



HAL
open science

Stopping criteria for nonlinear variational problems: iterative approach

Jules Djoko Kamdem, Toni Sayah

► **To cite this version:**

Jules Djoko Kamdem, Toni Sayah. Stopping criteria for nonlinear variational problems: iterative approach. Numerical Algorithms, 2024, 10.1007/s11075-024-01969-8 . hal-04802267

HAL Id: hal-04802267

<https://hal.science/hal-04802267v1>

Submitted on 25 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Stopping criteria for nonlinear variational problems: iterative approach

Jules Djoko Kamdem¹ · Toni Sayah²

Received: 20 December 2023 / Accepted: 21 October 2024

© The Author(s) 2024

Abstract

In this paper, we present and study a stopping criteria based on a priori error estimate for nonlinear variational problems. We show that for nonlinear variational problems, the a priori error estimate is divided into two sub errors namely: the discretization error and the linearization error. We then used that representation to devise a stopping criteria that help to reduce the computational time. The methodology is applied to three problems and the numerical results demonstrate the superiority of the new approach.

Keywords Nonlinear variational problems · Finite element method · Iterative scheme · A priori error estimates · Linearization error · Discretization error · Stopping criterion

Mathematics Subject Classification (2010) 65N30

1 Introduction

The a priori error analysis controls the overall discretization error of a problem by providing upper bounds that tend to zero with the mesh size. The upper bound of the error is not computable because depends most of the time of the exact solution of the problem. The present work investigates a priori error estimates of the finite element approximation of nonlinear problems. The goal of this analysis is to prove that for nonlinear variational problems, the overall error is decomposed into two parts; the

✉ Jules Djoko Kamdem
jules.djokokamdem@gmail.com

Toni Sayah
toni.sayah@usj.edu.lb

¹ School of Mathematical and Statistical Sciences, North West University, Potchefstroom, South Africa

² Laboratoire de Mathematiques et Applications, Unite de recherche “Mathematiques et Modelisation”, Faculté des Sciences, Université Saint-Joseph de Beyrouth, B.P 11-514 Riad El Solh, Beirut 1107 2050, Lebanon

discretization error and the linearization error. The discretization error is a result of the finite element approximation, while the linearization error is a consequence of solving the nonlinear finite element equations by iterative or incremental approach. From the theory, this approach is more appropriate than the traditional one where only the discretization error is measured (see [1–5] just to cite a few).

The idea of the stopping criteria taking into account the iterative solution was initially introduced and developed by in [6, 7] in the context of the *a posteriori* error estimates. In their works, the error between the exact and the approximated solutions is bounded by some local quantities called "indicators" which is divided in two parts namely; linearization and discretization indicators. It is shown that a good management of this two sources of error leads to a drastic reduction of the number of iterations. We note that this approach of the *a posteriori* error estimates combined with the stopping criteria based on the linearization and discretization indicators has proved to be very effective/competitive for mesh adaptation method, and it has been used in several works (see for instance [8–13]). In this work, our focus is on *a priori* error estimates, and we would like to see if the approach proposed in [6, 7] to reduce the iterations is applicable. We note that this exercise has not yet been explored in the literature. The task is not straightforward because the *a priori* analysis is a global analysis compared to the *a posteriori* analysis which is local analysis where the study had been performed. The *a priori* estimates depend only on the exact solution, but not the numerical solution and hence can be (in theory if not in practice) evaluated before computing the numerical solution. Conversely, *a posteriori* estimates depend on the numerical solution but not the exact solution, so they do require computing the solution but can actually be evaluated in practice. Hence from the theoretical analysis, there is a vast difference. From a mathematical point of view it is natural when dealing with a nonlinear problem to regard the source of the error as twofold: the discretization and linearization. In this work, we show that the idea proposed in [6, 7] can help to reduce the number of iterations for the nonlinear variational problems we will be dealing with. It is apparent that to reduce the computational time, we have to balance these two sources of error. Indeed, if the discretization error is more important than the linearization error, then we have less iterations. We illustrate the methodology by treating; a Nonlinear model problem in Section 2. In Section 3, we analyse the systems of equations of Darcy and Forchheimer, while in Section 4, we treat the Navier-Stokes system. Our approach of proof relies on the triangle's inequality to separate the discretization error to the linearization error, and then each error is estimated by exploiting the variational formulation. Concluding remarks and some perspectives are gathered in Section 5.

2 Nonlinear problem

Continuous problem We considered the following nonlinear problem:

$$-\Delta u + \lambda|u|^{2p}u = f \quad \text{in } \Omega, \quad \text{and } u = 0 \quad \partial\Omega, \quad (2.1)$$

where u is the unknown, λ and p are two positive real numbers, and $f \in H^{-1}(\Omega)$, the dual of the Sobolev space $H_0^1(\Omega)$. The variational model associated with (2.1) reads:

$$\begin{cases} \text{Find } u \in \mathbb{V} = H_0^1(\Omega) \text{ such that} \\ \text{for all } v \in \mathbb{V} \text{ , } \int_{\Omega} \nabla u \cdot \nabla v \, dx + \int_{\Omega} \lambda |u|^{2p} uv \, dx = \langle f, v \rangle . \end{cases} \tag{2.2}$$

We readily proved that the variational problem (2.2) is equivalent to the following minimization problem

$$\begin{cases} \text{Find } u \in \mathbb{V} = H_0^1(\Omega) \text{ such that} \\ \text{for all } v \in \mathbb{V} \text{ , } J(u) \leq J(v) \\ J(v) = \frac{1}{2} \int_{\Omega} |\nabla v|^2 \, dx + \frac{\lambda}{2p+2} \int_{\Omega} v^{2p+2} \, dx - \langle f, v \rangle . \end{cases} \tag{2.3}$$

Thus the existence theory of (2.2) is obtained via (2.3) by using the direct method of calculus of variations (see for instance [10, Theorem 2.3]).

The following well known relations will be used: for all $(a, b, p) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}$, with $p \geq 1$,

$$\begin{aligned} \left| |a|^p - |b|^p \right| &\leq p|a - b| \left(|a|^{p-1} + |b|^{p-1} \right) , \\ \left(|a|^{2p} a - |b|^{2p} b \right) (a - b) &\geq 0 \text{ .} \end{aligned} \tag{2.4}$$

2.1 Finite element approximation

For the discretization, we consider a regular family of triangulations $(\mathcal{T}_h)_h$ of Ω , as described in Ciarlet [5], which is a set of closed non-degenerate triangles for $d = 2$ or tetrahedra for $d = 3$, called elements, satisfying,

- for each h , $\overline{\Omega}$ is the union of all elements of \mathcal{T}_h ;
- the intersection of two distinct elements of \mathcal{T}_h is either empty, a common vertex, or an entire common edge (or face when $d = 3$);
- the ratio of the diameter h_K of an element $K \in \mathcal{T}_h$ to the diameter ρ_K of its inscribed circle when $d = 2$ or ball when $d = 3$ is bounded by a constant independent of h : there exists a positive constant σ independent of h such that,

$$\max_{K \in \mathcal{T}_h} \frac{h_K}{\rho_K} \leq \sigma .$$

As usual, h denotes the maximal diameter of all elements of \mathcal{T}_h . To define the finite element functions, let r be a non-negative integer. For each K in \mathcal{T}_h , we denote $\mathbb{P}_r(K)$ the space of polynomials in d variables, restricted to K , with a total degree at most r .

We introduce the following finite element space:

$$\mathbb{V}_h = \left\{ v_h \in \mathbb{V} \cap \mathcal{C}(\overline{\Omega}) : \forall K \in \mathcal{T}_h, \quad v_h|_K \in \mathbb{P}_1(K) \right\} .$$

The finite element approximation associated to (2.2) reads:

$$\begin{cases} \text{Find } u_h \in \mathbb{V}_h \text{ such that} \\ \text{for all } v_h \in \mathbb{V}_h, \int_{\Omega} \nabla u_h \cdot \nabla v_h \, dx + \int_{\Omega} \lambda |u_h|^{2p} u_h v_h \, dx = \langle f, v_h \rangle. \end{cases} \quad (2.5)$$

The unique existence solvability of (2.5) can be obtained by invoking the same direct method of calculus of variations. The similarity is drawn with [10, Theorem 2.3]. We state that

Proposition 2.1 [10] Let u be the solution of (2.2), and u_h the solution of (2.5). Assume that $u \in H^2(\Omega)$, there exists a real number $h_0 > 0$ such that for all $h \leq h_0$, the following *a priori* error estimate holds:

$$\|u - u_h\|_1 \leq Ch \|u\|_{2,\Omega}.$$

The nonlinear variational problem (2.5) can be solved by a fixed point strategy which is described next.

2.2 Iterative scheme: fixed point algorithm

Given $u_h^0 \in \mathbb{V}_h$, for $i \geq 0$, knowing $u_h^i \in \mathbb{V}_h$, we compute $u_h^{i+1} \in \mathbb{V}_h$ solution of

$$\text{for all } v_h \in \mathbb{V}_h, \int_{\Omega} \nabla u_h^{i+1} \cdot \nabla v_h \, dx + \int_{\Omega} \lambda |u_h^i|^{2p} u_h^{i+1} v_h \, dx = \langle f, v_h \rangle. \quad (2.6)$$

We readily checked by using Lax-Milgram theorem that (2.6) has a unique solution $u_h^i \in \mathbb{V}_h$ which satisfies the bound

$$\|\nabla u_h^i\| \leq \|f\|_{-1,\Omega}. \quad (2.7)$$

Concerning the convergence of the algorithm (2.6), we state that

Proposition 2.2 [10, Theorem 4.1] Let u_h^i be the solution of (2.6), and u_h the solution of (2.5). There exists a positive constant $C(f, \lambda, p)$ such that

$$\|u_h^{i+1} - u_h\|_1 \leq C(f, \lambda, p) \|u_h^i - u_h\|_1.$$

If we take f and λ such that $C(f, \lambda, p) < 1$ (smallness condition on the data), then $u_h^i \xrightarrow{i \rightarrow \infty} u_h$ in $H^1(\Omega)$.

Remark 2.1 The triangle inequality

$$\|u_h^{i+1} - u_h^i\|_1 \leq \|u_h^{i+1} - u_h\|_1 + \|u_h - u_h^i\|_1$$

and Proposition 2.2 implies that $\|u_h^{i+1} - u_h^i\|_1 \xrightarrow{i \rightarrow \infty} 0$.

The main result of this paragraph is the following

Theorem 2.1 *Let u be the solution of (2.2), with $u \in H^2(\Omega)$, and u_h^i given by (2.6). Then there exists a positive constant c independent of h and i , and such that*

$$\|u - u_h^i\|_1 \leq c(\|u_h^i - u_h^{i+1}\|_1 + h\|u\|_{2,\Omega}).$$

Proof We quickly observe that

$$\|u - u_h^{i+1}\|_1 \leq \|u - u_h\|_1 + \|u_h - u_h^{i+1}\|_1$$

with $\|u_h - u\|_1$ given by Proposition 2.1. Thus it remains to estimate $\|u_h - u_h^{i+1}\|_1$. The difference between (2.5) and (2.6) gives

$$\text{for all } v_h \in \mathbb{V}_h, \int_{\Omega} \nabla(u_h^{i+1} - u_h) \cdot \nabla v_h \, dx + \lambda \int_{\Omega} (|u_h^i|^{2p} u_h^{i+1} - |u_h|^{2p} u_h) v_h \, dx = 0. \tag{2.8}$$

We take $v_h = u_h - u_h^{i+1}$ in (2.8), we use (2.4), Cauchy-Schwarz’s inequality, Hölder’s inequality, Sobolev’s inequality, and (2.7) to obtain

$$\begin{aligned} \|\nabla(u_h^{i+1} - u_h)\|^2 + \lambda \int_{\Omega} \underbrace{(|u_h^{i+1}|^{2p} u_h^{i+1} - |u_h|^{2p} u_h)(u_h^{i+1} - u_h)}_{\geq 0} \, dx \\ = \lambda \int_{\Omega} (|u_h^i|^{2p} - |u_h^{i+1}|^{2p}) u_h^{i+1} (u_h - u_h^{i+1}) \, dx \\ \leq 2p\lambda \int_{\Omega} |u_h^{i+1}| (|u_h^{i+1}|^{2p-1} + |u_h^i|^{2p-1}) |u_h^i - u_h^{i+1}| |u_h - u_h^{i+1}| \, dx \\ \leq c \|\nabla u_h^{i+1}\| \| |u_h^{i+1}|^{2p-1} + |u_h^i|^{2p-1} \|_{L^8} \|\nabla(u_h^i - u_h^{i+1})\| \|\nabla(u_h - u_h^{i+1})\| \\ \leq c_1 \|\nabla(u_h^i - u_h^{i+1})\| \|\nabla(u_h - u_h^{i+1})\| \end{aligned}$$

from which we deduce that

$$\|\nabla(u_h^{i+1} - u_h)\| \leq c_1 \|\nabla(u_h^i - u_h^{i+1})\|$$

which ends the proof. □

Remark 2.2 If

$$\|u_h^i - u_h^{i-1}\|_1 \leq c_1 h,$$

where c_1 is a positive constant which balances the discretization error and the linearization error, then

$$\|u - u_h^i\|_1 \leq ch.$$

If

$$\|u_h^{i-1} - u_h^i\|_1 \geq c_1 h$$

then

$$\|u_h^i - u\|_1 \leq c \|u_h^{i-1} - u_h^i\|_1 \longrightarrow 0 \text{ geometrically when } i \rightarrow \infty.$$

2.3 Numerical experiments

In this paragraph we evaluate the performances of the method by introducing a new stopping condition. In fact, we assess the convergence properties of the method by performing a standard mesh refinement analysis. The numerical simulations in 2d are performed using Freefem++ (see [14]). The relative iterative error due to the linearization of the nonlinear problem (2.1) is

$$Er_i^1 = \frac{\|u_h^i - u_h^{i+1}\|_1}{\|u_h^{i+1}\|_1}.$$

We introduce the following conditions to stop the computations

$$\text{Classical one : } Er_i^1 < 10^{-5}, \tag{2.9}$$

$$\text{New one : } Er_i^1 < \gamma h, \tag{2.10}$$

where the parameter $\gamma \in (0, 1)$. In fact, for the new stopping criterion we consider $Er_i^1 < \gamma h$ instead of $Er_i^1 < \gamma \frac{h}{\|u_h^{i+1}\|_1}$ to be coherent with the relative error given by the classical one. Evidently, this new stopping criterion is efficient when $\gamma h \leq 10^{-5}$, otherwise we must use the classical stopping criterion.

We start the iterative scheme (2.6) by taking $u_h^0 = \mathbf{0}$, for the numerical investigations, we consider the domain $(x, y) \in \Omega = (0, 1)^2$ where each edge is divided into N equal segments. Then Ω is divided into N^2 equal squares and equivalently into $2N^2$ triangles. we consider $p = 2, \lambda = 1$ and the following exact solution:

$$u = 30xy(x - 1)(y - 1)(x^2 + y^2)e^{xy}.$$

So f is given by the first equation of (2.1).

First of all, we will study the convergence of Algorithm (2.6) with respect to the parameter γ in order to get a good precision with a fast algorithm. We introduce the following relative error

$$Ee_i^1 = \frac{\|u - u_h^i\|_1}{\|u\|_1}$$

where u_h^i is computed after the convergence of the iterative scheme (2.6) using one of the above stopping criteria. We denote by N_{iter} the number of iterations of Algorithm (2.6) to reach the desired stopping criterion.

We consider $N = 50$ and use the classical stopping criterion (2.10). Algorithm (2.6) gives $Ee_i^1 = 0.0524$ after $N_{iter} = 139$ iterations. Furthermore, Table 1 shows the error Ee_i^1 with respect to γ (new stopping criterion (2.10)). We deduce that, for this particular case and for $\gamma \leq 1.5$ we have the same error obtained with the classical stopping criterion (2.9), whereas for $\gamma = 1.5$ the number of iterations is $N_{iter} = 50$. Hence the advantage of the method proposed in this paper.

Table 1 2D non-linear model: Ee_i^1 and N_{iter} with respect to γ (New stopping criterion (2.10)) for $N = 50$

γ	0.1	1	1.3	1.5	1.7	2	5	10
Ee_i^1	0.0524	0.0524	0.0524	0.0524	0.0526	0.0527	0.0540	0.0576
N_{iter}	87	56	52	50	49	47	35	26

To complete this case, we show in Fig. 1 comparisons of the *a priori* error estimates between the classical criterion (2.9) and the new one (2.10) for $\gamma = 1.5$. The presented curves coincide and the corresponding slope is equal to 0.99 which is close to the theoretical one equal to 1.

Let us now compare our stopping criteria (2.10) with that introduced in [6, 7] corresponding to the *a posteriori* error estimates. We begin by recalling the indicators introduced in [10] and summarizing the corresponding error between the exact solution of Problem (2.2) and the iterative solution of Problem (2.6):

$$|u_h^{i+1} - u|_{1,\Omega} \leq C \left(\left(\sum_{K \in \mathcal{T}_h} ((\eta_{K,i}^{(D)})^2 + h_K^2 \|f - f_h\|_{L^2(K)}^2) \right)^{1/2} + \left(\sum_{K \in \mathcal{T}_h} (\eta_{K,i}^{(L)})^2 \right)^{1/2} \right),$$

where f_h is an approximation of the data f and,

$$\begin{aligned} \eta_{K,i}^{(L)} &= |u_h^{i+1} - u_h^i|_{1,K}, \\ (\eta_{K,i}^{(D)})^2 &= h_K^2 \|f_h + \Delta u_h^{i+1} - \lambda |u_h^i|^{2p} u_h^{i+1}\|_{L^2(K)}^2 + \frac{1}{2} \sum_{E \in \mathcal{E}_{h,\Omega}} h_E \left\| \left[\frac{\partial u_h^{i+1}}{\partial n} \right] \right\|_{L^2(E)}^2. \end{aligned}$$

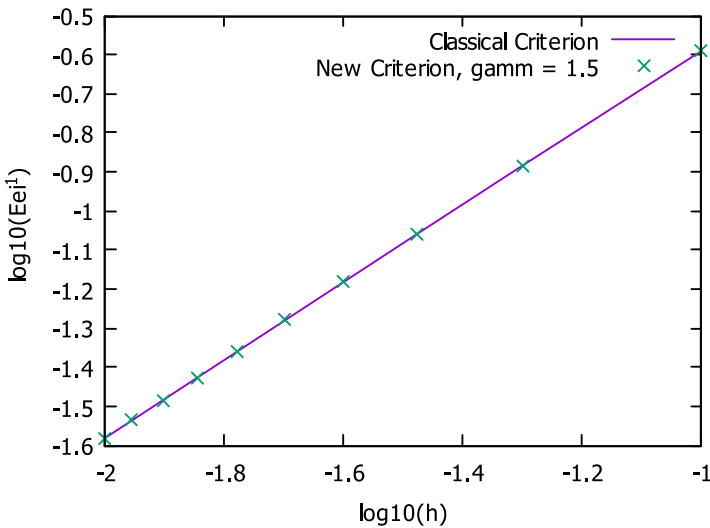


Fig. 1 2D non-linear model: A comparison of the *a priori* error between the classical criterion (2.9) and the new criterion (2.10) for $\gamma = 1.5$

The *a posteriori* stopping criterion is given by

$$\eta_i^{(L)} \leq \gamma_p \eta_i^{(D)}, \tag{2.11}$$

where γ_p is a given real parameter and

$$\eta_i^{(D)} = \left(\sum_{K \in \mathcal{T}_h} (\eta_{K,i}^{(D)})^2 \right)^{1/2} \quad \text{and} \quad \eta_i^{(L)} = \left(\sum_{K \in \mathcal{T}_h} (\eta_{K,i}^{(L)})^2 \right)^{1/2}.$$

Table 2 shows the error Ee_i^1 with respect to γ_p for $N = 50$ and for the *a posteriori* stopping criteria (2.11). We deduce that for $\gamma_p \leq 0.02$, we have the same error as that obtained by the classical stopping criteria (2.9), where for $\gamma_p = 0.02$ the corresponding number of iterations ($N_{iter} = 50$) is close to that obtained with the stopping criterion (2.10) for $\gamma = 1.5$. Thus, both stopping criteria (2.10) (for $\gamma = 1.5$) and (2.11) (for $\gamma_p = 0.02$) require similar number of iterations.

To complete this test, we show in Fig. 2 the comparison of the *a priori* error estimates obtained with the classical stopping criterion (2.9) and with the *a posteriori* stopping criterion (2.11) (for $\gamma_p = 0.02$). The slopes of the corresponding lines are equal to 0.99 and close to the theoretical one equals to 1.

We conclude that both stopping criteria (2.10) and (2.11) gives similar results with the following remarks:

- (2.10) requires an *a priori* error estimate while (2.11) is based on the *a posteriori* error estimate and needs to compute the indicators at each iteration of the iterative algorithm.
- (2.11) is very useful for the adaptive mesh method while (2.10) is restricted to the stopping criteria condition for the iterative algorithm.

3 Darcy-Forchheimer model

Formulation of the problem The Darcy-Forchheimer model is described by the following system:

$$\begin{cases} \frac{\mu}{\rho} K^{-1} \mathbf{u} + \frac{\beta}{\rho} |\mathbf{u}| \mathbf{u} + \nabla p = \mathbf{f}, & \text{in } \Omega \\ \operatorname{div} \mathbf{u} = 0, & \text{in } \Omega \\ \mathbf{u} \cdot \mathbf{n} = 0 & \text{on } \partial\Omega. \end{cases} \tag{3.1}$$

Table 2 2D non-linear model: Ee_i^1 and N_{iter} with respect to γ_p by using the *a posteriori* stopping criterion (2.11) for $N = 50$

γ_p [6, 7]	0.001	0.01	0.02	0.03	0.04	0.05	0.07	0.1
Ee_i^1	0.0524	0.0524	0.0524	0.0527	0.0531	0.0536	0.0545	0.0561
N_{iter}	88	58	51	44	41	37	33	29

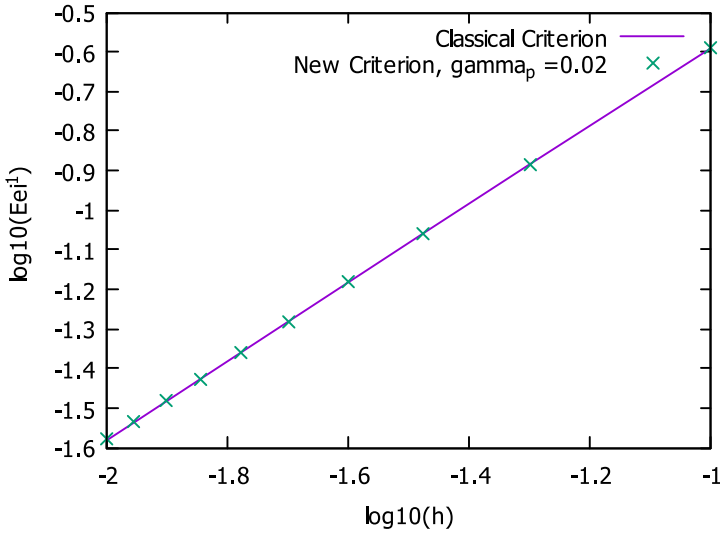


Fig. 2 2D non-linear model: A comparison of the *a priori* error between the classical stopping criterion (2.9) and the *a posteriori* stopping criterion (2.11) for $\gamma_p = 0.02$

In (3.1), $|\cdot|$ is the Euclidean vector norm $|\mathbf{u}|^2 = \mathbf{u} \cdot \mathbf{u}$, K is the permeability tensor assumed to be uniformly positive definite and bounded and Ω is a bounded open set in \mathbb{R}^d ($d=2,3$) with a Lipschitz-continuous boundary $\partial\Omega$. The data $\mathbf{f} : \Omega \rightarrow \mathbb{R}^d$ is the external body force per unit volume acting on the fluid. The unknown are the velocity \mathbf{u} the pressure and p . The parameter ρ is the density of the fluid, μ is its viscosity and β represent the Forchheimer number of the porous media, all assumed to be positive constants.

The weak formulation reads (see [15–18]):

$$\left\{ \begin{array}{l} \text{Find } (\mathbf{u}, p) \in L^3(\Omega)^d \times M, \text{ such that} \\ \text{and for all } (\mathbf{v}, q) \in L^3(\Omega)^d \times M, \\ a(\mathbf{u}, \mathbf{v}) + \frac{\beta}{\rho} \int_{\Omega} |\mathbf{u}| \mathbf{u} \cdot \mathbf{v} dx + b(\mathbf{v}, p) = \int_{\Omega} \mathbf{f} \cdot \mathbf{v} dx, \\ b(\mathbf{u}, q) = 0, \end{array} \right. \quad (3.2)$$

where

$$M = \left\{ q(x) \in W^{1,3/2}(\Omega); \int_{\Omega} q(x) dx = 0 \right\},$$

and

$$a(\mathbf{u}, \mathbf{v}) = \frac{\mu}{\rho} \int_{\Omega} K^{-1} \mathbf{u} \cdot \mathbf{v} dx, \quad b(\mathbf{v}, q) = \int_{\Omega} \nabla q \cdot \mathbf{v} dx.$$

The variational problem (3.2) is nonlinear of monotone type. Thus its existence is shown with the help monotone operator theory (see for instance [15, 18]).

3.1 Finite element approximation

The finite element approximation of (3.2) is standard, and keeping the notation used before, we approximate the velocity, and the pressure respectively in the spaces \mathbb{V}_h , M_h defined as follows

$$\begin{aligned} \mathbb{V}_h &= \left\{ \mathbf{v}_h \mid \mathbf{v}_h \in L^3(\Omega)^d, \text{ for all } K \in \mathcal{T}_h, \mathbf{v}_h|_K \in \mathbb{P}_0(K)^d \right\}, \\ M_h &= \left\{ q_h \mid q_h \in M \cap C^0(\overline{\Omega}), \text{ for all } K \in \mathcal{T}_h, q_h|_K \in \mathbb{P}_1(K) \right\}. \end{aligned}$$

With these spaces, we approximate (3.2) by the following finite element scheme

$$\left\{ \begin{aligned} &\text{Find } (\mathbf{u}_h, p_h) \in \mathbb{V}_h \times M_h, \text{ such that} \\ &\text{for all } (\mathbf{v}_h, q_h) \in \mathbb{V}_h \times M_h, \\ &a(\mathbf{u}_h, \mathbf{v}_h) + \frac{\beta}{\rho} \int_{\Omega} |\mathbf{u}_h| \mathbf{u}_h \cdot \mathbf{v}_h dx + b(\mathbf{v}_h, p_h) = \int_{\Omega} \mathbf{f} \cdot \mathbf{v}_h dx, \\ &b(\mathbf{u}_h, q_h) = 0. \end{aligned} \right. \tag{3.3}$$

For the existence and uniqueness of the solution of (3.3), we can refer to [18] where the authors have addressed a similar problem but using non-conforming finite element approach. The convergence of the finite element solution of (3.3) is well known and we refer the interested readers to [12], where the following is derived

Theorem 3.1 *Let (\mathbf{u}, p) the solution of (3.2), and (\mathbf{u}_h, p_h) the solution of (3.3). If $(\mathbf{u}, p) \in W^{1,4}(\Omega) \times H^2(\Omega)$, Then the following a priori errors holds*

$$\begin{aligned} \|p - p_h\|_{W^{1,3/2}(\Omega)^d} + \|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)^d} &\leq c_1 h \\ \|\mathbf{u} - \mathbf{u}_h\|_{L^3(\Omega)^d} &\leq c_2 h^{2/3}. \end{aligned}$$

Iterative scheme Knowing \mathbf{u}_h^0 and for $i \geq 1$, we compute $(\mathbf{u}_h^i, p_h^i) \in \mathbb{V}_h \times M_h$ as solution of

$$\left\{ \begin{aligned} &\text{for all } (\mathbf{v}_h, q_h) \in \mathbb{V}_h \times M_h, \\ &a(\mathbf{u}_h^i, \mathbf{v}_h) + \frac{\beta}{\rho} \int_{\Omega} |\mathbf{u}_h^{i-1}| \mathbf{u}_h^i \cdot \mathbf{v}_h dx + b(\mathbf{v}_h, p_h^i) = \int_{\Omega} \mathbf{f} \cdot \mathbf{v}_h dx, \\ &b(\mathbf{u}_h^i, q_h) = 0. \end{aligned} \right. \tag{3.4}$$

The linear system of equations (3.4) is well posed, and we have

$$\|\mathbf{u}_h^i\| + \|\mathbf{u}_h^i\|_{L^3(\Omega)^d} + \|\nabla p_h^i\|_{L^{3/2}(\Omega)^d} \leq c, \tag{3.5}$$

where c is a positive constant depending on \mathbf{f} . The main contribution of this paragraph is the following:

Theorem 3.2 *Let (\mathbf{u}, p) be the solution of (3.2) and (\mathbf{u}_h^i, p_h^i) given by (3.4). If $(\mathbf{u}, p) \in W^{1,4}(\Omega) \times H^2(\Omega)$, then there exists a positive constant c independent of h and i such that*

$$\|p - p_h^i\|_{W^{1,3/2}(\Omega)^d} + \|\mathbf{u} - \mathbf{u}_h^i\|_{L^3(\Omega)^d} \leq c_1 \|\mathbf{u}_h^i - \mathbf{u}_h^{i-1}\|_{L^3}^{2/3} + c_2 h^{2/3}.$$

Proof The triangle inequality gives

$$\begin{aligned} \|\mathbf{u} - \mathbf{u}_h^i\|_{L^3} &\leq \|\mathbf{u} - \mathbf{u}_h\|_{L^3} + \|\mathbf{u}_h - \mathbf{u}_h^i\|_{L^3} \\ \|p - p_h^i\|_{W^{1,3/2}} &\leq \|p - p_h\|_{W^{1,3/2}} + \|p_h - p_h^i\|_{W^{1,3/2}}, \end{aligned}$$

where $\|\mathbf{u}_h - \mathbf{u}\|_{L^3}$ and $\|p_h - p\|_{W^{1,3/2}}$ are given by theorem 3.1. Thus it remains to estimate $\|\mathbf{u}_h - \mathbf{u}_h^i\|_{L^3}$, and $\|p_h - p_h^i\|_{W^{1,3/2}}$.

The difference between (3.4) and (3.3) gives

$$\begin{cases} \text{for all } (\mathbf{v}_h, q_h) \in \mathbb{V}_h \times M_h, \\ a(\mathbf{u}_h^i - \mathbf{u}_h, \mathbf{v}_h) + \frac{\beta}{\rho} \int_{\Omega} (|\mathbf{u}_h^{i-1}| \mathbf{u}_h^i - |\mathbf{u}_h| \mathbf{u}_h) \cdot \mathbf{v}_h dx + b(\mathbf{v}_h, p_h^i - p_h) = 0, \\ b(\mathbf{u}_h^i - \mathbf{u}_h, q_h) = 0. \end{cases} \tag{3.6}$$

We take $\mathbf{v}_h = \mathbf{u}_h^i - \mathbf{u}_h$ and $q_h = p_h^i - p_h$ in (3.6), and obtain the following

$$\begin{aligned} &a(\mathbf{u}_h^i - \mathbf{u}_h, \mathbf{u}_h^i - \mathbf{u}_h) + \frac{\beta}{\rho} \int_{\Omega} (|\mathbf{u}_h^i| \mathbf{u}_h^i - |\mathbf{u}_h| \mathbf{u}_h) \cdot (\mathbf{u}_h^i - \mathbf{u}_h) dx \\ &= \frac{\beta}{\rho} \int_{\Omega} (|\mathbf{u}_h^i| \mathbf{u}_h^i - |\mathbf{u}_h^{i-1}| \mathbf{u}_h^i) \cdot (\mathbf{u}_h^i - \mathbf{u}_h) dx \\ &= \frac{\beta}{\rho} \int_{\Omega} (|\mathbf{u}_h^i| - |\mathbf{u}_h^{i-1}|) \mathbf{u}_h^i \cdot (\mathbf{u}_h^i - \mathbf{u}_h) dx \\ &\leq \frac{\beta}{\rho} \int_{\Omega} |\mathbf{u}_h^i - \mathbf{u}_h^{i-1}| |\mathbf{u}_h^i| |\mathbf{u}_h^i - \mathbf{u}_h| dx \\ &\leq \frac{\beta}{\rho} \|\mathbf{u}_h^i - \mathbf{u}_h^{i-1}\|_{L^3} \|\mathbf{u}_h^i\|_{L^3} \|\mathbf{u}_h^i - \mathbf{u}_h\|_{L^3}. \end{aligned} \tag{3.7}$$

We recall that (see [15])

$$\begin{aligned} &\text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \text{ there exists } c \text{ independent of } \mathbf{x}, \mathbf{y} \text{ such that} \\ &\text{for } s > 2, \quad c |\mathbf{y} - \mathbf{x}|^s \leq (|\mathbf{x}|^{s-2} \mathbf{x} - |\mathbf{y}|^{s-2} \mathbf{y}, \mathbf{x} - \mathbf{y}) \\ &\text{for } s \geq 2, \quad \left| |\mathbf{x}|^{s-2} \mathbf{x} - |\mathbf{y}|^{s-2} \mathbf{y} \right| \leq c |\mathbf{y} - \mathbf{x}| (|\mathbf{x}| + |\mathbf{y}|)^{s-2}. \end{aligned} \tag{3.8}$$

The left hand side of (3.7) combined with the first relation in (3.8) gives

$$C \|\mathbf{u}_h^i - \mathbf{u}_h\|^2 + c \|\mathbf{u}_h^i - \mathbf{u}_h\|_{L^3}^3 \leq a(\mathbf{u}_h^i - \mathbf{u}_h, \mathbf{u}_h^i - \mathbf{u}_h) + \int_{\Omega} (|\mathbf{u}_h^i| \mathbf{u}_h^i - |\mathbf{u}_h| \mathbf{u}_h) \cdot (\mathbf{u}_h^i - \mathbf{u}_h) dx$$

which together with (3.7), (3.5) and Young’s equality yields

$$\|\mathbf{u}_h^i - \mathbf{u}_h\|^2 + \|\mathbf{u}_h^i - \mathbf{u}_h\|_{L^3}^3 \leq C \|\mathbf{u}_h^i - \mathbf{u}_h^{i-1}\|_{L^3}^2. \tag{3.9}$$

Next, $p_h - p_h^i \in M_h$ and (M_h, \mathbb{V}_h) is inf-sup stable. Hence from (3.6) there exists $\gamma > 0$ such that

$$\begin{aligned} & \gamma \|p_h - p_h^i\|_{W^{1,3/2}} \\ & \leq \sup_{\mathbf{v} \in \mathbb{V}_h} \frac{b(\mathbf{v}, p_h - p_h^i)}{\|\mathbf{v}\|_{L^3}} \\ & \leq \sup_{\mathbf{v} \in \mathbb{V}_h} \frac{a(\mathbf{u}_h^i - \mathbf{u}_h, \mathbf{v}_h) + \frac{\beta}{\rho} \int_{\Omega} (|\mathbf{u}_h^{i-1}| |\mathbf{u}_h^i - |\mathbf{u}_h^i| |\mathbf{u}_h^i|) \cdot \mathbf{v}_h dx + \frac{\beta}{\rho} \int_{\Omega} (|\mathbf{u}_h^i| |\mathbf{u}_h^i - |\mathbf{u}_h| |\mathbf{u}_h|) \cdot \mathbf{v}_h dx}{\|\mathbf{v}\|_{L^3}} \\ & \leq C \|\mathbf{u}_h^i - \mathbf{u}_h\| + \|\mathbf{u}_h^{i-1} - \mathbf{u}_h^i\|_{L^3} \|\mathbf{u}_h^i\|_{L^3} + C \|\mathbf{u}_h^i - \mathbf{u}_h\|_{L^3} (\|\mathbf{u}_h^i\|_{L^3} + \|\mathbf{u}_h\|_{L^3}) \\ & \leq C \|\mathbf{u}_h^i - \mathbf{u}_h\| + \|\mathbf{u}_h^{i-1} - \mathbf{u}_h^i\|_{L^3} + C \|\mathbf{u}_h^i - \mathbf{u}_h\|_{L^3}, \end{aligned}$$

which together with (3.9) implies that

$$\|p_h - p_h^i\|_{W^{1,3/2}} \leq C \|\mathbf{u}_h^i - \mathbf{u}_h^{i-1}\|_{L^3} + \|\mathbf{u}_h^{i-1} - \mathbf{u}_h^i\|_{L^3} + C \|\mathbf{u}_h^i - \mathbf{u}_h^{i-1}\|_{L^3}^{2/3}. \tag{3.10}$$

(3.9) and (3.10) complete the proof. □

Remark 3.1 Adopting the similar approach as the proof of Theorem 3.2, and provide a condition on the data, we can show that $\lim_{i \rightarrow \infty} \|\mathbf{u}_h^i - \mathbf{u}_h^{i-1}\|_{L^3} = 0$.

3.2 Numerical experiments

For the numerical simulations we use the iterative scheme (3.4). The error due to linearization is

$$Er_i^2 = \frac{\|\mathbf{u}_h^i - \mathbf{u}_h^{i+1}\|_{L^3}}{\|\mathbf{u}_h^{i+1}\|_{L^3}}$$

and the stopping criterion will be one of the following tests:

1. Classical one: $Er_i^2 < 10^{-5}$,
2. New one: $Er_i^2 < \gamma h$.

Here also we mention that if $\gamma h \geq 10^{-5}$, we use simply the classical stopping criterion. We observe that both stopping criterion gives the following error:

$$\|p - p_h^i\|_{W^{1,3/2}(\Omega)^d} + \|\mathbf{u} - \mathbf{u}_h^i\|_{L^3(\Omega)^d} \leq c_1 h^{2/3}.$$

3.2.1 2D numerical simulations

We consider in this case $\Omega = (0, 1)^2$ and the following exact solution:

$$\psi(x, y) = x^2 y^2 (x - 1)^2 (y - 1)^2 \exp(xy),$$

$$\mathbf{u}(x, y) = (u_1, u_2) = c_u \left(\frac{\partial \psi}{\partial y}, -\frac{\partial \psi}{\partial x} \right),$$

$$p(x, y) = c_p xy(x - 2./3.)(y - 2./3.).$$

Then \mathbf{f} can be computed by using the first equation of (3.1) where $c_u = 50, c_p = 10, \beta = 5, \rho = 1, \mu = 1$.

We begin by studying the convergence of the iterative system (3.4) with respect to γ . We introduce the following relative error

$$Ee_i^2 = \frac{\|\mathbf{u} - \mathbf{u}_h^i\|_{L^3}}{\|\mathbf{u}\|_{L^3}} + \frac{\|p - p_h^i\|_{W^{1,3/2}}}{\|p\|_{W^{1,3/2}}}$$

where (\mathbf{u}_h^i, p_h^i) is computed after the convergence of the iterative scheme (3.4) using one of the above stopping criteria.

By taking $N = 50$ (the number of segments by edge) and using the classical stopping criterion, we get $N_{iter} = 54$ and $Ee_i^2 = 0.049$. Furthermore, Table 3 shows the error Ee_i^2 with respect to γ . We deduce that in this case and for $\gamma = 1.5$, we have (as in the previous case) the same error obtained with the classical stopping criterion $Ee_i^2 = 0.049$ and for $\gamma = 1.5$ the number of iterations is $N_{iter} = 19$. Hence, we also observe the advantage of the method proposed in this paper.

We give in Fig. 3 comparisons of the *a priori* error estimates between the classical stopping criterion and the new one for $\gamma = 1.5$. The slope of the curves is 0.95 which agrees (even better) with the theoretical slope $2/3$.

3.2.2 3D numerical simulations

In this part we consider $\Omega = (0, 1)^3$ and the following exact solution:

$$\psi(x, y) = x^2y^2z^2(x - 1)^2(y - 1)^2(z - 1)^2\exp(xyz),$$

$$\mathbf{u}(x, y) = (u_1, u_2, u_3) = c_u \text{curl}(\psi, \psi, \psi),$$

$$p(x, y) = c_p xyz(x - 2./3.)(y - 2./3.)(z - 2./3.).$$

Here also the mesh is such that every edge of the cube is divided into N segments, so the mesh contains $6N^3$ elements. We begin by studying the iterative convergence of (3.4) with respect to γ where $c_u = 10, c_p = 10, \beta = 10, \rho = 1, \mu = 1$.

By considering the classical stopping criterion and taking $N = 20$, the iterative scheme (3.4) converges after 20 iterations with an error $Ee_i^2 = 0.314$. Otherwise, by taking the new stopping criterion, Table 4 shows the error Ee_i^2 and the number of iterations N_{iter} with respect to γ . Here in the 3D case, $\gamma = 1.5$ seems to be the

Table 3 2D Darcy-Forchheimer model: Ee_i^2 and N_{iter} with respect to γ (New stopping criterion) for $N = 50$

γ	0.001	0.01	0.05	0.1	0.5	1	1.5	2	5
Ee_i^2	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.051	0.054
N_{iter}	49	39	33	30	23	20	19	17	14

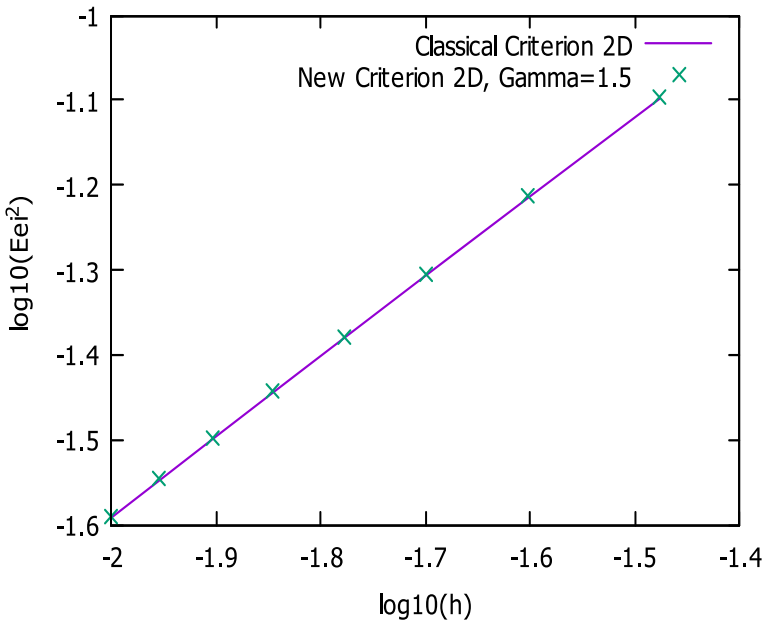


Fig. 3 2D Darcy-Forchheimer model: A comparison of the *a priori* error between the classical criterion and the new one for $\gamma = 1.5$

suitable value giving the same error obtained with the classical stopping criterion but with $N_{iter} = 6$ and showing the advantage of the new stopping criterion proposed in this paper.

Figure 4 shows comparisons of the *a priori* error estimates between the classical criterion and the new one for $\gamma = 1.5$. The slope of the curves is 0.71 which agrees (even better) with the theoretical slope $2/3$.

4 Navier-stokes equations

We consider for a positive constant ν , the following system:

$$\begin{cases} \nu \Delta \mathbf{u} + [\mathbf{u} \cdot \nabla] \mathbf{u} + \nabla p = \mathbf{f}, & \text{in } \Omega \\ \operatorname{div} \mathbf{u} = 0, & \text{in } \Omega \\ \mathbf{u} = 0 & \text{on } \partial \Omega, \end{cases} \quad (4.1)$$

Table 4 3D Darcy-Forchheimer model: Ee_i^2 and N_{iter} with respect to γ (New stopping criterion) for $N = 20$

γ	0.001	0.01	0.05	0.1	0.5	1	1.5	2	5
Ee_i^2	0.314	0.314	0.314	0.314	0.314	0.314	0.314	0.313	0.308
N_{iter}	18	14	11	10	8	7	6	6	4

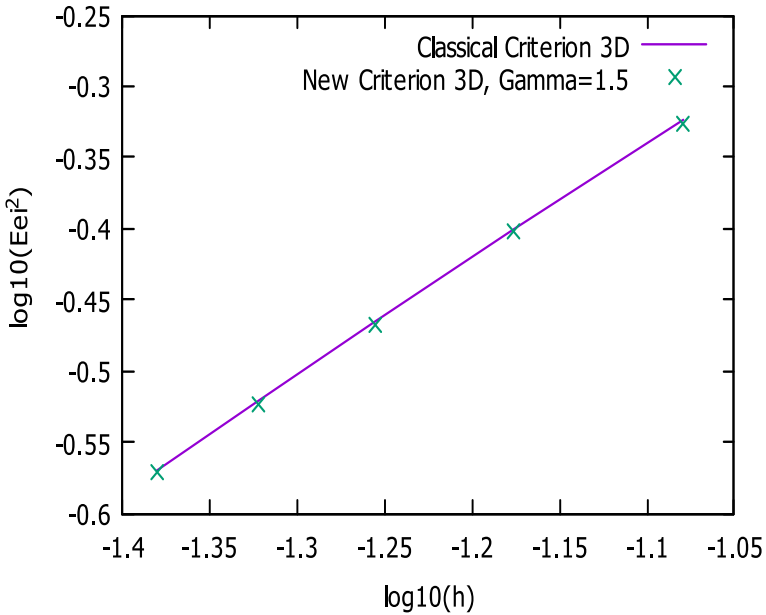


Fig. 4 3D textbfDarcy-Forchheimer model: A comparison of the *a priori* error between the classical criterion and the new one for $\gamma = 1.5$

where the unknowns are the velocity \mathbf{u} and the pressure p of the fluid. \mathbf{f} is an external forces acting on the fluid, assumed to be in $L^2(\Omega)^d$. The variational formulation of (4.1) reads;

$$\begin{cases} \text{Find } (\mathbf{u}, p) \in \mathbb{V} \times M, \text{ such that for all } (\mathbf{v}, q) \in \mathbb{V} \times M, \\ a(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) + c(\mathbf{u}, \mathbf{u}, \mathbf{v}) = \ell(\mathbf{v}), \\ b(\mathbf{u}, q) = 0, \end{cases} \tag{4.2}$$

with

$$\mathbb{V} = \left\{ \mathbf{v} \in H^1(\Omega)^d : \mathbf{v}|_{\Gamma} = 0 \right\},$$

$$M = L_0^2 = \left\{ p \in L^2(\Omega) : \int_{\Omega} p dx = 0 \right\}.$$

and

$$a(\mathbf{v}, \mathbf{u}) = \nu \int_{\Omega} \nabla \mathbf{u} : \nabla \mathbf{v} dx, \quad b(\mathbf{v}, q) = - \int_{\Omega} \text{div } \mathbf{v} q dx,$$

$$c(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \int_{\Omega} [\mathbf{u} \cdot \nabla] \mathbf{v} \cdot \mathbf{w} dx, \quad \ell(\mathbf{v}) = \int_{\Omega} \mathbf{f} \cdot \mathbf{v} dx.$$

The nonlinear boundary value problem (4.2) has been studied by several authors, and under smallness assumption on data (see [3, 4, 19]). In fact it is shown that there exists a positive constant C such that if

$$v^2 > C \|f\|, \tag{4.3}$$

then (4.2) has a unique solution $(u, p) \in \mathbb{V} \times M$ which moreover satisfies

$$\|u\|_1 + \|p\| \leq C \|f\|.$$

4.1 Finite element approximation

The finite element approximation of (4.2) is standard, and keeping the notation used before, we introduce the subspaces \mathbb{V}_h and M_h given such that the following inf-sup condition holds: there exists $\beta > 0$ (independent of h) such that

$$\text{for all } q_h \in M_h, \quad \beta \|q_h\| \leq \sup_{v \in \mathbb{V}_h} \frac{b(v, q_h)}{\|v\|_1}. \tag{4.4}$$

With these spaces, we approximate (4.2) by the following finite element scheme

$$\left\{ \begin{array}{l} \text{Find } (u_h, p_h) \in \mathbb{V}_h \times M_h, \text{ such that for all } (v_h, q_h) \in \mathbb{V}_h \times M_h, \\ a(u_h, v_h) + b(v_h, p_h) + \tilde{c}(u_h, u_h, v_h) = \ell(v_h), \\ b(u_h, q_h) = 0, \end{array} \right. \tag{4.5}$$

with

$$\tilde{c}(u_h, v_h, w_h) = c(u_h, v_h, w_h) + \frac{1}{2} \int_{\Omega} (\operatorname{div} u_h) v_h \cdot w_h dx.$$

As usual, the trilinear form $\tilde{c}(\cdot, \cdot, \cdot)$ is used to compensate for the fact that $\operatorname{div} u_h$ may not be equal zero. Of course, using Green's formula and the functions regularity, we get

$$\tilde{c}(u_h, v_h, w_h) = \frac{1}{2} \left(\int_{\Omega} (u_h \cdot \nabla) v_h \cdot w_h dx - \int_{\Omega} (u_h \cdot \nabla) w_h \cdot v_h dx \right).$$

Thus the trilinear form $\tilde{c}(\cdot, \cdot, \cdot)$ is anti-symmetric. The discrete problem (4.5) has at least one solution $(u_h, p_h) \in \mathbb{V}_h \times M_h$ (see [3, 4, 19]). Moreover under (4.3), the solution is unique, and satisfies the bound

$$\|u_h\|_1 + \|p_h\| \leq \frac{C}{\beta} \|f\|,$$

and the following error estimate holds

$$\|u - u_h\|_1 + \|p - p_h\| \leq C \left(\inf_{v_h \in \mathbb{V}_h} \|u - v_h\|_1 + \inf_{q_h \in M_h} \|q_h - p\| \right) \tag{4.6}$$

Remark 4.1 Following [3, 4], we can take one of this tow discretization:

1. Taylor-Hood elements: for some $k \geq 2$ we set

$$\begin{aligned} \mathbb{V}_h &= \left\{ \mathbf{v}_h \in \mathbb{V} \cap \mathcal{C}(\Omega)^d, \text{ for all } K \in \mathcal{T}_h, \mathbf{v}_h|_K \in \mathbb{P}_k(K)^d \right\}, \\ M_h &= \left\{ q_h \in L^2(\Omega) \cap \mathcal{C}(\Omega) : \text{ for all } K \in \mathcal{T}_h, q_h|_K \in \mathbb{P}_{k-1}(K) \right\}. \end{aligned}$$

Then the inf-sup condition (4.4) is satisfied, moreover for all $\mathbf{v} \in H^s(\Omega)^d, 1 \leq s \leq k + 1$, and for all $q \in H^t(\Omega), 0 \leq t \leq k$ we have

$$\inf_{\mathbf{v}_h \in \mathbb{V}_h} \|\mathbf{v} - \mathbf{v}_h\|_1 \leq Ch^{s-1} \|\mathbf{v}\|_s, \quad \inf_{q_h \in M_h} \|q_h - q\| \leq Ch^t \|q\|_t.$$

Thus if the solution of (4.2) satisfies $\mathbf{u} \in H^{k+1}(\Omega)^d$ and $p \in H^k(\Omega)$, (4.6) gives

$$\|\mathbf{u} - \mathbf{u}_h\|_1 + \|p - p_h\| \leq Ch^k \left(\|\mathbf{u}\|_{k+1} + \|p\|_k \right).$$

2. Mini-element:

$$\begin{aligned} \mathbb{V}_h &= \left\{ \mathbf{v}_h \in \mathbb{V} \cap \mathcal{C}(\Omega)^d, \text{ for all } K \in \mathcal{T}_h, \mathbf{v}_h|_K \in \mathcal{P}_1(K)^d \right\}, \\ M_h &= \left\{ q_h \in L^2(\Omega) \cap \mathcal{C}(\Omega) : \text{ for all } K \in \mathcal{T}_h, q_h|_K \in \mathbb{P}_1(K) \right\}. \end{aligned} \tag{4.7}$$

where $\mathcal{P}_1(K) = [\mathbb{P}_1(K) \oplus \text{span}\{\lambda_1, \dots, \lambda_d\}]^d$. We have

$$\|\mathbf{u} - \mathbf{u}_h\|_1 + \|p - p_h\| \leq Ch \left(\|\mathbf{u}\|_2 + \|p\|_1 \right).$$

Iterative algorithm Let \mathbf{u}_h^0 be an initial guess, We take, for $i \geq 0$, the following scheme:

$$\left\{ \begin{aligned} &\text{Find } (\mathbf{u}_h^{i+1}, p_h^{i+1}) \in \mathbb{V}_h \times M_h, \text{ such that for all } (\mathbf{v}_h, q_h) \in \mathbb{V}_h \times M_h \\ &a(\mathbf{u}_h^{i+1}, \mathbf{v}_h) + b(\mathbf{v}_h, p_h^{i+1}) + \tilde{c}(\mathbf{u}_h^i, \mathbf{u}_h^{i+1}, \mathbf{v}_h) = \ell(\mathbf{v}_h) \\ &b(\mathbf{u}_h^{i+1}, q_h) = 0. \end{aligned} \right. \tag{4.8}$$

Knowing \mathbf{u}_h^i , the scheme (4.8) is a linear variational problem which is theoretically solved using Lax-Milgram’s theorem. Taking the difference between (4.5) and (4.8) for $\mathbf{v}_h = \mathbf{u}_h^{i+1} - \mathbf{u}_h$, one gets

$$\begin{aligned} a(\mathbf{u}_h^{i+1} - \mathbf{u}_h, \mathbf{u}_h^{i+1} - \mathbf{u}_h) &= \tilde{c}(\mathbf{u}_h, \mathbf{u}_h, \mathbf{u}_h^{i+1} - \mathbf{u}_h) - \tilde{c}(\mathbf{u}_h^i, \mathbf{u}_h^{i+1}, \mathbf{u}_h^{i+1} - \mathbf{u}_h) \\ &= \tilde{c}(\mathbf{u}_h - \mathbf{u}_h^i, \mathbf{u}_h, \mathbf{u}_h^{i+1} - \mathbf{u}_h) \\ &\leq C \|\nabla(\mathbf{u}_h - \mathbf{u}_h^i)\| \|\nabla \mathbf{u}_h\| \|\nabla(\mathbf{u}_h^{i+1} - \mathbf{u}_h)\| \\ &\leq \frac{C}{\nu} \|\mathbf{f}\| \|\nabla(\mathbf{u}_h - \mathbf{u}_h^i)\| \|\nabla(\mathbf{u}_h^{i+1} - \mathbf{u}_h)\| \end{aligned}$$

which by the coercivity of $a(\cdot, \cdot)$ gives

$$\|\nabla(\mathbf{u}_h^{i+1} - \mathbf{u}_h)\| \leq \frac{C}{\nu^2} \|\mathbf{f}\| \|\nabla(\mathbf{u}_h^i - \mathbf{u}_h)\|. \tag{4.9}$$

From (4.9), we deduce that the algorithm (4.8) converges if (4.3) is verified. We next discuss a consistency notion introduced in [18, Remark 5]. We state that

Lemma 4.1 *The algorithm (4.8) is consistent with (4.5) in the sense that if there is $j \geq 1$ such that the solution (\mathbf{u}_h^j, p_h^j) coincides with the solution (\mathbf{u}_h, p_h) of (4.5), then for all $i > j$, (\mathbf{u}_h^i, p_h^i) coincides with (\mathbf{u}_h, p_h) .*

Proof It is done by induction. We assume that up to i , the solutions of (4.8) and (4.5) are the same. We take the difference between (4.5) and (4.8) and obtain that

$$\begin{cases} \text{for all } (\mathbf{v}, q) \in \mathbb{V}_h \times M_h, \\ a(\mathbf{u}_h^{i+1} - \mathbf{u}_h, \mathbf{v}) + b(\mathbf{v}, p_h^{i+1} - p_h) + \tilde{c}(\mathbf{u}_h, \mathbf{u}_h^{i+1}, \mathbf{v}) - \tilde{c}(\mathbf{u}_h, \mathbf{u}_h, \mathbf{v}) = 0, \\ b(\mathbf{u}_h^{i+1} - \mathbf{u}_h, q) = 0. \end{cases} \tag{4.10}$$

Thus taking $\mathbf{v} = \mathbf{u}_h^{i+1} - \mathbf{u}_h$, and using the linearity of $\tilde{c}(\cdot, \cdot, \cdot)$ with respect to the second component and the anti-symmetry property of $\tilde{c}(\cdot, \cdot, \cdot)$, we obtain

$$\begin{aligned} a(\mathbf{u}_h^{i+1} - \mathbf{u}_h, \mathbf{u}_h^{i+1} - \mathbf{u}_h) &= \tilde{c}(\mathbf{u}_h, \mathbf{u}_h, \mathbf{u}_h^{i+1} - \mathbf{u}_h) - \tilde{c}(\mathbf{u}_h, \mathbf{u}_h^{i+1}, \mathbf{u}_h^{i+1} - \mathbf{u}_h) \\ &= \tilde{c}(\mathbf{u}_h, \mathbf{u}_h - \mathbf{u}_h^{i+1}, \mathbf{u}_h^{i+1} - \mathbf{u}_h) = 0. \end{aligned}$$

We deduce using the coercivity of $a(\cdot, \cdot)$ that $\mathbf{u}_h^{i+1} = \mathbf{u}_h$. Returning to (4.10), one gets $b(\mathbf{v}, p_h^{i+1} - p_h) = 0$. From the inf-sup condition (4.4), one obtains $p_h^{i+1} = p_h$, which ends the proof. \square

The next result present an a priori error estimate and prove optimal convergence for $H^{k+1}(\Omega)^d \times H^k(\Omega)$ regularity of solution.

Theorem 4.1 *Let (\mathbf{u}, p) be the solution of (4.2) and (\mathbf{u}_h^i, p_h^i) given by (4.8). If (4.3) is verified, then there exists C independent of h and i such that*

$$\|\nabla(\mathbf{u} - \mathbf{u}_h^{i+1})\| + \|p - p_h^{i+1}\| \leq C \|\nabla(\mathbf{u}_h^i - \mathbf{u}_h^{i+1})\| + \|\nabla(\mathbf{u} - \mathbf{u}_h)\| + \|p - p_h\|.$$

Proof The triangle inequality gives

$$\begin{aligned} \|\nabla(\mathbf{u} - \mathbf{u}_h^{i+1})\| &\leq \|\nabla(\mathbf{u} - \mathbf{u}_h)\| + \|\nabla(\mathbf{u}_h - \mathbf{u}_h^{i+1})\| \\ \|p - p_h^{i+1}\| &\leq \|p - p_h\| + \|p_h - p_h^{i+1}\|, \end{aligned}$$

with $\|\nabla(\mathbf{u}_h - \mathbf{u})\| + \|p_h - p\|$ is given by (4.6). We next estimate $\|\nabla(\mathbf{u}_h - \mathbf{u}_h^{i+1})\| + \|p_h - p_h^{i+1}\|$. Considering (4.9), and using triangle's inequality, we get

$$\left(1 - \frac{C}{\nu^2} \|\mathbf{f}\|\right) \|\nabla(\mathbf{u}_h^{i+1} - \mathbf{u}_h)\| \leq \frac{C}{\nu^2} \|\mathbf{f}\| \|\nabla(\mathbf{u}_h^{i+1} - \mathbf{u}_h^i)\|. \tag{4.11}$$

The difference between (4.8) and (4.5) gives

$$b(\mathbf{v}_h, p_h - p_h^{i+1}) = a(\mathbf{u}_h^{i+1} - \mathbf{u}_h, \mathbf{v}_h) + \tilde{c}(\mathbf{u}_h^i, \mathbf{u}_h^{i+1}, \mathbf{v}_h) - \tilde{c}(\mathbf{u}_h, \mathbf{u}_h, \mathbf{v}_h).$$

Next, $p_h - p_h^i \in M_h$ and (M_h, \mathbb{V}_h) is inf-sup stable, thus (4.4) applies and

$$\begin{aligned} & \beta \|p_h - p_h^{i+1}\| \\ & \leq \sup_{\mathbf{v}_h \in \mathbb{V}_h} \frac{b(\mathbf{v}_h, p_h - p_h^{i+1})}{\|\mathbf{v}_h\|_1} \\ & \leq \sup_{\mathbf{v}_h \in \mathbb{V}_h} \frac{a(\mathbf{u}_h^{i+1} - \mathbf{u}_h, \mathbf{v}_h) + \tilde{c}(\mathbf{u}_h^i, \mathbf{u}_h^{i+1}, \mathbf{v}_h) - \tilde{c}(\mathbf{u}_h, \mathbf{u}_h, \mathbf{v}_h)}{\|\mathbf{v}_h\|_1} \\ & \leq \sup_{\mathbf{v}_h \in \mathbb{V}_h} \frac{a(\mathbf{u}_h^{i+1} - \mathbf{u}_h, \mathbf{v}_h) + \tilde{c}(\mathbf{u}_h^i - \mathbf{u}_h^{i+1}, \mathbf{u}_h^{i+1}, \mathbf{v}_h) + \tilde{c}(\mathbf{u}_h^{i+1}, \mathbf{u}_h^{i+1} - \mathbf{u}_h, \mathbf{v}_h) + \tilde{c}(\mathbf{u}_h^{i+1} - \mathbf{u}_h, \mathbf{u}_h, \mathbf{v}_h)}{\|\mathbf{v}_h\|_1} \\ & \leq \nu \|\nabla(\mathbf{u}_h^{i+1} - \mathbf{u}_h)\| + C \|\nabla(\mathbf{u}_h^i - \mathbf{u}_h^{i+1})\| \|\nabla \mathbf{u}_h^{i+1}\| + C \|\nabla \mathbf{u}_h^{i+1}\| \|\nabla(\mathbf{u}_h^{i+1} - \mathbf{u}_h)\| + \|\nabla(\mathbf{u}_h^{i+1} - \mathbf{u}_h)\| \|\nabla \mathbf{u}_h\| \\ & \leq \nu \|\nabla(\mathbf{u}_h^{i+1} - \mathbf{u}_h)\| + \frac{C}{\nu} \|\mathbf{f}\| \|\nabla(\mathbf{u}_h^i - \mathbf{u}_h^{i+1})\| + \frac{C}{\nu} \|\mathbf{f}\| \|\nabla(\mathbf{u}_h^{i+1} - \mathbf{u}_h)\| + \frac{C}{\nu} \|\mathbf{f}\| \|\nabla(\mathbf{u}_h^{i+1} - \mathbf{u}_h)\| \end{aligned}$$

which together with (4.11) implies that

$$\|p_h - p_h^{i+1}\| \leq \nu \|\nabla(\mathbf{u}_h^{i+1} - \mathbf{u}_h)\| + \frac{C}{\nu} \|\mathbf{f}\| \|\nabla(\mathbf{u}_h^i - \mathbf{u}_h^{i+1})\| + \frac{C}{\nu} \|\mathbf{f}\| \|\nabla(\mathbf{u}_h^{i+1} - \mathbf{u}_h)\| \tag{4.12}$$

Equations 4.11 and 4.12 complete the proof. □

4.2 Numerical experiments

In this part we will perform 2D and 3D numerical investigations based on the Mini-element discretization given by (4.7) and using Freefem++ (see [14]). In both cases, the domain $\Omega = (0, 1)^d$ and the relative iterative error due to the linearization of the nonlinear problem (4.2) is

$$Er_i^3 = \frac{\|\mathbf{u}_h^i - \mathbf{u}_h^{i+1}\|_1}{\|\mathbf{u}_h^{i+1}\|_1}$$

where the stopping criterion of Scheme (4.8) will be one of the following tests:

1. Classical one: $Er_i^3 < 10^{-5}$,
2. New one: $Er_i^3 < \gamma h$.

We observe that both stopping criterion gives the following error:

$$\|p - p_h^i\| + \|\mathbf{u} - \mathbf{u}_h^i\|_1 \leq c_1 h.$$

where (\mathbf{u}_h^i, p_h^i) is computed after the convergence of the iterative scheme (4.8). We introduce also the following relative error

$$Ee_i^3 = \frac{\|\mathbf{u} - \mathbf{u}_h^i\|_1}{\|\mathbf{u}\|_1} + \frac{\|p - p_h^i\|}{\|p\|}$$

Table 5 2D Navier-Stokes model: Ee_i^3 and N_{iter} with respect to γ (New stopping criterion) for $N = 50$

γ	0.001	0.01	0.05	0.1	0.5	1	1.5	5
Ee_i^3	0.135	0.135	0.135	0.135	0.135	0.135	0.135	0.138
N_{iter}	13	10	7	6	4	3	3	2

where (\mathbf{u}, p) is the exact solution of Problem (4.2) and (\mathbf{u}_h^i, p_h^i) is computed after the convergence of the iterative scheme (4.8) using one of the above stopping criteria.

4.2.1 2D numerical simulations

We consider in this case $\Omega = (0, 1)^2$ and the following exact solution:

$$\begin{aligned} \psi(x, y) &= x^2(x - 1)^2y^2(y - 1)^2 \cos(4\pi x) \cos(4\pi y), \\ \mathbf{u}(x, y) &= (u_1, u_2) = c_u \left(\frac{\partial \psi}{\partial y}, -\frac{\partial \psi}{\partial x} \right), \\ p(x, y) &= c_p xy(x - 2./3.)(y - 2./3.), \end{aligned}$$

where $c_u = 10, c_p = 10, \rho = 1, \nu = 0.01$. Thus, \mathbf{f} can be computed by using the first equation of (4.8).

We first take $N = 50$ and (the number of segments by edge) and use the classical stopping criterion, we get $N_{iter} = 15$ and $Ee_i^3 = 0.135$. Table 5 shows the dependency of the error Ee_i^3 and the number of iterations N_{iter} with respect to γ . In this case,

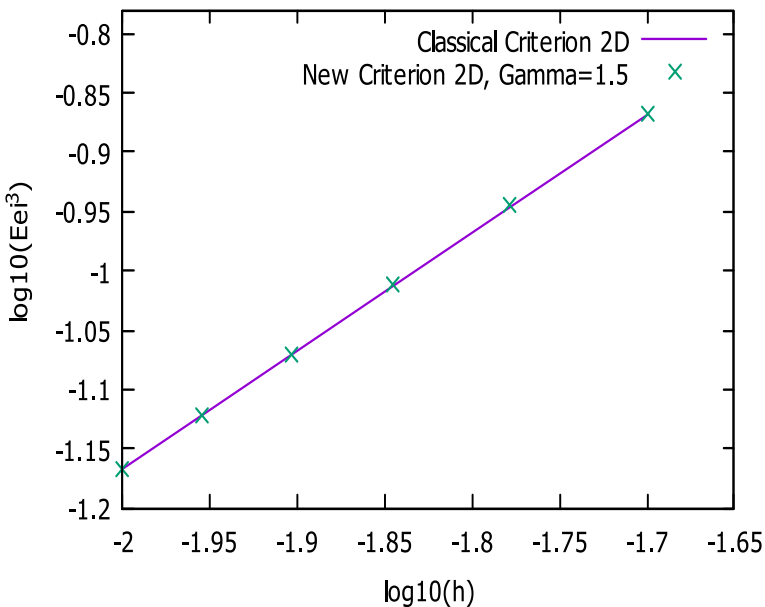


Fig. 5 2D case textbfNavier-Stokes model: A comparison of the *a priori* error between the classical criterion and the new one for $\gamma = 1.5$

Table 6 3D Navier-Stokes model: Ee_i^3 and N_{iter} with respect to γ (New stopping criterion) for $N = 20$

γ	0.001	0.01	0.05	0.1	0.5	1	1.5	5
Ee_i^3	0.183	0.183	0.183	0.183	0.183	0.183	0.183	0.188
N_{iter}	9	7	6	6	4	4	4	3

$\gamma = 1.5$ gives the same error obtained with the classical stopping criterion but with 3 iterations instead of 15 iterations obtained with the classical stopping criterion. This shows the advantage of this method.

Figure 5 shows comparisons of the *a priori* error estimates between the classical criterion and the new one for $\gamma = 1.5$. The slope of the curves is 0.99 which agrees with the theoretical slope equal to 1.

4.2.2 3D numerical simulations

We consider in this case $\Omega = (0, 1)^3$ and the following exact solution:

$$p(x, y) = c_p xyz(x - 2./3.)(y - 2./3.)(z - 2./3.),$$

$$\mathbf{u}(x, y, z) = (u_1, u_2, u_3) = c_u \text{curl } \Psi(x, y, z),$$

where

$$\Psi(x, y, z) = \psi(x, y, z), \psi(x, y, z), \psi(x, y, z),$$

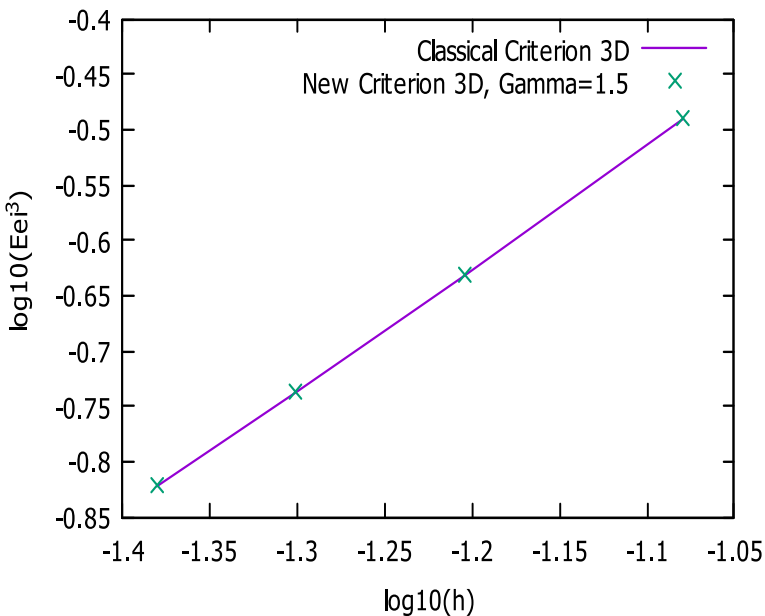


Fig. 6 3D Navier-Stokes model: A comparison of the *a priori* error between the classical criterion and the new one for $\gamma = 1.5$

and

$$\psi(x, y, z) = x^2(x-1)^2y^2(y-1)^2y^2(y-1)^2 \cos(2\pi xyz).$$

We take $c_u = 200$, $c_p = 20$, $\nu = 0.005$. Thus, f can be computed by using the first equation of (4.8).

We consider $N = 20$ (the number of segments by edge) and use the classical stopping criterion, we get $N_{iter} = 11$ and $Ee_i^3 = 0.183$. Table 6 shows the dependency of the error Ee_i^3 and the number of iterations N_{iter} with respect to γ . As in the previous cases, $\gamma = 1.5$ gives the same error obtained with the classical stopping criterion but with 4 iterations instead of 11 iterations obtained with the classical criterion.

Figure 6 shows comparisons of the *a priori* error estimates between the classical criterion and the new one for $\gamma = 1.5$. The slope of the curves is 1.09 which agrees with the theoretical slope equal to 1.

5 Concluding remarks

Our aim is was to introduce a new stopping criterion based on the iteration error in order to obtain a fast algorithm. This goal was achieved by avoiding the unnecessary iterations in comparison with the order of the *a priori* error estimates (the discretization error). This strategy was considered in three cases and the superiority of the novel approach was clearly demonstrated. The current methodology should be further developed, in particular by (i) extending the present analysis to variational inequalities of second kind; (ii) considering the case of a curved boundary.

Author Contributions conception: J.K Djoko analysis: J.K. Djoko and T. Sayah simulations: T. Sayah Writing: J.K. Djoko and T. Sayah

Funding Open access funding provided by North-West University.

Data Availability No datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Glowinski, R.: Numerical Methods for Nonlinear Variational Problems. Springer Series in Computational Physics, Springer-Verlag, Berlin Heidelberg, (2008)

2. Glowinski, R.: Finite element methods for incompressible viscous flow. In: Handbook of Numerical Analysis, Ciarlet, P.G, Lions, J.L. (eds.), vol. IX, North Holland; Amsterdam, pp. 3–1176, (2003)
3. Girault, V., Raviart, P.A.: Finite Element Methods for Navier-Stokes Equations: Theory and Algorithms. Springer-Verlag, Berlin, Heidelberg, New-York, Tokyo (1986)
4. Boffi, D., Brezzi, F., Fortin, M.: Mixed Finite Element Methods and Applications, Springer Series in Computational Mathematics, Springer Verlag Berlin, (2013)
5. Ciarlet, P.G.: Basic error estimates for elliptic problems. In Handbook of Numerical Analysis, Vol. II, Handbook of Numerical Analysis, North-Holland, Amsterdam, pp. 17–351 (1991)
6. El Alaoui, L., Ern, A., Vohralik, M.: Guaranteed and robust a posteriori error estimates and balancing of discretization and linearization errors for monotonic nonlinear problems. *Comput. Methods Appl. Mech. Eng.* **200**(37–40), 2782–2795 (2011)
7. Ern, A., Vohralik, M.: Inexact adaptive Newton methods with a posteriori stopping criteria for nonlinear diffusion PDEs. *SIAM J. Sci. Comput.* **35**(4), A1761–A1791 (2013)
8. Bernardi, C., Dakroub, J., Mansour, G., Sayah, T.: A posteriori analysis of iterative algorithms for Navier-Stokes problem. *ESAIM: M2AN.*, **50**(4), 1035–1055, (2016)
9. Dakroub, J., Faddoul, J., Sayah, T.: A posteriori analysis of the Newton method applied to the Navier-Stokes problem. *J. Comput. Appl. Math.* **63**, 411–437 (2020)
10. C. Bernardi, J. Dakroub, G. Mansour, T. Sayah. A posteriori analysis of iterative algorithms for a nonlinear problem. *J. Sci. Comput.* **65**, 672–697, 2015
11. Nassreddine, G., Omnes, P., Sayah, T.: A posteriori error estimates for the Large Eddy Simulation applied to stationary Navier-Stokes equations. *Numer. Methods Partial Differ. Equ.* **38**, 1468–1498 (2022)
12. Sayah, T., Semaan, G., Triki, F.: A Posteriori error estimates for the Darcy-Forchheimer problem. *CMAM* (2022). <https://doi.org/10.1515/cmam-2022-0047>
13. T. Sayah, G. Semaan F. Triki. A Posteriori error estimates for Darcy-Forchheimer’s problem coupled with the convection-diffusion-reaction equation. *IJNAM.* **21**(1), 65–103 (2024)
14. Hecht, F.: New development in FreeFem++. *J. Numer. Maths.* **20**, 251–266 (2012)
15. Fabrie, P.: Regularity of the solution of Darcy-Forchheimer’s equation. *Nonlinear Anal. Theory. Appl.* **13**(9), 1025–1049 (1989)
16. Achdou, Y., Bernardi, C., Coquel, F.: A priori and a posteriori analysis of finite volume discretizations of Darcy’s equations. *Numer. Math.* **96**, 17–42 (2003)
17. Bernardi, C., Maday, Y., Rapetti, F.: *Discretisations Variationelles de problemes aux limites Elliptiques*, Mathematiques Applications, vol. 45. Springer, Paris (2004)
18. Girault, V., Wheeler, M.F.: Numerical discretization of a Darcy-Forchheimer model. *Numer. Math.* **110**, 161–198 (2008)
19. Temam, R.: *Navier-Stokes equations: Theory and Numerical Analysis*, 2nd ed., AMS Chelsea publishing, (2001)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.