Japan J. Indust. Appl. Math., 26 (2009), 65–91 Area ⟨2⟩

An Explicit Method for Convection-Diffusion Equations

Vitoriano Ruas*, Antonio Brasil Jr.† and Paulo Trales‡

* *Institut Jean Le Rond d'Alembert, Université Paris 6, France and Visiting Researcher at Departamento de Análise Universidade Federal Fluminense, Rio de Janeiro, Brazil E-mail: vitoriano.ruas@upmc.fr*

† *Departamento de Engenharia Mecânica Universidade de Brasília, Brasília, Brazil E-mail: brasiljr@unb.br*

‡ *Departamento de Análise Universidade Federal Fluminense, Rio de Janeiro, Brazil E-mail: paulotrales@im.uff.br*

Received December 14, 2007
Revised October 31, 2008

Dedicated to the memory of Francisco Ruas Santos.

An explicit scheme for time-dependent convection-diffusion problems is presented. It is shown that convenient bounds for the time step value ensure L^∞ stability, in both space and time, for piecewise linear finite element discretizations in any space dimension. Convergence results in the same sense are also demonstrated under certain conditions. Numerical results certify the good performance of the scheme.

Key words: convection, diffusion, explicit, finite elements, Péclet number, stable, time-dependent

1. Introduction

This work deals with numerical methods based on variational formulations such as the finite element method, to solve convection-diffusion equations. In this framework, one of the first techniques employed to model convection was the so-called Lesaint–Raviart method (cf. [15]). However a little later the Japanese school gave relevant contributions to the subject, as it is well reported in [9]. In this respect we would like to quote the pioneer work of Tabata (cf. [20]) together with [13], among many others.

Since the mid-eighties, the most widespread manner to deal with dominant convection has been the use of stabilizing procedures based on the space mesh parameter, among which the streamline upwind Petrov–Galerkin (SUPG) technique introduced by Hughes and Brooks (cf. [3]) is one of the most popular. As far as time-dependent problems are concerned, it turns out that the time step plays a better stabilizing role, provided a formulation well suited to the equations to be solved is employed. A good illustration of this assertion in the case of the time-dependent Navier–Stokes equations can be found in [5] or yet in [17].

A good and more recent reference on finite element methods to deal with advection-diffusion phenomena is the book of Knabner and Angermann [14].

The authors intend to give a contribution in this direction, in the case of the convection-diffusion equations with dominant convection, discretized in space with piecewise linear finite elements, combined with a non standard explicit Euler scheme for the time integration, and a standard Galerkin approach. Our main theoretical result states that the numerical solution is stable in the L^∞-norm in both space and time, and even convergent under certain angle conditions, provided that roughly the time step is bounded by the space mesh parameter multiplied by a mesh-independent constant that we specify. As the authors should clarify, the scheme studied in this paper follows similar principles to the one long exploited by Kawahara and collaborators, for simulating convection dominated phenomena (see, e.g., [10], [11] and [12] among several other papers published by them before and later on). The originality of our contribution relies on the fact that we not only introduce a reliable scheme for any space dimension, but also exhibit rigorous conditions for it to provide converging sequences of approximations in the sense of the maximum norm. In addition to the theoretical results, numerical examples with known exact solution are given, in order to illustrate the adequacy of our numerical approach.

An outine of the paper is as follows: In Section 2 we recall the problem to solve and make some assumptions on the data. In Section 3 we describe the type of discretization corresponding to the new method, and more specially the weighted manner to deal with the mass matrices on both sides of the discrete equations. In Section 4 we give stability results for our method in the sense of the space and time maximum norm, applying to a non restrictive set of weights. Then in Section 5 we specify conditions to be satisfied by the weights allowing for optimal error estimates. A convergence result is also given in Section 5, applying to the particular case where the mesh is of the acute type (see, e.g., [21]). Next in Section 6 we give a particularly representative summary of the numerical results that we have obtained so far with our new scheme. Finally in Section 7 we draw some conclusions about the whole work.

2. Problem Statement

Let us consider a time-dependent convection-diffusion problem described as follows: Find a scalar valued function $u(\boldsymbol{x}, t)$ defined in $\bar{\Omega} \times [0, \infty)$, Ω being a bounded open subset of $\Re^N$ with boundary $\partial\Omega$, $N = 1$, 2 or 3, such that,

$$\begin{cases} u_t + \boldsymbol{a} \cdot \nabla u - \nu \Delta u = f & \text{in } \Omega \times (0, \infty), \\ u = g & \text{on } \partial\Omega \times (0, \infty), \\ u = u^0 & \text{in } \Omega, \quad \text{for } t = 0, \end{cases} \tag{1}$$

where u_t represents the first order derivative of u with respect to t, ν is a positive constant and $\boldsymbol{a}$ is a given solenoidal convective velocity at every time t, assumed to be uniformly bounded in $\Omega \times (0, \infty)$. The data f and g are respectively, a given forcing function belonging to $L^\infty[\Omega \times (0, \infty)]$, and a prescribed value on $L^\infty[\partial\Omega \times (0, \infty)]$. We further assume that $u^0 \in L^\infty(\Omega)$ and that for every $\boldsymbol{x} \in \partial\Omega$ $g(\boldsymbol{x}, \cdot\,)$ is of bounded variation in $(0, \infty)$.

We first set (1) in the following equivalent variational form:

$$\begin{cases} \text{For every } t \in (0,\infty), \text{ find } u(\,\cdot\,,t) \in H^1(\Omega) \text{ with } u_t(\,\cdot\,,t) \in L^2(\Omega), \\ u(\,\cdot\,,t) = g(\,\cdot\,,t) \text{ on } \partial\Omega, \text{ and } u = u^0 \text{ in } \Omega, \text{ for } t = 0, \text{ such that} \\ \displaystyle\int_\Omega (u_t + \boldsymbol{a}\cdot\nabla u)v + \nu\int_\Omega \nabla u\cdot\nabla v = \int_\Omega fv, \quad \forall v \in H_0^1(\Omega). \end{cases} \tag{2}$$

Henceforth we consider only normalized dimensionless lenghts, time and velocity. In so doing ν^{-1} represents the so-called Péclet number (see also [9]), and the convective dominant case corresponds to a low value of ν. Whenever ν is not so small, the diffusion will influence the phenomenon being modelled practically everywhere in the domain, and convection will not be dominant.

3. Space and time discretizations

In all the sequel, for the sake of simplicity, and without loss of essential aspects, we assume that Ω is an interval if $N = 1$, a polygon if $N = 2$ or a polyhedron if $N = 3$. In so doing we next consider a partition $\mathcal{T}_h$ of Ω into N-simplices, with maximum edge length equal to h. We assume that $\mathcal{T}_h$ satisfies the usual compatibility conditions for finite element meshes, and that it belongs to a quasi-uniform family of partitions. We further define a second mesh parameter $h_{\min}$ as the minimum height of all the elements of $\mathcal{T}_h$ if $N = 2$ or 3 and the minimum length of $K \in \mathcal{T}_h$ if $N = 1$.

Now for every $K \in \mathcal{T}_h$ we denote by $P_1(K)$ the space of polynomials of degree less than or equal to one defined in K. In so doing we introduce the following spaces or manifolds associated with $\mathcal{T}_h$:

$$\begin{aligned} V_h &:= \{v \mid v \in C^0(\bar{\Omega}) \text{ and } v|_K \in P_1(K), \forall K \in \mathcal{T}_h\}, \\ V_h^0 &:= V_h \cap H_0^1(\Omega). \end{aligned}$$

We further introduce for any function ϕ defined in $C^0(\partial\Omega)$ the following manifold of V_h:

$$V_h^\phi := \{v \in V_h \mid v(P) = \phi(P), \forall \text{ vertex } P \text{ of } \mathcal{T}_h \text{ on } \partial\Omega\}.$$

Now let u_h^0 be the field of $V_h^{g(\,\cdot\,,0)}$ satisfying $u_h^0(P) = u^0(P)$ for every vertex P of $\mathcal{T}_h$, and $\Delta t > 0$ be a given time step. Defining g^n on $\partial\Omega$ by $g^n(\,\cdot\,) = g(\,\cdot\,,n\Delta t)$, f^n in Ω by $f^n(\cdot) = f(\,\cdot\,,n\Delta t)$ and $\boldsymbol{a}^n$ in Ω by $\boldsymbol{a}^n(\,\cdot\,) = \boldsymbol{a}(\,\cdot\,,n\Delta t)$, for $n = 1,2,\ldots$, idealistically we wish to determine approximations $u_h^n(\,\cdot\,)$ of $u(\,\cdot\,,n\Delta t)$ for $n \in \mathbb{N}^*$, by solving the following finite element discrete set of equations, corresponding to the first order forward Euler scheme.

$$\begin{cases} \text{For } n \text{ successively equal to } 1,2,\ldots, \text{ find } u_h^n \in V_h^{g^n} \text{ satisfying } \forall v \in H_0^1(\Omega), \\ \displaystyle\int_\Omega u_h^n v = \int_\Omega u_h^{n-1}v + \Delta t\Big[\int_\Omega f^{n-1}v - \int_\Omega \boldsymbol{a}^{n-1}\cdot\nabla u_h^{n-1}v - \nu\int_\Omega \nabla u_h^{n-1}\cdot\nabla v\Big]. \end{cases} \tag{3}$$

Now we expand u_h^n into a sum of the form,

$$u_h^n = \sum_{j=1}^{N_h} u_j^n \varphi_i,$$

where φ_j is the canonical basis function of V_h associated with the j-th node of $\mathcal{T}_h$, say P_j, $u_j^n \in \Re$ is the value of u_h^n at P_j, and N_h is the dimension of V_h. We assume that the nodes P_j are numbered in such a manner that the first I_h nodes are located in the interior of Ω and the remaining $N_h - I_h$ nodes are located on $\partial\Omega$. Now we choose v successively equal to φ_i, for $i = 1, 2, \ldots, I_h$, and we approximate for every n, $\int_\Omega (\boldsymbol{a}^n \cdot \nabla)\varphi_j\varphi_i$ by $\int_\Omega (\boldsymbol{a}_i^n \cdot \nabla)\varphi_j\varphi_i$, and $\int_\Omega f^n\varphi_i$ by $\int_\Omega f_i^n\varphi_i$, where $\boldsymbol{a}_i^n := \boldsymbol{a}^n(P_i)$ and $f_i^n := f^n(P_i)$.

Still denoting the resulting values of $u_h^n(P_j)$ by u_j^n for $j = 1, 2, \ldots, N_h$, the unknown coefficients u_j^n for $j = 1, 2, \ldots, I_h$ and $n = 1, 2, \ldots$, are recursively determined by solving the following linear system of equations:

$$\sum_{j=1}^{N_h} m_{ij}^C u_j^n = \sum_{j=1}^{N_h} \left[m_{ij}^C - \Delta t\, a_{ij}^{n-1}\right] u_j^{n-1} + \Delta t\, b_i^n, \quad \text{for } i = 1, \ldots, I_h, \tag{4}$$

$n = 1, 2, \ldots$, where the coefficients m_{ij}^C, a_{ij}^n and b_i^n are given by

$$\begin{aligned} m_{ij}^C &= \int_\Omega \varphi_j\varphi_i, \\ a_{ij}^n &= \int_\Omega [(\boldsymbol{a}_i^n \cdot \nabla)\varphi_j\varphi_i + \nu\nabla\varphi_j \cdot \nabla\varphi_i], \\ b_i^n &= \int_\Omega f_i^{n-1}\varphi_i. \end{aligned} \tag{5}$$

Actually, since for every $i \in I_h$, φ_i vanishes on $\partial\Omega$ and $\boldsymbol{a}^n$ is solenoidal, by integration by parts we easily derive

$$\int_\Omega (\boldsymbol{a}_i^n \cdot \nabla)\varphi_i\varphi_i = 0.$$

Hence, denoting by δ_{ij} the Kronecker symbol, we may rewrite (5) as follows:

$$a_{ij}^n = \int_\Omega \{(1 - \delta_{ij})[(\boldsymbol{a}_i^n \cdot \nabla)\varphi_j\varphi_i] + \nu\nabla\varphi_j \cdot \nabla\varphi_i\}. \tag{6}$$

The L^∞-stability of the explicit scheme (3) is not ensured in general. Therefore stabilizing techniques have been introduced such as upwinding (cf. [20] and [2]), in which the integral corresponding to the convection term is computed only in the element(s) situated upwind to the node P_i, with respect to $\boldsymbol{a}_i^n$. Our strategy here is restricted to the use of two different quadrature formulae to compute the coefficients

m_{ij}^C, according to the side of equation (3). In the particular choice made in this work, on the left hand side of (4) m_{ij}^C is approximated by the trapezoidal rule, and on the right hand side the approximate value of m_{ij}^C denoted by m_{ij}^W is obtained by an asymmetric quadrature formula (at least for non uniform meshes) specified below. We recall that the trapezoidal rule gives an approximation of the integral of a continuous function ψ in Ω expressed by

$$\mathcal{J}_h(\psi) = \sum_{K \in \mathcal{T}_h} \frac{\text{meas}(K)}{N+1} \sum_{i=1}^{N+1} \psi(S_i^K),$$

where the S_i^K's denote the vertices of the N-simplex K, with $i = 1, 2, \ldots, N+1$.

Hence, letting S_i denote the support of φ_i, and Π_i represent its measure, the aproximation of m_{ii}^C on the left hand side of (4) is just the well-known lumped mass diagonal coefficient $m_{ii}^L = \frac{\Pi_i}{N+1}$, and the one of m_{ij}^C for $i \neq j$ equals zero.

In so doing, the unknown nodal values of u_h^n still denoted by u_j^n, for $j = 1, 2, \ldots, I_h$ and $n = 1, 2, \ldots,$ are determined by

$$u_i^n = \sum_{j=1}^{N_h} [\tilde{m}_{ij} - \Delta t\, \tilde{a}_{ij}^{n-1}] u_j^{n-1} + \Delta t\, \tilde{b}_i^n, \quad \text{for } i = 1, \ldots, I_h,\ n = 1, 2, \ldots, \tag{7}$$

where $\tilde{m}_{ij}$, $\tilde{a}_{ij}^n$ are just $N+1$ times the quotient between m_{ij}^W, a_{ij}^n and Π_i, and as we easily conclude, $\tilde{b}_i^n = f_i^{n-1}$.

Notice that this approach is a sort of compromise between the lumped mass scheme and the consistent mass scheme, both widespread among finite element users for about fourty years.

The approximate coefficients m_{ij}^W on the right hand side are determined as follows: Let M_i be the number of nodes different from P_i lying in the closure of S_i, i.e., $\bar{S}_i$, and P_{k_j} be such nodes for $j = 1, 2, \ldots, M_i$ with $1 \leq k_j \leq N_h$. Let also W_j^i be the measure fractions associated with P_{k_j} given by

$$W_j^i = \frac{\text{meas}(S_i \cap S_{k_j})}{N+1} \tag{8}$$

and ω_j^i be corresponding strictly positive weights satisfying

$$\sum_{j=1}^{M_i} \omega_j^i W_j^i = \frac{N\Pi_i}{(N+1)(N+2)}. \tag{9}$$

Notice that since each N-simplex in S_i appears in exactly N measure fractions W_j^i, we necessarily have

$$\sum_{j=1}^{M_i} W_j^i = \frac{N\Pi_i}{N+1}. \tag{10}$$

Now selecting the nodes P_{k_j} in $\bar{S}_i$ different from P_i, we define

$$m^W_{ik_j} = \frac{h_{\min}}{\nu + h_{\min}} \omega^i_j W^i_j \quad \text{for } i \neq k_j \tag{11}$$

together with

$$m^W_{ii} = \frac{\nu}{h_{\min} + \nu} m^L_{ii} + \frac{h_{\min}}{h_{\min} + \nu} m^C_{ii}. \tag{12}$$

Now recalling that (cf. [19]),

$$m^C_{ii} = \frac{2\Pi_i}{(N+1)(N+2)} \tag{13}$$

from (9) and (11) we have $\sum_{j=1}^{M_i} m^W_{ik_j} = \frac{N}{2} \frac{h_{\min}}{\nu + h_{\min}} m^C_{ii}$. On the other hand we know that $m^L_{ii} = \frac{(N+2)m^C_{ii}}{2}$. Then since (12) implies that $m^W_{ii} = m^L_{ii} + \frac{h_{\min}}{h_{\min}+\nu}(m^C_{ii} - m^L_{ii})$, it easily follows that we actually have $\forall i \in \{1, 2, \ldots, I_h\}$,

$$\begin{aligned} m^W_{ik_j} &= \frac{h_{\min}}{\nu + h_{\min}} \omega^i_j W^i_j \quad \text{for } P_{k_j} \in \bar{S}_i,\ i \neq k_j, \\ m^W_{ii} &= m^L_{ii} - \sum_{j=1}^{M_i} m^W_{ik_j}. \end{aligned} \tag{14}$$

Naturally enough, by definition, $m^W_{ij} = 0$ if P_j does not lie in $\bar{S}_i$. Typically we may choose $\omega^i_j = \frac{1}{N+2}$ for every j and for every node P_i, thereby generating a weighted combination of the lumped mass and the consistent mass matrix (cf. [19]) on the right hand side of (4), with weights equal to $\frac{\nu}{h_{\min}+\nu}$ and $\frac{h_{\min}}{h_{\min}+\nu}$, respectively. However, except for the case of uniform meshes, in principle this is not the choice to make, if one wishes to reach the best results in terms of accuracy, as seen in Section 5.

4. Stability results

In this section we show that, provided Δt is chosen conveniently small with respect to the spacial mesh parameter, the scheme (7) is stable in the sense of L^∞. First we have to define the following quantities:

- $A = \sup_{t \in (0,\infty)} \max_{1 \leq i \leq I_h} |\boldsymbol{a}_i(t)|$;
- $\omega = \min_{1 \leq i \leq I_h} \min_{1 \leq j \leq M_i} \omega^i_j$.

It is interesting to note that from (9) and (10) we easily establish that

$$\omega \leq \frac{1}{N+2}. \tag{15}$$

Next we prove the following lemma, which directly leads to the stability result stated in Theorem 4.2 hereafter.

LEMMA 4.1. *If Δt fulfills the condition*

$$\Delta t \leq \frac{\omega h_{\min}^3}{(\nu + h_{\min})[A h_{\min} + (N+1)\nu]}, \tag{16}$$

the $I_h \times N_h$ matrix $C^n = \{c_{ij}^n\}$ given by

$$c_{ij}^n = \tilde{m}_{ij} - \Delta t\, \tilde{a}_{ij}^{n-1} \tag{17}$$

is a non-negative matrix having a unit row-norm in the following sense:

$$\begin{cases} c_{ij}^n \geq 0 \quad \forall i \in \{1, 2, \ldots, I_h\} \text{ and } \forall j \in \{1, 2, \ldots, N_h\}, \\ \displaystyle\sum_{j=1}^{N_h} c_{ij}^n = 1 \quad \forall i \in \{1, 2, \ldots, I_h\}. \end{cases} \tag{18}$$

Proof. First we treat the coefficients c_{ii}^n, which are given by

$$c_{ii}^n = \frac{N+1}{\Pi_i}\big(m_{ii}^W - \Delta t\, a_{ii}^{n-1}\big), \tag{19}$$

where m_{ii}^W is defined by (12) and a_{ii}^n is given by

$$a_{ii}^n = \int_{S_i} \nu |\nabla \varphi_i|^2.$$

Straightforward calculations lead to the result:

$$a_{ii}^{n-1} \leq \nu h_{\min}^{-2} \Pi_i. \tag{20}$$

On the other hand from (12) and (13) we easily derive

$$m_{ii}^W = \frac{\nu(N+2) + 2h_{\min}}{(N+1)(N+2)(\nu + h_{\min})} \Pi_i. \tag{21}$$

It follows from (19)–(21) that $c_{ii}^n \geq 0\ \forall i \in \{1, 2, \ldots, I_h\}$ if

$$\Delta t \leq \frac{[2h_{\min} + \nu(N+2)] h_{\min}^2}{\nu(h_{\min} + \nu)(N+1)(N+2)}. \tag{22}$$

Now we switch to the coefficientes c_{ij}^n for $i \neq j$. Noticing that $c_{ij}^n = 0$ if P_j does not belong to $\bar{S}_i$, for those nodes P_{k_j} that do, $j = 1, 2, \ldots, M_i$, we have for $k_j \neq i$,

$$c_{ik_j}^n = \frac{N+1}{\Pi_i}\big(m_{ik_j}^W - \Delta t\, a_{ik_j}^{n-1}\big),$$

where $m_{ik_j}^W$ is defined by (11) and $a_{ik_j}^n$ is given by

$$a_{ik_j}^n = \int_{\Omega} \{(\boldsymbol{a}_i^n \cdot \nabla)\varphi_{k_j}\varphi_i + \nu \nabla \varphi_{k_j} \cdot \nabla \varphi_i\}. \tag{23}$$

From (11), (8) and the definition of ω we trivially have

$$m_{ik_j}^W \geq \sum_{K\in S_i\cap S_{k_j}} \frac{\omega h_{\min}}{(N+1)(\nu+h_{\min})}\,\mathrm{meas}(K).$$

Now in an element K belonging to $S_i \cap S_{k_j}$, let $\theta_{ik_j}^K$ be the angle between the gradients of φ_i and φ_{k_j} for $N=2$ or $N=3$, and $\theta_{ik_j}^K = \pi$ for $N=1$. Then after straightforward calculations we obtain for all $n \geq 0$ (see also [7]),

$$\left|a_{ik_j}^{n-1}\right| \leq \sum_{K\in S_i\cap S_{k_j}} \left[\frac{A}{h_{\min}(N+1)} + \nu\frac{\left|\cos\theta_{ik_j}^K\right|}{h_{\min}^2}\right]\mathrm{meas}(K).$$

It immediately follows that $c_{ik_j}^n \geq 0$ for $i \neq k_j$, if (16) holds.

On the other hand, according to well-known properties of the basis functions φ_j, we have $\sum_{j=1}^{N_h}\varphi_j = 1$ everywhere in Ω. Hence by linearity, we easily derive for every $i \in \{1,2,\ldots,I_h\}$ and for all $n \geq 0$,

$$\sum_{j=1}^{N_h} a_{ij}^n = 0. \tag{24}$$

Morevover from (14) we readily derive

$$\sum_{j=1}^{N_h} m_{ij}^W = m_{ii}^W + \sum_{j=1}^{M_i} m_{ik_j}^W = m_{ii}^L.$$

Thus for every $i \in \{1,2,\ldots,I_h\}$,

$$\sum_{j=1}^{N_h} m_{ij}^W = \frac{\Pi_i}{N+1}. \tag{25}$$

Then using the definitions $\tilde{a}_{ij}^n = (N+1)\frac{a_{ij}^n}{\Pi_i}$ and $\tilde{m}_{ij} = (N+1)\frac{m_{ij}^W}{\Pi_i}$, together with (17), (24) and (25), we readily conclude that for every $i \in \{1,2,\ldots,I_h\}$,

$$\sum_{j=1}^{N_h} c_{ij}^n = 1.$$

Finally recalling (15), we easily infer that, if Δt fulfills (16), then (22) also holds. This completes the proof. $\square$

Theorem 4.2. *Let Δt satisfy condition (16). Then the finite element solution sequence $\{u_h^n\}_n$ given by $u_h^n = \sum_{j=1}^{N_h} u_j^n\varphi_j$ generated by (7) for $n = 1,2,\ldots$ satisfies the following stability result for every $m \in \mathbb{N}$, whereby $\|F\|_{0,\infty,D}$ denotes* *the L^∞-norm of a function F defined in an open set D of $\Re^N$, and $BV[G]$ represents the standard norm of the space of functions $G(t)$ having bounded variation for $t \in (0,\infty)$,*

$$\|u_h^m\|_{0,\infty,\Omega} \leq \|u^0\|_{0,\infty,\Omega} + \max\left\{\max_{P\in\partial\Omega} BV[g(P,\,\cdot\,)], \Delta t\sum_{n=1}^{m}\|f^{n-1}\|_{\infty,\Omega}\right\}. \qquad (26)$$

Proof. Recalling (7)–(17), we may write for convenience, for $n = 1,2,\dots,$

$$u_i^n = \sum_{j=1}^{N_h} c_{ij}^n u_j^{n-1} + \Delta t\,\tilde{b}_i^n, \quad i = 1,2,\dots,I_h. \qquad (27)$$

Now we extend (27) to $i = I_h+1,\dots,N_h$, by setting for every $j\in\{1,2,\dots,N_h\}$, $c_{ij}^n = \delta_{ij}$ and $\tilde{b}_i^n = \frac{g^n(P_i)-g^{n-1}(P_i)}{\Delta t}$.

Since (16) holds, according to Lemma 4.1, (18) is also true. Therefore we have for $n = 1,2,\dots,$

$$u_i^n \leq \max_{1\leq j\leq N_h} u_j^{n-1} + \Delta t\,\tilde{b}_i^n, \quad \text{for } i = 1,2,\dots,N_h, \qquad (28)$$

together with

$$-u_i^n \leq \max_{1\leq j\leq N_h}\left(-u_j^{n-1}\right) - \Delta t\,\tilde{b}_i^n, \quad \text{for } i = 1,2,\dots,N_h. \qquad (29)$$

From (28)–(29) we infer that

$$\max_{1\leq i\leq N_h}|u_i^n| \leq \max_{1\leq i\leq N_h}|u_i^{n-1}| + \Delta t\max_{1\leq i\leq N_h}|\tilde{b}_i^n|. \qquad (30)$$

Then adding up (30) from $n = 1$ to $n = m$ we obtain for every m,

$$\max_{1\leq i\leq N_h}|u_i^m| \leq \max_{1\leq i\leq N_h}|u_i^0| + \Delta t\sum_{n=1}^{m}\max_{1\leq i\leq N_h}|\tilde{b}_i^n|. \qquad (31)$$

The remainder of the proof is a mere application of the definition of the variation of $g(P,t)$ for $P\in\partial\Omega$ and $t\in[0,\infty)$, together with the fact that $\max_{1\leq i\leq N_h}|u_i^0| \leq \|u^0\|_{0,\infty,\Omega}$. □

To conclude this section let us consider the particular case where the partition $\mathcal{T}_h$ is of the *acute type* (see, e.g., [21]), which means that $\forall i\in\{1,2,\dots,I_h\}$, $\pi/2 \leq \theta_{ik_j}^K \leq \pi$ $\forall j\in\{1,2,\dots,M_i\}$. In this case we can refine the stability result of the above theorem in the following manner.

THEOREM 4.3. *Assume that the partition $\mathcal{T}_h$ is of the acute type. Then if Δt satisfies the condition*

$$\Delta t \leq \frac{h_{\min}^2}{\nu + h_{\min}}\min\left[\frac{\omega}{A}, \frac{\nu(N+2)+2h_{\min}}{\nu(N+1)(N+2)}\right], \qquad (32)$$

the finite element solution sequence $\{u_h^n\}_n$ *given by* $u_h^n = \sum_{j=1}^{N_h} u_j^n \varphi_j$ *generated by* (7) *for* $n = 1, 2, \ldots$ *satisfies the stability condition* (26).

Proof. This theorem is a mere consequence of a modification of Lemma 4.1 in order to accomodate the replacement of bound (16) by (32). Recalling (23) we have

$$a_{ik_j}^{n-1} = \sum_{K \in S_i \cap S_{k_j}} \left[\boldsymbol{a}_i^{n-1} \cdot \nabla \varphi_{k_j} + \nu \cos \theta_{ik_j}^K |\nabla \varphi_i| \, |\nabla \varphi_{k_j}| (N+1) \right] \frac{\text{meas}(K)}{N+1}.$$

Therefore from the assumption that the partition $\mathcal{T}_h$ is of the acute type, we may assert that

$$a_{ik_j}^{n-1} \leq \sum_{K \in S_i \cap S_{k_j}} \boldsymbol{a}_i^{n-1} \cdot \nabla \varphi_{k_j} \frac{\text{meas}(K)}{N+1} \leq \frac{A W_j^i}{h_{\min}}.$$

It follows that $c_{ik_j}^n \geq 0$ if

$$\Delta t \leq \frac{\omega h_{\min}^2}{A(\nu + h_{\min})}.$$

On the other hand, (22) is still a condition for c_{ii}^n to be non negative in this case. Then the remainder of the proof goes in the very same way as for Theorem 4.2. □

5. Error estimates and convergence result

In this section we derive error estimates for the approximations of the solution of (1) generated by (7) under condition (16). We also prove that, provided the weights ω_j^i are suitably chosen, this scheme provides convergent approximations in the maximum norm, as both h and Δt go to zero, under the assumption (32) of Theorem 4.3. For this purpose we will mostly work with Sobolev spaces $W^{m,\infty}(D)$ equipped with the standand norm and seminorm denoted respectively by $\|\cdot\|_{m,\infty,D}$ and $|\cdot|_{m,\infty,D}$, where m is a non negative integer and D is a subset of $\Re^N$ (cf. [1]).

As usual we need here a suitable consistency result, which together with the stability results established in the previous section will lead to convergence. Besides standard arguments mostly borrowed from classical or celebrated works such as [21] and [20], the consistency of our scheme will be a consequence of the following lemma.

LEMMA 5.1. *Let* P_i *be a node of* $\mathcal{T}_h$, *for* $i \in \{1, 2, \ldots, I_h\}$, *and* $\boldsymbol{l}_j^i$ *be the vector leading from* P_i *to its neighbor* P_{k_j}, *that is, the* j*-th node belonging to* $\bar{S}_i$, $j = 1, 2, \ldots, M_i$. *Then there exists strictly positive weights* ω_j^i *satisfying* (9) *such that*

$$\sum_{j=1}^{M_i} \omega_j^i W_j^i \boldsymbol{l}_j^i = \mathbf{0}. \tag{33}$$

Proof. First we note that the set $L^i := \{\boldsymbol{l}^i_j \mid j = 1, \ldots, M_i\}$ generates the whole space $\Re^N$. Indeed, if it is not so all the vectors in L^i would belong to a certain proper subspace of $\Re^N$, and thus there would be elements in the mesh with zero measure, which is impossible. It follows that there can be no non zero vector orthogonal to all the vectors of L^i at a time.

Furthermore equation (33) corresponds to a homogeneous linear system of N equations with M_i unknowns, namely, the weights ω^i_j.

Now let us assume that there exists a non zero vector $\boldsymbol{d} = \{d_k\} \in \Re^N$ such that the quantities $\pi_j(\boldsymbol{d})$ given by $\sum_{k=1}^N d_k W^i_j \{\boldsymbol{l}^i_j\}_k$, satisfy

$$\pi_j(\boldsymbol{d}) \geq 0, \quad \text{for } j = 1, 2, \ldots, M_i. \tag{34}$$

Notice that, owing to the argument given in the first sentence of the proof, in this case it must hold that

$$\exists j_0 \in \{1, 2, \ldots, M_i\} \quad \text{such that} \quad \pi_{j_0}(\boldsymbol{d}) > 0. \tag{35}$$

This would mean that $\boldsymbol{d}$ makes angles less than or equal to $\pi/2$ with all the vectors $\boldsymbol{l}^i_j$. In this case the latter vectors would all belong to the same half space of $\Re^N$ defined by $\boldsymbol{d}$, but this is impossible because they are associated with an interior node of a finite element mesh. The non existence of a vector $\boldsymbol{d} \in \Re^N$ such that both (34) and (35) hold, is the necessary and sufficient condition given by Stiemke in his theorem of 1915 (cf. [6]) for the homogeneous system (33) to possess strictly positive solutions. □

Clearly enough, even by enforcing condition (9) the solution of (33) is not unique in general, except for the case $N = 1$. Nevertheless, we may obtain uniqueness by requiring some optimality condition on the weights besides those two and positiveness, such as the minimum distance to the consistence mass weights, that is,

$$\{\omega^i_1, \omega^i_2, \ldots, \omega^i_{M_i}\} = \arg\min_{\boldsymbol{z} \in \mathcal{Z}_\varepsilon} \left[\sum_{j=1}^{M_i} \left| z_j - \frac{1}{N+2} \right|^2 \right],$$

for some sufficiently small $\varepsilon > 0$, with

$$\mathcal{Z}_\varepsilon := \left\{ \boldsymbol{z} = \{z_j\} \in \Re^{M_i} \;\middle|\; z_j \geq \varepsilon \ \forall j, \ \sum_{j=1}^{M_i} z_j W^i_j \boldsymbol{l}^i_j = \boldsymbol{0}, \ \sum_{j=1}^{M_i} z_j W^i_j = \frac{N \Pi_i}{(N+1)(N+2)} \right\}.$$

According to well-known results (see, e.g., [8]) such problem is well posed and has a unique solution, since $\mathcal{Z}_\varepsilon$ is a closed convex set.

In all the sequel in this section we follow the main lines of [20].

To begin with we recall two operators denoted by $\Im_h$ and $\Re_h$ associated with $\mathcal{T}_h$. $\Im_h$ is the interpolation operator from $C^0(\bar{\Omega})$ onto V_h, that is, the operator defined by $\Im_h v(P) = v(P)$ for every node P of the partition $\mathcal{T}_h$. $\Re_h$ is the operator from $W^{2,\infty}(\Omega)$ onto V_h, defined as the unique solution of the following problem:

$$\begin{cases} \text{Given } v \in W^{2,\infty}(\Omega) \text{ find } \Re_h v \in V_h \text{ such that } \Re_h v = \Im_h v \text{ on } \partial\Omega, \text{ satisfying} \\ \displaystyle\int_\Omega \nabla \Re_h v \cdot \nabla w = -\int_\Omega \Delta v\, w, \quad \forall w \in V_h^0. \end{cases} \tag{36}$$

According to [16], we may assert that if $v \in W^{2,\infty}(\Omega)$ we have

$$\|\Re_h v - v\|_{0,\infty,\Omega} \le C_0 h^2 |\ln h| \cdot \|v\|_{2,\infty,\Omega}, \tag{37}$$

where C_0 is a constant independent of h. In so doing we further introduce an asymmetric averaging operator around the nodes P_i of $\mathcal{T}_h$, for $i \in \{1, 2, \ldots, I_h\}$, namely, $\mathcal{A}_h \colon C^0(\bar{\Omega}) \to V_h^0$ defined by

$$\mathcal{A}_h v(P_i) := \frac{(N+1)\sum_{j=1}^{M_i} \left[\frac{2}{N} v(P_i) + v(P_{k_j})\right] \omega_j^i W_j^i}{\Pi_i}. \tag{38}$$

We next establish the following preliminary results.

LEMMA 5.2. *Assume that for every $t \in (0, \infty)$ the solution u of (1) belongs to $W^{2,\infty}(\Omega)$, and let u^n be the function of $W^{2,\infty}(\Omega)$ corresponding to the value of u at time $n\Delta t$. Then setting $\tilde{u}_j^n := \Re_h u^n(P_j)$ for $j = 1, 2, \ldots, N_h$, for every $i \in \{1, 2, \ldots, I_h\}$, we have*

$$\tilde{u}_i^n - \sum_{j=1}^{N_h} c_{ij}^n \tilde{u}_j^{n-1} - \Delta t\, \tilde{b}_i^n = \frac{N+1}{\Pi_i} R_i^n(u), \tag{39}$$

where

$$\begin{cases} R_i^n(u) = \alpha_i(u^n) + \beta_i(u^n) + \gamma_i(u^n, u^{n-1}) \\ \qquad\qquad + \rho_i(u^n, u^{n-1}, \boldsymbol{a}^{n-1}, f^{n-1}) + \zeta_i(u^{n-1}, \boldsymbol{a}^{n-1}, f^{n-1}) \\ \text{with} \\ \alpha_i(v) := \displaystyle\int_{S_i} \frac{h_{\min}}{h_{\min} + \nu} [(\Im_h - \mathcal{A}_h)(\Re_h v - v)](P_i)\varphi_i, \\ \beta_i(v) := \displaystyle\int_{S_i} \frac{h_{\min}}{h_{\min} + \nu} [(\Im_h - \mathcal{A}_h) v](P_i)\varphi_i, \\ \gamma_i(v, w) := \displaystyle\int_{S_i} \frac{h_{\min}}{h_{\min} + \nu} \{\mathcal{A}_h[\Re_h(v - w)](P_i) - \mathcal{A}_h(v - w)(P_i)\}\varphi_i, \\ \rho_i(v, w, \boldsymbol{d}, e) := \displaystyle\int_{S_i} \Big\{ \Big[\frac{h_{\min}}{h_{\min} + \nu} \mathcal{A}_h(v - w)(P_i) + \frac{\nu}{h_{\min} + \nu} \Re_h(v - w)(P_i) \\ \qquad\qquad + \Delta t\, (\boldsymbol{d} \cdot \nabla \Re_h w - e) \Big] \varphi_i + \nu \Delta t\, \nabla \Re_h w \cdot \nabla \varphi_i \Big\}, \\ \zeta_i(w, \boldsymbol{d}, e) := \Delta t \displaystyle\int_{S_i} \{[\boldsymbol{d}(P_i) - \boldsymbol{d}] \cdot \nabla \Re_h w + [e - e(P_i)]\}\varphi_i. \end{cases} \tag{40}$$

Proof. First we note that according to (11) and (12), for a given continuous function v we have

$$\sum_{j=1}^{N_h} m_{ij}^W v(P_j) = \frac{h_{\min}}{\nu + h_{\min}} \left[m_{ii}^C v(P_i) + \sum_{j=1}^{M_i} \omega_j^i W_j^i v(P_{k_j}) \right] + \frac{\nu m_{ii}^L}{\nu + h_{\min}} v(P_i).$$

Then recalling (9) together with (13), we obtain,

$$\begin{aligned} \sum_{j=1}^{N_h} m_{ij}^W v(P_j) \; &= \; \frac{h_{\min}}{\nu + h_{\min}} \left\{ \sum_{j=1}^{M_i} \left[\frac{2}{N} v(P_i) + v(P_{k_j}) \right] \omega_j^i W_j^i \right\} \\ &\quad + \frac{\nu \Pi_i}{(\nu + h_{\min})(N+1)} v(P_i). \end{aligned}$$

It follows that $\sum_{j=1}^{N_h} \tilde{m}_{ij} v(P_j) = \frac{\nu}{\nu + h_{\min}} v(P_i) + \frac{h_{\min}}{\nu + h_{\min}} (\mathcal{A}_h v)(P_i)$, taking into account the definition (38) of $\mathcal{A}_h$. Then the proof of this lemma is just a matter of straightforward calculations starting from the definition of coefficients $\tilde{m}_{ij}$, $\tilde{a}_{ij}^n$ and $\tilde{b}_i^n$. $\square$

Lemma 5.3. *Assume again that for every $t \in (0, \infty)$ the solution u of* (1) *belongs to $C^0(\bar{\Omega})$, and let u_t^n denote u_t at time $n\Delta t$. Then recalling* (39) *and* (40), *and denoting by $\boldsymbol{I}$ the identity operator, for every $i \in \{1, 2, \ldots, I_h\}$, we have*

$$\begin{cases} \rho_i(u^n, u^{n-1}, \boldsymbol{a}^{n-1}, f^{n-1}) = \eta_i(u^n, u^{n-1}) + \chi_i(u^n, u^{n-1}) \\ \qquad\qquad\qquad + \sigma_i(u^{n-1}, \boldsymbol{a}^{n-1}) + \tau_i(u^n, u^{n-1}, u_t^{n-1}), \\ \textit{where} \\ \eta_i(v, w) := \displaystyle\int_{S_i} \frac{h_{\min}}{h_{\min} + \nu} [\mathcal{A}_h(v - w) - \Im_h(v - w)](P_i) \varphi_i, \\ \chi_i(v, w) := \displaystyle\int_{S_i} \frac{\nu}{h_{\min} + \nu} [\Re_h(v - w) - \Im_h(v - w)](P_i) \varphi_i, \\ \sigma_i(v, \boldsymbol{d}) := \Delta t \displaystyle\int_{S_i} \boldsymbol{d} \cdot \nabla(\Re_h - \boldsymbol{I}) v \varphi_i, \\ \tau_i(v, w, s) := \displaystyle\int_{S_i} [\Im_h(v - w)(P_i) - \Delta t\, s] \varphi_i. \end{cases} \tag{41}$$

Proof. Setting $v = \varphi_i$ in (2), for $t = (n-1)\Delta t$, and taking into account the definition of $\Re_h$, we obtain, $\forall i \in \{1, 2, \ldots, I_h\}$,

$$\left\{ \int_{S_i} (\partial_t^{n-1} u + \boldsymbol{a}^{n-1} \cdot \nabla u^{n-1} - f^{n-1}) \varphi_i + \nu \int_{\Omega} \nabla \Re_h u^{n-1} \cdot \nabla \varphi_i \right\} = 0. \tag{42}$$

Now we subtract from $\rho_i(u^n, u^{n-1}, \boldsymbol{a}^{n-1}, f^{n-1})$ the expression on the left hand side of (42), multiplied by Δt, and next we add to and subtract from the result the term E^n given by

$$E^n = \int_{S_i} (u^n - u^{n-1})(P_i)\varphi_i. \tag{43}$$

Now it suffices to cancel out identical terms and to group the remaining ones conveniently, to derive (41). □

Before concluding the consistency analysis we give the following technical result.

Lemma 5.4. *Let the weights ω_j^i fulfill* (33) *and* (9), *and v be a given function in $W^{2,\infty}(\Omega)$. Then there exists a constant C_A independent of h and v such that for $\forall i \in \{1, 2, \dots, I_h\}$,*

$$\left| \int_{S_i} [\mathcal{A}_h(\Im_h)(v) - \Im_h(v)](P_i)\varphi_i \right| \leq C_A h^2 \frac{\Pi_i}{N+1} \|v\|_{2,\infty,\Omega}. \tag{44}$$

Proof. Recalling the definition of $\mathcal{A}_h$ and (9) we may write

$$\begin{aligned}\frac{[\mathcal{A}_h(\Im_h)(v) - \Im_h(v)](P_i)\Pi_i}{N+1} &= \sum_{j=1}^{M_i} \omega_j^i W_j^i \left[\frac{2}{N} v(P_i) + v(P_{k_j})\right] \\ &\quad - \frac{N+2}{N} \sum_{j=1}^{M_i} \omega_j^i W_j^i v(P_i).\end{aligned}$$

Hence we have

$$\int_{S_i} [\mathcal{A}_h(\Im_h)(v) - \Im_h(v)](P_i)\varphi_i = \sum_{j=1}^{M_i} \omega_j^i W_j^i [v(P_{k_j}) - v(P_i)].$$

On the other hand $\forall j$ we have $v(P_{k_j}) - v(P_i) = \nabla v(P_i) \cdot \boldsymbol{l}_j^i + G_j^i(v)$, where $G_j^i(v)$ is a term whose absolute value is bounded above by $C_G h^2 \|v\|_{2,\infty,\Omega}$, C_G being a constant independent of h, i, j and v. It follows that

$$\begin{aligned}\left| \int_{S_i} [\mathcal{A}_h(\Im_h)(v) - \Im_h(v)](P_i)\varphi_i \right| &\leq \nabla v(P_i) \cdot \sum_{j=1}^{M_i} \omega_j^i W_j^i \boldsymbol{l}_j^i \\ &\quad + C_G h^2 \frac{N \Pi_i}{(N+1)(N+2)} \|v\|_{2,\infty,\Omega}.\end{aligned}$$

Recalling (33) the result follows with $C_A = C_G \frac{N}{N+2}$. □

Now we are ready to prove:

Theorem 5.5. *Assume that for a given finite time $T > 0$, for any integer l with $0 \leq l \leq 2$ and $\forall t \in [0,T]$, both the solution u of* (1) *and u_t belong to $W^{l,\infty}(\Omega)$ equipped with the standand norm and seminorm. Assume also that $\forall t \in [0,T]$, $\boldsymbol{a}(\,\cdot\,,t) \in [W^{1,\infty}(\Omega)]^N$ and $f(\,\cdot\,,t) \in W^{1,\infty}(\Omega)$, and $(u_t)_t$ belongs to $L^\infty(\Omega)$. Then Δt being given by $\frac{T}{m}$ for a strictly positive integer m, there exists a constant C_R independent of n, h and Δt such that the following estimate applies for every $n \in \{1,2,\dots,m\}$ and $i \in \{1,2,\dots,I_h\}$:*

$$|R_i^n(u)| \leq C_R \frac{\Pi_i}{N+1} \Delta t\, h B^n(u), \tag{45}$$

where

$$\begin{aligned} B^n(u) = {} & \frac{h_{\min}}{h_{\min}+\nu}\frac{h|\ln h|}{\Delta t}\|u^n\|_{2,\infty,\Omega} + \|u_t^{n-1}\|_{1,\infty,\Omega} \\ & + h|\ln h| \max_{s\in[(n-1)\Delta t, n\Delta t]} \|(u_t)_{/t=s}\|_{2,\infty,\Omega} \\ & + \frac{\Delta t}{h} \max_{s\in[(n-1)\Delta t, n\Delta t]} \|(u_{tt})_{/t=s}\|_{0,\infty,\Omega} \\ & + \|\boldsymbol{a}^{n-1}\|_{1,\infty,\Omega}(\|u^{n-1}\|_{1,\infty,\Omega} + h|\ln h|\cdot\|u^{n-1}\|_{2,\infty,\Omega}) \\ & + |\ln h|\cdot\|\boldsymbol{a}^{n-1}\|_{0,\infty,\Omega}\|u^{n-1}\|_{2,\infty,\Omega} + \|f^{n-1}\|_{1,\infty,\Omega}. \end{aligned} \tag{46}$$

Proof. Recalling the expression of $R_i^n(u)$ given by (40) together with (41), it suffices to derive proper bounds term by term with appropriate constants independent of n, h and Δt, denoted by the letter C with subscripts.

To begin with we consider the term α_i. From (37) and the obvious bounds of both $\Im_h w$ and $\mathcal{A}_h w$ by $\|w\|_{0,\infty,\Omega}$ for every $w \in L^\infty(\Omega)$, we have for $C_\alpha = 2C_0$,

$$|\alpha_i(u^n)| \leq C_\alpha h^2 |\ln h| \frac{h_{\min}\Pi_i}{(N+1)(\nu+h_{\min})}\|u^n\|_{2,\infty,\Omega}. \tag{47}$$

From Lemma 5.4, β_i is trivially bounded as follows with $C_\beta = C_A$:

$$|\beta_i(u^n)| \leq C_\beta h^2 \frac{h_{\min}\Pi_i}{(N+1)(\nu+h_{\min})}\|u^n\|_{2,\infty,\Omega}. \tag{48}$$

Again from the boundedness of $\mathcal{A}_h$, the application of $\Re_h$ to the function $u^n - u^{n-1}$ together with (37) yields

$$|\gamma_i(u^n,u^{n-1})| \leq C_0 h^2 |\ln h| \frac{h_{\min}\Pi_i}{(N+1)(\nu+h_{\min})}(\|u^n - u^{n-1}\|_{2,\infty,\Omega}). \tag{49}$$

On the other hand from Lemma 5.4 we have

$$|\eta_i(u^n, u^{n-1})| \leq C_A h^2 \frac{h_{\min} \Pi_i}{(N+1)(\nu + h_{\min})} (\|u^n - u^{n-1}\|_{2,\infty,\Omega}). \tag{50}$$

Moreover, similarly to the case of γ_i, we obtain

$$|\chi_i(u^n, u^{n-1})| \leq C_0 h^2 |\ln h| \frac{\nu \Pi_i}{(N+1)(\nu + h_{\min})} \|u^n - u^{n-1}\|_{2,\infty,\Omega}. \tag{51}$$

Now for an integer $l \in \{0,1,2\}$ we denote by D_{x^μ} the l-th order partial differential operator extended to $l = 0$, with respect to $x_{\lambda_1}^{\mu_1}$ and $x_{\lambda_2}^{\mu_2}$, where for two non negative integers μ_1 and μ_2, $\mu = (\mu_1, \mu_2)$ with $\mu_1 + \mu_2 = l$, and $\lambda = (\lambda_1, \lambda_2)$ with $1 \leq \lambda_1 \leq \lambda_2 \leq N$. In so doing, let $\xi_\mu^n(\boldsymbol{x})$ be a value of t in the interval $[(n-1)\Delta t, n\Delta t]$ such that for every $\boldsymbol{x} \in \Omega$ and for every μ,

$$D_{x^\mu}[u^n - u^{n-1}](\boldsymbol{x}) = \Delta t \, D_{x^\mu} u_t[\xi_\mu^n(\boldsymbol{x}), \boldsymbol{x}].$$

Then clearly enough, for $C_{\gamma\eta\chi} = C_0 + C_A$, we derive

$$\begin{aligned} &|\gamma_i(u^n, u^{n-1})| + |\eta_i(u^n, u^{n-1})| + |\chi_i(u^n, u^{n-1})| \\ &\leq C_{\gamma\eta\chi} \Delta t \, h^2 |\ln h| \frac{\Pi_i}{(N+1)} \max_{s \in [(n-1)\Delta t, n\Delta t]} \|(u_t)_{/t=s}\|_{2,\infty,\Omega}. \end{aligned} \tag{52}$$

Next, since by assumption $\boldsymbol{a}$ is solenoidal and φ_i vanishes on the boundary of S_i, applying integration by parts we easily derive

$$\sigma_i(u^{n-1}, \boldsymbol{a}^{n-1}) = -\Delta t \int_{S_i} \boldsymbol{a}^{n-1} \cdot \nabla \varphi_i (\Re_h - \boldsymbol{I}) u^{n-1}.$$

On the other hand we have $|\nabla \varphi_i| \leq h_{\min}^{-1}$ for every i, and moreover taking into account the assumptions on $\mathcal{T}_h$, there must exist a constant C_S independent of h such that

$$h \leq C_S h_{\min}. \tag{53}$$

It follows that for $C_\sigma = C_0 C_S$ we have

$$|\sigma_i(u^{n-1}, \boldsymbol{a}^{n-1})| \leq C_\sigma \Delta t \, h |\ln h| \frac{\Pi_i}{N+1} \|\boldsymbol{a}^{n-1}\|_{0,\infty,\Omega} \|u^{n-1}\|_{2,\infty,\Omega}. \tag{54}$$

As for τ_i we first observe that for every i, n and $\boldsymbol{x} \in S_i$ there exists a constant C_T such that

$$|u_t^{n-1}(\boldsymbol{x}) - u_t^{n-1}(P_i)| \leq C_T h \|u_t^{n-1}\|_{1,\infty,\Omega}.$$

On the other hand we have

$$|[u^n - u^{n-1}](P_i) - \Delta t \, u_t^{n-1}(P_i)| \leq \frac{\Delta t^2}{2} \max_{s \in [(n-1)\Delta t, n\Delta t]} \|(u_t)_{t/t=s}\|_{0,\infty,\Omega}.$$

It immediately follows the existence of $C_\tau = \max\{C_T, 0.5\}$ such that

$$\begin{aligned}
&|\tau_i(u^n, u^{n-1}, u_t^{n-1})| \\
&\leq C_\tau \frac{\Pi_i}{N+1} \Delta t \Big\{ h \|u_t^{n-1}\|_{1,\infty,\Omega} + \Delta t \max_{s \in [(n-1)\Delta t, n\Delta t]} \|(u_t)_{t/t=s}\|_{0,\infty,\Omega} \Big\}. \quad (55)
\end{aligned}$$

Finally the term ζ_i is handled as follows: First we note that, from the assumptions on $\mathcal{T}_h$, there exists a constant C_Z independent of h and i, such that for every n and $\forall \boldsymbol{x} \in S_i$ we have

$$|f^{n-1}(\boldsymbol{x}) - f^{n-1}(P_i)| \leq C_Z h \|f^{n-1}\|_{1,\infty,\Omega}$$

together with

$$|\boldsymbol{a}^{n-1}(P_i) - \boldsymbol{a}^{n-1}(\boldsymbol{x})| \leq C_Z h \|\boldsymbol{a}^{n-1}\|_{1,\infty,\Omega}.$$

Now we rewrite $\zeta_i(u_{n-1}, \boldsymbol{a}^{n-1}, f^{n-1})$ as the product with Δt of the integral over S_i of the sum of three terms Z_1, Z_2 and Z_3, namely,

- $Z_1 = [f^{n-1} - f^{n-1}(P_i)]\varphi_i$,
- $Z_2 = [\boldsymbol{a}^{n-1}(P_i) - \boldsymbol{a}^{n-1}] \cdot \nabla u^{n-1} \varphi_i$,
- $Z_3 = [\boldsymbol{a}^{n-1}(P_i) - \boldsymbol{a}^{n-1}] \cdot \nabla \varphi_i (\boldsymbol{I} - \Re_h) u^{n-1}$.

Applying arguments already exploited above it is easily seen that

$$\begin{aligned}
&|Z_1 + Z_2 + Z_3| \\
&\leq C_Z h [\|\boldsymbol{a}^{n-1}\|_{1,\infty,\Omega} (\|u^{n-1}\|_{1,\infty,\Omega} + C_0 C_S h |\ln h| \cdot \|u^{n-1}\|_{2,\infty,\Omega}) \\
&\qquad + \|f^{n-1}\|_{1,\infty,\Omega}]. \quad (56)
\end{aligned}$$

It immediately follows from (56) that for $C_\zeta = C_Z \max\{1, C_0 C_S\}$ it holds that

$$\begin{aligned}
&|\zeta_i(u^{n-1}, \boldsymbol{a}^{n-1}, f^{n-1})| \\
&\leq C_\zeta \frac{\Pi_i}{N+1} \Delta t \, h [\|\boldsymbol{a}^{n-1}\|_{1,\infty,\Omega} (\|u^{n-1}\|_{1,\infty,\Omega} + h |\ln h| \cdot \|u^{n-1}\|_{2,\infty,\Omega}) \\
&\qquad + \|f^{n-1}\|_{1,\infty,\Omega}]. \quad (57)
\end{aligned}$$

Putting together (47), (48), (52), (54), (55) and (57), and recalling (40) and (41), we readily derive (45) with (46) for a suitable constant C_R. □

We are now ready to derive error estimates for scheme (7).

THEOREM 5.6. *Let the strictly positive weights* ω_j^i, $\forall i \in \{1, 2, \ldots, I_h\}$ *and* $\forall j \in \{1, 2, \ldots, M_i\}$, *satisfy* (33) *and* (9). *Assume that for a given finite time* $T > 0$ *both the solution* u *of* (1) *and* u_t *belong to* $W^{2,\infty}(\Omega)$. *Assume also that* $\forall t \in [0, T]$, $\boldsymbol{a}(\cdot, t) \in [W^{1,\infty}(\Omega)]^N$ *and* $f(\cdot, t) \in W^{1,\infty}(\Omega)$, *and* $(u_t)_t$ *belongs to* $L^\infty(\Omega)$. *Finally let a strictly positive integer* k_T *be chosen as be the minimum of all integers* k *such* *that the quantity* $\Delta t := \frac{T}{k}$ *fulfills the condition* (16)*. Then there exists a constant* C_E *independent of* h *and* Δt *such that the following estimate applies for every* $m \in \{1, 2, \ldots, k_T\}$*:*

$$\begin{aligned}
&\|\Im_h u^m - u_h^m\|_{0,\infty,\Omega} \\
&\le C_E h \Delta t \sum_{n=1}^{m} \left\{ \frac{h^2 |\ln h|}{\Delta t\,(h_{\min} + \nu)} \|u^n\|_{2,\infty,\Omega} + \|u_t^{n-1}\|_{1,\infty,\Omega} \right. \\
&\qquad + h|\ln h| \max_{s \in [(n-1)\Delta t, n\Delta t]} \|(u_t)_{/t=s}\|_{2,\infty,\Omega} \\
&\qquad + \frac{\Delta t}{h} \max_{s \in [(n-1)\Delta t, n\Delta t]} \|(u_t)_{t/t=s}\|_{0,\infty,\Omega} \\
&\qquad + \|\boldsymbol{a}^{n-1}\|_{1,\infty,\Omega} (\|u^{n-1}\|_{1,\infty,\Omega} + h|\ln h| \cdot \|u^{n-1}\|_{2,\infty,\Omega}) \\
&\qquad \left. + |\ln h| \cdot \|\boldsymbol{a}^{n-1}\|_{0,\infty,\Omega} \|u^{n-1}\|_{2,\infty,\Omega} + \|f^{n-1}\|_{1,\infty,\Omega} \right\}. \qquad (58)
\end{aligned}$$

Proof. Let $\bar{u}_j^n := \tilde{u}_j^n - u_j^n$, for $n = 0, 1, 2, \ldots, k_T$ and $j \in \{1, 2, \ldots, N_h\}$. Clearly enough recalling (39) and (7) we have $\forall i \in \{1, 2, \ldots, I_h\}$,

$$\bar{u}_i^n - \sum_{j=1}^{N_h} c_{ij}^n \bar{u}_j^{n-1} - \frac{N+1}{\Pi_i} R_i^n(u) = 0.$$

Furthermore $\bar{u}_i^n = 0$ $\forall i \in \{I_h + 1, \ldots, N_h\}$. Thus since (16) holds, taking into account (53), application of Theorem 4.2 leads in a straightforward manner to a bound in all similar to (58), except for the discretization step size independent constant C_E, that becomes C_I, and for the left hand side, in which $\Re_h$ replaces $\Im_h$. On the other hand, (37) together with (53), trivially imply the existence of a constant C_1 independent of h such that for every m,

$$\|\Im_h u^m - u_h^m\|_{0,\infty,\Omega} \le \|\Re_h u^m - u_h^m\|_{0,\infty,\Omega} + C_1 \frac{h^2 |\ln h|}{h_{\min} + \nu} \|u^m\|_{2,\infty,\Omega}.$$

Hence the result follows with $C_E = C_I + C_1$. □

Owing to the first term in the summation on the right hand side of (58), we cannot assert that the method converges if only bound (16) holds. Indeed this term is bounded below by $(N+2)A\big[1+(N+1)\frac{\nu}{Ah_{\min}}\big]|\ln h_{\min}| \cdot \|u^n\|_{2,\infty,\Omega}$, which tends to ∞ as h goes to zero, asymptotically like $h^{-1}|\ln h|$. Nevertheless whenever ν is very small, say $\nu \ll A$, which means that convection is largely dominant, provided the mesh step size is also significantly greater than ν, we can expect that the scheme (7) will generate accurate numerical solutions, even if in the limiting process, i.e., as h approaches ν, we can expect that the scheme will fail to reduce discretization errors. However in principle, in practical situations such point is out of reach in the convection largely dominated case.

To conclude we give a convergence result for scheme (7).

Theorem 5.7. *Under the assumptions of* Theorem 5.6 *on the weights* ω_j^i *and on the regularity of the solution* u *of* (1)*, and also of* $\boldsymbol{a}(\,\cdot\,,t)$ *and* $f(\,\cdot\,,t)$ *in the interval* $[0,T]$*, let the strictly positive integer* k_T *be the minimum of all integers* k *such that the quantity* $\Delta t := \frac{T}{k}$ *fulfills the condition* (32)*. Assume that, besides belonging to a quasiuniform family of partitions* $\mathcal{T}_h$ *is of the acute type. Then there exists a constant* C *independent of* u*,* h *and* Δt*, such that the following estimate applies:*

$$\begin{aligned}
&\max_{1\le m\le k_T}\|u^m-u_h^m\|_{0,\infty,\Omega}\\
&\le Ch|\ln h|\max_{0\le s\le T}\{\|u(\,\cdot\,,s)\|_{2,\infty,\Omega}+\|u_t(\,\cdot\,,s)\|_{1,\infty,\Omega}\\
&\qquad\qquad+h\|u_t(\,\cdot\,,s)\|_{2,\infty,\Omega}+h\|(u_t)_t(\,\cdot\,,s)\|_{0,\infty,\Omega}\\
&\qquad\qquad+\|\boldsymbol{a}(\,\cdot\,,s)\|_{1,\infty,\Omega}(\|u(\,\cdot\,,s)\|_{1,\infty,\Omega}+h\|u(\,\cdot\,,s)\|_{2,\infty,\Omega})\\
&\qquad\qquad+\|\boldsymbol{a}(\,\cdot\,,s)\|_{0,\infty,\Omega}\|u(\,\cdot\,,s)\|_{2,\infty,\Omega}+\|f(\,\cdot\,,s)\|_{1,\infty,\Omega}\}.
\end{aligned}\tag{59}$$

Proof. First we note that, from well known results on interpolation theory ([4]), there exists a constant C_P independent of both h, Δt and u, such that for every m we have

$$\|u^m-u_h^m\|_{0,\infty,\Omega}\le\|\Im_h u^m-u_h^m\|_{0,\infty,\Omega}+C_Ph^2\|u^m\|_{2,\infty,\Omega}.$$

On the other hand, since by assumption (32) holds, from (53) we infer the existence of two mesh independent constants c_1 and c_2 such that $c_1h^2\le\Delta t\le c_2h^2$. Moreover in the case of a mesh of the acute type the result (58) trivially holds, as long as Δt satisfies (32). Hence the result (59) becomes a simple consequence of the arguments already employed in the proof of Theorem 5.6, together with straightforward calculations starting from (58). □

6. Numerical experiments

In [18] we give numerical results that, in spite of the simplicity of the test problems, illustrate the good performance of scheme (7), in particular in the explicit iterative solution of stationary problems. Here we apply our methodology to test problems with known analytic solution exhibiting exponential time decay.

6.1. One-dimensional computations

As a first test, we experimented both our scheme and the one corresponding to a combination of the classical lumped mass and consistent mass on the right hand side, with respective weights equal to $\frac{\nu}{\nu+h_{\min}}$ and $\frac{h_{\min}}{\nu+h_{\min}}$ (see, e.g., [10]) for a one-dimensional problem with sharp boundary layers.

More specifically, taking $\Omega = (0,1)$ we wish to solve problem (1) for $f = 1$, $g = 1 - e^{-t}$ and $\boldsymbol{a} = 1$, starting from $u_0 = v_0 + w_0$ where

$$v_0(x) := 1 + x - \frac{1 - e^{\frac{x}{\nu}}}{1 - e^{\frac{1}{\nu}}}$$

and

$$w_0(x) := \frac{1}{e^{-r_1} - e^{-r_2}}[(e^{-r_2} - 1)e^{r_1(x-1)} - (e^{-r_1} - 1)e^{r_2(x-1)}]$$

with $r_1 = \frac{1+\sqrt{1-4\nu}}{2\nu}$ and $r_2 = \frac{1-\sqrt{1-4\nu}}{2\nu}$.

The exact solution given by $u(x,t) = v_0(x) + e^{-t}w_0(x)$ presents a double boundary layer of width $\mathcal{O}(\nu)$ close to the point $x = 1$. The results given below are restricted to a given time, namely, $T = 1$, as they are sufficiently representative of the behavior of the numerical methods being experimented. This is due to the exponential decay of the solution.

In order to figure out the influence of the schemes in the numerical results, we use double precision and non uniform spacial meshes with N_h nodes, N_h being an odd number. The corresponding step sizes are h_i for $i = 1, 2, \ldots, N_h - 1$ where $h_{2k} = h$ and $h_{2k-1} = Rh$, for $k = 1, 2, \ldots, (N_h-1)/2$, R being a real number greater than one. In so doing $h_{\min} = \frac{h}{R} = \frac{2}{(N_h-1)(1+R)}$. Since in the one-dimensional case the mesh is necessarily of the acute type, we determine $\Delta t = T/k_T$ for each value of N_h, as the largest possible value that satisfies (32). Notice that the pair of weights for our method are necessarily either $R/3$ or $R^{-1}/3$, according to the parity of the node subscript. As for the classical method, both weights are equal to $1/3$ for every node.

First we take $\nu = 10^{-2}$ and $R = 4$. We display in Tables 1 and 2 below the relative errors for the indicated values of $N_h - 1$, in $L^\infty(0,1)$ and in $L^2(0,1)$, respectively. Notice that the maximum errors are attained near the abcissa $x = 1$.

Table 1. L^∞ errors for $T = 1$ and $\nu = 10^{-2}$.

$N_h - 1$	Present scheme	Classical scheme
64	26.357 %	27.152 %
128	22.409 %	22.850 %
256	14.263 %	18.610 %
512	9.832 %	14.998 %
1024	6.026 %	11.207 %

Table 2. L^2 errors for $T = 1$ and $\nu = 10^{-2}$.

$N_h - 1$	Present scheme	Classical scheme
64	9.434 %	9.373 %
128	4.397 %	5.243 %
256	2.902 %	4.067 %
512	1.748 %	2.953 %
1024	0.982 %	1.979 %

In order to give an idea of what one can be expected, as far as pointwise convergence is concerned, we also display in Table 3 the absolute errors of the computed approximations of $u(0.5, 1) = 0.89033610$, for different values of N_h.

Table 3. Errors of $u(0.5, T)$ for $T = 1$ and $\nu = 10^{-2}$.

$N_h - 1$	Present scheme	Classical scheme
64	0.03524	0.03541
128	0.01043	0.01496
256	0.00580	0.01019
512	0.00306	0.00656
1024	0.00157	0.00396

Next we take $\nu = 10^{-5}$ and again $R = 4$. We display in Tables 4 and 5 below the relative errors for the indicated values of $N_h - 1$, in $L^\infty(0, 1)$ and in $L^2(0, 1)$, respectively. In Table 6 we give corresponding absolute errors for the computed approximations of $u(0.5, T) = 0.89349008$.

Table 4. L^∞ errors for $T = 1$ and $\nu = 10^{-5}$.

$N_h - 1$	Present scheme	Classical scheme
512	32.617 %	22.494 %
1024	32.548 %	22.400 %
2048	33.277 %	24.640 %
4096	33.177 %	24.448 %
8182	32.667 %	23.740 %

Table 5. L^2 errors for $T = 1$ and $\nu = 10^{-5}$.

$N_h - 1$	Present scheme	Classical scheme
512	2.268 %	1.818 %
1024	1.571 %	1.259 %
2048	1.241 %	1.046 %
4096	0.854 %	0.721 %
8192	0.575 %	0.489 %

Table 6. Errors of $u(0.5, T)$ for $T = 1$ and $\nu = 10^{-5}$.

$N_h - 1$	Present scheme	Classical scheme
512	0.00198	0.00202
1024	0.00074	0.00080
2048	0.00059	0.00059
4096	0.00028	0.00029
8192	0.00018	0.00019

As one can infer from the above results, rather good numerical solutions can be generated even for problems with largely dominant convection, whose solutions present very sharp gradients.

In the case of a moderately dominant convection, i.e., for $\nu = 10^{-2}$, we can assert that the new scheme performs globally better than the classical one, as one might expect. Indeed convergence is observed for the former and significantly more weakly for the latter.

As for the convection largely dominant case with $\nu = 10^{-5}$, it is not possible to observe convergence in the maximum norm as h diminishes, for none of both schemes. However, as we should explain, the maximum errors occur at the grid point next to $x = 1$, and therefore this effect is not surprising at all. Indeed it is a well-known fact that the mesh must be even more refined locally, in order to reduce numerical errors in the interior of such narrow boundary layer, which was not done here. Notice that for this value of ν both schemes seem to converge in the sense of L^2, with a slight advantage of the classical scheme over the present one, at least up to the degree of refinement that we have attained.

6.2. Two-dimensional computations

We further tested our scheme and the one corresponding to the combination of the classical lumped mass and consistent mass on the right hand side, with the same weights as in our one-dimensional computations, for three two-dimensional problems. The domain Ω for all the test problems is the unit square $(0,1) \times (0,1)$, and as sufficiently representative of the computer results we display only those for time $t = 0.1$.

The type of mesh used in the computations is defined as follows: For a given abcissa value $d \in (0,1)$, we first subdivide Ω into four rectangular subdomains, namely,

$$\Omega_1 := (0,d) \times (0,d), \quad \Omega_2 := (d,1) \times (0,d),$$
$$\Omega_3 := (0,d) \times (d,1), \quad \Omega_4 := (d,1) \times (d,1).$$

Letting M_h be a given positive integer, we next subdivide each one of the intervals $(0,d)$ and $(d,1)$ in both directions of the plane into $M_h/2$ equal parts, thereby generating two systems of $M_h + 1$ parallel lines to each coordinate axis, and a corresponding partition of Ω into M_h^2 rectangles. The triangular mesh is then obtained by means of the subdivision of each rectangle of this partition into two triangles taking its diagonal with a negative slope. As for the weights ω_j^i, first we note that for all inner nodes $P_i = (a,b)$ we have $M_i = 6$. Next we take $\omega_j^i = \frac{1}{N+2}$ for all inner nodes such that $a \neq d$ and $b \neq d$. The weights corresponding to inner nodes such that $a = d$ or $b = d$ are determined by solving the minimisation problem stated in Section 5. Notice that there are essentially three types of nodes having d as an abcissa, and therefore we have to determine just three sets of six weights for the whole mesh, depending only on the mesh parameter r given by $R := d/(1-d)$. The particular case of a uniform mesh corresponds to the value $R = 1$, or yet to $d = 0.5$, in which the weights are necessarily equal to $\frac{1}{N+2}$ for all the inner nodes.

- *First test problem*: We consider a problem with linear exact solution in terms of the space variables. More precisely we have $u(x,y,t) = e^{-t}(x+y)$ and we take $\boldsymbol{a} = (1,1)$. For such solution the value of f is independent of ν and is computed accordingly. Here, provided the errors in the approximation of the time exponential together with round-off errors can be neglected, a really consistent scheme with linear finite elements must be able to reproduce the analytical solution exactly. In Tables 7 and 8 below we display the errors of the approximate solution in both the L^∞- and L^2-norms obtained with our scheme and the classical scheme, for $R = 4$ and $M_h = 16$ or $M_h = 32$, and two different values of ν, namely $\nu = 1$ and $\nu = 10^{-5}$, respectively. For both schemes we take the same time step Δt, which incidentally is here again the largest possible value of the form T/K_T for an integer K_T, with $T = 0.1$, satisfying the stability condition (4.3). In order to give an idea of the joint effect of the mesh and the Péclet number in the computational cost, we give in Table 9 the number of time steps K_T necessary to attain time T, corresponding to the different values of ν and M_h, for both schemes.

Table 7. Errors in the L^∞- and L^2-norms for $\nu = 1$.

Error type	Present scheme	Classical scheme
L^∞; $M_h = 16$	0.58922×10^{-6}	$0.13654 \times 10^{+0}$
L^2; $M_h = 16$	0.31387×10^{-6}	0.51440×10^{-1}
L^∞; $M_h = 32$	0.18831×10^{-6}	0.78026×10^{-1}
L^2; $M_h = 32$	0.36951×10^{-7}	0.28946×10^{-1}

Table 8. Errors in the L^∞- and L^2-norms for $\nu = 10^{-5}$.

Error type	Present scheme	Classical scheme
L^∞; $M_h = 16$	0.35978×10^{-4}	$0.41779 \times 10^{+0}$
L^2; $M_h = 16$	0.22387×10^{-4}	$0.13698 \times 10^{+0}$
L^∞; $M_h = 32$	0.24186×10^{-4}	$0.30262 \times 10^{+0}$
L^2; $M_h = 32$	0.14870×10^{-4}	0.86072×10^{-1}

Table 9. Values of K_T.

Case	K_T
$\nu = 1$ and $M_h = 16$	5120
$\nu = 1$ and $M_h = 32$	20480
$\nu = 10^{-5}$ and $M_h = 16$	80
$\nu = 10^{-5}$ and $M_h = 32$	160

As we can infer from Tables 7 and 8, the consistency of the new scheme is confirmed by the numerical results, whereas the lack of consistency of the classical scheme is clearly shown.

Notice that the higher is the Péclet number, the smaller is K_T as expected.

- *Second test problem*: In this test the exact solution is given by $u(x,y,t) = e^{-t}xy$ and we take $\boldsymbol{a} = (y,x)/2$. Here again the value of f is independent of ν and is determined accordingly. In Tables 10, 11, 12 and 13 below we display the relative errors of the approximate solution in either the L^∞- or the L^2-norms as indicated, obtained with our scheme and the classical scheme for $R = 2$ and increasing values of M_h. Tables 10 and 11 correspond to the value $\nu = 1$ and Tables 12 and 13 correspond to the value $\nu = 10^{-5}$. Moreover, in order to illustrate the effect of a non uniform mesh, we also give results obtained with $R = 1$, that is with a uniform mesh having the same number of nodes. Notice that in this case our scheme coincides with the classical one.

Table 10. L^∞ relative errors: $\nu = 1$, $R = 2$ or uniform mesh.

M_h	Present scheme	Classical scheme	$R = 1$
8	$0.10520 \times 10^{+0}$	$0.17385 \times 10^{+0}$	0.38972×10^{-1}
16	0.97713×10^{-1}	$0.10321 \times 10^{+0}$	0.11084×10^{-1}
32	0.47445×10^{-1}	0.56811×10^{-1}	0.57410×10^{-2}
64	0.25705×10^{-1}	0.29906×10^{-1}	0.27483×10^{-2}

Table 11. L^2 relative errors: $\nu = 1$, $R = 2$ or uniform mesh.

M_h	Present scheme	Classical scheme	$R = 1$
8	0.40587×10^{-1}	0.94531×10^{-1}	0.16150×10^{-1}
16	0.26945×10^{-1}	0.56323×10^{-1}	0.46988×10^{-2}
32	0.15733×10^{-1}	0.30942×10^{-1}	0.24474×10^{-2}
64	0.85408×10^{-2}	0.16249×10^{-1}	0.12479×10^{-2}

Table 12. L^∞ relative errors: $\nu = 10^{-5}$, $R = 4$ or uniform mesh.

M_h	Present scheme	Classical scheme	$R = 1$
32	$0.23769 \times 10^{+0}$	$0.40101 \times 10^{+0}$	0.54364×10^{-2}
64	$0.15305 \times 10^{+0}$	$0.28415 \times 10^{+0}$	0.23567×10^{-2}
128	0.89362×10^{-1}	$0.19934 \times 10^{+0}$	0.10837×10^{-2}
256	0.48039×10^{-1}	$0.13754 \times 10^{+0}$	0.50561×10^{-3}

Table 13. L^2 relative errors: $\nu = 10^{-5}$, $R = 4$ or uniform mesh.

M_h	Present scheme	Classical scheme	$R = 1$
32	0.75640×10^{-1}	$0.22298 \times 10^{+0}$	0.92090×10^{-2}
64	0.50552×10^{-1}	$0.13092 \times 10^{+0}$	0.48018×10^{-2}
128	0.30219×10^{-1}	0.74728×10^{-1}	0.24667×10^{-2}
256	0.16841×10^{-1}	0.42084×10^{-1}	0.12561×10^{-2}

Tables 10 through 13 confirm the predicted superiority of the new scheme over the classical one, for both small and large Péclet numbers. It is also clear from the above results that our scheme is much more accurate, in case uniform meshes are employed, but unfortunately this is seldom possible in practical situations.

- *Third test problem*: Finally we deal with a problem whose exact solution presents a sharp boundary layer, close to the boundary edges given by $x = 1$ and $y = 1$. More precisely we take $\boldsymbol{a} = (1,1)/\pi$ and $u(x,y,t) = e^{-t}[s(x)\sin(\pi y) + s(y)\sin(\pi x)]$, where, for $z \in [0,1]$,

$$s(z) := \frac{\left[e^{\frac{\pi^2\nu-1}{\pi\nu}} - 1\right]e^{\pi z} + (e^{\pi} - 1)e^{\frac{(\pi^2\nu-1)(1-z)}{\pi\nu}}}{e^{\frac{2\pi^2\nu-1}{\pi\nu}}}.$$

The corresponding value of the right hand side is given by $f(x,y,t) = e^{-t}[s(x)\cos(\pi y) + s(y)\cos(\pi x)]$, while the boundary values of u are given by $g(0,y,t) = g(1,y,t) = e^{-t}\sin(\pi y)$ and $g(x,0,t) = g(x,1,t) = e^{-t}\sin(\pi x)$. In Tables 14, 15, 16 and 17 below we display the relative errors of the approximate solution in either the L^∞- or the L^2-norms as indicated, obtained with our scheme and the classical scheme for $R = 4$, and increasing values of M_h. Tables 14 and 15 correspond to the value $\nu = 10^{-2}$ and Tables 16 and 17 correspond to the value $\nu = 10^{-5}$. Here again, we also give results obtained with a uniform mesh having the same number of nodes.

Table 14. L^∞ relative errors: $\nu = 10^{-2}$, $R = 4$ or uniform mesh.

M_h	Present scheme	Classical scheme	$R = 1$
32	$0.55253 \times 10^{+0}$	$0.53996 \times 10^{+0}$	$0.40427 \times 10^{+0}$
64	$0.33698 \times 10^{+0}$	$0.32935 \times 10^{+0}$	$0.21973 \times 10^{+0}$
128	$0.24904 \times 10^{+0}$	$0.24454 \times 10^{+0}$	$0.16351 \times 10^{+0}$
256	$0.15691 \times 10^{+0}$	$0.15601 \times 10^{+0}$	$0.10785 \times 10^{+0}$

Table 15. L^2 relative errors: $\nu = 10^{-2}$, $R = 4$ or uniform mesh.

M_h	Present scheme	Classical scheme	$R = 1$
32	$0.44118 \times 10^{+0}$	$0.44715 \times 10^{+0}$	$0.13929 \times 10^{+0}$
64	$0.15759 \times 10^{+0}$	$0.16228 \times 10^{+0}$	0.68628×10^{-1}
128	$0.10290 \times 10^{+0}$	$0.10733 \times 10^{+0}$	0.50245×10^{-1}
256	0.61345×10^{-1}	0.65283×10^{-1}	0.32450×10^{-1}

Table 16. L^∞ relative errors: $\nu = 10^{-5}$, $R = 4$ or uniform mesh.

M_h	Present scheme	Classical scheme	$R = 1$
32	$0.47320 \times 10^{+0}$	$0.46739 \times 10^{+0}$	$0.32787 \times 10^{+0}$
64	$0.47378 \times 10^{+0}$	$0.46936 \times 10^{+0}$	$0.32539 \times 10^{+0}$
128	$0.47439 \times 10^{+0}$	$0.47239 \times 10^{+0}$	$0.32244 \times 10^{+0}$
256	$0.47377 \times 10^{+0}$	$0.47363 \times 10^{+0}$	$0.32327 \times 10^{+0}$

Table 17. L^2 relative errors: $\nu = 10^{-5}$, $R = 4$ or uniform mesh.

M_h	Present scheme	Classical scheme	$R = 1$
32	$0.44118 \times 10^{+0}$	$0.44715 \times 10^{+0}$	$0.14754 \times 10^{+0}$
64	$0.33172 \times 10^{+0}$	$0.33134 \times 10^{+0}$	$0.12560 \times 10^{+0}$
128	$0.24608 \times 10^{+0}$	$0.24582 \times 10^{+0}$	0.99616×10^{-1}
256	$0.17891 \times 10^{+0}$	$0.17963 \times 10^{+0}$	0.74348×10^{-1}

A quick inspection of Tables 13 through 17 shows that, like in the one-dimensional case, the new scheme and the classical scheme behave very similarly in the presence of moderate to sharp boundary layers. This means that in this case, at least for meshes insufficiently refined such as those used in our computations, both schemes are roughly equivalent. On the other hand for sharp boundary layers, it is not surprising at all that the error in the maximum norm is much larger than the error in the L^2-norm (see Tables 4 and 5 too).

7. Conclusions

As a conclusion it is possible to assert that the method studied in this paper is a promising technique for solving, not only time-dependent convection-diffusion problems, but also stationary ones, and this is particularly true of the convection dominant case. First of all this assertion relies on the method's simplicity, since a piecewise linear finite element space discretization is employed. Moreover the method is based on a standard Galerkin formulation, thereby avoiding the addition of SUPG stabilizing terms or the use of upwinding techniques. Other important advantages of the method are low storage requirements, since it deals with explicit time integration. As for the computer time necessary to run the method, we can say that it remains fairly reasonable, as long as the mesh step size is not too small. Notice however that in real life problems, specially in higher dimension spaces, computations with discretization parameters as small as those used in the computations reported in Subsection 6.1 are generally out of reach. Nevertheless, whatever the case, the explicit time integration is a procedure well suited to parallel computations. which can be a good remedy for eventually excessive computational effort. Finally the authors would like to stress the fact that the method's reliability in terms of both stability and convergence is ensured, provided a simple non restrictive geometrical condition related to the spacial mesh, is satisfied by the time step.

Another interesting conclusion on the experiments carried out in this work is that the classical scheme of the type extensively exploited by Kawahara and collaborators (cf. [10], [11], [12]) seems sufficiently accurate in all cases, although it is probably not convergent in the strict mathematical sense for non uniform meshes.

Numerical experiments with our new scheme in situations with physical meaning are underway, and will be notified in due course.

Acknowledgement. The authors would like to thank their colleagues Masahisa Tabata from the Kyushu University and Mutsuto Kawahara from the Chuo University, for highly valuable comments on their research work. They are also very thankful to their colleagues Carlos Tomei and Marcos Azevedo da Siveira from PUC-Rio, the Catholic University of Rio de Janeiro, for fruitful discussions. The first author gratefully acknowledges the finacial support received from Rio de Janeiro state agency FAPERJ under the grant No. E26/152385/06.

References

[1] R.A. Adams, Sobolev Spaces. Academic Press, N.Y., 1975.

[2] M. Baba and M. Tabata, On a conservative upwind finite element scheme for convective-diffusion equations. RAIRO Analyse Numérique, **15-1** (1981), 3–25.

[3] A.N. Brooks and T.J.R. Hughes, The streamline upwind/Petrov–Galerkin formulation for convection dominated flows with particular emphasis on the Navier–Stokes equations. Computer Methods in Applied Mechanics and Engineering, **32** (1982), 199–259.

[4] P.G. Ciarlet, The Finite Element Method for Elliptic Problems. Noth-Holland, Amsterdam, 1978.

[5] R. Codina and O.C. Zienkiewicz, CBS versus GLS stabilization of the incompressible Navier–Stokes equations and the role of the time step as a stabilization parameter. Communications in Numerical Methods in Engineering, **18** (2002), 99–112.

[6] G.B. Dantzig, Linear Programming and Extensions. Princeton University Press, N.J., 1963.

[7] H. Fujii, Some remarks on finite element analysis of time-dependent field problems. Theory and Practice of Finite Element Structural Analysis, University of Tokyo Press, 1973, 91–106.

[8] R. Glowinski, Optimisation. Cours de DEA, Université Paris 6, Paris, 1972.

[9] T. Ikeda, Maximum Principle in Finite Element Models for Convection-Diffusion Phenomena. Lecture Notes in Numerical and Applied Analysis, Vol. 4, H. Fujita and M. Yamaguti eds., North-Holland Mathematical Studies **76**, Kinokunya, Tokyo, 1983.

[10] M. Kawahara, N. Takeuchi and T. Yoshida, Two step finite element methods for Tsunami wave propagation analysis. International Journal of Numerical Methods in Engineering, **12** (1978), 331–351.

[11] M. Kawahara, H. Hirano, K. Tsubota and K. Inagaki, Selective finite element methods for shallow water flow. International Journal of Numerical Methods in Engineering, **2** (1982), 89–112.

[12] M. Kawahara and H. Hirano, A finite element method for high Reynolds number viscous fluid flow using two step explicit scheme. International Journal of Numerical Methods in Fluids, **3** (1983), 137–163.

[13] F. Kikuchi and T. Ushijima, On finite element methods for convection dominated phenomena. Mathematical Methods in the Applied Sciences, **4** (1982), 98–122.

[14] P. Knabner and L. Angermann, Numerical Methods for Elliptic and Parabolic Partial Differential Equations. Springer Verlag, Berlin, 2003.

[15] P. Lesaint and P.A. Raviart, On a finite element method for solving the neutron transport equation. Mathematical Aspects of Finite Elements in Partial Differential Equations, C. de Boor ed., Academic Press, 1974, 89–123.

[16] J. Nitsche, L^∞-convergence of finite element approximations. Lecture Notes in Mathematics, **606**, Springer Verlag, Berlin-Heidelberg, 1977, 261–274.

[17] V. Ruas and A.P. Brasil Jr., Explicit solution of the incompressible Navier–Stokes equations with linear finite elements. Applied Mathematics Letters, **20** (2007), 1005–1010.

[18] V. Ruas and A.P. Brasil Jr., A stable explicit method for time-dependent convection-diffusion equations. Proceedings of ICNAAM, Corfu, Greece, 2007, 480–483.

[19] G. Strang and G.J. Fix, An Analysis of the Finite Element Method. Prentice-Hall, Englewood Cliffs, 1973.

[20] M. Tabata, Uniform convergence of the upwind finite element approximation for semilinear parabolic problems. Journal of Mathematics of the Kyoto University, **18-2** (1978), 327–351.

[21] V. Thomée, Galerkin Finite Element Methods for Parabolic Problems. Springer Series on Computational Mathematics, **25**, Springer Verlag, Berlin-Heidelberg, 1997.