\langle A_L u_{L,h} - A u_{L,h}, v \rangle_{V',V} &= (\boldsymbol{\sigma}_L(\nabla u_{L,h}) - \boldsymbol{\sigma}(\nabla u_{L,h}), \nabla v) \\
&\leqslant \left\{ \sum_{D \in \mathcal{D}_h} \eta_{L,D}^q \right\}^{\frac{1}{q}},
\end{aligned}$$

which completes the proof. □

3.3. Balancing discretization and linearization errors

We are now in a position to specify the stopping criterion for the iterative loop outlined in Section 2.3.

3.3.1. Global stopping criterion

Choose a positive parameter γ and stop the iterative loop whenever

$$\eta_L := \left\{ \sum_{D \in \mathcal{D}_h} \eta_{L,D}^q \right\}^{\frac{1}{q}} \leqslant \gamma \left\{ \sum_{D \in \mathcal{D}_h} (\eta_{R,D} + \eta_{DF,D})^q \right\}^{\frac{1}{q}} =: \gamma \eta_D. \tag{3.6}$$

This criterion equilibrates the *global linearization estimator* η_L and the *global discretization estimator* η_D, up to the constant γ. Practically, γ is of order 1; in the numerical experiments of Section 6, we have used $\gamma = 0.1$. This criterion is global since it is based on quantities defined for all mesh elements. As we will see in Theorem 4.8 (where a theoretical requirement on γ appears), (3.6) is sufficient for proving global efficiency.

3.3.2. Local stopping criterion

Choose a family of positive parameters $\{\gamma_D\}_{D \in \mathcal{D}_h}$ and stop the iterative loop whenever

$$\eta_{L,D} \leqslant \gamma_D (\eta_{R,D} + \eta_{DF,D}) \quad \forall D \in \mathcal{D}_h. \tag{3.7}$$

This criterion equilibrates the local linearization estimators $\eta_{L,D}$ and the local discretization estimators $\eta_{R,D} + \eta_{DF,D}$, up to the constants γ_D, where in practice, γ_D can be chosen as γ above. The criterion (3.7) is local since it is based on quantities defined for each mesh element separately. As we will see in Theorem 4.4 (where a theoretical requirement on γ_D appears), (3.7) implies local efficiency and hence suitability of our a posteriori error estimates to adaptive mesh refinement.

Please cite this article in press as: L. El Alaoui et al., Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems, Comput. Methods Appl. Mech. Engrg. (2010), doi:10.1016/j.cma.2010.03.024

4. Efficiency of the estimators

We examine in this section the local and global efficiencies of our estimates.

4.1. Preliminaries

Henceforth, we set for convenience

$\boldsymbol{\sigma}_{\mathrm{L},h} := \boldsymbol{\sigma}_{\mathrm{L}}(\nabla u_{\mathrm{L},h})$.

Taking into account the definitions (2.10) or (2.11) of the linearized flux function $\boldsymbol{\sigma}_{\mathrm{L}}$ and recalling that both $u_{\mathrm{L},h}$ and the function u_0 at which the linearization is performed are piecewise affine, it is inferred that $\boldsymbol{\sigma}_{\mathrm{L},h} \in [\mathbb{P}_0(\mathcal{T}_h)]^d$. The results presented in this section are valid more generally under the assumption $\boldsymbol{\sigma}_{\mathrm{L},h} \in [\mathbb{P}_k(\mathcal{T}_h)]^d$ for some fixed polynomial degree k. We also assume in this section that $f \in \mathbb{P}_k(\mathcal{T}_h)$.

In the sequel, $A \lesssim B$ stands for the inequality $A \leqslant CB$ with a generic constant C independent of the mesh size h, the nonlinear and linearized functions $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}_{\mathrm{L}}$, and the Lebesgue exponent p, but that can depend on the shape regularity parameter κ of the mesh family $(\mathcal{S}_h)_h$ and on the polynomial degree k.

Assumption 4.1 (*Approximation property of the reconstructed flux*). The reconstructed flux $\mathbf{t}_h$ is in $[\mathbb{P}_k(\mathcal{T}_h)]^d$ and there holds for all $D \in \mathcal{D}_h$,

$$\eta_{\mathrm{DF},D} \lesssim \eta_{\mathrm{res},D} := \left\{ \sum_{T\in\mathcal{S}_D} h_T^q \|f + \nabla\cdot\boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,T}^q + \sum_{F\in\mathcal{G}_D^{\mathcal{T}}} h_F \|[\![\boldsymbol{\sigma}_{\mathrm{L},h}\cdot\mathbf{n}]\!]\|_{q,F}^q \right\}^{\frac{1}{q}}, \tag{4.1}$$

where $\mathcal{G}_D^{\mathcal{T}} \subset \mathcal{G}_h^{\mathcal{T}}$ collects the faces of $\mathcal{G}_h^{\mathcal{T}}$ included in D; moreover, $[\![\cdot]\!]$ denotes the jump across a face.

We will verify Assumption 4.1 in Section 5. Under this assumption, we now prove a simple result, relating our estimates to classical residual ones; see, e.g., [35].

Lemma 4.2 (Upper bound by residual estimators). *Under Assumption 4.1, there holds*

$$\eta_{\mathrm{R},D} + \eta_{\mathrm{DF},D} \lesssim \eta_{\mathrm{res},D} \quad \forall D \in \mathcal{D}_h. \tag{4.2}$$

Proof. Taking into account (4.1), it remains to show the bound on $\eta_{\mathrm{R},D}$. Since $h_D \lesssim h_T$ for all $T \in \mathcal{S}_D$ by the construction of $\mathcal{D}_h$ and $\mathcal{S}_h$ and the shape regularity of $\mathcal{S}_h$, there holds

$$\eta_{\mathrm{R},D} \lesssim h_D \left\{ \sum_{T\in\mathcal{S}_D} \|f - \nabla\cdot\mathbf{t}_h\|_{q,T}^q \right\}^{\frac{1}{q}} \lesssim \left\{ \sum_{T\in\mathcal{S}_D} h_T^q \|f - \nabla\cdot\mathbf{t}_h\|_{q,T}^q \right\}^{\frac{1}{q}}.$$

As a result, using the triangle inequality and the inverse inequality (A.1) with $r = q$ leads to

$$\begin{aligned}\eta_{\mathrm{R},D} &\lesssim \left\{ \sum_{T\in\mathcal{S}_D} h_T^q \|f + \nabla\cdot\boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,T}^q \right\}^{\frac{1}{q}} + \left\{ \sum_{T\in\mathcal{S}_D} h_T^q \|\nabla\cdot(\boldsymbol{\sigma}_{\mathrm{L},h} + \mathbf{t}_h)\|_{q,T}^q \right\}^{\frac{1}{q}} \\ &\lesssim \left\{ \sum_{T\in\mathcal{S}_D} h_T^q \|f + \nabla\cdot\boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,T}^q \right\}^{\frac{1}{q}} + \eta_{\mathrm{DF},D},\end{aligned}$$

whence the result readily follows. □

4.2. Local efficiency

We address here the local efficiency of the estimators of Theorem 3.5. We first give a result relying on the techniques presented in [39]. The proof is postponed to Section A.2.

Lemma 4.3 (Local efficiency of residual estimators). *For all $D \in \mathcal{D}_h$, there holds*

$$\eta_{\mathrm{res},D} \lesssim \|\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}(\nabla u_{\mathrm{L},h})\|_{q,D} + \eta_{\mathrm{L},D}. \tag{4.3}$$

We are now ready to announce and prove the main result of this section.

Theorem 4.4 (Local efficiency). *Let (3.7), with γ_D small enough, and Assumption 4.1 hold true. Then, for all $D \in \mathcal{D}_h$,*

$$\eta_{\mathrm{L},D} + \eta_{\mathrm{R},D} + \eta_{\mathrm{DF},D} \lesssim \|\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}(\nabla u_{\mathrm{L},h})\|_{q,D}.$$

Proof. Using (4.2), (4.3), and (3.7), we infer

$$\begin{aligned}\eta_{\mathrm{R},D} + \eta_{\mathrm{DF},D} &\leqslant C\eta_{\mathrm{res},D} \leqslant \widetilde{C}(\|\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}(\nabla u_{\mathrm{L},h})\|_{q,D} + \eta_{\mathrm{L},D}) \\ &\leqslant \widetilde{C}\|\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}(\nabla u_{\mathrm{L},h})\|_{q,D} + \widetilde{C}\gamma_D(\eta_{\mathrm{R},D} + \eta_{\mathrm{DF},D}).\end{aligned}$$

Thus, choosing $\gamma_D = 1/(2\widetilde{C})$,

$$\frac{1}{2}(\eta_{\mathrm{R},D} + \eta_{\mathrm{DF},D}) \leqslant \widetilde{C}\|\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}(\nabla u_{\mathrm{L},h})\|_{q,D}.$$

Consequently, using once again (3.7),

$$\begin{aligned}\eta_{\mathrm{L},D} + \eta_{\mathrm{R},D} + \eta_{\mathrm{DF},D} &\leqslant (1+\gamma_D)(\eta_{\mathrm{R},D} + \eta_{\mathrm{DF},D}) \\ &\leqslant (2\widetilde{C}+1)\|\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}(\nabla u_{\mathrm{L},h})\|_{q,D}.\end{aligned}$$ □

Remark 4.5 (*Local efficiency*). Whereas the estimates are derived for the error measure $\mathcal{J}_u(u_{\mathrm{L},h})$, the local efficiency of Theorem 4.4 uses the L^q-norm of the difference of fluxes. This is not fully satisfactory, but it seems to be the price to obtain local efficiency and not only global, as presented in the next section.

Remark 4.6 (*Local efficiency on the given element only*). Owing to the fact that only (a subset of) the interior faces of $\mathcal{S}_D$ appear(s) in (4.1), the local efficiency result of Theorem 4.4 is stated on the given dual volume D only (no neighbors appear). This is in correspondence with the fact that the stopping criterion (3.7) is local on each $D \in \mathcal{D}_h$.

4.3. Global efficiency

We address here the global efficiency of the estimators of Theorem 3.5, relying on the techniques presented in [39,40]. We begin with the following lemma, whose proof is postponed to Section A.3.

Lemma 4.7 (Global efficiency of residual estimators). *There holds*

$$\begin{aligned}\eta_{\mathrm{res}} &:= \left\{ \sum_{T\in\mathcal{S}_h} h_T^q \|f + \nabla\cdot\boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,T}^q + \sum_{F\in\mathcal{G}_h^{\mathcal{T}}} h_F \|[\![\boldsymbol{\sigma}_{\mathrm{L},h}\cdot\mathbf{n}]\!]\|_{q,F}^q \right\}^{\frac{1}{q}} \\ &\lesssim \|Au - Au_{\mathrm{L},h}\|_{V'} + \eta_{\mathrm{L}}.\end{aligned}$$

We are now ready to announce the main result of this section. The proof is skipped since it follows by the same arguments as in Theorem 4.4. (A similar bound on γ as that on γ_D in Theorem 4.4 needs to be used.)

Theorem 4.8 (Global efficiency). *Let (3.6), with γ small enough, and Assumption 4.1 hold true. Then,*

$$\eta \lesssim \mathcal{J}_u(u_{\mathrm{L},h}).$$

Please cite this article in press as: L. El Alaoui et al., Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems, Comput. Methods Appl. Mech. Engrg. (2010), doi:10.1016/j.cma.2010.03.024

Remark 4.9 (*Robustness*). Theorem 4.8 means that the a posteriori error estimates of Theorem 3.5 are fully robust with respect to the "size" of the nonlinearity. Indeed, the lower bound of Theorem 4.8 is independent of the nonlinear function a, as well as of the Lebesgue exponent p.

5. Flux reconstruction

In this section, we give two examples on how to reconstruct a flux $\mathbf{t}_h$ satisfying Assumptions 3.4 and 4.1 in the context of the finite element method (2.8). In both cases, the reconstructed flux $\mathbf{t}_h$ belongs to the lowest-order Raviart–Thomas–Nédélec finite element space $\mathbf{RTN}(\mathcal{S}_h)$ associated with the simplicial submesh $\mathcal{S}_h$.

Recall the notation $\boldsymbol{\sigma}_{\mathrm{L},h} := \boldsymbol{\sigma}_{\mathrm{L}}(\nabla u_{\mathrm{L},h})$ so that (2.8) becomes $(\boldsymbol{\sigma}_{\mathrm{L},h}, \nabla v_h) = (f, v_h)$ for all $v_h \in V_h$, and observe that by assumption $\boldsymbol{\sigma}_{\mathrm{L},h}$ is piecewise constant on $\mathcal{T}_h$. We first assume that f is piecewise constant on $\mathcal{T}_h$. The general case is postponed to Remark 5.5. Under these assumptions, following [4, Lemma 3] (cf. also [46]), we infer that (2.8) is equivalent to finding $u_{\mathrm{L},h} \in V_h$ such that

$$-(\boldsymbol{\sigma}_{\mathrm{L},h} \cdot \mathbf{n}_D, 1)_{\partial D} = (f, 1)_D \quad \forall D \in \mathcal{D}_h^{\mathrm{int}}, \tag{5.1}$$

where $\mathbf{n}_D$ denotes the outward unit normal vector to D. The scheme defined by (5.1) is also nothing but the vertex-centered finite volume method for approximating (2.7). In the vertex-centered finite volume framework, f does not need to be piecewise constant on $\mathcal{T}_h$.

The degrees of freedom of the reconstructed flux $\mathbf{t}_h \in \mathbf{RTN}(\mathcal{S}_h)$ are its normal components on all the faces $F \in \mathcal{G}_h$. In the sequel, for $F \in \mathcal{G}_h$, $\mathbf{n}_F$ denotes a unit normal vector to F (its orientation is fixed but irrelevant). We first set

$$\mathbf{t}_h \cdot \mathbf{n}_F := -\boldsymbol{\sigma}_{\mathrm{L},h} \cdot \mathbf{n}_F \quad \forall F \in \mathcal{G}_h^{\mathcal{D}}. \tag{5.2}$$

An immediate and important consequence of (5.1) and (5.2) is the following lemma.

Lemma 5.1 (Local conservation). *For $\mathbf{t}_h$ satisfying (5.2), Assumption 3.4 holds true.*

Following Vohralík [46], there exist several ways of defining $\mathbf{t}_h \cdot \mathbf{n}_F$ on the remaining faces of $\mathcal{G}_h$ (that is, those located inside some $D \in \mathcal{D}_h$ and those located on the boundary $\partial\Omega$, i.e., $F \in (\mathcal{G}_h^{\mathrm{int}} \setminus \mathcal{G}_h^{\mathcal{D}}) \cup \mathcal{G}_h^{\mathrm{ext}}$). We present here two methods.

- Direct prescription: Firstly, we can directly prescribe

$$\mathbf{t}_h \cdot \mathbf{n}_F := -\{\{\boldsymbol{\sigma}_{\mathrm{L},h} \cdot \mathbf{n}_F\}\} \tag{5.3}$$

  on all the remaining faces. Here, $\{\{\cdot\}\}$ denotes the mean value on interior faces and the actual value on boundary faces.
- Prescription by local linear system solves: Secondly, following [5,21,46], we can solve local Neumann problems: for a given $D \in \mathcal{D}_h$, let

$$\mathbf{RTN}_{\mathrm{N}}(\mathcal{S}_D) = \{\mathbf{v}_h \in \mathbf{RTN}(\mathcal{S}_D); \mathbf{v}_h \cdot \mathbf{n}_F = -\boldsymbol{\sigma}_{\mathrm{L},h} \cdot \mathbf{n}_F \ \forall F \in \mathcal{G}_h^{\mathcal{D}}, F \subset \partial D\}.$$

  Define $\mathbf{RTN}_{\mathrm{N},0}(\mathcal{S}_D)$ as $\mathbf{RTN}_{\mathrm{N}}(\mathcal{S}_D)$ but with the normal flux condition $\mathbf{v}_h \cdot \mathbf{n}_F = 0$ for all $F \in \mathcal{G}_h^{\mathcal{D}}$, $F \subset \partial D$. Let $\mathbb{P}_0^*(\mathcal{S}_D)$ be spanned by piecewise constants on $\mathcal{S}_D$ with zero mean on D when $D \in \mathcal{D}_h^{\mathrm{int}}$; when $D \in \mathcal{D}_h^{\mathrm{ext}}$, the mean value condition is not imposed. The local problem consists in finding $\mathbf{t}_h \in \mathbf{RTN}_{\mathrm{N}}(\mathcal{S}_D)$ and $q_h \in \mathbb{P}_0^*(\mathcal{S}_D)$, the mixed finite element approximations (cf. [8]) of local Neumann problems on $D \in \mathcal{D}_h^{\mathrm{int}}$ and local Neumann/Dirichlet problems on $D \in \mathcal{D}_h^{\mathrm{ext}}$:

$$(\mathbf{t}_h + \boldsymbol{\sigma}_{\mathrm{L},h}, \mathbf{v}_h)_D - (q_h, \nabla \cdot \mathbf{v}_h)_D = 0 \quad \forall \mathbf{v}_h \in \mathbf{RTN}_{\mathrm{N},0}(\mathcal{S}_D), \tag{5.4a}$$

$$(\nabla \cdot \mathbf{t}_h, \phi_h)_D = (f, \phi_h)_D \quad \forall \phi_h \in \mathbb{P}_0^*(\mathcal{S}_D). \tag{5.4b}$$

Note in particular that the function $-\boldsymbol{\sigma}_{\mathrm{L},h} \cdot \mathbf{n}_F$, used to impose the Neumann boundary condition on ∂D for each $D \in \mathcal{D}_h^{\mathrm{int}}$, satisfies by (5.1) the compatibility condition with the datum f, whence the existence and uniqueness of the solution to (5.4a) and (5.4b).

Remark 5.2 (*Comparison of the two approaches*). A solution of a local linear system on each $D \in \mathcal{D}_h$ corresponding to (5.4a) and (5.4b) is necessary in the second case, but the advantage over the first case is twofold. Firstly, the effectivity indices are close to the optimal value of one, as observed in our numerical experiments. Secondly, owing to the assumption that f is piecewise constant on $\mathcal{T}_h$ and to (5.4b), $\|f - \nabla \cdot \mathbf{t}_h\|_{q,T} = 0$ for all $T \in \mathcal{S}_h$, so that $\eta_{\mathrm{R},D} = 0$ for all $D \in \mathcal{D}_h$ and one does not need to evaluate the constants $C_{\mathrm{P},p,D}$ and $C_{\mathrm{F},p,D,\partial\Omega}$ from (3.4) and (3.5), respectively. For more details, we refer to [46].

We now consider Assumption 4.1. The proof is given in Section A.4.

Lemma 5.3 (Approximation property). *For $\mathbf{t}_h$ prescribed by (5.2) and either directly through (5.3) or by solving the local problems (5.4a) and (5.4b), Assumption 4.1 holds true.*

Remark 5.4 (*Other flux equilibration procedures*). The procedure (5.4a) and (5.4b) is closely related to that of Luce and Wohlmuth [31]. Many other approaches rely on equilibration on the given mesh $\mathcal{T}_h$. We cite in particular [1,7,17,27]. All these procedures can likewise be used to produce a discrete $\mathbf{W}^q(\mathrm{div}, \Omega)$-conforming vector field satisfying the local conservation property (Assumption 3.4). It is unknown to the authors whether these vector fields also satisfy the approximation property (Assumption 4.1). If it is the case, then they can be used in the present framework. Following Remark 4.6, a slight modification of the stopping criterion (3.7) would be necessary for the approaches relying on the given mesh $\mathcal{T}_h$.

Remark 5.5 (*General f*). When f is not piecewise constant on $\mathcal{T}_h$, we can proceed as follows: denote $f_h \in \mathbb{P}_0(\mathcal{T}_h)$ the function given by elementwise mean values of f. Consider $\bar{u}$, the solution to (1.3) with f replaced by f_h, and $u_{\mathrm{L},h}$, the approximation of (2.8) with f likewise replaced by f_h. Then, by the triangle inequality,

$$\mathcal{J}_u(u_{\mathrm{L},h}) = \|Au - Au_{\mathrm{L},h}\|_{V'} \leqslant \|Au - A\bar{u}\|_{V'} + \|A\bar{u} - Au_{\mathrm{L},h}\|_{V'}.$$

The second term fits the framework of Sections 4 and 5 and is therefore bounded by the estimator η (with f replaced by f_h). The first term can be bounded as in the proof of Theorem 3.5 by the data oscillation term

$$\eta_{\mathrm{Osc}} := \left\{ \sum_{T \in \mathcal{T}_h} (\|f - f_h\|_{q,T} C_{\mathrm{P},p,T} h_T)^q \right\}^{\frac{1}{q}}.$$

Note that this term is localized on the mesh $\mathcal{T}_h$ and not on $\mathcal{D}_h$.

6. Numerical results

We present in this section numerical experiments for the p-Laplacian. We first recall that the error measure $\mathcal{J}_u(u_{\mathrm{L},h}) = \|Au - Au_{\mathrm{L},h}\|_{V'}$ involves a dual norm that cannot be evaluated explicitly even if u is known. Following Chaillou and Suri [12], however, we deduce from (3.1) the following computable upper and lower bounds for $\mathcal{J}_u(u_{\mathrm{L},h})$:

$$\mathcal{J}_u(u_{\mathrm{L},h}) \leqslant \mathcal{J}_u^{\mathrm{up}}(u_{\mathrm{L},h}) := \|\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}(\nabla u_{\mathrm{L},h})\|_q, \tag{6.1a}$$

$$\mathcal{J}_u(u_{\mathrm{L},h}) \geqslant \mathcal{J}_u^{\mathrm{low}}(u_{\mathrm{L},h}) := \frac{|(\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}(\nabla u_{\mathrm{L},h}), \nabla(u - u_{\mathrm{L},h}))|}{\|\nabla(u - u_{\mathrm{L},h})\|_p}. \tag{6.1b}$$

Please cite this article in press as: L. El Alaoui et al., Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems, Comput. Methods Appl. Mech. Engrg. (2010), doi:10.1016/j.cma.2010.03.024

As we will see below, our estimate η defined in Theorem 3.5 turns out to be very close to $\mathcal{J}_u^{\rm up}(u_{{\rm L},h})$, whence our estimates are indeed very close to the error in the fluxes, which are often the physically most interesting quantity. We will use below the corresponding upper and lower effectivity indices, defined respectively as

$$\mathcal{I}^{\rm up} := \frac{\eta}{\mathcal{J}_u^{\rm up}(u_{{\rm L},h})} \quad \text{and} \quad \mathcal{I}^{\rm low} := \frac{\eta}{\mathcal{J}_u^{\rm low}(u_{{\rm L},h})}.$$

Consequently, the effectivity index for the error $\mathcal{J}_u(u_{{\rm L},h})$, defined as

$$\mathcal{I} := \frac{\eta}{\mathcal{J}_u(u_{{\rm L},h})},$$

lies between $\mathcal{I}^{\rm up}$ and $\mathcal{I}^{\rm low}$. We observe that $\mathcal{I}^{\rm up}$ can become smaller than one, which does not contradict that our estimates give a guaranteed upper bound; only $\mathcal{I}$ has to be greater than or equal to one.

We use continuous, piecewise affine finite elements and the Newton–Raphson method, which corresponds to solving (2.8) with $\boldsymbol{\sigma}_{\rm L}(\nabla u_{{\rm L},h})$ defined by (2.11). In order to construct the equilibrated flux $\mathbf{t}_h$, we use the solution of local Neumann problems by the mixed finite element method (5.4a) and (5.4b). We use the local stopping criterion (3.7) with $\gamma_D = 0.1$ for all $D \in \mathcal{D}_h$. On an initial mesh, we start the nonlinear iterative solver from the interpolation of the exact solution. On refined meshes, we interpolate the approximate solution from the next coarser mesh. We consider the following overall strategy:

(1) On the given mesh, perform the iterative linearization of Section 2.3 until the local stopping criterion (3.7) is satisfied;
(2) If η is less than the desired precision, then stop, else refine the mesh adaptively, on the basis of η, and go to step (1).

In practice, all the elements where the estimate exceeds 50% of the maximal error are marked for refinement. Every marked triangular element is refined regularly into four sub-elements and then the so-called longest edge refinement is used so as to recover a matching mesh (without hanging nodes). The error stemming from inhomogeneous boundary conditions is not taken into account.

Table 1
Flux and potential regularities and experimental orders of convergence, case 1.

p	Flux				Potential	
	Regularity	Error up	Error low	Estimate	Regularity	Error energy
1.4	$\mathbf{W}^{1.57,q}$	1.01	1.01	1.00	$W^{4.93,p}$	1.01
3	$\mathbf{W}^{2.33,q}$	1.01	1.01	1.01	$W^{2.17,p}$	1.00
10	$\mathbf{W}^{2.80,q}$	1.00	1.68	1.00	$W^{1.31,p}$	0.31

All the linear systems within the nonlinear iterative procedure are solved by a direct solver. The present approach can also be combined with a linear iterative solver, and to further save computational effort, the latter can be stopped whenever the algebraic error does not contribute significantly to the overall error, following Jiránek et al. [25].

6.1. A first test case

In this first test case, $\Omega := (0,1)\times(0,1)$, $f = 2$, and the Dirichlet boundary condition is given by the exact solution

$$u(x,y) = -\frac{p-1}{p}|(x,y)-(0.5,0.5)|^{\frac{p}{p-1}} + \frac{p-1}{p}\left(\frac{1}{2}\right)^{\frac{p}{p-1}}.$$

This is a two-dimensional extension of a test case from Chaillou and Suri [11]. The Sobolev regularity for the potential u and the flux $-\boldsymbol{\sigma}(\nabla u)$ is reported in Table 1 for the different values of p considered in the experiments. More generally, $u \in W^{s_u,p}(\Omega)$ and $\boldsymbol{\sigma}(\nabla u) \in \mathbf{W}^{s_{\boldsymbol{\sigma}},q}(\Omega) := [W^{s_{\boldsymbol{\sigma}},q}(\Omega)]^d$ with $s_u < p/(p-1) + 2/p$ and $s_{\boldsymbol{\sigma}} < 3 - 2/p$. In particular, the flux is always sufficiently regular for all values of $p \in (1,+\infty)$ (that is, $s_{\boldsymbol{\sigma}} > 1$), whereas the potential can be less regular than $W^{2,p}$ for values of p larger than $\sim$3.73. In Table 1, we also report the experimental orders of convergence

$$\text{e.o.c.} := \frac{\log(e_N) - \log(e_{N-1})}{\frac{1}{2}\log|\mathcal{V}_{N-1}| - \frac{1}{2}\log|\mathcal{V}_N|},$$

where e_N is the error on the last mesh, e_{N-1} is the error on the last but one mesh, and $|\mathcal{V}_N|$ and $|\mathcal{V}_{N-1}|$ denote the corresponding number of vertices. For the errors e_i, we consider the error upper bound $\mathcal{J}_u^{\rm up}(u_{{\rm L},h})$ from (6.1a), the error lower bound $\mathcal{J}_u^{\rm low}(u_{{\rm L},h})$ from (6.1b), the total estimator η, and the energy error $\|\nabla(u - u_{{\rm L},h})\|_p$. We only consider here uniformly refined meshes; since the flux $-\boldsymbol{\sigma}(\nabla u)$ is always regular and our a posteriori error estimates are related to the error in the flux, adaptive refinement does not lead here to improved convergence rates.

Figs. 2–4 present the error upper and lower bounds $\mathcal{J}_u^{\rm up}(u_{{\rm L},h})$ and $\mathcal{J}_u^{\rm low}(u_{{\rm L},h})$, the total estimators η, and the corresponding upper and lower effectivity indices $\mathcal{I}^{\rm up}$ and $\mathcal{I}^{\rm low}$ as a function of the number of mesh vertices. We can see from Figs. 2–4 that $\mathcal{I}^{\rm up}$ is very close to one for all the considered values of p. We have proven in Sections 3 and 4 that $\eta \sim \mathcal{J}_u(u_{{\rm L},h})$, i.e., that η and $\mathcal{J}_u(u_{{\rm L},h})$ are equivalent up to a constant independent of the nonlinear function a and of p. Thus, since numerically $\eta \sim \mathcal{J}_u^{\rm up}(u_{{\rm L},h})$, we deduce that in the present case, $\mathcal{J}_u(u_{{\rm L},h}) \sim \mathcal{J}_u^{\rm up}(u_{{\rm L},h})$. This is not the case for the computable lower bound $\mathcal{J}_u^{\rm low}(u_{{\rm L},h})$ for the larger values of p. The increase of the lower effectivity index $\mathcal{I}^{\rm low}$ for $p = 10$ (see Fig. 4, right) can be attributed to the fact that in this case, the potential

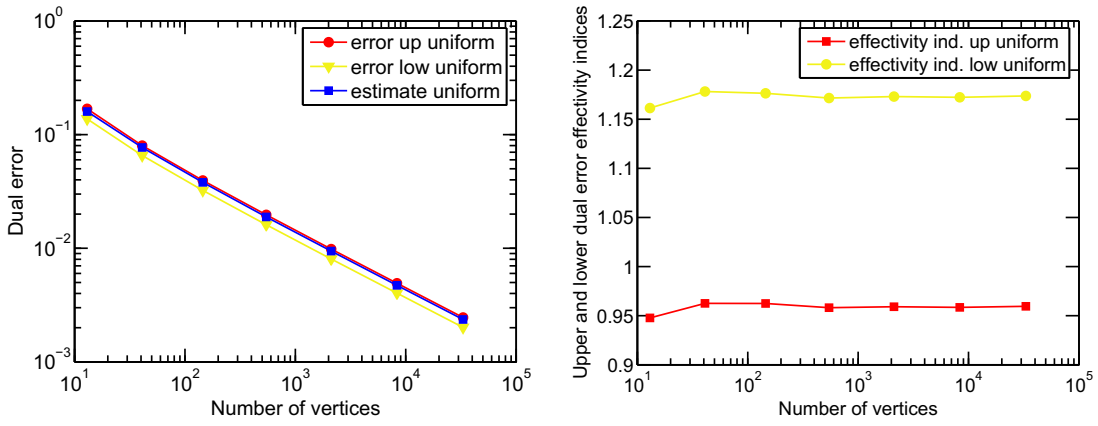

Fig. 2. Estimated and actual dual errors (left) and corresponding effectivity indices (right) for $p = 1.4$, case 1.

Please cite this article in press as: L. El Alaoui et al., Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems, Comput. Methods Appl. Mech. Engrg. (2010), doi:10.1016/j.cma.2010.03.024

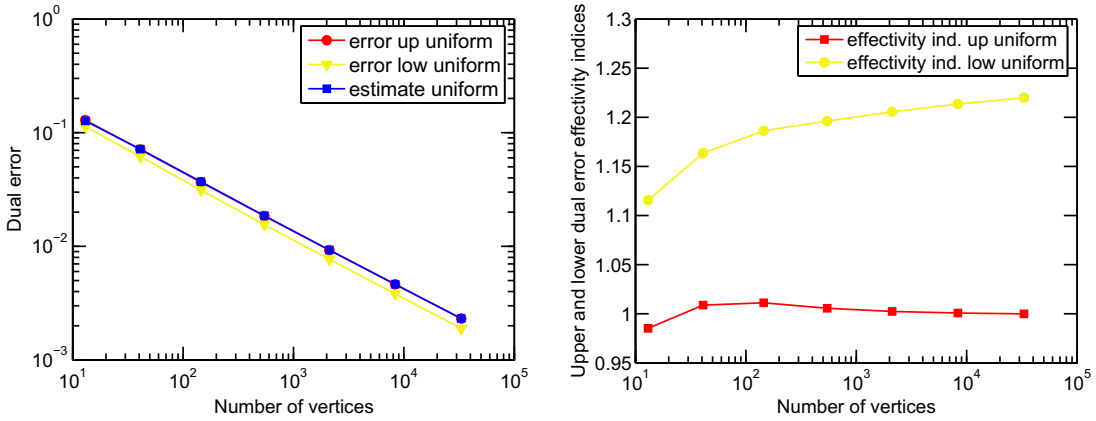

Fig. 3. Estimated and actual dual errors (left) and corresponding effectivity indices (right) for $p = 3$, case 1.

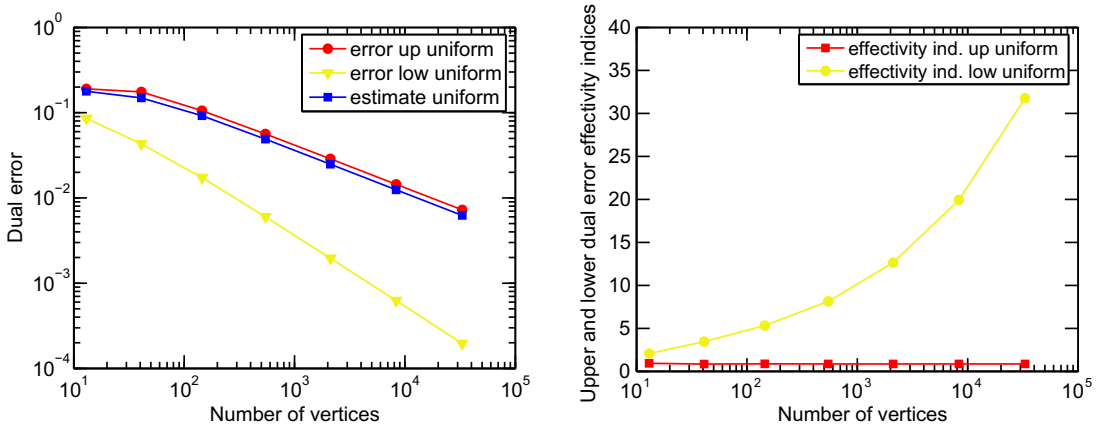

Fig. 4. Estimated and actual dual errors (left) and corresponding effectivity indices (right) for $p = 10$, case 1.

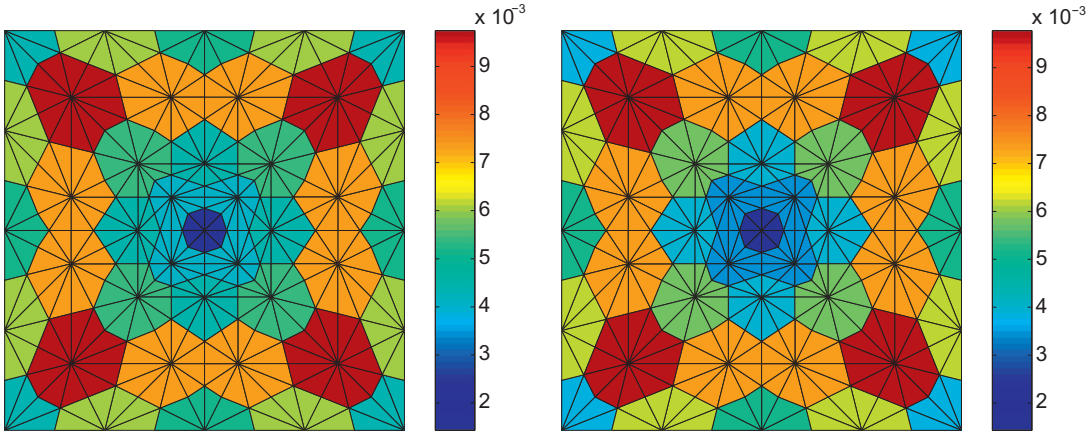

Fig. 5. Estimated (left) and actual (right) error distribution for $p = 3$, case 1.

has low regularity, as reflected in Table 1, and that $\mathcal{J}_u^{\text{low}}(u_{L,h})$ is scaled by the reciprocal of $\|\nabla(u - u_{L,h})\|_p$.

Fig. 5 compares the error distribution predicted by our estimator η (we show the quantity $\eta_{R,D} + \eta_{DF,D} + \eta_{L,D}$ on each dual volume) with the actual error represented by the localized version of its upper bound, namely $\|\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}(\nabla u_{L,h})\|_{q,D}$. By Remark 5.2, $\eta_{R,D} = 0$ for all $D \in \mathcal{D}_h$ in the present case, while $\eta_{L,D}$ is bounded by (3.7); hence, the only significative contribution to η stems from the diffusive flux estimators $\eta_{DF,D}$. We stress that all the above results are presented with the Newton–Raphson method not fully converged but stopped following (3.7) with $\gamma_D = 0.1$.

Finally, Figs. 6 and 7 illustrate the performance of our stopping criterion for the Newton–Raphson iteration by comparing it to a more classical stopping criterion, namely $\eta_L \leqslant 10^{-8}$. Firstly, Fig. 6

Please cite this article in press as: L. El Alaoui et al., Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems, Comput. Methods Appl. Mech. Engrg. (2010), doi:10.1016/j.cma.2010.03.024

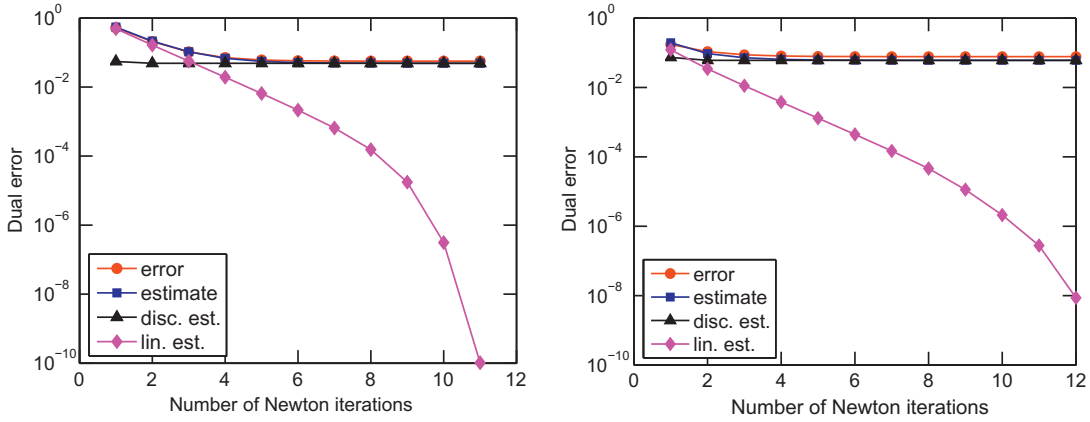

Fig. 6. Discretization and linearization estimators, total estimator, and total error as a function of Newton–Raphson iterations for $p = 10$ and the 4th level uniformly refined mesh (left) and $p = 50$ and the 5th level uniformly refined mesh (right), case 1.

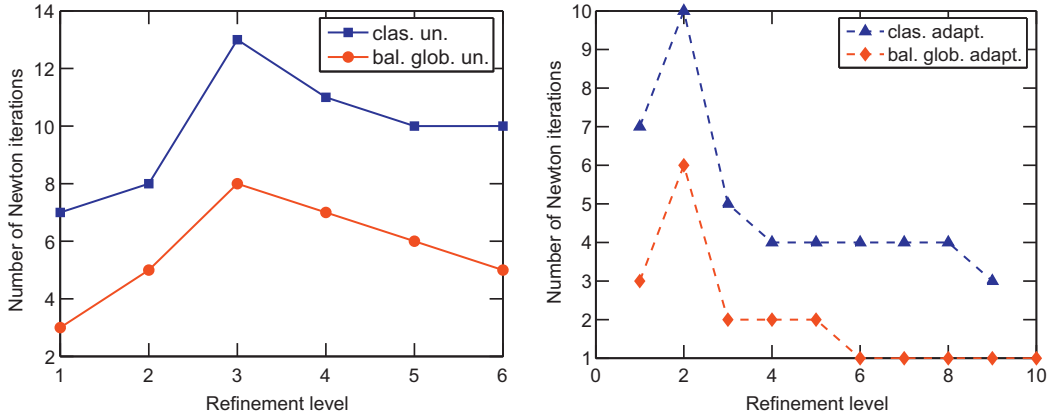

Fig. 7. Number of Newton–Raphson iterations for the classical stopping criterion and the global stopping criterion (3.6) on uniformly (left)/adaptively (right) refined meshes, $p = 10$, case 1.

presents the discretization estimator η_D, the linearization estimator η_L, the total estimator η, and the error upper bound $\mathcal{J}_u^{up}(u_{L,h})$ as a function of Newton–Raphson iterations on a fixed mesh. We clearly see that the linearization estimator dominates the discretization one only at the first few iterations, and then becomes negligible while the total error stagnates. This confirms that the Newton–Raphson iteration can be safely stopped rather early. This effect becomes more pronounced as p is increased and the mesh is refined. For instance, in the left part of Fig. 6, the global stopping criterion (3.6) is reached after five iterations and the classical one after 11 iterations; in the right part of the figure, these numbers become respectively 4 and 12. Secondly, the left part of Fig. 7 presents the Newton–Raphson iteration numbers on a series of uniformly refined meshes in two situations, namely the global stopping criterion (3.6) and the classical one. It appears that around 50% of Newton–Raphson iterations can be spared. The advantage of the present algorithm shows more noticeably in the right part of Fig. 7. Here, we refine the mesh in an adaptive way. As we remarked before, this does not lead to increased experimental orders of convergence, as the flux possesses enough regularity. However, it appears that the elements slowing down the Newton–Raphson convergence are suitably refined, so that now the number of necessary iterations with the global stopping criterion (3.6) with $\gamma = 0.1$ on refined meshes drops down to one or two. This should be compared with the upper curve in the left part of Fig. 7, showing that the number of iterations using the classical stopping criterion in combination with uniform mesh refinement is about 10. A similar behavior is observed for the case $p = 1.4$ (not shown). This effect is less pronounced for $p = 3$ since in this case, the Newton–Raphson method always converges in a few iterations.

6.2. A second test case

We consider here a test case taken from Carstensen and Klose [9, Example 3]. We consider the L-shaped domain $\Omega := (-1,1) \times (-1,1) \backslash [0,1] \times [-1,0]$ and prescribe the Dirichlet boundary condition and the source term f by the exact solution

$$u(r,\theta) = r^{\alpha} \sin(\alpha\theta).$$

Here (r,θ) are the polar coordinates and $\alpha = 7/8$. We consider the value $p = 4$ for which the regularity of the potential and the flux, as defined above, are $s_u = 1.38$ and $s_\sigma = 1.13$.

For this second test case, f is not piecewise constant and hence, following Remark 5.5, the final a posteriori error estimate is given by $\eta + \eta_{Osc}$, with f replaced by f_h in η. The stopping criteria of Section 3.3 are not modified, that is, they rely on η and its components only. Likewise, we only use η and not $\eta + \eta_{Osc}$ to drive mesh

Please cite this article in press as: L. El Alaoui et al., Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems, Comput. Methods Appl. Mech. Engrg. (2010), doi:10.1016/j.cma.2010.03.024

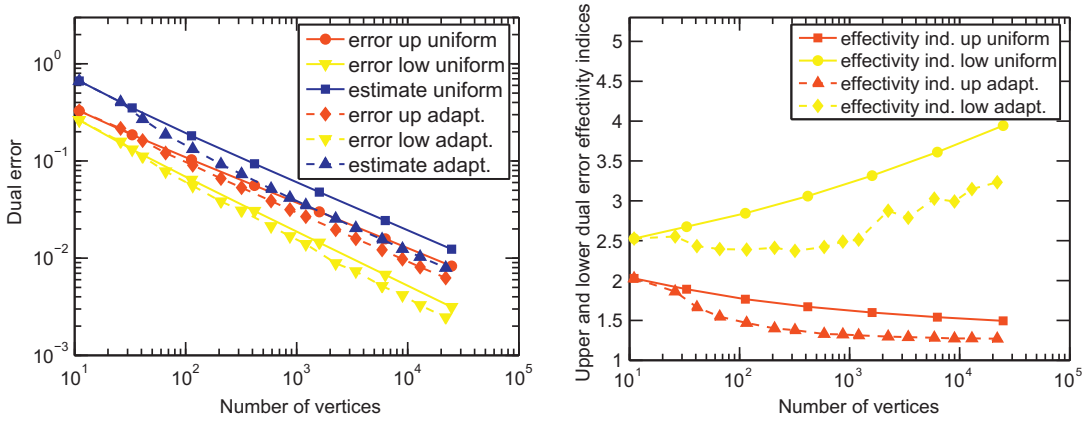

Fig. 8. Estimated and actual dual errors (left) and corresponding effectivity indices (right) for $p = 4$, case 2.

adaptivity. Fig. 8 presents the error upper and lower bounds $\mathcal{J}_u^{\mathrm{up}}(u_{\mathrm{L},h})$ and $\mathcal{J}_u^{\mathrm{low}}(u_{\mathrm{L},h})$, the total estimators $\eta + \eta_{\mathrm{Osc}}$, and the effectivity indices $(\eta + \eta_{\mathrm{Osc}})/\mathcal{J}_u^{\mathrm{up}}(u_{\mathrm{L},h})$ and $(\eta + \eta_{\mathrm{Osc}})/\mathcal{J}_u^{\mathrm{low}}(u_{\mathrm{L},h})$ as a function of the number of mesh vertices. In particular the effectivity indices $(\eta + \eta_{\mathrm{Osc}})/\mathcal{J}_u^{\mathrm{up}}(u_{\mathrm{L},h})$ are dominated by the data oscillation $\eta_{\mathrm{Osc}}/\mathcal{J}_u^{\mathrm{up}}(u_{\mathrm{L},h})$ for rough meshes, since the source term f is singular here. They only tend to the optimal value of one when η_{Osc} becomes insignificant. The effectivity index $\eta/\mathcal{J}_u^{\mathrm{up}}(u_{\mathrm{L},h})$ is close to one on all meshes, uniformly or adaptively refined, in agreement with the theory and similarly to case 1. The experimental orders of convergence are close to one for the error upper and lower bounds and for the error estimate using either uniform and adaptively refined meshes. Furthermore, Fig. 9 compares the energy error $\|u - u_{\mathrm{L},h}\|_V$ decrease on uniform and adaptively refined meshes. The experimental order of convergence is 0.38 on uniform meshes, in agreement with the regularity of the exact solution, and 0.89 on adaptively refined meshes. Finally, Fig. 10 shows the comparison of the predicted error distribution given by our estimator η and of the actual error, both quantities being localized as before, at the 5th adaptively refined mesh. We can see that the estimated and actual error distributions match extremely well.

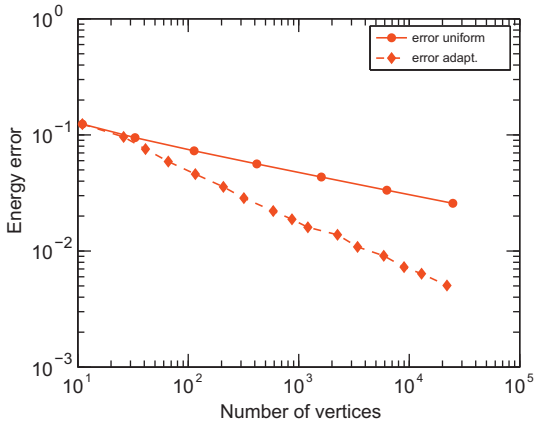

Fig. 9. Energy errors $\|u - u_{\mathrm{L},h}\|_V$ on uniformly/adaptively refined meshes for $p = 4$, case 2.

Appendix A. Technical results

In this appendix we collect various technical results. Henceforth, C_κ denotes a generic constant that only depends on the shape regularity parameter κ of the mesh family $(\mathcal{S}_h)_h$ and whose value can change at each occurrence.

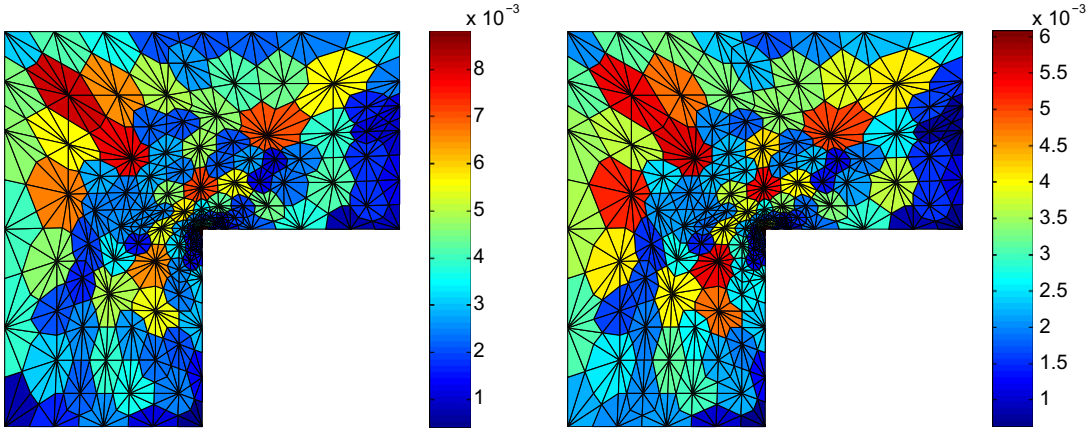

Fig. 10. Estimated (left) and actual (right) error distribution for $p = 4$, case 2.

Please cite this article in press as: L. El Alaoui et al., Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems, Comput. Methods Appl. Mech. Engrg. (2010), doi:10.1016/j.cma.2010.03.024

A.1. Inverse inequality

Let $T \in \mathcal{S}_h$ with diameter denoted by h_T. For an integer $k \geqslant 1$, let $\mathbb{P}_k(T)$ denote the vector space of polynomials of total degree $\leqslant k$ on T.

Lemma 6.1. *For all $v \in \mathbb{P}_k(T)$ and for all $r \in [1,+\infty]$, there holds*

$$\|\nabla v\|_{r,T} \leqslant C h_T^{-1} \|v\|_{r,T}, \tag{A.1}$$

where the constant C can depend on d, k, and κ, but is independent of the Lebesgue exponent r.

Proof. Following Verfürth [42], for a set $K \subset \Omega$ with positive d-dimensional measure, we consider the constants

$$C^{\sharp}_{d,k,r}(K) := \sup_{v \in \mathbb{P}_k(K)\setminus\{0\}} \frac{|K|^{\frac{1}{2}-\frac{1}{r}} \|v\|_{r,K}}{\|v\|_K}, \tag{A.2a}$$

$$C^{\flat}_{d,k,r}(K) := \sup_{v \in \mathbb{P}_k(K)\setminus\{0\}} \frac{\|v\|_K}{|K|^{\frac{1}{2}-\frac{1}{r}} \|v\|_{r,K}}, \tag{A.2b}$$

where $|K|$ denotes the d-dimensional Lebesgue measure of K. It is proven in [42] that

$$C^{\sharp}_{d,k,r}(K) \leqslant \begin{cases} 1 & \text{if } 1 \leqslant r \leqslant 2, \\ C^{\sharp}_{d,k,\infty}(K)^{1-\frac{2}{r}} & \text{if } 2 < r \leqslant +\infty, \end{cases} \tag{A.3a}$$

$$C^{\flat}_{d,k,r}(K) \leqslant \begin{cases} C^{\sharp}_{d,k,\infty}(K)^{\frac{2}{r}-1} & \text{if } 1 \leqslant r \leqslant 2, \\ 1 & \text{if } 2 < r \leqslant +\infty, \end{cases} \tag{A.3b}$$

so that

$$C^{\flat}_{d,k,r}(K) C^{\sharp}_{d,k,r}(K) \leqslant C^{\sharp}_{d,k,\infty}(K)^{\left|1-\frac{2}{r}\right|}.$$

Moreover, for a d-dimensional simplex T, there holds

$$C^{\sharp}_{d,k,\infty}(T) \leqslant C^{\star}_{d,k} := (2k+2)^{\frac{1}{2}} (4k+2)^{\frac{d-1}{2}}. \tag{A.4}$$

Hence, since $C^{\star}_{d,k} \geqslant 1$ and $\left|1-\frac{2}{r}\right| \leqslant 1$,

$$C^{\flat}_{d,k,r}(T) C^{\sharp}_{d,k,r}(T) \leqslant C^{\star}_{d,k}.$$

As a result, using the constant C_κ from the usual inverse inequality in the L^2-setting [20] leads to

$$\begin{aligned} \|\nabla v\|_{r,T} &\leqslant C^{\sharp}_{d,k,r}(T) |T|^{\frac{1}{r}-\frac{1}{2}} \|\nabla v\|_T \leqslant C_\kappa h_T^{-1} C^{\sharp}_{d,k,r}(T) |T|^{\frac{1}{r}-\frac{1}{2}} \|v\|_T \\ &\leqslant C_\kappa h_T^{-1} C^{\sharp}_{d,k,r}(T) C^{\flat}_{d,k,r}(T) \|v\|_{r,T} \leqslant C^{\star}_{d,k} C_\kappa h_T^{-1} \|v\|_{r,T}, \end{aligned}$$

which completes the proof. □

A.2. Proof of Lemma 4.3

We will use the following inequalities (see Ref. [37] for the L^2 framework, Refs. [39,42] for the extension to the L^p framework, and Lemma 6.1): for all $T \in \mathcal{S}_h$, for all $v \in \mathbb{P}_k(T)$, for all $F \in \mathcal{G}_h^{\mathrm{int}}$, and for all $\phi \in \mathbb{P}_k(F)$, there holds

$$\|v\|_{q,T} \lesssim \sup_{w \in \mathbb{P}_k(T),\ \|w\|_{p,T}=1} (v, \Psi_T w)_T, \tag{A.5a}$$

$$\|\nabla(\Psi_T v)\|_{p,T} \lesssim h_T^{-1} \|v\|_{p,T}, \tag{A.5b}$$

$$\|\phi\|_{q,F} \lesssim \sup_{w \in \mathbb{P}_k(F),\ \|w\|_{p,F}=1} (\phi, \Psi_F w)_F, \tag{A.5c}$$

$$\|\nabla(\Psi_F \phi)\|_{p,\omega_F} \lesssim h_F^{-\frac{1}{q}} \|\phi\|_{p,F}, \tag{A.5d}$$

$$\|\Psi_F \phi\|_{p,\omega_F} \lesssim h_F^{\frac{1}{p}} \|\phi\|_{p,F}, \tag{A.5e}$$

where Ψ_T is the usual element bubble associated with T, Ψ_F the usual extension operator on the mesh $\mathcal{S}_h$ associated with the face bubble on F, and $\omega_F \subset \Omega$ denotes the union of the two simplices of $\mathcal{S}_h$ sharing F. The proof is decomposed into three parts. Let $D \in \mathcal{D}_h$.

We first prove that

$$h_T \|f + \nabla \cdot \boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,T} \lesssim \|\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,T} \quad \forall T \in \mathcal{S}_D. \tag{A.6}$$

Let $T \in \mathcal{S}_D$, set $v := f + \nabla \cdot \boldsymbol{\sigma}_{\mathrm{L},h}$ and observe that $v \in \mathbb{P}_k(T)$ by the assumptions on f and $\boldsymbol{\sigma}_{\mathrm{L},h}$. Thus, using (A.5a) and (1.3), the Green theorem, the fact that Ψ_T vanishes on ∂T, the Hölder inequality, and (A.5b) leads to

$$\begin{aligned} \|v\|_{q,T} &\lesssim \sup_{w \in \mathbb{P}_k(T),\ \|w\|_{p,T}=1} (f + \nabla \cdot \boldsymbol{\sigma}_{\mathrm{L},h}, \Psi_T w)_T \\ &= \sup_{w \in \mathbb{P}_k(T),\ \|w\|_{p,T}=1} (\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}_{\mathrm{L},h}, \nabla(\Psi_T w))_T \\ &\lesssim \sup_{w \in \mathbb{P}_k(T),\ \|w\|_{p,T}=1} \|\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,T} h_T^{-1} \|w\|_{p,T} \\ &= h_T^{-1} \|\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,T}, \end{aligned}$$

whence (A.6) follows.

We next show that

$$h_F^{\frac{1}{q}} \|[\![\boldsymbol{\sigma}_{\mathrm{L},h} \cdot \mathbf{n}]\!]\|_{q,F} \lesssim \|\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,\omega_F} \quad \forall F \in \mathcal{G}_D^{\mathcal{T}}. \tag{A.7}$$

Set $\phi := [\![\boldsymbol{\sigma}_{\mathrm{L},h} \cdot \mathbf{n}]\!]$ and observe that $\phi \in \mathbb{P}_k(F)$. Let $w \in \mathbb{P}_k(F)$ with $\|w\|_{p,F} = 1$. Using (1.3), the Green theorem, the properties of the bubble functions, the Hölder inequality, (A.5d), (A.5e), (A.6), and the fact that $1/q = 1 - 1/p$ then yields

$$\begin{aligned} (\phi, \Psi_F w)_F &= (f + \nabla \cdot \boldsymbol{\sigma}_{\mathrm{L},h}, \Psi_F w)_{\omega_F} - (\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}_{\mathrm{L},h}, \nabla(\Psi_F w))_{\omega_F} \\ &\leqslant \|f + \nabla \cdot \boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,\omega_F} \|\Psi_F w\|_{p,\omega_F} + \|\boldsymbol{\sigma}(\nabla u) \\ &\quad - \boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,\omega_F} \|\nabla(\Psi_F w)\|_{p,\omega_F} \lesssim h_F^{\frac{1}{p}} \|f + \nabla \cdot \boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,\omega_F} \|w\|_{p,F} \\ &\quad + h_F^{-\frac{1}{q}} \|\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,\omega_F} \|w\|_{p,F} \lesssim h_F^{-\frac{1}{q}} \|\boldsymbol{\sigma}(\nabla u) \\ &\quad - \boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,\omega_F} \|w\|_{p,F} \\ &= h_F^{-\frac{1}{q}} \|\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,\omega_F}, \end{aligned}$$

whence (A.7) follows from (A.5c). Finally, combining the above results leads to $\eta_{\mathrm{res},D} \lesssim \|\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,D}$, and using the triangle inequality yields

$$\|\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,D} \leqslant \|\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}(\nabla u_{\mathrm{L},h})\|_{q,D} + \eta_{\mathrm{L},D},$$

whence (4.3) follows.

A.3. Proof of Lemma 4.7

We first give a technical extension of (A.5a). Let $\mathbb{P}_k(\mathcal{S}_h)$ be spanned by piecewise polynomials of total degree $\leqslant k$ on the mesh $\mathcal{S}_h$. Then, for all $v \in \mathbb{P}_k(\mathcal{S}_h)$,

$$\left\{ \sum_{T \in \mathcal{S}_h} h_T^q \|v\|_{q,T}^q \right\}^{\frac{1}{q}} \lesssim \sup_{w \in \mathbb{P}_k(\mathcal{S}_h),\ \|w\|_p=1} \sum_{T \in \mathcal{S}_h} (v, h_T \Psi_T w)_T. \tag{A.8}$$

Indeed, for a given $T \in \mathcal{S}_h$, multiplying (A.5a) by h_T yields

$$\{h_T^q \|v\|_{q,T}^q\}^{\frac{1}{q}} \lesssim \sup_{w \in \mathbb{P}_k(T),\ \|w\|_{p,T}=1} (v, h_T \Psi_T w)_T.$$

Thus,

$$h_T^q \|v\|_{q,T}^q = \{h_T^q \|v\|_{q,T}^q\}^{\frac{1}{q}} \{h_T^q \|v\|_{q,T}^q\}^{\frac{1}{p}} \lesssim \sup_{w \in \mathbb{P}_k(T),\ \|w\|_{p,T}=\{h_T^q \|v\|_{q,T}^q\}^{\frac{1}{p}}} (v, h_T \Psi_T w)_T.$$

Consequently, since the restrictions of functions in $\mathbb{P}_k(\mathcal{S}_h)$ to the elements of $\mathcal{S}_h$ can be chosen independently,

$$\sum_{T\in\mathcal{S}_h} h_T^q \|v\|_{q,T}^q \lesssim \sup_{w\in\mathbb{P}_k(\mathcal{S}_h),\ \|w\|_{p,T}=\{h_T^q\|v\|_{q,T}^q\}^{\frac{1}{p}}\ \forall T\in\mathcal{S}_h} \sum_{T\in\mathcal{S}_h} (v, h_T\Psi_T w)_T,$$

whence (A.8) follows by extending the supremum set and rescaling the argument w. The following extension of (A.5c) is proven similarly: For all $\phi\in\mathbb{P}_k(\mathcal{G}_h^{\mathcal{T}})$, the vector space spanned by piecewise polynomials of total degree $\leqslant k$ on each face of $\mathcal{G}_h^{\mathcal{T}}$,

$$\left\{\sum_{F\in\mathcal{G}_h^{\mathcal{T}}} h_F\|\phi\|_{q,F}^q\right\}^{\frac{1}{q}} \lesssim \sup_{w\in\mathbb{P}_k(\mathcal{G}_h^{\mathcal{T}}),\ \|w\|_{p,\mathcal{G}_h^{\mathcal{T}}}=1} \sum_{F\in\mathcal{G}_h^{\mathcal{T}}} (\phi, h_F^{\frac{1}{q}}\Psi_F w)_F, \tag{A.9}$$

where $\|w\|_{p,\mathcal{G}_h^{\mathcal{T}}} := \{\sum_{F\in\mathcal{G}_h^{\mathcal{T}}}\|w\|_{p,F}^p\}^{\frac{1}{p}}$.

We now prove that

$$\left\{\sum_{T\in\mathcal{S}_h} h_T^q\|f+\nabla\cdot\boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,T}^q\right\}^{\frac{1}{q}} \lesssim \|Au - A_{\mathrm{L}}u_{\mathrm{L},h}\|_{V'}. \tag{A.10}$$

Set $v|_T := (f+\nabla\cdot\boldsymbol{\sigma}_{\mathrm{L},h})|_T$ for all $T\in\mathcal{S}_h$, and observe that $v\in\mathbb{P}_k(\mathcal{S}_h)$. Let $w\in\mathbb{P}_k(\mathcal{S}_h)$ with $\|w\|_p = 1$. Set $\lambda|_T := h_T\Psi_T w|_T$ for all $T\in\mathcal{S}_h$, and observe that $\lambda\in V$ since Ψ_T vanishes on ∂T for all $T\in\mathcal{S}_h$, and that owing to (A.5b), $\|\lambda\|_V = \|\nabla\lambda\|_p \lesssim \|w\|_p = 1$. Then, proceeding as in Section A.2,

$$\begin{aligned}\sum_{T\in\mathcal{S}_h}(v, h_T\Psi_T w)_T &= \sum_{T\in\mathcal{S}_h}(v,\lambda)_T = (\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}_{\mathrm{L},h}, \nabla\lambda)\\ &\leqslant \|Au - A_{\mathrm{L}}u_{\mathrm{L},h}\|_{V'}\|\lambda\|_V \lesssim \|Au - A_{\mathrm{L}}u_{\mathrm{L},h}\|_{V'},\end{aligned}$$

whence (A.10) follows from (A.8).

We next show that

$$\left\{\sum_{F\in\mathcal{G}_h^{\mathcal{T}}} h_F\|[\![\boldsymbol{\sigma}_{\mathrm{L},h}\cdot\mathbf{n}]\!]\|_{q,F}^q\right\}^{\frac{1}{q}} \lesssim \|Au - A_{\mathrm{L}}u_{\mathrm{L},h}\|_{V'}. \tag{A.11}$$

Set $\phi|_F := [\![\boldsymbol{\sigma}_{\mathrm{L},h}\cdot\mathbf{n}]\!]_F$ for all $F\in\mathcal{G}_h^{\mathcal{T}}$ and observe that $\phi\in\mathbb{P}_k(\mathcal{G}_h^{\mathcal{T}})$. Let $w\in\mathbb{P}_k(\mathcal{G}_h^{\mathcal{T}})$ with $\|w\|_{p,\mathcal{G}_h^{\mathcal{T}}} = 1$. Set $\lambda|_F := h_F^{\frac{1}{q}}\Psi_F w|_F$ for all $F\in\mathcal{G}_h^{\mathcal{T}}$. This defines the function λ on the set $\bigcup_{F\in\mathcal{G}_h^{\mathcal{T}}}\omega_F$, and the function λ is extended by zero outside this set. We first observe that $\lambda\in V$. Moreover, since for $F, F'\in\mathcal{G}_h^{\mathcal{T}}$, $F\neq F'$, $\omega_F\cap\omega_{F'}$ has zero measure, it is inferred using (A.5e) that

$$\left\{\sum_{T\in\mathcal{S}_h} h_T^{-p}\|\lambda\|_{p,T}^p\right\}^{\frac{1}{p}} \lesssim \left\{\sum_{F\in\mathcal{G}_h^{\mathcal{T}}} h_F^{-p}\|\lambda\|_{p,\omega_F}^p\right\}^{\frac{1}{p}} \lesssim \|w\|_{p,\mathcal{G}_h^{\mathcal{T}}} = 1.$$

Similarly, owing to (A.5d), $\|\lambda\|_V = \|\nabla\lambda\|_p \lesssim 1$. As a result,

$$\begin{aligned}\sum_{F\in\mathcal{G}_h^{\mathcal{T}}}(\phi, h_F^{\frac{1}{q}}\Psi_F w)_F &= \sum_{F\in\mathcal{G}_h^{\mathcal{T}}}(\phi,\lambda)_F\\ &= \sum_{T\in\mathcal{S}_h}\left\{(f+\nabla\cdot\boldsymbol{\sigma}_{\mathrm{L},h},\lambda)_T - (\boldsymbol{\sigma}(\nabla u) - \boldsymbol{\sigma}_{\mathrm{L},h}, \nabla\lambda)_T\right\}\\ &\leqslant \left\{\sum_{T\in\mathcal{S}_h} h_T^q\|f+\nabla\cdot\boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,T}^q\right\}^{\frac{1}{q}}\left\{\sum_{T\in\mathcal{S}_h} h_T^{-p}\|\lambda\|_{p,T}^p\right\}^{\frac{1}{p}}\\ &\quad + \|Au - A_{\mathrm{L}}u_{\mathrm{L},h}\|_{V'}\|\lambda\|_V \lesssim \|Au - A_{\mathrm{L}}u_{\mathrm{L},h}\|_{V'},\end{aligned}$$

owing to (A.10), whence (A.11) follows from (A.9).

Finally, (A.10) and (A.11) together with the triangle inequality imply that $\eta_{\mathrm{res}} \lesssim \|Au - A_{\mathrm{L}}u_{\mathrm{L},h}\|_{V'}$, and using a further triangle inequality leads to

$$\|Au - A_{\mathrm{L}}u_{\mathrm{L},h}\|_{V'} \leqslant \|Au - Au_{\mathrm{L},h}\|_{V'} + \|Au_{\mathrm{L},h} - A_{\mathrm{L}}u_{\mathrm{L},h}\|_{V'}.$$

Using the Hölder inequality, the second term in the right-hand side is bounded by η_{L} as defined by (3.6). This concludes the proof.

A.4. Proof of Lemma 5.3

Consider first the case of direct prescription. Let $D\in\mathcal{D}_h$ and let $T\in\mathcal{S}_D$ with outward unit normal $\mathbf{n}_T$. Then, for all $\mathbf{v}_h\in\mathbf{RTN}(T)$, the lowest-order Raviart–Thomas–Nédélec finite element space on T, owing to the usual equivalence result in the L^2-setting for $\mathbf{RTN}(T)$ functions (which can be shown using the equivalence of norms in finite dimension, the Piola transformation, and the shape regularity of $\mathcal{S}_h$),

$$\|\mathbf{v}_h\|_{q,T} \leqslant C_{d,1,q}^\sharp(T)|T|^{\frac{1}{q}-\frac{1}{2}}\|\mathbf{v}_h\|_T \leqslant C_\kappa h_T^{\frac{1}{2}}C_{d,1,q}^\sharp(T)|T|^{\frac{1}{q}-\frac{1}{2}}\|\mathbf{v}_h\cdot\mathbf{n}_T\|_{\partial T},$$

where the generic constant C_κ depends on the shape regularity of $\mathcal{S}_h$. Furthermore, since $\mathbf{v}_h\cdot\mathbf{n}_T$ is piecewise constant on ∂T, it easily verified that

$$\|\mathbf{v}_h\cdot\mathbf{n}_T\|_{\partial T} \leqslant C_\kappa|\partial T|^{\frac{1}{2}-\frac{1}{q}}\|\mathbf{v}_h\cdot\mathbf{n}_T\|_{q,\partial T},$$

where $|\partial T|$ denotes the $(d-1)$-dimensional Lebesgue measure of ∂T. As a result, since $C_{d,1,q}^\sharp(T) \leqslant \max\{1, C_{d,1}^\star\}$and using the shape regularity of $\mathcal{S}_h$, it is inferred that

$$\|\mathbf{v}_h\|_{q,T} \leqslant C_{\kappa,d}(|T||\partial T|^{-1})^{\frac{1}{q}}\|\mathbf{v}_h\cdot\mathbf{n}_T\|_{q,\partial T},$$

where $C_{\kappa,d}$ denotes a generic constant only depending on κ and d. Finally, since $|T| \leqslant C_\kappa h_T|\partial T|$ and since $\frac{1}{q}\leqslant 1$, this yields

$$\|\mathbf{v}_h\|_{q,T} \leqslant C_{\kappa,d}h_T^{\frac{1}{q}}\|\mathbf{v}_h\cdot\mathbf{n}_T\|_{q,\partial T}.$$

We now apply this estimate to $\mathbf{v}_h := \boldsymbol{\sigma}_{\mathrm{L},h} + \mathbf{t}_h$. Then, owing to (5.2) and (5.3),

$$\|\boldsymbol{\sigma}_{\mathrm{L},h} + \mathbf{t}_h\|_{q,T}^q \lesssim h_T \sum_{F\subset\partial T,\ F\in\mathcal{G}_D^{\mathcal{T}}} \|[\![\boldsymbol{\sigma}_{\mathrm{L},h}\cdot\mathbf{n}]\!]\|_{q,F}^q.$$

Hence, $\eta_{\mathrm{DF},D}^q \lesssim \sum_{F\in\mathcal{G}_D^{\mathcal{T}}} h_F\|[\![\boldsymbol{\sigma}_{\mathrm{L},h}\cdot\mathbf{n}]\!]\|_{q,F}^q$, whence (4.1) follows.

We now consider the case of local linear system solves. Let $D\in\mathcal{D}_h$. Using the approach of [44, Section 4.1] (cf. also [2,3]), there exists a postprocessing $\tilde{q}_h\in M(\mathcal{S}_D)$ of q_h such that

$$\begin{aligned}&-\nabla\tilde{q}_h = \boldsymbol{\sigma}_{\mathrm{L},h} + \mathbf{t}_h \quad \forall T\in\mathcal{S}_D,\\ &\frac{(\tilde{q}_h, 1)_T}{|T|} = q_h|_T \quad \forall T\in\mathcal{S}_D.\end{aligned}$$

Here, $M(\mathcal{S}_D)$ is a space of particular piecewise polynomials on $\mathcal{S}_D$ of total degree $\leqslant 2$ whose means of traces on interior faces of $\mathcal{S}_D$ are continuous and whose mean value over D is zero when $D\in\mathcal{D}_h^{\mathrm{int}}$ and whose mean values over faces lying in $\partial\Omega\cap\partial D$ are zero when $D\in\mathcal{D}_h^{\mathrm{ext}}$. Then,

$$\|\boldsymbol{\sigma}_{\mathrm{L},h} + \mathbf{t}_h\|_{q,D} \lesssim \sup_{m_h\in M(\mathcal{S}_D),\ \|\nabla m_h\|_{p,D}=1} (\boldsymbol{\sigma}_{\mathrm{L},h} + \mathbf{t}_h, \nabla m_h)_D. \tag{A.12}$$

Indeed, $\|\boldsymbol{\sigma}_{\mathrm{L},h} + \mathbf{t}_h\|_{q,D} = (\nabla\tilde{q}_h, \nabla\tilde{m}_h)_D$ with $\tilde{m}_h = (\|\nabla\tilde{q}_h\|_{q,D}/\|\nabla\tilde{q}_h\|_D^2)\tilde{q}_h$, so that introducing constants as in (A.2b) yields

$$\|\nabla\tilde{m}_h\|_{p,D} = \frac{\|\nabla\tilde{q}_h\|_{q,D}\|\nabla\tilde{q}_h\|_{p,D}}{\|\nabla\tilde{q}_h\|_D^2} \leqslant C_{d,2,p}^\sharp(D)C_{d,2,q}^\sharp(D),$$

and bounding the constants by $C_{d,2,\infty}(D)$ and using the shape regularity of the submesh of the dual volume D leads, as before, to the bound $\|\nabla\tilde{m}_h\|_{p,D} \lesssim 1$. We now develop the right-hand side of (A.12),

Please cite this article in press as: L. El Alaoui et al., Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems, Comput. Methods Appl. Mech. Engrg. (2010), doi:10.1016/j.cma.2010.03.024

$$
\begin{aligned}
(\boldsymbol{\sigma}_{\mathrm{L},h}+\mathbf{t}_h,\nabla m_h)_D &= \sum_{T\in\mathcal{S}_D}\left\{-(m_h,\nabla\cdot(\boldsymbol{\sigma}_{\mathrm{L},h}+\mathbf{t}_h))_T+((\boldsymbol{\sigma}_{\mathrm{L},h}+\mathbf{t}_h)\cdot\mathbf{n},m_h)_{\partial T}\right\}\\
&= -\sum_{T\in\mathcal{S}_D}(m_h,f+\nabla\cdot\boldsymbol{\sigma}_{\mathrm{L},h})_T+\sum_{F\in\mathcal{G}_D^{\mathcal{T}}}\left([\![\boldsymbol{\sigma}_{\mathrm{L},h}\cdot\mathbf{n}]\!],m_h\right)_F\\
&\le \left\{\sum_{T\in\mathcal{S}_D}h_T^{-p}\|m_h\|_{p,T}^p\right\}^{\frac{1}{p}}\left\{\sum_{T\in\mathcal{S}_D}h_T^q\|f+\nabla\cdot\boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,T}^q\right\}^{\frac{1}{q}}\\
&\quad+\left\{\sum_{F\in\mathcal{G}_D^{\mathcal{T}}}h_F^{-\frac{p}{q}}\|m_h\|_{p,F}^p\right\}^{\frac{1}{p}}\left\{\sum_{F\in\mathcal{G}_D^{\mathcal{T}}}h_F\|[\![\boldsymbol{\sigma}_{\mathrm{L},h}\cdot\mathbf{n}]\!]\|_{q,F}^q\right\}^{\frac{1}{q}}\\
&\lesssim h_D^{-1}\|m_h\|_{p,D}\left\{\sum_{T\in\mathcal{S}_D}h_T^q\|f+\nabla\cdot\boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,T}^q\right.\\
&\quad\left.+\sum_{F\in\mathcal{G}_D^{\mathcal{T}}}h_F\|[\![\boldsymbol{\sigma}_{\mathrm{L},h}\cdot\mathbf{n}]\!]\|_{q,F}^q\right\}^{\frac{1}{q}},
\end{aligned}
$$

using the Green theorem, the fact that $\nabla\cdot\mathbf{t}_h=f$ for all $T\in\mathcal{S}_D$ owing to (5.4b), the fact that $[\![\mathbf{t}_h\cdot\mathbf{n}]\!]_F=0$ for all $F\in\mathcal{G}_D^{\mathcal{T}}$ since $\mathbf{t}_h\in\mathbf{RTN}_{\mathrm{N}}(\mathcal{S}_D)$, the Hölder inequality, and the inverse inequality

$$\|m_h\|_{p,F}\lesssim h_F^{-\frac{1}{p}}\|m_h\|_{p,T},$$

which can be proven by proceeding as above using the usual inverse inequality in the L^2-setting and the norm equivalence constants. We now use the discrete Poincaré/Friedrichs inequality (recall that $(m_h,1)_D=0$ or that $(m_h,1)_{\partial D\cap\partial\Omega}=0$ since $m_h\in M(\mathcal{S}_D)$)

$$\|m_h\|_{p,D}\lesssim h_D\|\nabla m_h\|_{p,D},$$

which can be proven by proceeding as above using the usual discrete Poincaré/Friedrichs inequality in the L^2-setting (for nonconvex D, an upper bound only depending on κ can be inferred from Vohralík [43]) and the norm equivalence constants. Then,

$$(\boldsymbol{\sigma}_{\mathrm{L},h}+\mathbf{t}_h,\nabla m_h)_D\lesssim\left\{\sum_{T\in\mathcal{S}_D}h_T^q\|f+\nabla\cdot\boldsymbol{\sigma}_{\mathrm{L},h}\|_{q,T}^q+\sum_{F\in\mathcal{G}_D^{\mathcal{T}}}h_F\|[\![\boldsymbol{\sigma}_{\mathrm{L},h}\cdot\mathbf{n}]\!]\|_{q,F}^q\right\}^{\frac{1}{q}},$$

and (4.1) now follows from (A.12).

References

[1] M. Ainsworth, J.T. Oden, A posteriori error estimation in finite element analysis, Pure Appl. Math. (2000).
[2] T. Arbogast, Z. Chen, On the implementation of mixed methods as nonconforming methods for second-order elliptic problems, Math. Comp. 64 (211) (1995) 943–972.
[3] D.N. Arnold, F. Brezzi, Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates, RAIRO Modél. Math. Anal. Numér. 19 (1) (1985) 7–32.
[4] R.E. Bank, D.J. Rose, Some error estimates for the box method, SIAM J. Numer. Anal. 24 (4) (1987) 777–787.
[5] R.E. Bank, A. Weiser, Some a posteriori error estimators for elliptic partial differential equations, Math. Comp. 44 (170) (1985) 283–301.
[6] J.W. Barrett, W.B. Liu, Finite element approximation of the p-Laplacian, Math. Comp. 61 (204) (1993) 523–537.
[7] D. Braess, J. Schöberl, Equilibrated residual error estimator for edge elements, Math. Comp. 77 (262) (2008) 651–672.
[8] F. Brezzi, M. Fortin, Mixed and Hybrid Finite Element Methods, Springer Series in Computational Mathematics, vol. 15, Springer-Verlag, New York, 1991.
[9] C. Carstensen, R. Klose, A posteriori finite element error control for the p-Laplace problem, SIAM J. Sci. Comput. 25 (3) (2003) 792–814.
[10] C. Carstensen, W. Liu, N. Yan, A posteriori FE error control for p-Laplacian by gradient recovery in quasi-norm, Math. Comp. 75 (256) (2006) 1599–1616.
[11] A. Chaillou, M. Suri, A posteriori estimation of the linearization error for strongly monotone nonlinear operators, J. Comput. Appl. Math. 205 (1) (2007) 72–87.
[12] A.L. Chaillou, M. Suri, Computable error estimators for the approximation of nonlinear problems by linearized models, Comput. Methods Appl. Mech. Engrg. 196 (1–3) (2006) 210–224.
[13] I. Cheddadi, R. Fučík, M.I. Prieto, M. Vohralík, Guaranteed and robust a posteriori error estimates for singularly perturbed reaction–diffusion problems, M2AN Math. Model. Numer. Anal. 43 (5) (2009) 867–888.
[14] S.-K. Chua, R.L. Wheeden, Estimates of best constants for weighted Poincaré inequalities on convex domains, Proc. London Math. Soc. (3) (2006) 197–226. 93, 1.
[15] P.G. Ciarlet, The Finite Element Method for Elliptic Problems, Classics in Applied Mathematics, vol. 40, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002 (Reprint of the 1978 original, North-Holland, Amsterdam, MR0520174 (58 #25001)).
[16] A. Cohen, W. Dahmen, R. Devore, Adaptive wavelet schemes for nonlinear variational problems, SIAM J. Numer. Anal. 41 (5) (2003) 1785–1823.
[17] P. Destuynder, B. Métivet, Explicit error bounds in a conforming finite element method, Math. Comp. 68 (228) (1999) 1379–1396.
[18] L. Diening, F. Ettwein, Fractional estimates for non-differentiable elliptic systems with general growth, Forum Math. 20 (3) (2008) 523–556.
[19] L. Diening, C. Kreuzer, Linear convergence of an adaptive finite element method for the p-Laplacian equation, SIAM J. Numer. Anal. 46 (2) (2008) 614–638.
[20] A. Ern, J.-L. Guermond, Theory and Practice of Finite Elements, Applied Mathematical Sciences, vol. 159, Springer-Verlag, New York, NY, 2004.
[21] A. Ern, M. Vohralík, Flux reconstruction and a posteriori error estimation for discontinuous Galerkin methods on general nonmatching grids, C.R. Math. Acad. Sci. Paris 347 (2009) 441–444.
[22] A. Ern, M. Vohralík, A posteriori error estimation based on potential and flux reconstruction for the heat equation, SIAM J. Numer. Anal. 48 (1) (2010) 198–223.
[23] R. Glowinski, A. Marrocco, Sur l'approximation, par éléments finis d'ordre 1, et la résolution, par pénalisation-dualité, d'une classe de problèmes de Dirichlet non linéaires, C.R. Acad. Sci. Paris Sér. A 278 (1974) 1649–1652.
[24] R. Glowinski, A. Marrocco, Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité, d'une classe de problèmes de Dirichlet non linéaires, Rev. Française Automat. Informat. Recherche Opérationnelle (RAIRO Analyse Numérique) 9 (R-2) (1975) 41–76.
[25] P. Jiránek, Z. Strakoš, M. Vohralík, A posteriori error estimates including algebraic error and stopping criteria for iterative solvers, SIAM J. Sci. Comput. 32 (3) (2010) 1567–1590.
[26] K.Y. Kim, A posteriori error estimators for locally conservative methods of nonlinear elliptic problems, Appl. Numer. Math. 57 (2007) 1065–1080.
[27] P. Ladevèze, D. Leguillon, Error estimate procedure in the finite element method and applications, SIAM J. Numer. Anal. 20 (3) (1983) 485–509.
[28] W. Liu, N. Yan, Quasi-norm a priori and a posteriori error estimates for the nonconforming approximation of p-Laplacian, Numer. Math. 89 (2) (2001) 341–378.
[29] W. Liu, N. Yan, Quasi-norm local error estimators for p-Laplacian, SIAM J. Numer. Anal. 39 (1) (2001) 100–127.
[30] W. Liu, N. Yan, On quasi-norm interpolation error estimation and a posteriori error estimates for p-Laplacian, SIAM J. Numer. Anal. 40 (5) (2002) 1870–1895.
[31] R. Luce, B.I. Wohlmuth, A local a posteriori error estimator based on equilibrated fluxes, SIAM J. Numer. Anal. 42 (4) (2004) 1394–1414.
[32] A. Veeser, Convergent adaptive finite elements for the nonlinear Laplacian, Numer. Math. 92 (4) (2002) 743–770.
[33] R. Verfürth, A posteriori error estimates for nonlinear problems. Finite element discretizations of elliptic equations, Math. Comp. 62 (206) (1994) 445–475.
[34] R. Verfürth, A Review of a Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques, Teubner-Wiley, Stuttgart, 1996.
[35] R. Verfürth, A posteriori error estimates for nonlinear problems. L^r-estimates for finite element discretizations of elliptic equations, RAIRO Modél. Math. Anal. Numér. 32 (7) (1998) 817–842.
[36] R. Verfürth, A note on polynomial approximation in Sobolev spaces, M2AN Math. Model. Numer. Anal. 33 (4) (1999) 715–719.
[37] R. Verfürth, On the constants in some inverse inequalities for finite element functions, Technical Report, Ruhr-Universität Bochum, 1999.
[38] R. Verfürth, A posteriori error estimates for finite element discretizations of the heat equation, Calcolo 40 (3) (2003) 195–212.
[39] R. Verfürth, A posteriori error estimates for non-linear parabolic equations, Technical Report, Ruhr-Universität Bochum, 2004.
[40] R. Verfürth, Robust a posteriori error estimates for stationary convection-diffusion equations, SIAM J. Numer. Anal. 43 (4) (2005) 1766–1782.
[41] R. Verfürth, A note on constant-free a posteriori error estimates, SIAM J. Numer. Anal. 47 (4) (2009) 3180–3194.
[42] R. Verfürth, Personnal communication, 2009.
[43] M. Vohralík, On the discrete Poincaré–Friedrichs inequalities for nonconforming approximations of the Sobolev space H^1, Numer. Funct. Anal. Optim. 26 (7–8) (2005) 925–952.
[44] M. Vohralík, A posteriori error estimates for lowest-order mixed finite element discretizations of convection–diffusion–reaction equations, SIAM J. Numer. Anal. 45 (4) (2007) 1570–1599.
[45] M. Vohralík, A posteriori error estimation in the conforming finite element method based on its local conservativity and using local minimization, C.R. Math. Acad. Sci. Paris 346 (11–12) (2008) 687–690.
[46] M. Vohralík, Guaranteed and fully robust a posteriori error estimates for conforming discretizations of diffusion problems with discontinuous coefficients, J. Sci. Comput., doi:10.1007/s10915-010-9410-1.

Please cite this article in press as: L. El Alaoui et al., Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems, Comput. Methods Appl. Mech. Engrg. (2010), doi:10.1016/j.cma.2010.03.024