Report no. 97/04

A Posteriori Error Analysis for Systems of Nonlinear Convection–Diffusion Equations

Paul Houston Endre Süli

In this paper we derive an *a posteriori* error estimate for the Lagrange–Galerkin discretisation of a nonlinear system of unsteady convection–diffusion problems in two space–dimensions. The proof of the error estimate is based on strong stability estimates of an associated *linearised* dual problem, together with the Galerkin orthogonality of the finite element method. Based on this *a posteriori* error estimate, we design the corresponding adaptive algorithm to ensure global control of each component of the discretisation error, with respect to a fixed tolerance.

Key words and phrases: A posteriori error analysis, Lagrange–Galerkin finite element methods, adaptive algorithms, nonlinear systems

The first author would like to acknowledge the financial support of the EPSRC.

The work reported here forms part of the research programme of the Oxford–Reading Institute for Computational Fluid Dynamics.

Oxford University Computing Laboratory
Numerical Analysis Group
Wolfson Building
Parks Road
Oxford, England OX1 3QD

May, 1997

Contents

1 Introduction

The aim of this paper is to extend our work on adaptive Lagrange–Galerkin finite element methods for convection–diffusion problems, developed in [7, 8, 9] for linear problems, to a system of nonlinear equations of the form

$$\begin{aligned} \mathbf{u}_t + (\mathbf{a}(\mathbf{x}, t, \mathbf{u}) \cdot \boldsymbol{\nabla})\mathbf{u} - \nu\Delta\mathbf{u} &= \mathbf{f}, \\ \mathbf{u}(\mathbf{x}, 0) &= \mathbf{u}_0(\mathbf{x}), \end{aligned}$$

where $\nu > 0$, in two space–dimensions.

As in the linear case, the *a posteriori* error estimate will be constructed using strong stability estimates of an associated *linearised* dual problem, together with the Galerkin orthogonality of the finite element method. However, the linearisation used to construct the dual problem must be chosen in such a way that an error representation formula, depending only on the numerical solution $\mathbf{u}_h$ and the solution of the dual problem, may be established. Furthermore, we note here that the coefficients of the dual problem will now depend on both the analytical solution $\mathbf{u}$ and the Galerkin approximation $\mathbf{u}_h$; and therefore the strong stability constants C^s arising in the *a posteriori* error estimate will depend on both $\mathbf{u}$ and $\mathbf{u}_h$, i.e. $C^s = C^s(\mathbf{u}, \mathbf{u}_h)$. For related work on nonlinear problems, see Larson [12], Sandboge [14] and Johnson *et al.* [3, 5, 10], for example.

The outline of this paper is as follows: in Section 2 we summarise some of the notational conventions that we shall use. In Section 3 we formally state the model problem to be considered and formulate the Lagrange–Galerkin method for this problem. Then, in Section 4 we derive a global *a posteriori* error estimate for our model problem in the $L^2(0, T; L^2(\Omega)^2)$ norm. Based on this error estimate, in Section 5 we design an adaptive algorithm to ensure global control of the error with respect to a pre-determined tolerance, TOL. Next, in Section 6 we briefly comment on the practical estimation of the strong stability constants arising in the *a posteriori* error estimate. Finally, in Section 7 we summarise the work presented in this paper.

2 Notation and basic definitions

Let $\mathbf{Z}$ denote the set of integers, $\mathbf{N}$ the set of positive integers, $\mathbf{N}_0$ the set of non-negative integers, $\mathbf{R}$ the set of real numbers and $\mathbf{R}^+$ the set of positive real numbers.

Let ω be a bounded open subset of $\mathbf{R}^d$ ($d \in \mathbf{N}$) with boundary $\partial\omega$. For $1 \le p \le \infty$, let $L^p(\omega)$ denote the usual Lebesgue space of real-valued functions with norm $\| \cdot \|_{L^p(\omega)}$. For $p = 2$ and for $u, v \in L^2(\omega)$ we denote by $(\cdot, \cdot)_\omega$ the

$L^2(\omega)$ inner product defined as

$$(u,v)_\omega := \int_\omega u(\mathbf{x})v(\mathbf{x})\mathrm{d}\mathbf{x}.$$

For $\omega = \Omega$, where Ω will be specified later, we denote $\|\cdot\|_{L^2(\Omega)}$ by $\|\cdot\|$, and $(\cdot,\cdot)_\Omega$ by $(\cdot,\cdot)$.

Let $\alpha = (\alpha_1,\ldots,\alpha_d)$, with each $\alpha_i \in \mathbf{N}_0$, $i = 1,\ldots,d$. Further, let $D^\alpha := D_1^{\alpha_1}\ldots D_d^{\alpha_d}$ and $D_j = \partial/\partial x_j$ for $1 \le j \le d$. For $m \in \mathbf{N}_0$, we denote by $C^m(\omega)$ the set of all continuous real-valued functions defined on ω such that $D^\alpha u$ is continuous on ω for all $|\alpha| \le m$. $C^m(\bar{\omega})$ will denote the set of all u in $C^m(\omega)$ such that $D^\alpha u$ can be extended from ω to a continuous function on $\bar{\omega}$ for all $|\alpha| \le m$. In particular, when $m = 0$ we simply write $C(\bar{\omega})$ instead of $C^0(\bar{\omega})$. The subspace $C_0^m(\bar{\omega})$ will denote the set of functions in $C^m(\bar{\omega})$ which have compact support in ω.

For $m \in \mathbf{R}^+ \cup \{0\}$, let $W^{m,p}(\omega)$ denote the classical Sobolev space endowed with the norm $\|\cdot\|_{W^{m,p}(\omega)}$ and the semi-norm $|\cdot|_{W^{m,p}(\omega)}$ (cf. Adams [1]). Further, $W_0^{m,p}(\omega)$ will denote the closure of $C_0^\infty(\omega)$ in the norm of $W^{m,p}(\omega)$. For $p = 2$ we write $H^m(\omega)$ and $H_0^m(\omega)$ for $W^{m,2}(\omega)$ and $W_0^{m,2}(\omega)$, respectively.

Let X be any of the spaces just defined. Then X^2 will denote the topological product $X \times X$.

3 Model problem and discretisation

Given a final time $T > 0$, we consider the following problem: given $\mathbf{f} = (f,g)$ and $\mathbf{u}_0 = (u_0, v_0)$, find $\mathbf{u} = (u,v)$ such that

$$\begin{aligned}
\mathbf{u}_t + (\mathbf{a}(\mathbf{x},t,\mathbf{u})\cdot\boldsymbol{\nabla})\mathbf{u} - \nu\Delta\mathbf{u} &= \mathbf{f}, && \mathbf{x}\in\Omega,\quad t\in(0,T], && (3.1\text{a})\\
\mathbf{u}(\cdot,t) &= \mathbf{0}, && \mathbf{x}\in\partial\Omega,\ t\in[0,T], && (3.1\text{b})\\
\mathbf{u}(\mathbf{x},0) &= \mathbf{u}_0(\mathbf{x}), && \mathbf{x}\in\Omega, && (3.1\text{c})
\end{aligned}$$

where Ω is a bounded convex polygonal domain in $\mathbf{R}^2$ with boundary $\partial\Omega$ and $\nu > 0$. Further, we assume that $\mathbf{a}$, $\mathbf{f}$ and $\mathbf{u}_0$ satisfy the following hypotheses:

H1. $\mathbf{a}(\cdot,\cdot,\cdot)$ is a locally Lipschitz continuous function of all its arguments.

H2. There exists a locally Lipschitz continuous function $\vartheta(\cdot,\cdot)$ such that

$$\begin{aligned}
|\mathbf{a}(\mathbf{x},t,\mathbf{v})| &\le \vartheta(|\mathbf{v}|,t);\\
|\mathbf{a}(\mathbf{x},t,\mathbf{u}) - \mathbf{a}(\mathbf{x},t,\mathbf{v})| &\le \vartheta(|\mathbf{u}|+|\mathbf{v}|,t)|\mathbf{u}-\mathbf{v}|.
\end{aligned}$$

H3. $\mathbf{f} \in C^1(\bar{\Omega}\times[0,T])^2$.

H4. $\mathbf{u}_0 \in [H_0^1(\Omega) \cap H^\eta(\Omega)]^2$ for some $\eta > 1$.

We may now state the following existence and uniqueness result for problem (3.1).

Theorem 3.1 *Suppose that hypotheses H1 to H4 hold; then there exists a unique solution to (3.1), such that*

1. $\mathbf{u} \in C([0,T] \times \bar{\Omega})^2 \cap C(0,T; C^2(\Omega)^2)$ *and*
2. $\mathbf{u}_t \in C((0,T) \times \Omega)^2$.

Proof This result is a special case of the more general existence and uniqueness theorem proved by Hill [6] (see Theorem 3.38) for $\mathbf{a} = \mathbf{a}(\mathbf{x}, t, \mathbf{u}, \boldsymbol{\nabla}\mathbf{u})$. □

We note that under the above hypotheses on $\mathbf{a}$, there exists a 2×2 'quotient' matrix $\mathcal{Q} = (\mathcal{Q})_{ij}$ for $i, j = 1, 2$, satisfying

$$\mathbf{a}(\mathbf{x}, t, \mathbf{v}) - \mathbf{a}(\mathbf{x}, t, \mathbf{w}) = \mathcal{Q}(\mathbf{x}, t, \mathbf{v}, \mathbf{w})(\mathbf{v} - \mathbf{w}) \quad \forall \mathbf{x} \in \Omega \;\; \forall t \in [0, T], \qquad (3.2)$$

for any $\mathbf{v}$ and $\mathbf{w}$, where

$$(\mathcal{Q}(\cdot, \cdot, \mathbf{v}, \mathbf{w}))_{ij} = \int_0^1 \frac{\partial \mathbf{a}_i}{\partial \mathbf{u}_j}(\cdot, \cdot, \theta\mathbf{v} + (1-\theta)\mathbf{w}) \mathrm{d}\theta,$$

and $\mathbf{v}_i$ denotes the ith component of the vector $\mathbf{v}$.

Remark 3.1 *To emphasise the dependence of the convection term* $\mathbf{a}$ *on the solution* $\mathbf{u}$*, we shall sometimes write* $\mathbf{a}(\mathbf{u})$ *in lieu of* $\mathbf{a}(\cdot, \cdot, \mathbf{u})$.

We now formulate the Lagrange–Galerkin method for problem (3.1). However, let us first introduce the following notation.

Let $0 = t^0 < t^1 < \ldots < t^M < t^{M+1} = T$ be a subdivision (not necessarily uniform) of $[0, T]$, with corresponding time intervals $I^n = (t^{n-1}, t^n]$ and time steps $k_n = t^n - t^{n-1}$. For each n, $0 \le n \le M+1$, let $\mathcal{T}^n = \{\kappa\}$ be an admissible subdivision of Ω into closed triangles κ, with corresponding mesh function h_n satisfying

$$c_1 h_\kappa^2 \;\le\; \mathrm{meas}(\kappa) \quad \forall \kappa \in \mathcal{T}^n, \qquad (3.3a)$$

$$c_2 h_\kappa \le h_n(\mathbf{x}) \;\le\; h_\kappa \qquad \forall \mathbf{x} \in \kappa \;\; \forall \kappa \in \mathcal{T}^n, \qquad (3.3b)$$

where $h_\kappa = \mathrm{diam}(\kappa)$ and c_1 and c_2 are positive constants independent of h_n. Further, h is defined to be the global mesh function given by $h(\mathbf{x}, t) = h_n(\mathbf{x})$, for $(\mathbf{x}, t) \in \Omega \times I^n$ and we define the corresponding time step function $k = k(t)$ by $k(t) = k_n,\; t \in I^n$.

For some $n \in \mathbf{N}_0$, we associate with $\mathcal{T}^n$ the set $E^n = \{\tau\}$ consisting of those line segments in $\mathbf{R}^2$ which appear as an edge of some $\kappa \in \mathcal{T}^n$. We also denote by E_i^n, those τ in E^n which are interior to $\bar{\Omega}$ (i.e. not part of $\partial\Omega$).

Let $S^n = \Omega \times I^n$; for $r \in \mathbf{N}$ and $s \in \mathbf{N}_0$ we define the following finite element spaces

$$\begin{aligned} S^{h_n} &= \{v \in C_0(\Omega) : v|_\kappa \in \mathbf{P}_r(\kappa)\ \forall \kappa \in \mathcal{T}^n\}, \\ V^{h_n} &= \{v : v(\mathbf{x}, t)|_{S^n} = \sum_{j=0}^{s} t^j v_j,\ v_j \in S^{h_n}\}, \\ V^h &= \{v : v(\mathbf{x}, t)|_{S^n} \in V^{h_n},\ n = 1, \ldots, M+1\}. \end{aligned}$$

In the following we shall assume that $s = 0$. We note that if $v \in V^h$, then v is continuous in space at any time, but may be discontinuous in time at the discrete time levels t^n. To account for this, we introduce the notation

$$v_\pm^n := \lim_{s \to 0^+} v(t^n \pm s) \quad \text{and} \quad [v^n] := v_+^n - v_-^n.$$

The construction of the Lagrange–Galerkin method involves writing problem (3.1) in a Lagrangian form. To this end, we assume for simplicity that

$$\begin{aligned} \mathbf{u} &\in C\left([0,T]; W_0^{1,\infty}(\Omega)^2\right); \\ \mathbf{a}(\mathbf{x}, t, \mathbf{0}) &= \mathbf{0}, \quad \text{for } \mathbf{x} \in \partial\Omega,\ t \in (0, T]. \end{aligned}$$

Then, also $\mathbf{a}(\mathbf{u}) \in C([0,T]; W_0^{1,\infty}(\Omega)^2)$.

We may now define the particle trajectories, $\mathbf{X}_\mathbf{u}(\mathbf{x}, s; \cdot)$ for $\mathbf{x} \in \Omega$ and $s \in (0, T]$, associated with problem (3.1) as the solution of the following initial value problem

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{X}_\mathbf{u}(\mathbf{x}, s; t) &= \mathbf{a}(\mathbf{X}_\mathbf{u}(\mathbf{x}, s; t), t, \mathbf{u}(\mathbf{X}_\mathbf{u}(\mathbf{x}, s; t), t)), \\ \mathbf{X}_\mathbf{u}(\mathbf{x}, s; s) &= \mathbf{x}. \end{aligned}$$

Further, the material derivative $D_t\mathbf{u}$ may be defined as

$$\begin{aligned} D_t\mathbf{u}(\mathbf{x}, t) &:= \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}(\mathbf{X}_\mathbf{u}(\mathbf{x}, s; t), t)\ |_{s=t} \\ &= \frac{\partial}{\partial t}\mathbf{u}(\mathbf{x}, t) + (\mathbf{a}(\mathbf{x}, t, \mathbf{u}) \cdot \boldsymbol{\nabla})\mathbf{u}(\mathbf{x}, t) \quad \forall \mathbf{x} \in \Omega,\ t \in (0, T]. \end{aligned}$$

Hence, using the material derivative, equation (3.1) may be rewritten in the following (weak) form: find $\mathbf{u}(t) \in V$, such that

$$(D_t\mathbf{u}(\cdot, t), \mathbf{v}) + (\nu\boldsymbol{\nabla}\mathbf{u}(\cdot, t), \boldsymbol{\nabla}\mathbf{v}) = (\mathbf{f}(\cdot, t), \mathbf{v}) \quad \forall \mathbf{v} \in V, \tag{3.4a}$$

$$(\mathbf{u}(\cdot, 0), \mathbf{v}) = (\mathbf{u}_0(\cdot), \mathbf{v}) \quad \forall \mathbf{v} \in V, \tag{3.4b}$$

where $V = H_0^1(\Omega)^2$. The Lagrange–Galerkin time–discretisation involves approximating the material derivative $D_t\mathbf{u}$ by a backward Euler method, giving for $n = 0, \ldots, M$:

$$\begin{aligned}\left(\frac{\mathbf{u}(\cdot, t^{n+1}) - \mathbf{u}(\mathbf{X_u}(\cdot, t^{n+1}; t^n), t^n)}{k_{n+1}}, \mathbf{v}\right) & \\ +(\nu \boldsymbol{\nabla}\mathbf{u}(\cdot, t^{n+1}), \boldsymbol{\nabla}\mathbf{v}) &\approx (\mathbf{f}(\cdot, t^{n+1}), \mathbf{v}) \quad \forall \mathbf{v} \in V, \quad (3.5a)\\ (\mathbf{u}(\cdot, 0), \mathbf{v}) &= (\mathbf{u}_0(\cdot), \mathbf{v}) \quad \forall \mathbf{v} \in V. \quad (3.5b)\end{aligned}$$

If we now let $\mathbf{u}_h^n = \mathbf{u}_h(\cdot, t^n)$, where $\mathbf{u}_h = (u_h, v_h)$ denotes the Galerkin finite element approximation to $\mathbf{u}$; then applying the finite element method to (3.5) yields the Lagrange–Galerkin discretisation of (3.1) as follows: find $\mathbf{u}_h^{n+1} \in \left(S^{h_{n+1}}\right)^2$ for $0 \leq n \leq M$ such that

$$\begin{aligned}\left(\frac{\mathbf{u}_h^{n+1} - \mathbf{u}_h^n(\mathbf{X}_{\mathbf{u}_h}(\cdot, t^{n+1}; t^n))}{k_{n+1}}, \mathbf{v}\right) & \\ +(\nu \boldsymbol{\nabla}\mathbf{u}_h^{n+1}, \boldsymbol{\nabla}\mathbf{v}) &= (\mathbf{f}^{n+1}, \mathbf{v}) \quad \forall \mathbf{v} \in \left(S^{h_{n+1}}\right)^2, \quad (3.6a)\\ (\mathbf{u}_h^0, \mathbf{v}) &= (\mathbf{u}_0, \mathbf{v}) \quad \forall \mathbf{v} \in \left(S^{h_0}\right)^2, \quad (3.6b)\end{aligned}$$

where we define $\mathbf{X}_{\mathbf{u}_h}|_{S^{n+1}}$ for $0 \leq n \leq M$ by

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{X}_{\mathbf{u}_h}(\mathbf{x}, s; t) &= \mathbf{a}(\mathbf{X}_{\mathbf{u}_h}(\mathbf{x}, s; t), t, \mathbf{u}_h(\mathbf{X}_{\mathbf{u}_h}(\mathbf{x}, s; t), t^n)), \quad (3.7a)\\ \mathbf{X}_{\mathbf{u}_h}(\mathbf{x}, s; s) &= \mathbf{x}, \quad (3.7b)\end{aligned}$$

and $\mathbf{f}^{n+1}(\cdot) := \mathbf{f}(\cdot, t^{n+1})$. This is the same approach as that used by Bercovier & Pironneau [2], Pironneau [13] and Süli [15, 16], for example.

Further, integrating (3.6a) with respect to t over I^{n+1}, we obtain the following equivalent formulation: find $\mathbf{u}_h$ such that, for $n = 0, 1, \ldots, M$, $\mathbf{u}_h|_{S^{n+1}} \in \left(V^{h_{n+1}}\right)^2$ and satisfies

$$\begin{aligned}(D_t^h\mathbf{u}_h, \mathbf{v})_{n+1} + (\nu \boldsymbol{\nabla}\mathbf{u}_h, \boldsymbol{\nabla}\mathbf{v})_{n+1} &= (\bar{\mathbf{f}}, \mathbf{v})_{n+1} \quad \forall \mathbf{v} \in \left(V^{h_{n+1}}\right)^2, \quad (3.8a)\\ (\mathbf{u}_{h-}^0, \mathbf{v}) &= (\mathbf{u}_0, \mathbf{v}) \quad \forall \mathbf{v} \in \left(S^{h_0}\right)^2, \quad (3.8b)\end{aligned}$$

where

$$D_t^h\mathbf{u}_h|_{S^{n+1}} := (\mathbf{u}_{h-}(\mathbf{X}_{\mathbf{u}_h}(\mathbf{x}, t^{n+1}; t^{n+1}), t^{n+1}) - \mathbf{u}_{h-}(\mathbf{X}_{\mathbf{u}_h}(\mathbf{x}, t^{n+1}; t^n), t^n))/k_{n+1},$$

and $\bar{\mathbf{f}}|_{S^{n+1}} := \mathbf{f}(\cdot, t^{n+1})$. Further, for $v, w \in L^2(I^{n+1}; L^2(\Omega))$, we have used

$$(v, w)_{n+1} := \int_{t^n}^{t^{n+1}} (v, w)\mathrm{d}t.$$

4 A posteriori error analysis

In this section we shall derive an *a posteriori* estimate for the error $\mathbf{e} = \mathbf{u} - \mathbf{u}_h$, in the $L^2(0,T;L^2(\Omega)^2)$ norm, where $\mathbf{u}$ and $\mathbf{u}_h$ are the solutions of (3.1) and (3.8), respectively. However, before we proceed, we shall first introduce the following notation: given a vector function $\mathbf{v} = (v_1, v_2)$ we define the 2×2 Jacobi matrix $\boldsymbol{\nabla}_m \mathbf{v} = (\boldsymbol{\nabla}_m \mathbf{v})_{ij}$ for $i,j = 1,2$, where

$$(\boldsymbol{\nabla}_m \mathbf{v})_{ij} = \frac{\partial v_i}{\partial x_j};$$

in addition, we write $\boldsymbol{\nabla}_m^T \mathbf{v}$ to denote the transpose of $\boldsymbol{\nabla}_m \mathbf{v}$. Further, for $v, w \in L^2(0,T;L^2(\Omega))$ we define

$$\begin{aligned}(v,w)_Q &:= \sum_{n=0}^{M} \int_{t^n}^{t^{n+1}} (v,w)\mathrm{d}t,\\ \|v\|_Q &:= \left((v,v)_Q\right)^{1/2},\end{aligned}$$

where $Q = \Omega \times (0,T)$.

4.1 Error representation

The linearised (backward) dual problem takes the form: find $\boldsymbol{\phi} = (\phi, \psi)$ such that

$$\begin{aligned}-\boldsymbol{\phi}_t - \boldsymbol{\nabla}\cdot(\mathbf{a}(\mathbf{u})\boldsymbol{\phi}) + \left[\mathcal{Q}^T \boldsymbol{\nabla}_m^T \mathbf{u}_h\right]\boldsymbol{\phi} - \nu\Delta\boldsymbol{\phi} &= \mathbf{e}, \quad \mathbf{x}\in\Omega, \quad t\in[0,T), &(4.1a)\\ \boldsymbol{\phi}(\mathbf{x},t) &= \mathbf{0}, \quad \mathbf{x}\in\partial\Omega,\ t\in[0,T], &(4.1b)\\ \boldsymbol{\phi}(\mathbf{x},T) &= \mathbf{0}, \quad \mathbf{x}\in\Omega, &(4.1c)\end{aligned}$$

where

$$\boldsymbol{\nabla}\cdot(\mathbf{a}\boldsymbol{\phi}) := \begin{pmatrix} \boldsymbol{\nabla}\cdot(\mathbf{a}\phi) \\ \boldsymbol{\nabla}\cdot(\mathbf{a}\psi) \end{pmatrix}.$$

Alternatively, if we let $\mathcal{Q}_1^c$ and $\mathcal{Q}_2^c$ denote the columns of the matrix $\mathcal{Q}$, respectively, then we may write (4.1a) in the following component form,

$$\begin{aligned}-\phi_t - \boldsymbol{\nabla}\cdot(\mathbf{a}(\mathbf{u})\phi) + (\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}u_h)\,\phi + (\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}v_h)\,\psi - \nu\Delta\phi &= u - u_h, &(4.2a)\\ -\psi_t - \boldsymbol{\nabla}\cdot(\mathbf{a}(\mathbf{u})\psi) + (\mathcal{Q}_2^c\cdot\boldsymbol{\nabla}u_h)\,\phi + (\mathcal{Q}_2^c\cdot\boldsymbol{\nabla}v_h)\,\psi - \nu\Delta\psi &= v - v_h, &(4.2b)\end{aligned}$$

for $\mathbf{x}\in\Omega$ and $t\in[0,T)$.

In the following theorem, we shall establish the existence, uniqueness and regularity of the solution $\boldsymbol{\phi}$ to the dual problem (4.1); for completeness, we shall explicitly indicate the regularity of the data of the dual problem (4.1) needed for the statement of the theorem to hold, although, we note that these conditions are automatically satisfied.

Theorem 4.1 *Suppose that the following regularity assumptions hold:*

$$\begin{aligned} \mathbf{e} &\in L^1(0,T;L^2(\Omega)^2); \\ |\mathbf{a}(\mathbf{u})|^2 &\in L^2(0,T;L^2(\Omega)); \\ |\mathcal{Q}^T \boldsymbol{\nabla}_m^T \mathbf{u}_h|_F &\in L^2(0,T;L^2(\Omega)); \end{aligned}$$

where $|\cdot|$ *denotes the vector two–norm,* $|\cdot|_F$ *the Frobenius norm for matrices and* Ω *is an open bounded domain in* $\mathbf{R}^2$*. Then, (4.1) has a unique (weak) solution* $\boldsymbol{\phi} \in C([0,T];L^2(\Omega)^2) \cap L^2(0,T;H_0^1(\Omega)^2)$*. Further, if* Ω *is convex and*

$$\begin{aligned} \mathbf{e} &\in L^2(0,T;L^2(\Omega)^2); \\ \mathbf{a}(\mathbf{u}) &\in L^\infty(0,T;W^{1,\infty}(\Omega)^2); \\ \left[\mathcal{Q}^T \boldsymbol{\nabla}_m^T \mathbf{u}_h\right]_{ij} &\in L^\infty(0,T;L^\infty(\Omega)) \quad \text{for } i,j = 1,2; \end{aligned}$$

then $\boldsymbol{\phi} \in H^1(0,T;L^2(\Omega)^2) \cap L^2(0,T;\left[H_0^1(\Omega) \cap H^2(\Omega)\right]^2)$.

Proof The first part of the theorem follows from Ladyženskaya *et al.* [11] (Theorem 1.1, Chapter VII). The remaining part of the proof is essentially identical to the proof of Theorem 4.1 in [7]; here, we follow the steps in the proof of Theorem 4.1 for each component of $\boldsymbol{\phi}$ separately, and make use of Lemma 4.7 (see Section 4.3). □

We shall now proceed to prove the following error representation: multiplying (4.2a) by $u - u_h$ and integrating by parts in both space and time, we get

$$\begin{aligned} &\|u - u_h\|_Q^2 \\ &= \sum_{n=0}^{M} \int_{t^n}^{t^{n+1}} (u - u_h, -\phi_t - \boldsymbol{\nabla} \cdot (\mathbf{a}(\mathbf{u})\phi) - \nu \Delta \phi)\, \mathrm{d}t \\ &\quad + \sum_{n=0}^{M} \int_{t^n}^{t^{n+1}} (u - u_h, (\mathcal{Q}_1^c \cdot \boldsymbol{\nabla} u_h)\, \phi + (\mathcal{Q}_1^c \cdot \boldsymbol{\nabla} v_h)\, \psi)\, \mathrm{d}t \\ &= \sum_{n=0}^{M} \int_{t^n}^{t^{n+1}} (u_t + \mathbf{a}(\mathbf{u}) \cdot \boldsymbol{\nabla}(u - u_h), \phi) \mathrm{d}t + \sum_{n=0}^{M} \int_{t^n}^{t^{n+1}} (\nu \boldsymbol{\nabla}(u - u_h), \boldsymbol{\nabla}\phi) \mathrm{d}t \\ &\quad - \sum_{n=0}^{M} ([u_h^n], \phi(t^n)) + (u_0 - u_{h-}^0, \phi(0)) \\ &\quad + \sum_{n=0}^{M} \int_{t^n}^{t^{n+1}} (u - u_h, (\mathcal{Q}_1^c \cdot \boldsymbol{\nabla} u_h)\, \phi + (\mathcal{Q}_1^c \cdot \boldsymbol{\nabla} v_h)\, \psi)\, \mathrm{d}t \\ &= \sum_{n=0}^{M} \int_{t^n}^{t^{n+1}} (f - \mathbf{a}(\mathbf{u}_h) \cdot \boldsymbol{\nabla} u_h, \phi) \mathrm{d}t - \sum_{n=0}^{M} \int_{t^n}^{t^{n+1}} (\nu \boldsymbol{\nabla} u_h, \boldsymbol{\nabla}\phi) \mathrm{d}t \end{aligned}$$

$$
\begin{aligned}
&-\sum_{n=0}^{M}([u_h^n],\phi(t^n)) + (u_0 - u_{h-}^0,\phi(0)) - \sum_{n=0}^{M}\int_{t^n}^{t^{n+1}}([\mathbf{a}(\mathbf{u})-\mathbf{a}(\mathbf{u}_h)]\cdot\boldsymbol{\nabla} u_h,\phi)\,\mathrm{d}t \\
&+\sum_{n=0}^{M}\int_{t^n}^{t^{n+1}}(u-u_h,(\mathcal{Q}_1^c\cdot\boldsymbol{\nabla} u_h)\,\phi + (\mathcal{Q}_1^c\cdot\boldsymbol{\nabla} v_h)\,\psi)\,\mathrm{d}t, \qquad (4.3)
\end{aligned}
$$

where we have used (3.4a). Similarly, multiplying (4.2b) by $v-v_h$ and integrating by parts in both space and time, we get

$$
\begin{aligned}
&\|v-v_h\|_Q^2 \\
&= \sum_{n=0}^{M}\int_{t^n}^{t^{n+1}}(g-\mathbf{a}(\mathbf{u}_h)\cdot\boldsymbol{\nabla} v_h,\psi)\mathrm{d}t - \sum_{n=0}^{M}\int_{t^n}^{t^{n+1}}(\nu\boldsymbol{\nabla} v_h,\boldsymbol{\nabla}\psi)\mathrm{d}t \\
&\quad -\sum_{n=0}^{M}([v_h^n],\psi(t^n)) + (v_0 - v_{h-}^0,\psi(0)) - \sum_{n=0}^{M}\int_{t^n}^{t^{n+1}}([\mathbf{a}(\mathbf{u})-\mathbf{a}(\mathbf{u}_h)]\cdot\boldsymbol{\nabla} v_h,\psi)\,\mathrm{d}t \\
&\quad +\sum_{n=0}^{M}\int_{t^n}^{t^{n+1}}(v-v_h,(\mathcal{Q}_2^c\cdot\boldsymbol{\nabla} u_h)\,\phi + (\mathcal{Q}_2^c\cdot\boldsymbol{\nabla} v_h)\,\psi)\,\mathrm{d}t. \qquad (4.4)
\end{aligned}
$$

Hence, adding (4.3) and (4.4), and using (3.2) gives

$$
\begin{aligned}
&\|\mathbf{u}-\mathbf{u}_h\|_Q^2 \\
&= (f-\mathbf{a}(\mathbf{u}_h)\cdot\boldsymbol{\nabla} u_h,\phi)_Q - (\nu\boldsymbol{\nabla} u_h,\boldsymbol{\nabla}\phi)_Q - \sum_{n=0}^{M}([u_h^n],\phi(t^n)) + (u_0-u_{h-}^0,\phi(0)) \\
&\quad +(g-\mathbf{a}(\mathbf{u}_h)\cdot\boldsymbol{\nabla} v_h,\psi)_Q - (\nu\boldsymbol{\nabla} v_h,\boldsymbol{\nabla}\psi)_Q - \sum_{n=0}^{M}([v_h^n],\psi(t^n)) + (v_0-v_{h-}^0,\psi(0)) \\
&\quad + ([\mathcal{Q}(\mathbf{u}-\mathbf{u}_h)]\cdot\boldsymbol{\nabla} u_h,\phi)_Q - ([\mathbf{a}(\mathbf{u})-\mathbf{a}(\mathbf{u}_h)]\cdot\boldsymbol{\nabla} u_h,\phi)_Q \\
&\quad + ([\mathcal{Q}(\mathbf{u}-\mathbf{u}_h)]\cdot\boldsymbol{\nabla} v_h,\psi)_Q - ([\mathbf{a}(\mathbf{u})-\mathbf{a}(\mathbf{u}_h)]\cdot\boldsymbol{\nabla} v_h,\psi)_Q \\
&= (f-\mathbf{a}(\mathbf{u}_h)\cdot\boldsymbol{\nabla} u_h,\phi)_Q - (\nu\boldsymbol{\nabla} u_h,\boldsymbol{\nabla}\phi)_Q - \sum_{n=0}^{M}([u_h^n],\phi(t^n)) + (u_0-u_{h-}^0,\phi(0)) \\
&\quad +(g-\mathbf{a}(\mathbf{u}_h)\cdot\boldsymbol{\nabla} v_h,\psi)_Q - (\nu\boldsymbol{\nabla} v_h,\boldsymbol{\nabla}\psi)_Q - \sum_{n=0}^{M}([v_h^n],\psi(t^n)) + (v_0-v_{h-}^0,\psi(0)).
\end{aligned}
$$

If we now let $\boldsymbol{\phi}_h = (\phi_h,\psi_h) \in \left(V^h\right)^2$, then using (3.8) we have

$$
\begin{aligned}
\|\mathbf{e}\|_Q^2 &= \|\mathbf{u}-\mathbf{u}_h\|_Q^2 \\
&= \sum_{n=0}^{M}\int_{t^n}^{t^{n+1}}\sum_{\kappa\in\mathcal{T}^{n+1}}([u_h^n]/k_{n+1} + \mathbf{a}(\mathbf{u}_h)\cdot\boldsymbol{\nabla} u_h - \nu\Delta u_h - f,\phi_h-\phi)_\kappa\,\mathrm{d}t \\
&\quad +\sum_{n=0}^{M}\int_{t^n}^{t^{n+1}}\left(\sum_{\kappa\in\mathcal{T}^{n+1}}(\nu\Delta u_h,\phi_h-\phi)_\kappa + (\nu\boldsymbol{\nabla} u_h,\boldsymbol{\nabla}(\phi_h-\phi))\right)\mathrm{d}t
\end{aligned}
$$

$$
\begin{aligned}
&+ \sum_{n=0}^{M} \int_{t^n}^{t^{n+1}} (D_t^h u_h - ([u_h^n]/k_{n+1} + \mathbf{a}(\mathbf{u}_h)\cdot\boldsymbol{\nabla} u_h), \phi_h)\mathrm{d}t \\
&+ \sum_{n=0}^{M} \int_{t^n}^{t^{n+1}} ([u_h^n]/k_{n+1}, \phi - \phi(t^n))\mathrm{d}t + (f - \bar{f}, \phi_h)_Q + (u_0 - u_{h-}^0, \phi(0)) \\
&+ \sum_{n=0}^{M} \int_{t^n}^{t^{n+1}} \sum_{\kappa\in\mathcal{T}^{n+1}} ([v_h^n]/k_{n+1} + \mathbf{a}(\mathbf{u}_h)\cdot\boldsymbol{\nabla} v_h - \nu\Delta v_h - g, \psi_h - \psi)_\kappa \,\mathrm{d}t \\
&+ \sum_{n=0}^{M} \int_{t^n}^{t^{n+1}} \left(\sum_{\kappa\in\mathcal{T}^{n+1}} (\nu\Delta v_h, \psi_h - \psi)_\kappa + (\nu\boldsymbol{\nabla} v_h, \boldsymbol{\nabla}(\psi_h - \psi)) \right) \mathrm{d}t \\
&+ \sum_{n=0}^{M} \int_{t^n}^{t^{n+1}} (D_t^h v_h - ([v_h^n]/k_{n+1} + \mathbf{a}(\mathbf{u}_h)\cdot\boldsymbol{\nabla} v_h), \psi_h)\mathrm{d}t \\
&+ \sum_{n=0}^{M} \int_{t^n}^{t^{n+1}} ([v_h^n]/k_{n+1}, \psi - \psi(t^n))\mathrm{d}t + (g - \bar{g}, \psi_h)_Q + (v_0 - v_{h-}^0, \psi(0)) \\
\equiv\ & \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV} + \mathrm{V} + \mathrm{VI} + \mathrm{VII} + \mathrm{VIII} + \mathrm{IX} + \mathrm{X} + \mathrm{XI} + \mathrm{XII} \qquad (4.5)
\end{aligned}
$$

$\forall \boldsymbol{\phi}_h \in \left(V^h\right)^2$.

4.2 Interpolation/projection estimates for the dual problem

We shall now define $\boldsymbol{\phi}_h \in \left(V^h\right)^2$ in (4.5) to be the quasi-interpolant of $\boldsymbol{\phi}$ in space and the L^2-projection of $\boldsymbol{\phi}$ in time, cf. [9]; i.e. we first define the spatial operator

$$\tilde{\mathcal{I}}_n : L^1(\Omega) \to S^{h_n},$$

to be the quasi-interpolation operator introduced in [9]. Secondly, we define the operator

$$\pi_n : L^2(I^n) \to \mathbf{P}_0(I^n),$$

to be the temporal L^2-projection defined by

$$\int_{t^{n-1}}^{t^n} (\pi_n \phi - \phi) v \mathrm{d}t = 0 \quad \forall v \in \mathbf{P}_0(I^n). \qquad (4.6)$$

Then, we can define (locally) $\boldsymbol{\phi}_h|_{S^n} \in \left(V^{h_n}\right)^2$ by letting

$$\boldsymbol{\phi}_h|_{S^n} = \tilde{\mathcal{I}}_n \pi_n \boldsymbol{\phi} = \pi_n \tilde{\mathcal{I}}_n \boldsymbol{\phi} \in \left(V^{h_n}\right)^2,$$

where $\boldsymbol{\phi} = \boldsymbol{\phi}|_{S^n}$. Further, we introduce $\tilde{\mathcal{I}}$ and π by

$$(\tilde{\mathcal{I}}\boldsymbol{\phi})|_{S^n} = \tilde{\mathcal{I}}_n(\boldsymbol{\phi}|_{S^n}), \qquad (4.7a)$$

$$(\pi\boldsymbol{\phi})|_{S^n} = \pi_n(\boldsymbol{\phi}|_{S^n}), \qquad (4.7b)$$

and we let $\boldsymbol{\phi}_h \in \left(V^h\right)^2$ be

$$\boldsymbol{\phi}_h = \tilde{\mathcal{I}}\pi\boldsymbol{\phi} = \pi\tilde{\mathcal{I}}\boldsymbol{\phi} \in \left(V^h\right)^2.$$

We now give the following error estimates for the operators $\tilde{\mathcal{I}}$ and π in order to estimate $\phi - \phi_h = \phi - \tilde{\mathcal{I}}\pi\phi$; we refer to [9] for the proofs of these estimates. However, let us first introduce the following notation: for the rest of this section, we shall write Φ and Φ_h to denote either ϕ and ϕ_h, respectively, or ψ and ψ_h, respectively. Further, for each edge $\tau \in E_i^n$, let $\mathbf{n}_\tau$ denote the unit normal to τ in the outward direction to κ, and define for $v \in S^{h_n}$ (for some $n \in \mathbf{N}_0$),

$$\left[\frac{\partial v}{\partial \mathbf{n}_\tau}\right] = \lim_{s\to 0^+} \left(\boldsymbol{\nabla} v(\mathbf{x} + s\mathbf{n}_\tau) - \boldsymbol{\nabla} v(\mathbf{x} - s\mathbf{n}_\tau)\right) \cdot \mathbf{n}_\tau, \quad \mathbf{x} \in \tau,$$

that is, $[\partial v/\partial \mathbf{n}_\tau]$ is the jump across τ in the normal component of $\boldsymbol{\nabla} v$. Finally, we introduce the discrete second derivatives

$$D_h^2 v|_\kappa = \sum_{\tau\in\partial\kappa\cap E_i^n} \left\| \left[\frac{\partial v}{\partial \mathbf{n}_\tau}\right] \right\|_{L^\infty(\tau)} \frac{1}{h_\kappa}, \quad \kappa \in \mathcal{T}^n.$$

Lemma 4.2 *Suppose that $\mathcal{T}^n$ satisfies conditions (3.3) for $n = 0, 1, \ldots, M+1$; then there exists a positive constant $C_1^{\tilde{i}}$ such that*

$$\|\Phi_h\|_Q \le C_1^{\tilde{i}} \|\Phi\|_Q.$$

Moreover, for $w \in H_0^1(\Omega) \cap H^2(\Omega)$, there exists positive constants $C_2^{\tilde{i}}$ and $C_3^{\tilde{i}}$ such that

$$\left(\sum_{\kappa\in\mathcal{T}^n} h_\kappa^{-4} \|w - \tilde{\mathcal{I}}_n w\|_{L^2(\kappa)}^2\right)^{1/2} \le C_2^{\tilde{i}} |w|_{H^2(\Omega)},$$

$$\left(\sum_{\kappa\in\mathcal{T}^n} h_\kappa^{-2} |w - \tilde{\mathcal{I}}_n w|_{H^1(\kappa)}^2\right)^{1/2} \le C_3^{\tilde{i}} |w|_{H^2(\Omega)},$$

respectively.

Lemma 4.3 *Suppose that $R \in L^2(0,T;L^2(\Omega))$ and $w \in V^h$ then*

$$|(R, \tilde{\mathcal{I}}\Phi - \Phi)_Q| \le C_1^{\tilde{p}} \|h^2 R\|_Q \|\Delta\Phi\|_Q,$$

$$\left| \sum_{n=0}^{M} \int_{t^n}^{t^{n+1}} \left(\sum_{\kappa\in\mathcal{T}^{n+1}} (\nu\Delta w, \tilde{\mathcal{I}}_{n+1}\Phi - \Phi)_\kappa + (\nu\boldsymbol{\nabla} w, \boldsymbol{\nabla}(\tilde{\mathcal{I}}_{n+1}\Phi - \Phi)) \right) \mathrm{d}t \right| \le C_2^{\tilde{p}} \|h^2 D_h^2 w\|_Q \|\nu\Delta\Phi\|_Q,$$

where $C_1^{\tilde{p}} = C_2^{\tilde{i}} c_2^{-2}$, $C_2^{\tilde{p}} = C^t(C_2^{\tilde{i}} c_2^{-1} + C_3^{\tilde{i}})/(2c_2^2)$ and $C^t = 4\sqrt{2}/c_1$.

Lemma 4.4 *Suppose that $R \in L^2(0,T;L^2(\Omega))$ and $w \in V^h$ then*

$$
\begin{aligned}
|(R, \tilde{\mathcal{I}}(\pi\Phi - \Phi))_Q| &\leq C_3^{\tilde{p}} \|kR\|_Q \|\Phi_t\|_Q, \\
\left| \sum_{n=0}^{M} \int_{t^n}^{t^{n+1}} (R, \Phi^n - \Phi) \mathrm{d}t \right| &\leq C_4^{\tilde{p}} \|kR\|_Q \|\Phi_t\|_Q, \\
(\nu \boldsymbol{\nabla} w, \boldsymbol{\nabla} \tilde{\mathcal{I}}(\pi\Phi - \Phi))_Q &= 0, \\
\sum_{n=0}^{M} \int_{t^n}^{t^{n+1}} \left(\sum_{\kappa \in \mathcal{T}^{n+1}} (\nu \Delta w, \tilde{\mathcal{I}}_{n+1}(\pi_{n+1}\Phi - \Phi))_\kappa \right) \mathrm{d}t &= 0,
\end{aligned}
$$

where $C_3^{\tilde{p}} = C_1^{\tilde{i}} C_1^{i}$, $C_1^{i} = 1/\sqrt{2}$, $C_4^{\tilde{p}} = 1$ and $\Phi^n = \Phi(\mathbf{x}, t^n)$.

4.3 Strong stability of the dual problem

In this section we derive strong stability estimates for the dual problem (4.1); to simplify the notation, throughout this section we shall suppress the dependence of $\mathbf{a}$ on $\mathbf{u}$ by writing $\mathbf{a}$ in lieu of $\mathbf{a}(\mathbf{u})$.

Lemma 4.5 *Let $\boldsymbol{\phi}$ be the solution of (4.1); then there exists a constant $C_1^s = C_1^s(\mathbf{u}, \mathbf{u}_h, T, \mathbf{a})$ such that*

$$
\|\boldsymbol{\phi}\|^2_{L^\infty(0,T;L^2(\Omega))} + 2\|\nu^{1/2} \boldsymbol{\nabla} \boldsymbol{\phi}\|_Q^2 \leq C_1^s \|\mathbf{e}\|_Q^2,
$$

where

$$
\|\nu^{1/2} \boldsymbol{\nabla} \boldsymbol{\phi}\|_Q^2 := \|\nu^{1/2} \boldsymbol{\nabla} \phi\|_Q^2 + \|\nu^{1/2} \boldsymbol{\nabla} \psi\|_Q^2,
$$

$C_1^s = 2\,exp\left(\int_0^T m_1(t)\mathrm{d}t\right)$ *and*

$$
\begin{aligned}
& m_1(t) \\
& = 1 + \|\boldsymbol{\nabla} \cdot \mathbf{a}(t) - 2\left(\mathcal{Q}_1^c \cdot \boldsymbol{\nabla} u_h\right)(t)\|_{L^\infty(\Omega)} + \|\boldsymbol{\nabla} \cdot \mathbf{a}(t) - 2\left(\mathcal{Q}_2^c \cdot \boldsymbol{\nabla} v_h\right)(t)\|_{L^\infty(\Omega)} \\
& \quad + \|\left(\mathcal{Q}_2^c \cdot \boldsymbol{\nabla} u_h\right)(t) + \left(\mathcal{Q}_1^c \cdot \boldsymbol{\nabla} v_h\right)(t)\|_{L^\infty(\Omega)}. \qquad (4.8)
\end{aligned}
$$

Proof Multiply (4.2a) by ϕ and integrate over Ω to obtain,

$$
\begin{aligned}
& -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\phi(t)\|^2 - \frac{1}{2}(\boldsymbol{\nabla} \cdot \mathbf{a}(t)\phi(t), \phi(t)) + \left(\left(\mathcal{Q}_1^c \cdot \boldsymbol{\nabla} u_h\right)(t)\phi(t), \phi(t)\right) \\
& \quad + \left(\left(\mathcal{Q}_1^c \cdot \boldsymbol{\nabla} v_h\right)(t)\psi(t), \phi(t)\right) + \|\nu^{1/2}\boldsymbol{\nabla}\phi(t)\|^2 = (u(t) - u_h(t), \phi(t)). \quad (4.9)
\end{aligned}
$$

Similarly, multiplying (4.2b) by ψ and integrating over Ω we obtain,

$$
\begin{aligned}
& -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\psi(t)\|^2 - \frac{1}{2}(\boldsymbol{\nabla} \cdot \mathbf{a}(t)\psi(t), \psi(t)) + \left(\left(\mathcal{Q}_2^c \cdot \boldsymbol{\nabla} u_h\right)(t)\phi(t), \psi(t)\right) \\
& \quad + \left(\left(\mathcal{Q}_2^c \cdot \boldsymbol{\nabla} v_h\right)(t)\psi(t), \psi(t)\right) + \|\nu^{1/2}\boldsymbol{\nabla}\psi(t)\|^2 = (v(t) - v_h(t), \psi(t)). \quad (4.10)
\end{aligned}
$$

Adding (4.9) and (4.10) together and applying the Cauchy–Schwarz inequality and Hölder's inequality, gives

$$
\begin{aligned}
-\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{\phi}(t)\|^2 &+ \|\nu^{1/2}\boldsymbol{\nabla}\boldsymbol{\phi}(t)\|^2 \\
&\le \|u(t)-u_h(t)\|\|\phi(t)\| + \|v(t)-v_h(t)\|\|\psi(t)\| \\
&\quad + \left\|\frac{1}{2}\boldsymbol{\nabla}\cdot\mathbf{a}(t) - \left(\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}u_h\right)(t)\right\|_{L^\infty(\Omega)}\|\phi(t)\|^2 \\
&\quad + \left\|\frac{1}{2}\boldsymbol{\nabla}\cdot\mathbf{a}(t) - \left(\mathcal{Q}_2^c\cdot\boldsymbol{\nabla}v_h\right)(t)\right\|_{L^\infty(\Omega)}\|\psi(t)\|^2 \\
&\quad + \left\|\left(\mathcal{Q}_2^c\cdot\boldsymbol{\nabla}u_h\right)(t) + \left(\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}v_h\right)(t)\right\|_{L^\infty(\Omega)}\|\phi(t)\|\|\psi(t)\| \\
&\le \frac{1}{2}\|\mathbf{e}(t)\|^2 \\
&\quad + \frac{1}{2}\Big(1 + \|\boldsymbol{\nabla}\cdot\mathbf{a}(t) - 2\left(\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}u_h\right)(t)\|_{L^\infty(\Omega)} \\
&\qquad\qquad + \|\left(\mathcal{Q}_2^c\cdot\boldsymbol{\nabla}u_h\right)(t) + \left(\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}v_h\right)(t)\|_{L^\infty(\Omega)}\Big)\|\phi(t)\|^2 \\
&\quad + \frac{1}{2}\Big(1 + \|\boldsymbol{\nabla}\cdot\mathbf{a}(t) - 2\left(\mathcal{Q}_2^c\cdot\boldsymbol{\nabla}v_h\right)(t)\|_{L^\infty(\Omega)} \\
&\qquad\qquad + \|\left(\mathcal{Q}_2^c\cdot\boldsymbol{\nabla}u_h\right)(t) + \left(\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}v_h\right)(t)\|_{L^\infty(\Omega)}\Big)\|\psi(t)\|^2.
\end{aligned}
$$

Now, integrating with respect to time over the interval (t,T) and using (4.1c), we get

$$
\|\boldsymbol{\phi}(t)\|^2 + 2\int_t^T \|\nu^{1/2}\boldsymbol{\nabla}\boldsymbol{\phi}(s)\|^2\mathrm{d}s \le \|\mathbf{e}\|_Q^2 + \int_t^T m_1(s)\|\boldsymbol{\phi}(s)\|^2\mathrm{d}s,
$$

where $m_1(s)$ is as defined by (4.8). Applying Gronwall's lemma (cf. Girault & Raviart [4], Lemma 1.8, p.167.), we have

$$
\|\boldsymbol{\phi}\|^2_{L^\infty(0,T;L^2(\Omega))} + 2\|\nu^{1/2}\boldsymbol{\nabla}\boldsymbol{\phi}\|_Q^2 \le 2\mathrm{e}^{\int_0^T m_1(t)\mathrm{d}t}\|\mathbf{e}\|_Q^2,
$$

as required. □

Lemma 4.6 *Let* $\boldsymbol{\phi}$ *be the solution of (4.1); then there exists a constant* $C_2^s = C_2^s(\mathbf{u},\mathbf{u}_h,T,\mathbf{a},\nu)$ *such that*

$$
\|\nu^{1/2}\boldsymbol{\nabla}\boldsymbol{\phi}\|^2_{L^\infty(0,T;L^2(\Omega))} + \|\nu\Delta\boldsymbol{\phi}\|_Q^2 \le C_2^s\|\mathbf{e}\|_Q^2,
$$

where

$$
\begin{aligned}
C_2^s &= 2\min\Big\{\|m_2\|_{L^\infty(0,T)}\,exp\left(4\|\mathbf{a}\|^2_{L^2(0,T;L^\infty(\Omega))}/\nu\right), \\
&\qquad\qquad\qquad \left(\|m_2\|_{L^\infty(0,T)} + 2C_1^s\|\mathbf{a}\|^2_{L^\infty(0,T;L^\infty(\Omega))}/\nu\right)\Big\}, \\
m_2(t) &= 4\Big[1 + C_1^s\Big(\|\boldsymbol{\nabla}\cdot\mathbf{a}(t) - \left(\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}u_h\right)(t)\|^2_{L^\infty(\Omega)} + \|\left(\mathcal{Q}_2^c\cdot\boldsymbol{\nabla}u_h\right)(t)\|^2_{L^\infty(\Omega)} \\
&\quad + \|\boldsymbol{\nabla}\cdot\mathbf{a}(t) - \left(\mathcal{Q}_2^c\cdot\boldsymbol{\nabla}v_h\right)(t)\|^2_{L^\infty(\Omega)} + \|\left(\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}v_h\right)(t)\|^2_{L^\infty(\Omega)}\Big)\Big] \qquad (4.11)
\end{aligned}
$$

and C_1^s *is as defined in Lemma 4.5.*

Proof Multiply (4.2a) by $-\nu\Delta\phi$ and integrate over Ω to obtain,

$$-\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\nu^{1/2}\boldsymbol{\nabla}\phi(t)\|^2+\|\nu\Delta\phi(t)\|^2$$
$$=-\left(u(t)-u_h(t)+\boldsymbol{\nabla}\cdot(\mathbf{a}(t)\phi(t))-\left(\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}u_h\right)(t)\phi(t)-\left(\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}v_h\right)(t)\psi(t),\nu\Delta\phi(t)\right).$$

Using the Cauchy–Schwarz inequality, gives

$$-\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\nu^{1/2}\boldsymbol{\nabla}\phi(t)\|^2+\|\nu\Delta\phi(t)\|^2$$
$$\leq\|u(t)-u_h(t)+\boldsymbol{\nabla}\cdot(\mathbf{a}(t)\phi(t))-\left(\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}u_h\right)(t)\phi(t)-\left(\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}v_h\right)(t)\psi(t)\|\,\|\nu\Delta\phi(t)\|$$
$$\leq\frac{1}{2}\|u(t)-u_h(t)+\boldsymbol{\nabla}\cdot(\mathbf{a}(t)\phi(t))-\left(\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}u_h\right)(t)\phi(t)-\left(\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}v_h\right)(t)\psi(t)\|^2$$
$$+\frac{1}{2}\|\nu\Delta\phi(t)\|^2.$$

In addition, applying the triangle inequality and Hölder's inequality, gives

$$\begin{aligned}-\frac{\mathrm{d}}{\mathrm{d}t}&\|\nu^{1/2}\boldsymbol{\nabla}\phi(t)\|^2+\|\nu\Delta\phi(t)\|^2\\&\leq4\|u(t)-u_h(t)\|^2+4\left\|\boldsymbol{\nabla}\cdot\mathbf{a}(t)-\left(\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}u_h\right)(t)\right\|^2_{L^\infty(\Omega)}\|\phi(t)\|^2\\&\quad+\frac{4}{\nu}\|\mathbf{a}(t)\|^2_{L^\infty(\Omega)}\|\nu^{1/2}\boldsymbol{\nabla}\phi(t)\|^2+4\left\|\left(\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}v_h\right)(t)\right\|^2_{L^\infty(\Omega)}\|\psi(t)\|^2.\end{aligned}\tag{4.12}$$

Similarly, it can be shown that

$$\begin{aligned}-\frac{\mathrm{d}}{\mathrm{d}t}&\|\nu^{1/2}\boldsymbol{\nabla}\psi(t)\|^2+\|\nu\Delta\psi(t)\|^2\\&\leq4\|v(t)-v_h(t)\|^2+4\left\|\boldsymbol{\nabla}\cdot\mathbf{a}(t)-\left(\mathcal{Q}_2^c\cdot\boldsymbol{\nabla}v_h\right)(t)\right\|^2_{L^\infty(\Omega)}\|\psi(t)\|^2\\&\quad+\frac{4}{\nu}\|\mathbf{a}(t)\|^2_{L^\infty(\Omega)}\|\nu^{1/2}\boldsymbol{\nabla}\psi(t)\|^2+4\left\|\left(\mathcal{Q}_2^c\cdot\boldsymbol{\nabla}u_h\right)(t)\right\|^2_{L^\infty(\Omega)}\|\phi(t)\|^2.\end{aligned}\tag{4.13}$$

Adding (4.12) and (4.13) together and using Lemma 4.5, gives

$$\begin{aligned}-\frac{\mathrm{d}}{\mathrm{d}t}&\|\nu^{1/2}\boldsymbol{\nabla}\boldsymbol{\phi}(t)\|^2+\|\nu\Delta\boldsymbol{\phi}(t)\|^2\\&\leq4\|\mathbf{e}(t)\|^2+\frac{4}{\nu}\|\mathbf{a}(t)\|^2_{L^\infty(\Omega)}\|\nu^{1/2}\boldsymbol{\nabla}\boldsymbol{\phi}(t)\|^2\\&\quad+4\left[\|\boldsymbol{\nabla}\cdot\mathbf{a}(t)-\left(\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}u_h\right)(t)\|^2_{L^\infty(\Omega)}+\|\left(\mathcal{Q}_2^c\cdot\boldsymbol{\nabla}u_h\right)(t)\|^2_{L^\infty(\Omega)}\right]\|\phi(t)\|^2\\&\quad+4\left[\|\boldsymbol{\nabla}\cdot\mathbf{a}(t)-\left(\mathcal{Q}_2^c\cdot\boldsymbol{\nabla}v_h\right)(t)\|^2_{L^\infty(\Omega)}+\|\left(\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}v_h\right)(t)\|^2_{L^\infty(\Omega)}\right]\|\psi(t)\|^2\\&\leq m_2(t)\|\mathbf{e}(t)\|^2+\frac{4}{\nu}\|\mathbf{a}(t)\|^2_{L^\infty(\Omega)}\|\nu^{1/2}\boldsymbol{\nabla}\boldsymbol{\phi}(t)\|^2,\end{aligned}$$

where $m_2(t)$ is as defined by (4.11). Now, integrating with respect to time over the interval (t,T) and using (4.1c) and Hölder's inequality, we get

$$\begin{aligned}\|\nu^{1/2}\boldsymbol{\nabla}\boldsymbol{\phi}(t)\|^2&+\int_t^T\|\nu\Delta\boldsymbol{\phi}(s)\|^2\mathrm{d}s\\&\leq\int_t^T m_2(s)\|\mathbf{e}(s)\|^2\mathrm{d}s+\frac{4}{\nu}\int_t^T\|\mathbf{a}(s)\|^2_{L^\infty(\Omega)}\|\nu^{1/2}\boldsymbol{\nabla}\boldsymbol{\phi}(s)\|^2\mathrm{d}s\\&\leq\|m_2\|_{L^\infty(0,T)}\|\mathbf{e}\|^2_Q+\frac{4}{\nu}\int_t^T\|\mathbf{a}(s)\|^2_{L^\infty(\Omega)}\|\nu^{1/2}\boldsymbol{\nabla}\boldsymbol{\phi}(s)\|^2\mathrm{d}s.\end{aligned}$$

Applying Gronwall's lemma gives

$$\|\nu^{1/2}\boldsymbol{\nabla}\boldsymbol{\phi}\|^2_{L^\infty(0,T;L^2(\Omega))} + \|\nu\Delta\boldsymbol{\phi}\|^2_Q \le 2\|m_2\|_{L^\infty(0,T)}\|\mathbf{e}\|^2_Q\, \mathrm{e}^{(4/\nu)\|\mathbf{a}\|^2_{L^2(0,T;L^\infty(\Omega))}}. \quad (4.14)$$

Alternatively, using Hölder's inequality and Lemma 4.5, we have

$$\|\nu^{1/2}\boldsymbol{\nabla}\boldsymbol{\phi}\|^2_{L^\infty(0,T;L^2(\Omega))} + \|\nu\Delta\boldsymbol{\phi}\|^2_Q \le 2\left(\|m_2\|_{L^\infty(0,T)} + 2C_1^s\|\mathbf{a}\|^2_{L^\infty(0,T;L^\infty(\Omega))}/\nu\right)\|\mathbf{e}\|^2_Q. \quad (4.15)$$

The lemma now follows from (4.14) and (4.15). □

Lemma 4.7 *Let $\boldsymbol{\phi}$ be the solution of (4.1); then there exists a constant $C_3^s = C_3^s(\mathbf{u}, \mathbf{u}_h, T, \mathbf{a}, \nu)$ such that*

$$\|\boldsymbol{\phi}_t\|^2_Q + \|\nu^{1/2}\boldsymbol{\nabla}\boldsymbol{\phi}(0)\|^2 \le C_3^s\|\mathbf{e}\|^2_Q,$$

where $C_3^s = \left(\|m_2\|_{L^\infty(0,T)} + (2/\nu)\min\left\{C_1^s\|\mathbf{a}\|^2_{L^\infty(0,T;L^\infty(\Omega))}, 2C_2^s\|\mathbf{a}\|^2_{L^2(0,T;L^\infty(\Omega))}\right\}\right)$, C_1^s is as defined in Lemma 4.5 and C_2^s and $m_2(t)$ are as defined in Lemma 4.6.

Proof Multiply (4.2a) by $-\phi_t$ and integrate over Ω to obtain,

$$\begin{aligned}&\|\phi_t(t)\|^2 - \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\nu^{1/2}\boldsymbol{\nabla}\phi(t)\|^2 \\ &= -\left(u(t) - u_h(t) + \boldsymbol{\nabla}\cdot(\mathbf{a}(t)\phi(t)) - (\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}u_h)(t)\phi(t) - (\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}v_h)(t)\psi(t), \phi_t(t)\right).\end{aligned}$$

Using the Cauchy–Schwarz inequality, gives

$$\begin{aligned}&\|\phi_t(t)\|^2 - \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\nu^{1/2}\boldsymbol{\nabla}\phi(t)\|^2 \\ &\le \|u(t) - u_h(t) + \boldsymbol{\nabla}\cdot(\mathbf{a}(t)\phi(t)) - (\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}u_h)(t)\phi(t) - (\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}v_h)(t)\psi(t)\|\,\|\phi_t(t)\| \\ &\le \frac{1}{2}\|u(t) - u_h(t) + \boldsymbol{\nabla}\cdot(\mathbf{a}(t)\phi(t)) - (\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}u_h)(t)\phi(t) - (\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}v_h)(t)\psi(t)\|^2 \\ &\quad + \frac{1}{2}\|\phi_t(t)\|^2.\end{aligned}$$

In addition, using the triangle inequality and Hölder's inequality, gives

$$\begin{aligned}&\|\phi_t(t)\|^2 - \frac{\mathrm{d}}{\mathrm{d}t}\|\nu^{1/2}\boldsymbol{\nabla}\phi(t)\|^2 \\ &\qquad \le 4\|u(t) - u_h(t)\|^2 + 4\|\boldsymbol{\nabla}\cdot\mathbf{a}(t) - (\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}u_h)(t)\|^2_{L^\infty(\Omega)}\|\phi(t)\|^2 \\ &\qquad\quad + \frac{4}{\nu}\|\mathbf{a}(t)\|^2_{L^\infty(\Omega)}\|\nu^{1/2}\boldsymbol{\nabla}\phi(t)\|^2 + 4\|(\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}v_h)(t)\|^2_{L^\infty(\Omega)}\|\psi(t)\|^2. \quad (4.16)\end{aligned}$$

Similarly, it can be shown that

$$\begin{aligned}&\|\psi_t(t)\|^2 - \frac{\mathrm{d}}{\mathrm{d}t}\|\nu^{1/2}\boldsymbol{\nabla}\psi(t)\|^2 \\ &\qquad \le 4\|v(t) - v_h(t)\|^2 + 4\|\boldsymbol{\nabla}\cdot\mathbf{a}(t) - (\mathcal{Q}_2^c\cdot\boldsymbol{\nabla}v_h)(t)\|^2_{L^\infty(\Omega)}\|\psi(t)\|^2 \\ &\qquad\quad + \frac{4}{\nu}\|\mathbf{a}(t)\|^2_{L^\infty(\Omega)}\|\nu^{1/2}\boldsymbol{\nabla}\psi(t)\|^2 + 4\|(\mathcal{Q}_2^c\cdot\boldsymbol{\nabla}u_h)(t)\|^2_{L^\infty(\Omega)}\|\phi(t)\|^2. \quad (4.17)\end{aligned}$$

Adding (4.16) and (4.17) together and using Lemma 4.5, gives

$$
\begin{aligned}
&\|\boldsymbol{\phi}_t(t)\|^2 - \frac{\mathrm{d}}{\mathrm{d}t}\|\nu^{1/2}\boldsymbol{\nabla}\boldsymbol{\phi}(t)\|^2 \\
&\qquad \le 4\|\mathbf{e}(t)\|^2 + \frac{4}{\nu}\|\mathbf{a}(t)\|^2_{L^\infty(\Omega)}\|\nu^{1/2}\boldsymbol{\nabla}\boldsymbol{\phi}(t)\|^2 \\
&\qquad\quad +4\left[\|\boldsymbol{\nabla}\cdot\mathbf{a}(t) - (\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}u_h)(t)\|^2_{L^\infty(\Omega)} + \|(\mathcal{Q}_2^c\cdot\boldsymbol{\nabla}u_h)(t)\|^2_{L^\infty(\Omega)}\right]\|\phi(t)\|^2 \\
&\qquad\quad +4\left[\|\boldsymbol{\nabla}\cdot\mathbf{a}(t) - (\mathcal{Q}_2^c\cdot\boldsymbol{\nabla}v_h)(t)\|^2_{L^\infty(\Omega)} + \|(\mathcal{Q}_1^c\cdot\boldsymbol{\nabla}v_h)(t)\|^2_{L^\infty(\Omega)}\right]\|\psi(t)\|^2 \\
&\qquad \le m_2(t)\|\mathbf{e}(t)\|^2 + \frac{4}{\nu}\|\mathbf{a}(t)\|^2_{L^\infty(\Omega)}\|\nu^{1/2}\boldsymbol{\nabla}\boldsymbol{\phi}(t)\|^2,
\end{aligned}
$$

where $m_2(t)$ is as defined in Lemma 4.6. Now, integrating with respect to time over the interval $(0,T)$ and using (4.1c), Hölder's inequality and Lemma 4.5 again, we get

$$
\begin{aligned}
\|\boldsymbol{\phi}_t\|^2_Q + \|\nu^{1/2}\boldsymbol{\nabla}\boldsymbol{\phi}(0)\|^2 &\le \|m_2\|_{L^\infty(0,T)}\|\mathbf{e}\|^2_Q + \frac{4}{\nu}\|\mathbf{a}\|^2_{L^\infty(0,T;L^\infty(\Omega))}\|\nu^{1/2}\boldsymbol{\nabla}\boldsymbol{\phi}\|^2_Q \\
&\le \left(\|m_2\|_{L^\infty(0,T)} + \frac{2}{\nu}C_1^s\|\mathbf{a}\|^2_{L^\infty(0,T;L^\infty(\Omega))}\right)\|\mathbf{e}\|^2_Q. \qquad (4.18)
\end{aligned}
$$

Alternatively, integrating with respect to time over the interval $(0,T)$ and using (4.1c), Hölder's inequality and Lemma 4.6, gives

$$
\begin{aligned}
&\|\boldsymbol{\phi}_t\|^2_Q + \|\nu^{1/2}\boldsymbol{\nabla}\boldsymbol{\phi}(0)\|^2 \\
&\qquad \le \|m_2\|_{L^\infty(0,T)}\|\mathbf{e}\|^2_Q + \frac{4}{\nu}\|\mathbf{a}\|^2_{L^2(0,T;L^\infty(\Omega))}\|\nu^{1/2}\boldsymbol{\nabla}\boldsymbol{\phi}\|^2_{L^\infty(0,T;L^2(\Omega))} \\
&\qquad \le \left(\|m_2\|_{L^\infty(0,T)} + \frac{4}{\nu}C_2^s\|\mathbf{a}\|^2_{L^2(0,T;L^\infty(\Omega))}\right)\|\mathbf{e}\|^2_Q. \qquad (4.19)
\end{aligned}
$$

The lemma now follows from (4.18) and (4.19). $\square$

4.4 Completion of the proof of the a posteriori error estimate

We shall now proceed to estimate the terms I-XII on the right–hand side of (4.5). For the first term I, we have

$$
\begin{aligned}
\mathrm{I} &= \sum_{n=0}^{M}\int_{t^n}^{t^{n+1}}\sum_{\kappa\in\mathcal{T}^{n+1}}([u_h^n]/k_{n+1} + \mathbf{a}(\mathbf{u}_h)\cdot\boldsymbol{\nabla}u_h - \nu\Delta u_h - f, \phi_h - \phi)_\kappa\,\mathrm{d}t \\
&\equiv (R_1(u_h,f), \phi_h - \phi)_Q \\
&= (R_1(u_h,f), \tilde{\mathcal{I}}\phi - \phi)_Q + (R_1(u_h,f), \tilde{\mathcal{I}}(\pi\phi - \phi))_Q \\
&\equiv \mathrm{I}_1 + \mathrm{I}_2,
\end{aligned}
$$

where $\tilde{\mathcal{I}}$ and π are as defined by (4.7), and

$$
R_1(u_h,f)|_\kappa = [u_h^n]/k_{n+1} + \mathbf{a}(\mathbf{u}_h)\cdot\boldsymbol{\nabla}u_h - \nu\Delta u_h - f, \quad \text{for } \kappa\in\mathcal{T}^{n+1}.
$$

By Lemma 4.3 and Lemma 4.6, it follows that

$$|\mathrm{I}_1| \le C_1^{\tilde{p}} \|h^2 R_1(u_h, f)\|_Q \|\Delta\phi\|_Q \le \frac{C_1^{\tilde{p}}\sqrt{C_2^s}}{\nu} \|h^2 R_1(u_h, f)\|_Q \|\mathbf{e}\|_Q.$$

Similarly, using Lemma 4.4 and Lemma 4.7, we have

$$|\mathrm{I}_2| \le C_3^{\tilde{p}} \|k R_1(u_h, f)\|_Q \|\phi_t\|_Q \le C_3^{\tilde{p}} \sqrt{C_3^s}\, \|k R_1(u_h, f)\|_Q \|\mathbf{e}\|_Q.$$

Hence,

$$|\mathrm{I}| \le \frac{C_1^{\tilde{p}}\sqrt{C_2^s}}{\nu} \|h^2 R_1(u_h, f)\|_Q \|\mathbf{e}\|_Q + C_3^{\tilde{p}} \sqrt{C_3^s}\, \|k R_1(u_h, f)\|_Q \|\mathbf{e}\|_Q.$$

Analogously,

$$\begin{aligned}
\mathrm{II} &= \sum_{n=0}^{M} \int_{t^n}^{t^{n+1}} \left(\sum_{\kappa \in \mathcal{T}^{n+1}} (\nu \Delta u_h, \phi_h - \phi)_\kappa + (\nu \boldsymbol{\nabla} u_h, \boldsymbol{\nabla}(\phi_h - \phi)) \right) \mathrm{d}t \\
&= \sum_{n=0}^{M} \int_{t^n}^{t^{n+1}} \left(\sum_{\kappa \in \mathcal{T}^{n+1}} (\nu \Delta u_h, \tilde{\mathcal{I}}_{n+1}\phi - \phi)_\kappa + (\nu \boldsymbol{\nabla} u_h, \boldsymbol{\nabla}(\tilde{\mathcal{I}}_{n+1}\phi - \phi)) \right) \mathrm{d}t \\
&\quad + \sum_{n=0}^{M} \int_{t^n}^{t^{n+1}} \Bigg(\sum_{\kappa \in \mathcal{T}^{n+1}} (\nu \Delta u_h, \tilde{\mathcal{I}}_{n+1}(\pi_{n+1}\phi - \phi))_\kappa \\
&\qquad\qquad + (\nu \boldsymbol{\nabla} u_h, \boldsymbol{\nabla} \tilde{\mathcal{I}}_{n+1}(\pi_{n+1}\phi - \phi)) \Bigg) \mathrm{d}t \\
&\equiv \mathrm{II}_1 + \mathrm{II}_2.
\end{aligned}$$

By Lemma 4.3 and Lemma 4.6, we have

$$|\mathrm{II}_1| \le C_2^{\tilde{p}} \|h^2 D_h^2 u_h\|_Q \|\nu \Delta\phi\|_Q \le C_2^{\tilde{p}} \sqrt{C_2^s}\, \|h^2 D_h^2 u_h\|_Q \|\mathbf{e}\|_Q.$$

Also, by Lemma 4.4,

$$\mathrm{II}_2 = 0.$$

Thus, we have that

$$|\mathrm{II}| \le C_2^{\tilde{p}} \sqrt{C_2^s}\, \|h^2 D_h^2 u_h\|_Q \|\mathbf{e}\|_Q.$$

Next, we consider term III: applying the Cauchy–Schwarz inequality, Lemma 4.2 and Lemma 4.5, it follows that

$$\begin{aligned}
|\mathrm{III}| &\le \|k R_3(u_h)\|_Q \|\phi_h\|_Q \le C_1^{\tilde{i}} \|k R_3(u_h)\|_Q \|\phi\|_Q \\
&\le C_1^{\tilde{i}} \sqrt{T}\, \|k R_3(u_h)\|_Q \|\phi\|_{L^\infty(0,T;L^2(\Omega))} \\
&\le C_1^{\tilde{i}} \sqrt{C_1^s T}\, \|k R_3(u_h)\|_Q \|\mathbf{e}\|_Q,
\end{aligned}$$

where

$$R_3(u_h)|_{S^{n+1}} = (D_t^h u_h - ([u_h^n]/k_{n+1} + \mathbf{a}(\mathbf{u}_h) \cdot \boldsymbol{\nabla} u_h))/k_{n+1}.$$

Now, we consider term IV: using Lemma 4.4 and Lemma 4.7, we get

$$|\mathrm{IV}| \leq C_4^{\tilde{p}} \|kR_4(u_h)\|_Q \|\phi_t\|_Q \leq C_4^{\tilde{p}} \sqrt{C_3^s}\, \|kR_4(u_h)\|_Q \|\mathbf{e}\|_Q,$$

where

$$R_4(u_h)|_{S^{n+1}} = [u_h^n]/k_{n+1} = (u_h^{n+1} - u_h^n)/k_{n+1}.$$

Next, we consider term V: using the Cauchy–Schwarz inequality, Lemma 4.2 and Lemma 4.5, we get

$$\begin{aligned} |\mathrm{V}| &\leq \|f - \bar{f}\|_Q \|\phi_h\|_Q \leq C_1^{\tilde{i}} \|f - \bar{f}\|_Q \|\phi\|_Q \\ &\leq C_1^{\tilde{i}} \sqrt{T}\, \|f - \bar{f}\|_Q \|\phi\|_{L^\infty(0,T;L^2(\Omega))} \\ &\leq C_1^{\tilde{i}} \sqrt{C_1^s T}\, \|f - \bar{f}\|_Q \|\mathbf{e}\|_Q. \end{aligned}$$

Let us consider term VI: using the Cauchy–Schwarz inequality and Lemma 4.5, we have

$$\begin{aligned} |\mathrm{VI}| &\leq \|u_0 - u_{h-}^0\| \|\phi(0)\| \leq \|u_0 - u_{h-}^0\| \|\phi\|_{L^\infty(0,T;L^2(\Omega))} \\ &\leq \sqrt{C_1^s}\, \|u_0 - u_{h-}^0\| \|\mathbf{e}\|_Q. \end{aligned}$$

The terms VII-XII, relating to the second component v_h of the finite element solution $\mathbf{u}_h$, are treated in exactly the same way. Hence, we may now state the following *a posteriori* error estimate:

Theorem 4.8 *Let $\mathbf{u}$ and $\mathbf{u}_h = (u_h, v_h)$ be solutions of (3.1) and (3.8), respectively. If $\mathcal{T}^n$, $0 \leq n \leq M+1$, satisfies conditions (3.3); then*

$$\|\mathbf{e}\|_Q = \|\mathbf{u} - \mathbf{u}_h\|_Q \leq \mathring{\mathcal{E}}(\mathbf{u}_h, h, k, \mathbf{f}), \tag{4.20}$$

where

$$\begin{aligned} \mathring{\mathcal{E}}(\mathbf{u}_h, h, k, \mathbf{f}) &= \mathcal{E}(\mathbf{u}_h, h, k, \mathbf{f}) + \mathcal{E}_0(\mathbf{u}_0, \mathbf{u}_{h-}^0, h), \\ \mathcal{E}(\mathbf{u}_h, h, k, \mathbf{f}) &= C_1\|h^2 R_1(u_h, f)\|_Q + C_2\|kR_1(u_h, f)\|_Q + C_3\|h^2 R_2(u_h)\|_Q \\ &\quad + C_4\|kR_3(u_h)\|_Q + C_5\|kR_4(u_h)\|_Q + C_6\|kR_5(f)\|_Q \\ &\quad + C_1\|h^2 R_1(v_h, g)\|_Q + C_2\|kR_1(v_h, g)\|_Q + C_3\|h^2 R_2(v_h)\|_Q \\ &\quad + C_4\|kR_3(v_h)\|_Q + C_5\|kR_4(v_h)\|_Q + C_6\|kR_5(g)\|_Q, \\ \mathcal{E}_0(\mathbf{u}_0, \mathbf{u}_{h-}^0, h) &= C_7\|u_0 - u_{h-}^0\| + C_7\|v_0 - v_{h-}^0\|, \end{aligned}$$

and, for $w_h = u_h$ *or* $w_h = v_h$ *and* $f' = f$ *or* $f' = g$*, respectively, we define*

$$
\begin{aligned}
R_1(w_h, f')|_\kappa &= [w_h^n]/k_{n+1} + \mathbf{a}(\mathbf{u}_h) \cdot \boldsymbol{\nabla} w_h - \nu \Delta w_h - f', \quad \textit{for } \kappa \in \mathcal{T}^{n+1}, \\
R_2(w_h) &= D_h^2 w_h, \\
R_3(w_h)|_{S^{n+1}} &= (D_t^h w_h - ([w_h^n]/k_{n+1} + \mathbf{a}(\mathbf{u}_h) \cdot \boldsymbol{\nabla} w_h))/k_{n+1}, \\
R_4(w_h)|_{S^{n+1}} &= [w_h^n]/k_{n+1}, \\
R_5(f') &= (f' - \bar{f}')/k, \\
C_1 &= \frac{C_1^{\tilde{p}} \sqrt{C_2^s}}{\nu} = \frac{C_2^{\tilde{i}} \sqrt{C_2^s}}{c_2^2 \nu}, \\
C_2 &= C_3^{\tilde{p}} \sqrt{C_3^s} = C_1^{\tilde{i}} C_1^i \sqrt{C_3^s}, \\
C_3 &= C_2^{\tilde{p}} \sqrt{C_2^s} = C^t \sqrt{C_2^s}\, (C_2^{\tilde{i}} c_2^{-1} + C_3^{\tilde{i}})/(2c_2^2), \\
C_4 &= C_1^{\tilde{i}} \sqrt{C_1^s T}, \\
C_5 &= C_4^{\tilde{p}} \sqrt{C_3^s}, \\
C_6 &= C_4 = C_1^{\tilde{i}} \sqrt{C_1^s T}, \\
C_7 &= \sqrt{C_1^s}.
\end{aligned}
$$

5 Adaptive algorithm

For a given tolerance TOL, we now consider the problem of finding a discretisation in space and time $\mathcal{S}^h = \{(\mathcal{T}^n, t^n)\}_{n \geq 0}$ such that:

1. $\|\mathbf{u} - \mathbf{u}_h\|_Q \leq \text{TOL}$;
2. $\mathcal{S}^h$ is optimal in the sense that the number of degrees of freedom is minimal.

In order to satisfy these criteria we shall use the *a posteriori* error estimate (4.20) to choose $\mathcal{S}^h$ such that:

1. $\overset{\circ}{\mathcal{E}}(\mathbf{u}_h, h, k, \mathbf{f}) \leq \text{TOL}$;
2. The number of degrees of freedom of $\mathcal{S}^h$ is minimal.

The term $\mathcal{E}_0(\mathbf{u}_0, \mathbf{u}_{h-}^0, h)$ is easily controlled at the start of a computation; so here we shall only consider the problem of constructing $\mathcal{S}^h$ in an efficient way to ensure that

$$\mathcal{E}(\mathbf{u}_h, h, k, \mathbf{f}) \leq \text{TOL}',$$

where $\mathrm{TOL} = \mathrm{TOL}' + \mathcal{E}_0(\mathbf{u}_0, \mathbf{u}^0_{h-}, h)$. As in [7, 8], we split the tolerance TOL' into a spatial part, TOL_h, and a temporal part, TOL_k.

We propose the following adaptive algorithm for choosing $\mathcal{S}^h$, assuming that the final time T is fixed: for each $n = 1, 2, \ldots, M+1$, with $\mathcal{T}^n_0$ a given initial mesh and $k_{n,0}$ an initial time step, determine meshes $\mathcal{T}^n_j$ with N^n_j elements of size $h_{n,j}(\mathbf{x})$ and time steps $k_{n,j}$ and corresponding approximate solution $\mathbf{u}^n_{h,j} = (u^n_{h,j}, v^n_{h,j})$ defined on I^n_j such that, for $j = 0, 1, \ldots, \hat{n}-1$,

$$\begin{aligned} C_1\|h^2_{n,j+1}R_1(u^n_{h,j}, f)\|_{L^2(\kappa)} + C_3\|h^2_{n,j+1}R_2(u^n_{h,j})\|_{L^2(\kappa)} \\ +C_1\|h^2_{n,j+1}R_1(v^n_{h,j}, g)\|_{L^2(\kappa)} + C_3\|h^2_{n,j+1}R_2(v^n_{h,j})\|_{L^2(\kappa)} &= \frac{\mathrm{TOL}_h}{\sqrt{N^n_j T}} \quad \forall \kappa \in \mathcal{T}^n_j, \\ C_2\|k_{n,j+1}R_1(u^n_{h,j}, f)\| + C_4\|k_{n,j+1}R_3(u^n_{h,j})\| \\ +C_5\|k_{n,j+1}R_4(u^n_{h,j})\| + C_6\|k_{n,j+1}R_5(f)\| \\ +C_2\|k_{n,j+1}R_1(v^n_{h,j}, g)\| + C_4\|k_{n,j+1}R_3(v^n_{h,j})\| \\ +C_5\|k_{n,j+1}R_4(v^n_{h,j})\| + C_6\|k_{n,j+1}R_5(g)\| &= \frac{\mathrm{TOL}_k}{\sqrt{T}}, \end{aligned}$$

where $I^n_j = (t^{n-1}, t^{n-1} + k_{n,j}]$ and $\mathrm{TOL}' = \mathrm{TOL}_h + \mathrm{TOL}_k$. We define $\mathcal{T}^n = \mathcal{T}^n_{\hat{n}}$, $k_n = k_{n,\hat{n}}$ and $h_n = h_{n,\hat{n}}$, where for each n, the number of trials $\hat{n}$ is the smallest integer such that for $j = \hat{n}$, the stopping condition

$$\begin{aligned} C_1\|h^2_{n,\hat{n}}R_1(u^n_{h,\hat{n}}, f)\|_{L^2(\kappa)} + C_3\|h^2_{n,\hat{n}}R_2(u^n_{h,\hat{n}})\|_{L^2(\kappa)} \\ +C_1\|h^2_{n,\hat{n}}R_1(v^n_{h,\hat{n}}, g)\|_{L^2(\kappa)} + C_3\|h^2_{n,\hat{n}}R_2(v^n_{h,\hat{n}})\|_{L^2(\kappa)} &\le \frac{\mathrm{TOL}_h}{\sqrt{N^n_{\hat{n}} T}} \quad \forall \kappa \in \mathcal{T}^n_{\hat{n}}, \\ C_2\|k_{n,\hat{n}}R_1(u^n_{h,\hat{n}}, f)\| + C_4\|k_{n,\hat{n}}R_3(u^n_{h,\hat{n}})\| \\ +C_5\|k_{n,\hat{n}}R_4(u^n_{h,\hat{n}})\| + C_6\|k_{n,\hat{n}}R_5(f)\| \\ +C_2\|k_{n,\hat{n}}R_1(v^n_{h,\hat{n}}, g)\| + C_4\|k_{n,\hat{n}}R_3(v^n_{h,\hat{n}})\| \\ +C_5\|k_{n,\hat{n}}R_4(v^n_{h,\hat{n}})\| + C_6\|k_{n,\hat{n}}R_5(g)\| &\le \frac{\mathrm{TOL}_k}{\sqrt{T}}, \end{aligned}$$

is satisfied.

Remark 5.1 *We may improve the accuracy of the calculation of the particle trajectories $\mathbf{X}_{\mathbf{u}_h}|_{S^n}$ by using information from the adaptive algorithm. For instance, for $j = 0$ we simply calculate $\mathbf{X}_{\mathbf{u}_h}|_{S^n}$ using only information about $\mathbf{u}^{n-1}_h$, cf. (3.7); but for $j = 1, \ldots, \hat{n}$, the calculation of $\mathbf{X}_{\mathbf{u}_h}|_{S^n}$ may involve information about both $\mathbf{u}^{n-1}_h$ and $\mathbf{u}^n_{h,j-1}$, where $\mathbf{u}^n_{h,j-1}$ is the solution calculated on the mesh $\mathcal{T}^n_{j-1}$ in the previous iteration of the adaptive algorithm.*

6 Computational estimation of the strong stability constants

In this section we shall briefly comment on the computational estimation of the strong stability constants $C_i^s(\mathbf{u}, \mathbf{u}_h)$ for $i = 1, \ldots, 3$, arising in the adaptive algorithm formulated in the previous section; in the following we shall simply write $C^s(\mathbf{u}, \mathbf{u}_h)$ to denote these stability factors.

The natural extension of the approach described in [7] would be to replace $\mathbf{u}$ and $\mathbf{u}_h$ in the dual problem by $\mathbf{u}_h^{\text{fine}}$ and $\mathbf{u}_h^{\text{coarse}}$, respectively, and calculate the corresponding stability factor $C^s(\mathbf{u}_h^{\text{fine}}, \mathbf{u}_h^{\text{coarse}})$ by solving the dual problem numerically with the right–hand side function $\mathbf{e}$ replaced by $\mathbf{e}_h = \mathbf{u}_h^{\text{fine}} - \mathbf{u}_h^{\text{coarse}}$; here, we note that $\mathbf{u}_h^{\text{fine}}$ and $\mathbf{u}_h^{\text{coarse}}$ are numerical approximations to $\mathbf{u}$ calculated on fine and coarse meshes, respectively.

An alternative approach proposed by Larson [12], Sandboge [14] and Johnson *et al.* [3, 10] is to calculate $C^s(\mathbf{u}_h, \mathbf{u}_h)$ as an approximation to $C^s(\mathbf{u}, \mathbf{u}_h)$, since if the error $\mathbf{e}$ is sufficiently small, then we may expect that $C^s(\mathbf{u}, \mathbf{u}_h) \sim C^s(\mathbf{u}_h, \mathbf{u}_h)$. We note, however, that this approach still requires the calculation of $\mathbf{u}_h^{\text{coarse}}$ for the right–hand side function of the dual problem.

However, irrespective of which of the above approaches is adopted, the task of computing these stability factors for time–dependent problems is further complicated by the fact that the amount of computer memory required to store the data for the dual problem (i.e. $\mathbf{u}_h^{\text{fine}}$ and $\mathbf{u}_h^{\text{coarse}}$, for example) for the entire length of the computation is immense, cf. [7]. Furthermore, the reliability of the adaptive algorithm can no longer be fully guaranteed, since the stability factors are numerically estimated; the degree of reliability will depend on the amount of computational effort devoted to calculating these constants.

7 Closing remarks

In this paper we have derived an *a posteriori* error estimate for the Lagrange–Galerkin discretisation of a system of nonlinear convection–diffusion equations in two space–dimensions. Moreover, based on this error bound, we have designed an adaptive algorithm to ensure global control of the error (i.e. reliability) in the $L^2(0, T; L^2(\Omega)^2)$ norm with respect to a pre-determined tolerance TOL. This algorithm will be implemented into an adaptive Lagrange–Galerkin finite element code as part of our program of future research.

We note that the theory developed in this paper may equally be applied to the incompressible Navier–Stokes equations to establish an error representation formula for the velocity error in terms of the residual of the Galerkin approximation and the solution to the associated (linearised) dual problem; in this case, $\mathcal{Q}$ will simply be the identity matrix. Furthermore, based on this error representation formula, an *a posteriori* error estimate bounding the velocity error in the $L^2(0,T;L^2(\Omega)^2)$ norm may be derived. Finally, we note that the extension of the results presented in this paper to three space dimensions is easily arrived at following the arguments presented here; in this case, we let Ω be a convex polyhedral domain in $\mathbf{R}^3$ and assume that $\mathbf{u}_0 \in [H_0^1(\Omega) \cap H^\eta(\Omega)]^3$ for some $\eta > 3/2$. Furthermore, we note that the minimal regularity conditions prescribed on the data of the dual problem in the first part of Theorem 4.1 are slightly modified to: $|\mathbf{a}(\mathbf{u})|^2 \in L^2(0,T;L^3(\Omega))$ and $|\mathcal{Q}^T \boldsymbol{\nabla}_m^T \mathbf{u}_h|_F \in L^2(0,T;L^3(\Omega))$; although, again we note that these conditions are automatically satisfied.

References

[1] R.A. Adams. *Sobolev Spaces.* Academic Press, New York, 1975.

[2] M. Bercovier and O. Pironneau. Characteristics and the finite element method. In T. Kawai, editor, *Proceedings of the Fourth International Symposium on Finite Element Methods in Flow Problems*, pages 67–73. North–Holland, 1982.

[3] K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems IV: Nonlinear problems. *SIAM Journal of Numerical Analysis*, 32(6):1729–1749, 1995.

[4] V. Girault and P.-A. Raviart. *Finite Element Approximation of the Navier–Stokes Equations*, Volume 749 of *Lecture Notes in Mathematics.* Springer–Verlag, 1979.

[5] P. Hansbo and C. Johnson. Streamline diffusion finite element methods for fluid flow. von Karman Institute Lectures, 1995.

[6] A. T. Hill. *Attractors for Convection–Diffusion Equations and their Numerical Approximation.* PhD thesis, University of Oxford, 1992.

[7] P. Houston and E. Süli. Adaptive Lagrange–Galerkin methods for unsteady convection–dominated diffusion problems. Technical Report NA95/24, Oxford University Computing Laboratory, Wolfson Building, Parks Road, Oxford OX1 3QD, 1995.

[8] P. Houston and E. Süli. On the design of an artificial diffusion model for the Lagrange–Galerkin method on unstructured triangular grids. Technical Report NA96/07, Oxford University Computing Laboratory, Wolfson Building, Parks Road, Oxford OX1 3QD, 1996.

[9] P. Houston and E. Süli. A posteriori error analysis for linear convection–diffusion problems under weak mesh regularity assumptions. Technical Report NA97/03, Oxford University Computing Laboratory, Wolfson Building, Parks Road, Oxford OX1 3QD, 1997.

[10] C. Johnson, R. Rannacher, and M. Boman. Numerics and hydrodynamic stability: toward error control in computational fluid mechanics. *SIAM Journal of Numerical Analysis*, 32(4):1058–1079, 1995.

[11] O.A. Ladyženskaya, V.A. Solonnikov, and N.N. Ural'ceva. *Linear and Quasilinear Equations of Parabolic Type*, Volume 23 of *Translations of Mathematical Monographs.* American Mathematical Society, 1968.

[12] M.G. Larson. Adaptive finite element methods for models of compressible fluid flow. Technical Report 1995-01, Department of Mathematics, Chalmers University of Technology, Göteborg, Sweden, 1995.

[13] O. Pironneau. On the transport–diffusion algorithm and its applications to the Navier–Stokes equations. *Numerische Mathematik*, 38:309–332, 1982.

[14] R. Sandboge. *Adaptive Finite Element Methods for Reactive Flow Problems.* PhD thesis, Chalmers University of Technology, 1997.

[15] E. Süli. Lagrange–Galerkin mixed finite element approximation of the Navier–Stokes equations. In K.W. Morton and M.J. Baines, editors, *Numerical Methods for Fluid Dynamics II*, pages 439–448. Clarendon Press, 1986.

[16] E. Süli. Convergence and nonlinear stability of the Lagrange–Galerkin method for the Navier–Stokes equations. *Numerische Mathematik*, 53:459–483, 1988.