ADAPTIVE INEXACT NEWTON METHODS WITH A POSTERIORI STOPPING CRITERIA

Alexandre Ern
***Martin Vohralík**

Université Paris-Est, CERMICS
Ecole des Ponts ParisTech
77455 Marne la Vallée, France
ern@cermics.enpc.fr

*Université Paris 6
Laboratoire Jacques-Louis Lions
75005 Paris, France
vohralik@ann.jussieu.fr

Abstract

We consider nonlinear algebraic systems arising from numerical discretizations of nonlinear partial differential equations of diffusion type. In order to solve them, some iterative nonlinear solver, and, on each step of this solver, some iterative linear solver are used. We propose adaptive stopping criteria for both these solvers, based on an a posteriori error estimate which distinguishes the different error components, namely the discretization, linearization, and algebraic ones. Our estimates give a guaranteed error upper bound and also a robust error lower bound. Numerical experiments for the nonlinear Laplace equation, nonconforming finite element discretization, Newton linearization, and conjugate gradients algebraic solver illustrate the theory.

Keywords: Nonlinear algebraic system; adaptive linearization, adaptive algebraic solution; stopping criterion; a posteriori error estimate.

Introduction

Consider a system of nonlinear algebraic equations written in the form: find a vector $U \in \mathbb{R}^N$, $N \geq 1$, such that

$$\mathscr{A}(U) = F, \tag{1}$$

where $\mathscr{A} : \mathbb{R}^N \to \mathbb{R}^N$ is a nonlinear operator and $F \in \mathbb{R}^N$ a given vector. We describe in this contribution an adaptive version of the inexact Newton method, cf. [2], for problem (1).

Our method is driven by stopping criteria based on a posteriori error estimators distinguishing three main error components, namely the discretization, linearization, and algebraic ones. On a nonlinear solver step k, $k \geq 1$, and linear solver step i, $i \geq 1$, these estimators are respectively denoted by $\eta_{\mathrm{disc}}^{k,i}$, $\eta_{\mathrm{lin}}^{k,i}$, and $\eta_{\mathrm{alg}}^{k,i}$. A notion of an algebraic remainder estimator $\eta_{\mathrm{rem}}^{k,i}$ also appears. All the estimators are fully computable quantities on each iteration step; their precise form depends on the nonlinear problem, numerical discretization, and nonlinear solver at hand, but is independent of the linear solver. Examples are given below, whereas the detailed forms can be found in [1].

Let γ_{rem}, γ_{alg}, and γ_{lin} be positive user-given weights, typically of order 0.1, related to the maximum percentage part of the given error component in the total error. The algorithm reads:

Algorithm 1 (Adaptive inexact Newton method)**.**

1. *Choose an initial vector $U^0 \in \mathbb{R}^N$. Set $k := 1$.*

2. *From U^{k-1}, define a matrix $\mathbb{A}^k \in \mathbb{R}^{N,N}$ and a vector $F^k \in \mathbb{R}^N$. Consider the following system of linear algebraic equations:*

$$\mathbb{A}^k U^k = F^k. \tag{2}$$

3. (a) *Define $U^{k,0} := U^{k-1}$ and set $i := 1$.*

   (b) *Perform a step of a chosen iterative linear solver for the solution of the linear system* (2)*, starting from the vector $U^{k,i-1}$. This yields an approximation $U^{k,i}$ to U^k which satisfies*

$$\mathbb{A}^k U^{k,i} = F^k - R^{k,i}, \tag{3}$$

   where $R^{k,i} \in \mathbb{R}^N$ is the algebraic residual vector on step i.

   (c) *Perform $\nu > 0$ additional steps of the iterative linear solver yielding an approximation $U^{k,i+\nu}$ to U^k which satisfies*

$$\mathbb{A}^k U^{k,i+\nu} = F^k - R^{k,i+\nu}, \tag{4}$$

   where $R^{k,i+\nu} \in \mathbb{R}^N$ is the algebraic residual vector on step $i+\nu$. The parameter ν is progressively increased until

$$\eta^{k,i}_{\mathrm{rem}} \le \gamma_{\mathrm{rem}} \max\{\eta^{k,i}_{\mathrm{disc}}, \eta^{k,i}_{\mathrm{lin}}, \eta^{k,i}_{\mathrm{alg}}\}. \tag{5}$$

   (d) *Check the convergence criterion for the linear solver in the form*

$$\eta^{k,i}_{\mathrm{alg}} \le \gamma_{\mathrm{alg}} \max\{\eta^{k,i}_{\mathrm{disc}}, \eta^{k,i}_{\mathrm{lin}}\}. \tag{6}$$

   If satisfied, set $U^k := U^{k,i}$. If not, set $i := i+\nu$ and go back to step 3b.

4. *Check the convergence criterion for the nonlinear solver in the form*

$$\eta^{k,i}_{\mathrm{lin}} \le \gamma_{\mathrm{lin}} \eta^{k,i}_{\mathrm{disc}}. \tag{7}$$

   If satisfied, finish. If not, set $k := k+1$ and go back to step 2.

A prominent example of linearization is the Newton one, giving (2) with

$$\mathbb{A}^k_{ij} := \frac{\partial \mathscr{A}_i}{\partial U_j}(U^{k-1}), \qquad F^k := F - \mathscr{A}(U^{k-1}) + \mathbb{A}^k U^{k-1}, \tag{8}$$

but other linearizations like the fixed point one are also allowed. As for the iterative algebraic solver yielding (3), we also do not make any requirement.

1 A Nonlinear Partial Differential Equation and its Numerical Approximation

The nonlinear systems (1) typically arise from some numerical approximation of a nonlinear partial differential equation. Let $\Omega \subset \mathbb{R}^d$, $d \ge 2$, be a polygonal (polyhedral) domain (open, bounded, and connected set). We consider the following model partial differential equation: find $u : \Omega \to \mathbb{R}$ such that

$$-\nabla\cdot\boldsymbol{\sigma}(u,\nabla u) = f \qquad \text{in } \Omega, \tag{9a}$$

$$u = 0 \qquad \text{on } \partial\Omega, \tag{9b}$$

where $\boldsymbol{\sigma} : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ is a nonlinear flux function and $f : \Omega \to \mathbb{R}$ a source term. The scalar-valued unknown function u is termed the *potential*, and, given a potential u, the vector-valued function $-\boldsymbol{\sigma}(u, \nabla u)$ is termed the *flux*. We assume that $f \in L^q(\Omega)$, $q > 1$, and set $p := \frac{q}{q-1}$ so that $\frac{1}{p} + \frac{1}{q} = 1$. The energy space is $V := W_0^{1,p}(\Omega)$, i.e., the space of $L^p(\Omega)$ functions whose weak derivatives are in $L^p(\Omega)$, with the zero trace on $\partial\Omega$.

The exact solution u lies in the space V. Let $u_h^{k,i}$ be a numerical approximation on a mesh $\mathcal{T}_h$ of Ω, linearization step $k \geq 1$, and algebraic solver step $i \geq 1$, corresponding to the algebraic vector $U^{k,i}$ of (3). We suppose $u_h^{k,i} \in V(\mathcal{T}_h)$, where

$$V(\mathcal{T}_h) := \{v \in L^p(\Omega),\, v|_K \in W^{1,p}(K) \quad \forall K \in \mathcal{T}_h\}. \tag{10}$$

Remark that $V(\mathcal{T}_h) \not\subset V$, so that $u_h^{k,i}$ can be nonconforming. Let $\mathcal{E}_K$ regroup the faces e of an element $K \in \mathcal{T}_h$, denote by h_e the diameter of the face e, and by $[\![\cdot]\!]$ the jump operator, yielding the difference of (the traces of) the argument from the two mesh elements that share e on interfaces and the actual trace if e is a boundary face. The error between the exact solution u of (9) and the approximate solution $u_h^{k,i}$ is measured as

$$\mathscr{J}_u(u_h^{k,i}) := \mathscr{J}_{u,\mathrm{F}}(u_h^{k,i}) + \mathscr{J}_{u,\mathrm{NC}}(u_h^{k,i}), \tag{11}$$

where

$$\mathscr{J}_{u,\mathrm{F}}(u_h^{k,i}) := \sup_{\varphi \in V;\, \|\nabla\varphi\|_p = 1} \big(\boldsymbol{\sigma}(u, \nabla u) - \boldsymbol{\sigma}(u_h^{k,i}, \nabla u_h^{k,i}), \nabla\varphi\big), \tag{12a}$$

$$\mathscr{J}_{u,\mathrm{NC}}(u_h^{k,i}) := \left\{ \sum_{K \in \mathcal{T}_h} \sum_{e \in \mathcal{E}_K} h_e^{1-q} \|[\![u - u_h^{k,i}]\!]\|_{q,e}^q \right\}^{\frac{1}{q}}. \tag{12b}$$

The quantity $\mathscr{J}_{u,\mathrm{F}}(u_h^{k,i})$ measures the error in the approximation of the exact flux $-\boldsymbol{\sigma}(u, \nabla u)$ by the approximate one $-\boldsymbol{\sigma}(u_h^{k,i}, \nabla u_h^{k,i})$ and represents the dual norm of the residual of (9); $\mathscr{J}_{u,\mathrm{NC}}(u_h^{k,i})$ then measures the nonconformity of the discrete potential $u_h^{k,i}$, i.e., the departure of $u_h^{k,i}$ from the space V. Therein $\|\cdot\|_r$ stands for the Lebesgue norm in L^r. Importantly, there holds $\mathscr{J}_u(u_h^{k,i}) = 0$ if and only if $u_h^{k,i} = u$. The error measure $\mathscr{J}_u(u_h^{k,i})$ is not easily computable for a known exact solution u; the Hölder inequality, however, gives

$$\mathscr{J}_u(u_h^{k,i}) \leq \mathscr{J}_u^{\mathrm{up}}(u_h^{k,i}) := \|\boldsymbol{\sigma}(u, \nabla u) - \boldsymbol{\sigma}(u_h^{k,i}, \nabla u_h^{k,i})\|_q + \mathscr{J}_{u,\mathrm{NC}}(u_h^{k,i}), \tag{13}$$

which is simple to evaluate, being based on the $[L^q(\Omega)]^d$-difference of the fluxes, plus the nonconformity term. Our numerical experiments indicate that both error measures $\mathscr{J}_u(u_h^{k,i})$ and $\mathscr{J}_u^{\mathrm{up}}(u_h^{k,i})$ exhibit a very close behavior.

2 A posteriori Error Estimates and their Efficiency

Recall the estimators $\eta_{\mathrm{disc}}^{k,i}$, $\eta_{\mathrm{lin}}^{k,i}$, $\eta_{\mathrm{alg}}^{k,i}$, and $\eta_{\mathrm{rem}}^{k,i}$ introduced in the Introduction. These are quantities that are fully computable from $u_h^{k,i}$. Their precise forms for the model problem (9), various numerical discretizations, and various linearizations are given in [1]. Let moreover $\eta_{\mathrm{quad}}^{k,i}$ be a quadrature and $\eta_{\mathrm{osc}}^{k,i}$ a data oscillation estimator. The following theorem has been shown in [1]:

Theorem 1 (A posteriori error estimate distinguishing the different error components)**.** *Let $u \in V$ solve* (9) *and let $u_h^{k,i} \in V(\mathscr{T}_h)$. Then*

$$\mathscr{J}_u(u_h^{k,i}) \le \eta_{\mathrm{disc}}^{k,i} + \eta_{\mathrm{lin}}^{k,i} + \eta_{\mathrm{alg}}^{k,i} + \eta_{\mathrm{rem}}^{k,i} + \eta_{\mathrm{quad}}^{k,i} + \eta_{\mathrm{osc}}^{k,i}. \tag{14}$$

Theorem 1 gives an overall error control on each step k of the linearization and i of the algebraic solver. This control is tight in the sense of the following result, shown in [1]:

Theorem 2 (Global efficiency and robustness)**.** *Let $u \in V$ solve* (9) *and let $u_h^{k,i} \in V(\mathscr{T}_h)$. Let the global stopping and balancing criteria* (5), (6), (7) *be satisfied. Then there exists a generic constant C independent of the mesh size h, the domain Ω, the nonlinear function $\boldsymbol{\sigma}$, and the Lebesgue exponent q such that*

$$\eta_{\mathrm{disc}}^{k,i} + \eta_{\mathrm{lin}}^{k,i} + \eta_{\mathrm{alg}}^{k,i} + \eta_{\mathrm{rem}}^{k,i} \le C(\mathscr{J}_u(u_h^{k,i}) + \eta_{\mathrm{quad}}^{k,i} + \eta_{\mathrm{osc}}^{k,i}). \tag{15}$$

3 Numerical Illustration

This section gives a quick numerical illustration of the theoretical developments. We consider (9) with $\boldsymbol{\sigma}(u,\nabla u) = |\nabla u|^{p-2}\nabla u$, which is the so-called p-Laplacian. Here we only consider the value $p = 10$. We take $\Omega := (0,1)\times(0,1)$, $f := 2$, and prescribe an inhomogeneous Dirichlet boundary condition by the exact solution

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

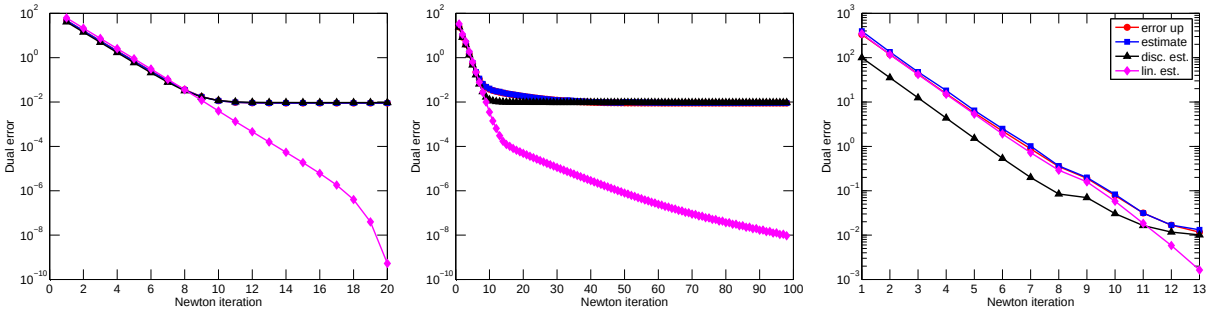
We employ the Crouzeix–Raviart nonconforming finite element method for the discretization, the Newton linearization (8), and the conjugate gradient algebraic solver with a diagonal preconditioning.

We compare three different stopping criteria in Algorithm 1, leading to three different solution approaches:

- In the *Full Newton (FN) method*, both the nonlinear and linear solvers are iterated to "almost" convergence, with the global stopping criteria $\eta_{\mathrm{alg}}^{k,i} \le 10^{-8}$ and $\eta_{\mathrm{lin}}^{k,i} \le 10^{-8}$. The balancing criterion (5) is employed with $\gamma_{\mathrm{rem}} = 0.1$.
- In the *Inexact Newton (IN) method*, the only difference with FN is that a fixed number of preconditioned CG iterations is performed on each Newton linearization step. These values were chosen respectively as 2, 3, 5, 8, 10, 15 on each level of the uniform mesh refinement.
- Finally, in the *Adaptive Inexact Newton (AIN) method* that we propose, we rely on the global stopping criteria (5), (6), and (7) with $\gamma_{\mathrm{rem}} = \gamma_{\mathrm{alg}} = \gamma_{\mathrm{lin}} = 0.3$.

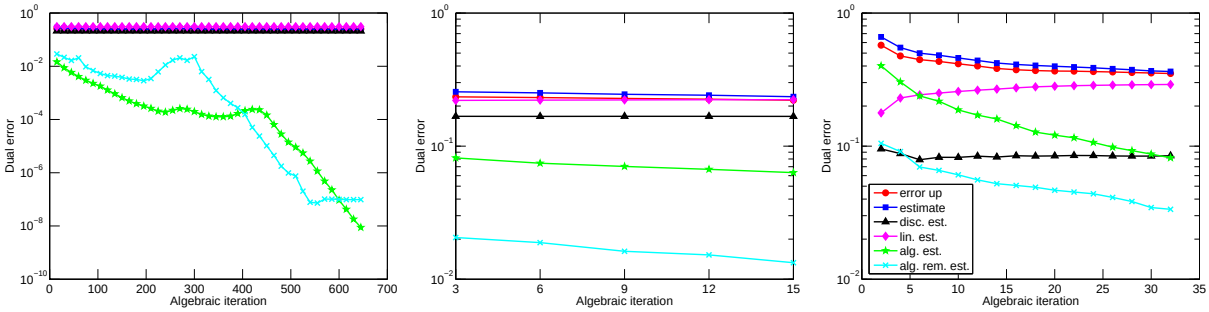
Figure 1 focuses on the 6th level uniformly refined mesh and tracks the dependence of the error measure $\mathscr{J}_u^{\mathrm{up}}(u_h^{k,i})$, the overall error estimator, and the discretization and linearization estimators $\eta_{\mathrm{disc}}^{k,i}$ and $\eta_{\mathrm{lin}}^{k,i}$ of Theorem 1 on the Newton iterations. Typically, the error and all the estimators except $\eta_{\mathrm{lin}}^{k,i}$ start to stagnate after the linearization error ceases to dominate. This is precisely the point where the nonlinear iteration is stopped in AIN using (7), whereas both FN and IN perform many unnecessary additional iterations.

Figure 2 further analyzes the situation on one chosen Newton iteration from Figure 1. To be in a region with similar error measure $\mathscr{J}_u^{\mathrm{up}}(u_h^{k,i})$, we have chosen the 6th iteration for FN and



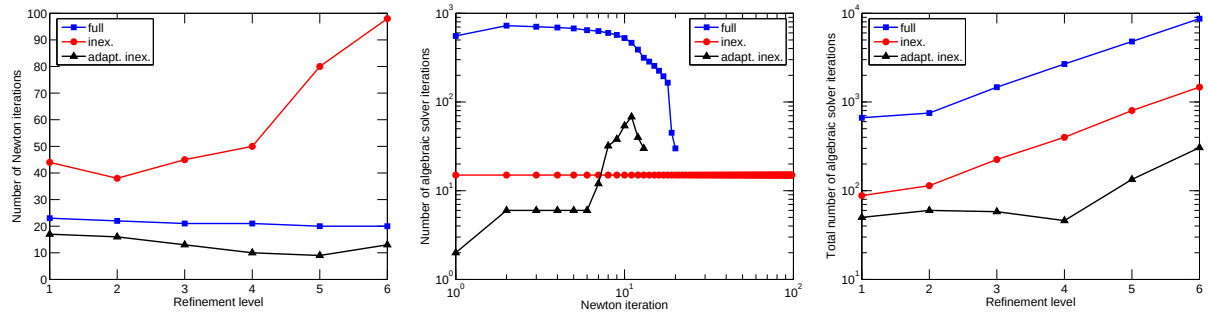
Source: Own

Fig. 1. *Error and estimators as a function of Newton iterations, 6th level mesh. Newton (left), inexact Newton (middle), and adaptive inexact Newton (right)*



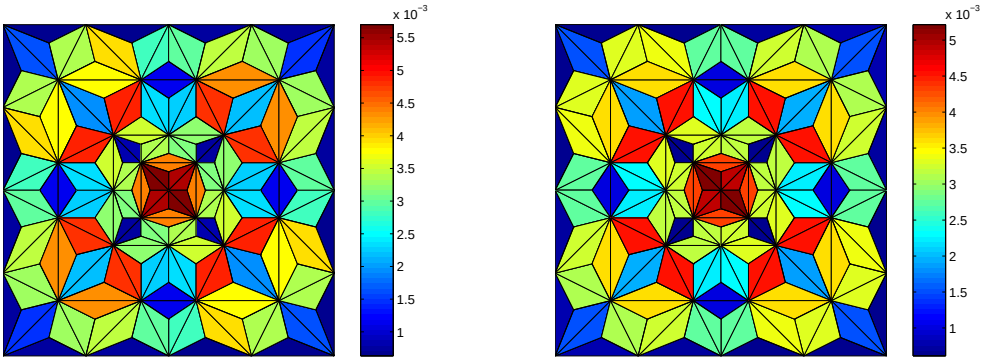
Source: Own

Fig. 2. *Error and estimators as a function of preconditioned CG iterations, 6th level mesh. Newton, 6th step (left), inexact Newton, 6th step (middle), and adaptive inexact Newton, 8th step (right)*



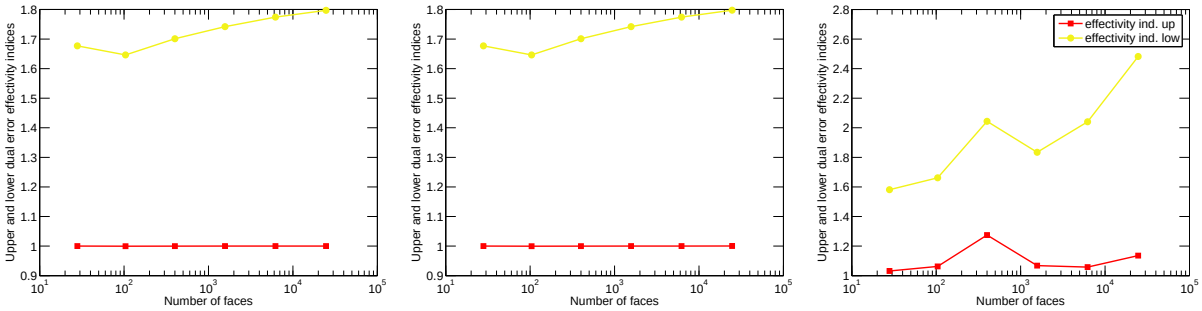
Source: Own

Fig. 3. *Number of Newton iterations per refinement level (left), number of linear solver iterations per Newton step on 6th level mesh (middle), and total number of linear solver iterations per refinement level (right)*



Source: Own

Fig. 4. *Estimated (left) and actual (right) error distribution, 2nd level uniformly refined mesh, adaptive inexact Newton*



Source: Own

Fig. 5. *Upper and lower effectivity indices. Newton (left), inexact Newton (middle), and adaptive inexact Newton (right)*

IN and the 8th iteration for AIN. We see that almost no decrease of the error measure $\mathscr{J}_u^{\mathrm{up}}(u_h^{k,i})$ can be observed during the almost 650 iterations of the preconditioned CG method in the FN case. The fixed 15 CG iterations in the IN case are, on the contrary, not sufficient to decrease significantly the error. In our approach, just the sufficient, "online-decided" number of CG iterations is performed and timely stopped using (6).

Figure 3 illustrates the overall performance of three approaches. We can see that the number of Newton iterations per refinement level is stable around 20 for FN. It increases significantly for IN, whereas it is still reduced for AIN. On one Newton iteration, the number of CG iterations also varies significantly among three approaches. Many iterations are necessary in the FN case and fixed 15 iterations in the IN case, whereas AIN picks up the number that is "just necessary." Remark that this number is equal to two on the first Newton step; from here, the error is "lagged" as a function of Newton iterations in the AIN case, cf. Figure 1. The total number of necessary CG iterations per refinement level is displayed in the right part of Figure 3. On the last mesh, AIN only needs 306 total iterations, whereas IN needs 1470, and FN 8690 iterations. Thus, our approach yields an economy by a factor of roughly 5 with respect to IN and roughly 30 with respect to FN in terms of total iterations.

Figure 4 displays the distribution of the overall error estimator and of the error measure $\mathscr{J}_u^{\mathrm{up}}(u_h^{k,i})$ on the 2nd level uniformly refined mesh for AIN. We see that even in presence of algebraic and linearization errors, the overall error distribution is very well predicted. Finally,

let $\mathscr{J}_u^{\mathrm{low}}(u_h^{k,i})$ be an easily computable lower bound for the error measure $\mathscr{J}_u(u_h^{k,i})$ obtained by estimating the supremum in (12a) just with one φ. We define the upper and lower effectivity indices respectively as $\mathscr{I}^{\mathrm{up}} := \eta^{k,i} / \mathscr{J}_u^{\mathrm{up}}(u_h^{k,i})$ and $\mathscr{I}^{\mathrm{low}} := \eta^{k,i} / \mathscr{J}_u^{\mathrm{low}}(u_h^{k,i})$ and observe that the effectivity index for the original error measure $\mathscr{J}_u(u_h^{k,i})$, defined as $\mathscr{I} := \eta / \mathscr{J}_u(u_h^{k,i})$, lies between $\mathscr{I}^{\mathrm{up}}$ and $\mathscr{I}^{\mathrm{low}}$. For the three methods (FN, IN, and AIN), all these indices are close to the optimal value of 1 and in particular $\mathscr{I}^{\mathrm{up}}$ takes values very close to 1, see Figure 5.

Conclusion

In this work, we have presented an inexact Newton method with a posteriori error control and adaptive stopping criteria. Numerical experiments illustrate that tight error bound and important computational gains can be achieved by our approach. Details on all the presented developments can be found in [1].

Acknowledgments

This work was partly supported by the Groupement MoMaS (PACEN/CNRS, ANDRA, BRGM, CEA, EdF, IRSN) and by the ERT project “Enhanced oil recovery and geological sequestration of CO_2: mesh adaptivity, a posteriori error control, and other advanced techniques” (LJLL/IFPEN).

Literature

[1] ERN, A.; VOHRALÍK, M.: Adaptive inexact Newton methods with a posteriori stopping criteria for nonlinear diffusion PDEs. *HAL Preprint 00681422v2*, submitted for publication, 2012.

[2] QUARTERONI, A.; VALLI, A.: *Numerical approximation of partial differential equations*. Springer-Verlag, Berlin, 1994.

prof. Alexandre Ern, Ph.D., Ing. Martin Vohralík, Ph.D.

Adaptivní nepřesné Newtonovy metody s a posteriorními zastavovacími kritérii

V této práci uvažujeme systémy nelineárních algebraických rovnic vznikající při numerické diskretizaci nelineárních parciálních diferenciálních rovnic difúzního typu. K jejich (přibližnému) řešení uvažujeme nelineární iterační metodu a, na každém jejím kroku, iterační řešič systému lineárních algebraických rovnic. Navrhujeme adaptivní uzpůsobení počtu kroků obou iteračních řešičů. Obě zastavovací kritéria jsou založena na a posteriorních odhadech, které rozlišují různé složky celkové chyby, v daném případě algebraickou chybu, linearizační chybu a diskretizační chybu. Naše a posteriorní odhady poskytují zaručenou horní hranici na celkovou chybu mezi přibližným a přesným řešením a zároveň robustnostní hranici spodní. Numerické experimenty pro nelineární Laplaceovu rovnici, nekonformní metodu konečných prvků, Newtonovu linearizaci a metodu sdružených gradientů pro řešení soustav lineárních algebraických rovnic ilustrují teoretické výsledky.

Die adaptive ungenaue Newton-Methode mit A-posteriori-Stopp-Kriterien

In dieser Arbeit betrachten wir die Systeme nichtlinearer algebraischer Gleichungen, die bei der numerischen Diskretisation nichtlinearer partieller Differenzialgleichungen entstehen. Zu deren - annäherungsweisen - Lösung betrachten wir die nicht lineare iterative Methode und - auf jedem ihrer Schritte - den iterativen Löser des Systems der linearen algebraischen Gleichungen. Wir schlagen eine adaptive Angleichung der Schrittzahl beider iterativer Löser vor. Beide Stopp-Kriterien gründen sich auf A-posteriori-Schätzungen, welche verschiedene Bestandteile des Gesamtfehlers unterscheiden, im vorliegenden Fall einen algebraischen Fehler, einen linearisierenden Fehler und einen Diskreditisationsfehler. Unsere A-posteriori-Schätzungen gewährleisten eine garantierte Obergrenze für den Gesamtfehler zwischen der Näherungs- und der genauen Lösung und gleichzeitig eine Robustitätsgrenze nach unten. Die numerischen Experimente für die nichtlineare Laplace-Gleichung, die nichtkonforme Methode der Endelemente, die Newton'sche Linearisation sowie die Methode der dualen Gradienten zur Lösung der Systeme linearer algebraischer Gleichungen illustrieren die theoretischen Ergebnisse.

Adaptacyjne niedokładne metody Newtona z kryteriami stopującymi a posteriori

W niniejszym opracowaniu przedstawiono systemy nieliniowych równań algebraicznych powstające podczas numerycznej dyskretyzacji nieliniowych parcjalnych równań różniczkowych o charakterze dyfuzyjnym. Do ich (przybliżonego) rozwiązania zastosowano nieliniową metodę iteracji a na każdym jej etapie iteracyjny sposób rozwiązania układu liniowych równań algebraicznych. Zaproponowano adaptacyjne dostosowanie liczby etapów obu sposobów. Oba kryteria stopujące oparte są na szacunkach a posteriori, które odróżniają różne elementy ogólnego błędu, w tym przypadku algebraiczne, linearyzacyjne i dyskretyzacyjne. Nasze oraz a posteriori szacunki zapewniają górną granicę ogólnego błędu pomiędzy przybliżonym a dokładnym rozwiązaniem oraz dolną stałą granicę błędu. Eksperymenty numeryczne dla nieliniowego równania Laplace'a, niekonformiczną metodę elementów skończonych, linearyzację Newtona oraz metodę gradientów sprzężonych do rozwiązywania układów liniowych równań algebraicznych przedstawiono w postaci teoretycznych wyników.