Numer Algor (2006) 43:355–383
DOI 10.1007/s11075-007-9066-6

ORIGINAL PAPER

A fully discrete H^1-Galerkin method with quadrature for nonlinear advection–diffusion–reaction equations

M. Ganesh · K. Mustapha

Received: 9 May 2006 / Accepted: 16 January 2007 /
Published online: 15 February 2007

Abstract We propose and analyze a fully discrete H^1-Galerkin method with quadrature for nonlinear parabolic advection–diffusion–reaction equations that requires only linear algebraic solvers. Our scheme applied to the special case heat equation is a fully discrete quadrature version of the least-squares method. We prove second order convergence in time and optimal H^1 convergence in space for the computer implementable method. The results of numerical computations demonstrate optimal order convergence of scheme in H^k for $k = 0, 1, 2$.

Keywords advection–diffusion–reaction equations · parabolic equations · Galerkin method · quadrature

Mathematics Subject Classifications (2000) Primary 65M12 · 65M60

Support of the Australian Research Council is gratefully acknowledged.

M. Ganesh (✉)
Department of Mathematical and Computer Sciences, Colorado School of Mines,
Golden, CO 80401, USA
e-mail: mganesh@mines.edu

K. Mustapha
Department of Mathematical Sciences, King Fahd University of Petroleum and Minerals,
Dharhan 31261, Saudi Arabia
e-mail: kassem@kfupm.edu.sa

1 Introduction

We consider the nonlinear parabolic advection–diffusion–reaction equation

$$\frac{\partial u}{\partial t}(\mathbf{x}, t) = a(\mathbf{x}, u)\,\Delta u + f(\mathbf{x}, t, u, \nabla u), \qquad (\mathbf{x}, t) \in \Omega \times (0, T], \tag{1.1}$$

$$u(\mathbf{x}, 0) = g(\mathbf{x}), \qquad \mathbf{x} \in \overline{\Omega}, \tag{1.2}$$

$$u(\mathbf{x}, t) = 0, \qquad (\mathbf{x}, t) \in \partial\Omega \times (0, T], \tag{1.3}$$

where $\Omega \subset \mathbb{R}^2$ is a bounded domain. The functions f and g are defined respectively on $\overline{\Omega} \times [0, T] \times \mathbb{R}^3$ and $\overline{\Omega}$, and for some positive constant $a_{\min}$, the nonlinear diffusion coefficient a, defined on $\Omega \times \mathbb{R}$, satisfies

$$a(\mathbf{x}, \alpha) \geq a_{\min}, \quad \mathbf{x} \in \Omega, \quad \alpha \in \mathbb{R}. \tag{1.4}$$

The algorithm and analysis in this paper are applicable for a large class of linear and nonlinear functions (including polynomials and exponentials) in the unknown variables. Throughout the paper, we assume the following *mild* Lipschitz continuity conditions on a and f: *there exist positive constants* C and s *such that for* $\mathbf{x} \in \Omega$, $t \in [0, T]$, *and* $\alpha_1, \alpha_2, \beta, \gamma \in \mathbb{R}$,

$$|a(\mathbf{x}, \alpha_1) - a(\mathbf{x}, \alpha_2)| \leq C\left[e^{|\alpha_1|} + e^{|\alpha_2|})\right]^s |\alpha_1 - \alpha_2|\,, \tag{1.5}$$

$$|f(\mathbf{x}, t, \alpha_1, \beta, \gamma) - f(\mathbf{x}, t, \alpha_2, \beta, \gamma)| \leq C\left[e^{|\alpha_1|} + e^{|\alpha_2|} + e^{|\beta|} + e^{|\gamma|})\right]^s |\alpha_1 - \alpha_2|\,, \tag{1.6}$$

$$|f(\mathbf{x}, t, \beta, \alpha_1, \gamma) - f(\mathbf{x}, t, \beta, \alpha_2, \gamma)| \leq C\left[e^{|\alpha_1|} + e^{|\alpha_2|} + e^{|\beta|} + e^{|\gamma|})\right]^s |\alpha_1 - \alpha_2|\,, \tag{1.7}$$

$$|f(\mathbf{x}, t, \beta, \gamma, \alpha_1) - f(\mathbf{x}, t, \beta, \gamma, \alpha_2)| \leq C\left[e^{|\alpha_1|} + e^{|\alpha_2|} + e^{|\beta|} + e^{|\gamma|})\right]^s |\alpha_1 - \alpha_2|\,. \tag{1.8}$$

A brief, but stronger, version of the above assumptions is that we require each entry in the Jacobian of the nonlinear terms in (1.1) has at most exponential growth. The model problem (1.1), (1.2) and (1.3) includes the non-divergence diffusion processes in biological applications, and the standard divergence form $\nabla \cdot [a(\mathbf{x}, u(\mathbf{x}, t))\nabla u(\mathbf{x}, t)]$. In the next section, we describe further regularity and domain assumptions required for analysis, after introducing appropriate function spaces.

The main aim of the paper is to propose and analyze a generalized fully discrete quadrature version of the least-squares H^1-Galerkin method for the nonlinear problem. The least-squares method and analysis have been investigated by many authors for linear and semi-linear problems, following the seminal work [4, 5], on smooth domains. The *smoothness* ($\mathcal{C}^\infty$) assumption is essential for the convergence analysis of the standard least-squares method [3, Theorem 2.1]. In this work we do not require such smoothness assumptions on the domain and our algorithm and analysis are applicable even for domains with corners (satisfying the assumptions described in the next section).

The finite element test functions in the standard least-squares method are obtained by applying the given strongly elliptic operator on the trial functions [3–5]. In particular, for a model second-order semi-linear elliptic problem $Lw = \psi(w)$, with linear elliptic operator L and homogeneous Dirichlet boundary condition the standard least-squares method is to find $w_h \in S_h$ such that

$$(Lw_h, L\Psi) = (\psi(w_h), L\Psi)\,, \qquad \Psi \in S_h, \tag{1.9}$$

where for $v, z \in H^0(\Omega) = L^2(\Omega)$,

$$(v, z) = \int_\Omega v(\mathbf{x}) z(\mathbf{x})\, d\mathbf{x}, \qquad \|v\|_0^2 = (v, v), \tag{1.10}$$

and $\mathcal{S}_h \subset H^2(\Omega) \cap H_0^1(\Omega)$ denotes the space of all splines that are polynomials of degree at most $r \geq 3$ on each element of a quasi-uniform triangulation $\mathcal{T}_h$ of Ω (with largest mesh size h) such that for $\phi \in H^k(\Omega) \cap H_0^1(\Omega),\ 2 \leq k \leq r+1$

$$\inf_{\chi \in \mathcal{S}_h} \|\phi - \chi\|_\ell \leq C h^{k-\ell} \|\phi\|_k, \qquad \ell = 0, 1, 2. \tag{1.11}$$

Such finite elements are usually constructed by starting with a reference element $\widehat{\rho}$ and taking the boundary vertices of the triangulation to be on $\partial\Omega$. (If the interior Ω_h of the union of the finite elements is not equal to Ω, then the functions in $\mathcal{S}_h$ vanish on $\overline{\Omega} \setminus \Omega_h$. Details of construction of $\mathcal{C}^1$ splines on arbitrary topology are in [15, 16]).

Throughout the paper, for a nonnegative integer k, the standard norm in the Sobolev space $H^k(\Omega)$ is denoted by $\|\cdot\|_k$, $L^2(\Omega) = H^0(\Omega)$; $H_0^1(\Omega)$ denotes the space of all functions $\phi \in H^1(\Omega)$ with $\phi = 0$ on $\partial\Omega$; C denotes a *generic* positive constant which may depend on r, but which is independent of h, the time-discretization parameter τ and the exact solution of the partial differential equation.

We remark that the H^1-Galerkin method for second-order parabolic problems requires an H^2 trial space (spanned by $\mathcal{C}^1$ basis functions). Hence this method is considered to be less practical compared to the standard Galerkin method based on an H^1 trial space, for linear problems described in divergence form that are suitable for weak formulations, allowing reduction in the smoothness requirement. However, for our nonlinear model problem (1.1) in non-divergence form with diffusion coefficient depending linearly or nonlinearly on the unknown solution, the choice of H^2 trial space is natural. Construction of $\mathcal{C}^1$ splines (and hence H^2 trial spaces) is no longer considered to be difficult even on arbitrary topology, in view of the recent work [15, 16] and references therein. In comparison with $\mathcal{C}^0$ finite element Galerkin methods, $\mathcal{C}^1$ smoothness of the H^1-Galerkin approximate solutions leads to significantly smaller linear systems.

Trial spaces spanned by $\mathcal{C}^1$ splines are essential for collocation finite element methods, see the survey paper [7] and the recent work [19] and references therein. The finite element Galerkin method for biharmonic problems based on a weak formulation requires an H^2 trial space, see the recent work [1, 2] and references therein. A byproduct of the analysis in this work in Section 3 yields optimal order convergence of a quadrature finite element Galerkin solution for a nonlinear biharmonic problem, extending the analysis in [2]. Further, the scheme in this paper has the advantage of allowing a wider class of nonlinear processes in the advection–diffusion–reaction model (1.1), (1.2) and (1.3).

In (1.1), (1.2) and (1.3), the variable coefficients of the second order elliptic operator depend on the unknown solution u, and hence the standard H^1-Galerkin method (with test space obtained by applying the given elliptic operator) is not appropriate for the model problem. Further, the integral in (1.10) with a non-polynomial integrand cannot in general be evaluated exactly and hence a computer implementable version of (1.9) requires a quadrature approximation for the inner product $(\cdot, \cdot)$.

Since the integrand in (1.9) for the stiffness matrix is a polynomial of degree at most $2r-2$ on each triangular finite element, we consider a quadrature rule with

degree of precision $2r-2$ over the reference element $\widehat{\rho}$ with positive weights and quadrature points in the interior of $\widehat{\rho}$. This rule induces a quadrature formula $(\cdot,\cdot)_{\rho,h}$ on each element $\rho \in \mathcal{T}_h$ so that

$$\int_\rho v(\mathbf{x})z(\mathbf{x})\,d\mathbf{x} = (v,z)_\rho \approx (v,z)_{\rho,h}, \quad (\Psi_1,\Psi_2)_\rho = (\Psi_1,\Psi_2)_{\rho,h}, \quad \Psi_1\Psi_2 \in \mathbb{P}_{2r-2}, \tag{1.12}$$

where $\mathbb{P}_j$ is the space of all polynomials of degree at most j. Let

$$(v,z)_h = \sum_{\rho\in\mathcal{T}_h}(v,z)_{\rho,h}, \qquad \|v\|_h^2 = (v,v)_h. \tag{1.13}$$

Replacing the continuous inner product $(\cdot,\cdot)$ in (1.9) by $(\cdot,\cdot)_h$, we get a fully discrete computer implementable quadrature approximation version of the standard least–square scheme in (1.9). However, a similar approach for the time–dependent problem, using stable implicit rules for discretization of the time derivative operator, will lead to the expensive requirement of solving a *nonlinear* algebraic system at each time step.

Now we are ready to describe a new fully discrete H^1-Galerkin method with quadrature to solve the parabolic problem (1.1), (1.2) and (1.3), that requires solving only linear system at each time step. In our new scheme, the finite element trial and test spaces for (1.1), (1.2) and (1.3) are chosen to be the same as that used in the least-squares method for the Poisson equation.

For a positive integer N^t, let $\Pi^t = \{t_n\}_{n=0}^{N^t}$ be a uniform partition of the time interval $[0,T]$ such that $t_n = n\tau$, where $\tau = T/N^t$, and let $t_{n+\frac{1}{2}} = t_n + \tau/2$. Throughout the paper, we use the following notation for a function ϕ.

$$\phi^n = \phi(t_n), \quad \partial_t\phi^n = \frac{\phi^{n+1}-\phi^n}{\tau}, \quad \phi^{n+\frac{1}{2}} = \frac{\phi^{n+1}+\phi^n}{2}, \quad \widehat{\phi}^{n+\frac{1}{2}} = \frac{3\phi^n-\phi^{n-1}}{2}. \tag{1.14}$$

Our fully discrete quadrature scheme to solve (1.1), (1.2) and (1.3) is: find $U : \Pi^t \to S_h$, such that

$$(\partial_t U^n,\Delta\Psi)_h = (\mathcal{A}\widehat{U}^{n+\frac{1}{2}}\Delta U^{n+\frac{1}{2}},\Delta\Psi)_h + (\mathcal{F}(t_{n+\frac{1}{2}})\widehat{U}^{n+\frac{1}{2}},\Delta\Psi)_h, \Psi\in S_h,\ n=1,\ldots,N^t-1, \tag{1.15}$$

where $\mathcal{A}$, $\mathcal{F}(t)$, $t\in[0,T]$ are defined for $\psi \in \mathcal{C}^1(\overline{\Omega})$ by

$$[\mathcal{A}\psi](\mathbf{x}) = a(\mathbf{x},\psi(\mathbf{x})), \qquad [\mathcal{F}(t)\psi](\mathbf{x}) = f(\mathbf{x},t,\psi(\mathbf{x}),\nabla\psi(\mathbf{x})), \qquad \mathbf{x}\in\Omega. \tag{1.16}$$

The linear system in (1.15) requires selection of $U^0, U^1 \in S_h$. Given $U^0 \in S_h$, depending on the initial data u^0 in (1.2), we select $U^1 \in S_h$ by solving the following predictor-corrector linear systems

$$(\partial_t V^0,\Delta\Psi)_h = (\mathcal{A}V^0\Delta V^{\frac{1}{2}},\Delta\Psi)_h + (\mathcal{F}(t_{\frac{1}{2}})V^0,\Delta\Psi)_h, \quad \Psi\in S_h, \tag{1.17}$$

$$(\partial_t U^0,\Delta\Psi)_h = (\mathcal{A}V^{\frac{1}{2}}\Delta U^{\frac{1}{2}},\Delta\Psi)_h + (\mathcal{F}(t_{\frac{1}{2}})V^{\frac{1}{2}},\Delta\Psi)_h, \quad \Psi\in S_h, \tag{1.18}$$

where $V^0 = U^0$ and $V^1 \in S_h$.

The purpose of rest of the paper is to establish second order in time and optimal order H^1 norm error bounds in space for the computer implementable scheme (1.15), (1.16), (1.17) and (1.18), and to demonstrate convergence of the algorithm in H^k norms for all $k = 0, 1, 2$ with numerical experiments.

The outline of this paper is as follows. In the next section, we discuss preliminary results including the derivation of a non-standard approximation property in S_h. The convergence analysis is based on analyzing the error in two stages, through an elliptic projection of the exact solution, for each fixed time. In Section 3, for each fixed time, we investigate the stability and convergence of a finite element elliptic projection with quadrature. In Section 4, using the elliptic projections as comparison functions, we prove the uniform boundedness, second order in time and optimal order H^1 norm convergence of the approximate solutions satisfying (1.15), (1.16), (1.17) and (1.18). Numerical experiments in Section 5 confirm the theoretical results, and demonstrate optimal order convergence of the algorithm for a nonlinear parabolic problem in H^k norm for all $k = 0, 1, 2$.

2 Preliminaries

For a nonnegative integer k, in addition to the Sobolev space $H^k(\Omega)$ (with norm $\|\cdot\|_k$), we use the standard norm $\|\cdot\|_{k,\infty}$ in the Banach space $\mathcal{C}^k(\overline{\Omega})$. (Note that $H^0(\Omega) = L^2(\Omega)$, $(.,.) = \|\cdot\|_0^2 = \|\cdot\|_{L^2(\Omega)}^2$, $\mathcal{C}^0(\overline{\Omega}) = \mathcal{C}(\overline{\Omega})$, $\|\cdot\|_{0,\infty} = \|\cdot\|_\infty$.) We define the "broken" H^m and $\mathcal{C}^m$ norms $\|v\|_{m,\mathcal{T}_h}$ and $\|v\|_{m,\infty,\mathcal{T}_h}$ on the quasi-uniform triangulation $\mathcal{T}_h$ by

$$\|v\|_{m,\mathcal{T}_h}^2 = \sum_{\rho\in\mathcal{T}_h} \|v\|_{m,\rho}^2\,, \qquad \|v\|_{m,\infty,\mathcal{T}_h}^2 = \max_{\rho\in\mathcal{T}_h} \|v\|_{m,\infty,\rho}^2\,,$$

where $\|\cdot\|_{m,\rho} = \|\cdot\|_{H^m(\rho)}$ and $\|\cdot\|_{m,\infty,\rho} = \|\cdot\|_{\mathcal{C}^m(\rho)}$.

We assume throughout the paper that u is a solution of (1.1), (1.2) and (1.3) with

$$u \in \mathcal{C}^5(\overline{\Omega}\times[0,T]), \quad u,\, u_t \in \mathcal{C}(H^{r+3}(\Omega),[0,T]),$$

(where $u_t = \frac{\partial u}{\partial t}$), and that the coefficient a in (1.1) is such that

$$a \in \mathcal{C}^5(\overline{\Omega}\times[-\delta,\delta]), \quad \delta = \max_{0\le t\le T}\|u\|_\infty. \tag{2.1}$$

Throughout the paper, for $\phi, \phi_t \in \mathcal{C}(\overline{\Omega})$, $C(\phi)$ denotes a *generic* positive constant depending only on r, ϕ, ϕ_t, and $C_i(\phi)$ is a specific $C(\phi)$, for $i = 1, 2, 3$.

For each fixed $t \in [0, T]$, optimal regularity requirement of the exact solution in the standard H^1-Galerkin method analysis (without quadrature), using splines of degree r, is $H^{r+1}(\Omega)$. It is common in *quadrature* finite element analysis to assume extra regularity [1, 2, 6, 7, 10, 19]. We require $H^{r+3}(\Omega)$ regularity, mainly due to technical details involved in analysis of the quadrature error in Lemma 3.1. Such extra regularity assumption on the exact solution is essential for analysis of $\mathcal{C}^1$ spline collocation methods that do not involve integrals, but require evaluation of functions at certain quadrature points [7, 10, 19]. The analysis in [1, 2] for finite element Galerkin method with quadrature for a linear biharmonic problem, restricted to the $\mathcal{C}^1$ cubic spline ($r = 3$) case, requires $H^{r+5}(\Omega)$ regularity. As demonstrated in [6, Section 5], the extra regularity assumptions on the solution and coefficients may be

required only in *quadrature* finite element analysis, and may not affect convergence rates in practical computations. It is useful to recall that implementation of the scheme (1.15), (1.16), (1.17) and (1.18) does not require such regularity conditions.

Using the approximation property (1.11), we obtain the following result.

Lemma 2.1 *If $\phi \in H^k(\Omega) \cap H_0^1(\Omega),\ 2 \le k \le r+1$, then there exists $\Phi \in S_h$ such that*

$$\|\phi - \Phi\|_\ell \le Ch^{k-\ell}\|\phi\|_k, \qquad \ell = 0, 1, 2.$$

The result of the next lemma is well known in the approximation theory [9].

Lemma 2.2 *If $\phi \in H^k(\Omega),\ k \ge 2$, then there exists a spline Ψ of degree $k-1$ defined on $\mathcal{T}_h$ such that*

$$\|\phi - \Psi\|_j \le Ch^{k-j}\|\phi\|_k, \qquad j = 0, \cdots, k-1.$$

For each $\Phi \in S_h$, $\Delta\Phi$ is a polynomial of degree $r-2$ on $\rho \in \mathcal{T}_h$, and hence the degree precision $2r-2$ of the quadrature (see (1.12)), integration by parts, the Poincaré inequality, and results in [9, Chapter 4], yield properties in the following two lemmas. (The constant $C = 1$ in the first two inequalities in Lemma 2.3 for triangular elements.)

Lemma 2.3 *For any $\Phi, \Psi \in S_h$,*

$$\|\Delta\Phi\|_h \le C\|\Delta\Phi\|_0 \le C\|\Delta\Phi\|_h, \quad \|\Phi\|_1^2 \le C(\Phi, -\Delta\Phi)_h, \text{ and } (\Phi, \Delta\Psi)_h = (\Delta\Phi, \Psi)_h.$$

Lemma 2.4 *If $\mathcal{X}$ is a piecewise polynomial of degree r on $\mathcal{T}_h$ and $g \in H^2(\rho),\ \rho \in \mathcal{T}_h$, then*

$$\left|(\mathcal{X}, g)_\rho - (\mathcal{X}, g)_{\rho,h}\right| \le Ch^2\|g\|_{2,\rho}\|\mathcal{X}\|_{0,\rho}$$

Remark 2.5 Depending on the shape of the domain Ω, it may be convenient to use a rectangular partition $\mathcal{T}_h$ of Ω, to define the finite dimensional approximation space S_h spanned by splines that are polynomials of degree at most r (with respect to *each* variable) on each rectangular element ρ.

In case of rectangular elements, if Ψ is a polynomial of degree r on ρ, $\Delta\Psi$ is also a polynomial of degree r on ρ. Hence, unlike in the triangular elements case, choosing a quadrature formula $(\cdot,\cdot)_{\rho,h}$ on each element ρ with degree of precision $2r-2$ and (1.13) need not directly imply properties stated in Lemmas 2.3 and 2.4. We need to be more precise about the choice of quadrature rule and proof of Lemmas 2.3 and 2.4 in this case. We discuss the rectangular elements and quadrature in Appendix (see Section 6.1).

We use the following bounds throughout the paper. Using the H^2 regularity of the unique solution of the homogeneous Dirichlet problem for the Poisson equation on a convex or a $\mathcal{C}^2$ domain Ω [14, 18], we get the equivalence relation

$$\|\Delta\phi\|_0 \le \|\phi\|_2 \le C\|\Delta\phi\|_0, \qquad \phi \in H^2(\Omega) \cap H_0^1(\Omega). \tag{2.2}$$

The Cauchy–Schwarz inequality applied to the quadrature sum yields

$$|(\phi, \psi)_h| \le C\|\phi\|_h\|\psi\|_h, \quad \text{for any } \phi,\ \psi \text{ defined on all quadrature points in } \Omega. \tag{2.3}$$

Next we derive a non-standard approximation and stability property in S_h using the H^4 regularity of the classical biharmonic problem:

$$\textit{If } \psi \in L^2(\Omega), \textit{ there exists } \chi \in H^4(\Omega) \textit{ such that } \Delta^2\chi = \psi, \textit{ and } \|\chi\|_4 \le C\|\psi\|_0. \tag{A1}$$

The assumption (A1) holds for $\Omega \subset \mathbb{R}^2$ with a *piecewise* smooth boundary satisfying a maximum interior angle condition at the corners [8, Theorem 2] (for example rectangles), and also for $\mathcal{C}^4$ domains [17, 18]. A direct application of the assumption (A1) is required only in the proof of Lemma 2.6. An important relation connecting the quadrature finite element analysis with a variable coefficient biharmonic problem is given in the next section.

Thus, we assume the following sufficient domain condition for (2.2) and (A1) to hold: *The domain Ω is either $\mathcal{C}^4$ or it is convex and Lipschitz continuous with interior angles at the corner points do not exceed* 126 *degrees.*

Lemma 2.6 *Let* (A1) *hold and let* $\widetilde{\phi} \in L^2(\Omega)$. *Then there exist* $\phi, \phi^u \in H^2(\Omega) \cap H_0^1(\Omega)$, *and* $\Phi \in S_h$ *such that*

$$\Delta\phi = \widetilde{\phi}, \quad \phi^u = \phi/\mathcal{A}u, \quad \|\phi^u - \Delta\Phi\|_0 \le C(u)h^2\|\widetilde{\phi}\|_0, \quad \|\Delta\Phi\|_{2,\mathcal{T}_h} \le C(u)\|\widetilde{\phi}\|_0. \tag{2.4}$$

Proof The existence of solution $\phi \in H^2(\Omega) \cap H_0^1(\Omega)$ of the Poisson equation satisfying $\|\phi\|_2 \le C\|\widetilde{\phi}\|_0$ is well known [18]. Hence ϕ^u in (2.4) exists and $\|\phi^u\|_2 \le C(u)\|\widetilde{\phi}\|_0$.

Since $\Delta\phi^u \in L^2(\Omega)$, using (A1) there exists $\psi \in H^4(\Omega)$ such that

$$\Delta^2\psi = \Delta\phi^u, \qquad \|\psi\|_4 \le C\|\Delta\phi^u\|_0 \le C\|\phi^u\|_2 \le C(u)\|\widetilde{\phi}\|_0. \tag{2.5}$$

Hence using $r \ge 3$, Lemma 2.1, and (2.2), there exists $\Phi \in S_h$, such that

$$\|\psi - \Phi\|_\ell \le Ch^{4-\ell}\|\psi\|_4 \le C(u)h^{4-\ell}\|\widetilde{\phi}\|_0 \quad \ell = 0, 1, 2. \tag{2.6}$$

Using (2.5) and (2.6),

$$\|\phi^u - \Delta\Phi\|_0 = \|\Delta(\psi - \Phi)\|_0 \le C\|\psi - \Phi\|_2 \le C(u)h^2\|\widetilde{\phi}\|_0. \tag{2.7}$$

Lemma 2.1 and the last equality in (2.5) yield a spline $\Psi \in S_h$ such that

$$\|\phi^u - \Psi\|_\ell \le C(u)h^{2-\ell}\|\widetilde{\phi}\|_0, \quad \ell = 0, 1, 2. \tag{2.8}$$

The triangle and inverse inequalities, (2.8), and (2.7) yield

$$\begin{aligned}\|\Delta\Phi\|_{2,\mathcal{T}_h} &\le \|\Delta\Phi - \Psi\|_{2,\mathcal{T}_h} + \|\Psi - \phi^u\|_2 + \|\phi^u\|_2 \le C(u)\left(h^{-2}\|\Delta\Phi - \Psi\|_0 + \|\phi^u\|_2\right)\\ &\le C(u)\left(h^{-2}\|\Delta\Phi - \phi^u\|_0 + h^{-2}\|\phi^u - \Psi\|_0 + \|\phi^u\|_2\right) \le C(u)\|\widetilde{\phi}\|_0. \qquad \square\end{aligned}$$

Remark 2.7 Our main analysis in the rest of the paper is based only on the properties stated in Lemmas 2.1 to 2.6. Accordingly, the algorithm and analysis in this paper can be extended for (1.1), (1.2) and (1.3) defined on any domain $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3, \cdots$

with corresponding approximation space and quadrature rule satisfying results in this section. For notational convenience, throughout the paper we restrict to the case $d = 2$.

3 Properties of fully discrete elliptic projections

Following a traditional approach, our convergence analysis is based on splitting the error between the exact solution u and the computable U as $u - U = (u - W) + (W - U)$, where W is a comparison function obtained using a fully discrete elliptic projection of the exact solution u in S_h, for each fixed time. In this section we analyze the first stage error and stability of the comparison function, by choosing $W : [0, T] \to S_h$ to be the solution of

$$(\mathcal{A}u\,\Delta W, \Delta\Psi)_h = (\mathcal{A}u\,\Delta u, \Delta\Psi)_h\,, \quad \Psi \in S_h, \quad t \in [0, T]. \tag{3.1}$$

For each fixed $t \in [0, T]$, the comparison function $W(t) \in S_h$ in (3.1) may also be considered as a quadrature finite element Galerkin solution of the biharmonic-type problem $\Delta(\alpha(\mathbf{x})\Delta w) = \chi(\mathbf{x})$ for the unknown solution w subject to the boundary conditions $w = Bw = 0$ on $\partial\Omega$, where $B = \Delta$ or ∂_n, the outer normal derivative. Hence convergence analysis in this section generalizes that in the recent work [2] (for the constant coefficient biharmonic operator) to the variable coefficient biharmonic problem with less regularity requirement on the solution.

We bound the comparison function error, denoted throughout the paper by $\eta = u - W$, after studying the effect of quadrature approximations of various integrals that will occur in our main analysis.

Lemma 3.1 *Let $\mathcal{D}$ be the linear differential operator, defined for $\psi \in H^2(\Omega)$ by*

$$\mathcal{D}\psi = (\mathcal{D}_2 + \mathcal{D}_1 + \mathcal{D}_0)\psi,\ \mathcal{D}_2\psi = \alpha\Delta\psi,\ \mathcal{D}_1\psi = \boldsymbol{\beta}.\nabla\psi,\ \mathcal{D}_0\psi = \gamma\psi, \quad \alpha, \gamma \in \mathbb{R},\ \boldsymbol{\beta} \in \mathbb{R}^2. \tag{3.2}$$

Let $\Phi \in S_h$ and let $\widetilde{\phi} = \phi - \widetilde{\Phi}$, for some $\phi \in H^{r+3}(\Omega)$ and $\widetilde{\Phi} \in S_h$.

(a) *If $\|\mathcal{D}\widetilde{\phi}\|_0 \le C(\phi)h^{r-1}$, then for any $v \in \mathcal{C}^2(\overline{\Omega})$,*

$$\left|\left(\Delta\Phi, v\mathcal{D}\widetilde{\phi}\right) - \left(\Delta\Phi, v\mathcal{D}\widetilde{\phi}\right)_h\right| \le C(v)C(\phi)h^{r+1}\|\Delta\Phi\|_{2,\mathcal{T}_h}, \tag{3.3}$$

(b) *If $\|\nabla_{x_i}\widetilde{\phi}\|_0 \le C(\phi)h^{r-1}$, and $w_i \in \mathcal{C}^2(\Omega),\ i = 1, 2$ then with $\mathbf{w} = (w_1, w_2)$,*

$$\left|\left(\Delta\Phi, \mathbf{w}.\nabla\widetilde{\phi}\right) - \left(\Delta\Phi, \mathbf{w}.\nabla\widetilde{\phi}\right)_h\right| \le C(\mathbf{w})C(\phi)h^{r+1}\|\Delta\Phi\|_{2,\mathcal{T}_h}. \tag{3.4}$$

(c) *For $j = 0, 1, 2$, if $\|\mathcal{D}_j\widetilde{\phi}\|_0 \le C(\phi)h^{k-j}$, for some $3 \le k \le r + 1$, then*

$$\|\mathcal{D}_j\widetilde{\phi}\|_h \le C(\phi)h^{k-j}, \qquad j = 0, 1, 2. \tag{3.5}$$

(d) *If $\|\mathcal{D}\widetilde{\Phi}\|_0 \le C(\phi)h^{r-1}$, then for any $v \in \mathcal{C}^2(\Omega)$*

$$\left|\left(\Delta\Phi, v\mathcal{D}\widetilde{\Phi}\right) - \left(\Delta\Phi, v\mathcal{D}\widetilde{\Phi}\right)_h\right| \le C(v)C(\phi)h^{r+1}\|\Delta\Phi\|_{2,\mathcal{T}_h}. \tag{3.6}$$

Proof Since $\mathcal{D}\phi \in H^{r+1}(\Omega)$, using Lemma 2.1, there exists a spline $\Psi \in S_h$ of degree r such that

$$\|\mathcal{D}\phi - \Psi\|_\ell \le Ch^{r+1-\ell}\|\phi\|_{r+3}, \quad \ell = 0, 1, 2. \tag{3.7}$$

Using the triangle inequality,

$$\begin{aligned}
&\left|(\Delta\Phi, v\mathcal{D}\widetilde{\phi}) - \left(\Delta\Phi, v\mathcal{D}\widetilde{\phi}\right)_h\right| \\
&\quad \le \left|(\Delta\Phi, v[\mathcal{D}\phi - \Psi]) - (\Delta\Phi, v[\mathcal{D}\phi - \Psi])_h\right| + \left|\left(\Delta\Phi, v[\Psi - \mathcal{D}\widetilde{\Phi}]\right)\right. \\
&\qquad \left. - \left(\Delta\Phi, v\left[\Psi - \mathcal{D}\widetilde{\Phi}\right]\right)_h\right| \equiv I_1 + I_2.
\end{aligned} \tag{3.8}$$

Using Lemma 2.4 (with $g = v[\mathcal{D}\phi - \Psi]$, $\mathcal{X} = \Delta\Phi$), Leibniz's rule, the Cauchy–Schwarz inequality, and (3.7), we have

$$\begin{aligned}
I_1 &\le Ch^2 \sum_{\rho\in\mathcal{T}_h} \|v[\mathcal{D}\phi - \Psi]\|_{2,\rho}\|\Delta\Phi\|_{0,\rho} \le C(v)h^2 \sum_{\rho\in\mathcal{T}_h} \|\mathcal{D}\phi - \Psi\|_{2,\rho}\|\Delta\Phi\|_{0,\rho} \\
&\le C(v)h^2\|\mathcal{D}\phi - \Psi\|_2\|\Delta\Phi\|_0 \le C(v)h^{r+1}\|\phi\|_{r+3}\|\Delta\Phi\|_0.
\end{aligned} \tag{3.9}$$

To bound I_2, we use Lemma 2.4 (with $g = v\Delta\Phi$, $\mathcal{X} = [\Psi - \mathcal{D}\widetilde{\Phi}]$), Leibniz's rule, the Cauchy–Schwarz inequality and triangle inequalities, assumption on $\widetilde{\phi}$ in (a), and (3.7) to obtain

$$\begin{aligned}
I_2 &\le Ch^2 \sum_{\rho\in\mathcal{T}_h} \|v\Delta\Phi\|_{2,\rho}\|\mathcal{D}\widetilde{\Phi} - \Psi\|_{0,\rho} \le C(v)h^2\|\Delta\Phi\|_{2,\mathcal{T}_h}\|\mathcal{D}\widetilde{\Phi} - \Psi\|_0 \\
&\le C(v)h^2\|\Delta\Phi\|_{2,\mathcal{T}_h}\left(\|\mathcal{D}\widetilde{\phi}\|_0 + \|\mathcal{D}\phi - \Psi\|_0\right) \le C(v)C(\phi)h^{r+1}\|\phi\|_{r+3}\|\Delta\Phi\|_{2,\mathcal{T}_h}.
\end{aligned} \tag{3.10}$$

Hence, (3.3) follows from (3.8), (3.9) and (3.10). The bound (3.4) can be proved similarly.

For $j = 0, 1, 2$, since $\mathcal{D}_j\phi \in H^{r+3-j}(\Omega)$ with $r \ge 3$, using Lemma 2.2, there exists a spline Ψ_j of degree $r + 2 - j$ defined on $\mathcal{T}_h$ such that

$$\|\mathcal{D}_j\phi - \Psi_j\|_\ell \le Ch^{r+3-j-\ell}\|\phi\|_{r+3}, \quad j = 0, 1, 2 \ \ \ell = 0, 1, 2, 3. \tag{3.11}$$

The triangle inequality yields

$$\|\mathcal{D}_j\widetilde{\phi}\|_h^2 \le C\left(J_1^j + J_2^j + \|\mathcal{D}_j\phi - \Psi_j\|_0^2 + \|\Psi_j - \mathcal{D}_j\widetilde{\Phi}\|_0^2\right), \qquad j = 0, 1, 2, \tag{3.12}$$

where

$$J_1^j = \|\mathcal{D}_j\phi - \Psi_j\|_h^2 - \|\mathcal{D}_j\phi - \Psi_j\|_0^2, \quad J_2^j = \|\Psi_j - \mathcal{D}_j\widetilde{\Phi}\|_h^2 - \|\Psi_j - \mathcal{D}_j\widetilde{\Phi}\|_0^2. \tag{3.13}$$

Using the triangle inequality, (3.11), and assumption on $\widetilde{\phi}$ in (b), we obtain for $j = 0, 1, 2$, and $3 \le k \le r + 1$,

$$\|\mathcal{D}_j\phi - \Psi_j\|_0^2 + \|\Psi_j - \mathcal{D}_j\widetilde{\Phi}\|_0^2 \le C\left(\|\mathcal{D}_j\phi - \Psi_j\|_0^2 + \|\mathcal{D}_j\widetilde{\phi}\|_0^2\right) \le C(\phi)h^{2k-2j}\|\phi\|_{r+3}^2. \tag{3.14}$$

Using the Bramble–Hilbert lemma (see [9]) for the quadrature formula, Leibniz's rule, Cauchy–Schwarz inequalities, and (3.11), for $j = 0, 1, 2$, and $3 \le k \le r+1$, we get

$$\begin{aligned} J_1^j &\le Ch^3 \sum_{\rho\in\mathcal{T}_h} \sum_{\alpha_1+\alpha_2=3} \int_\rho \left| \frac{\partial^{\alpha_1+\alpha_2}}{\partial x^{\alpha_1}\partial y^{\alpha_2}} [\mathcal{D}_j\phi - \Psi_j]^2(x, y) \right| dxdy \\ &\le Ch^3 \sum_{\rho\in\mathcal{T}_h} \left(\|\mathcal{D}_j\phi - \Psi_j\|_{0,\rho} \|\mathcal{D}_j\phi - \Psi_j\|_{3,\rho} + \|\mathcal{D}_j\phi - \Psi_j\|_{1,\rho} \|\mathcal{D}_j\phi - \Psi_j\|_{2,\rho} \right) \\ &\le Ch^3 \left(\|\mathcal{D}_j\phi - \Psi_j\|_0 \|\mathcal{D}_j\phi - \Psi_j\|_3 + \|\mathcal{D}_j\phi - \Psi_j\|_1 \|\mathcal{D}_j\phi - \Psi_j\|_2 \right) \\ &\le Ch^{2r+6-2j} \|\phi\|_{r+3}^2 \le Ch^{2k-2j} \|\phi\|_{r+3}^2. \end{aligned} \tag{3.15}$$

Similarly, we obtain

$$J_2^j \le Ch^3 \sum_{\rho\in\mathcal{T}_h} \left(\|\Psi_j - \mathcal{D}_j\widetilde{\Phi}\|_{0,\rho} \|\Psi_j - \mathcal{D}_j\widetilde{\Phi}\|_{3,\rho} + \|\Psi_j - \mathcal{D}_j\widetilde{\Phi}\|_{1,\rho} \|\Psi_j - \mathcal{D}_j\widetilde{\Phi}\|_{2,\rho} \right).$$

Hence, the inverse inequality and (3.14), for $j = 0, 1, 2$, and $3 \le k \le r+1$, yield

$$J_2^j \le C \sum_{\rho\in\mathcal{T}_h} \|\Psi_j - \mathcal{D}_j\widetilde{\Phi}\|_{0,\rho}^2 \le Ch^{2k-2j} \|\phi\|_{r+3}^2. \tag{3.16}$$

The bound (3.5) now follows from using (3.14), (3.15) and (3.16) in (3.12).

Finally, using Lemma 2.4 (with $g = v\Delta\Phi$, $\mathcal{X} = \Delta\widetilde{\Phi}$) and the assumption on $\widetilde{\Phi}$, we get

$$|(\Delta\Phi, v\mathcal{D}\widetilde{\Phi}) - (\Delta\Phi, v\mathcal{D}\widetilde{\Phi})_h| \le Ch^2 \|v\Delta\Phi\|_{2,\mathcal{T}_h} \|\mathcal{D}\widetilde{\Phi}\|_0 \le C(v)C(\phi)h^{r+1} \|\Delta\Phi\|_{2,\mathcal{T}_h}. \quad \square$$

For each $t \in [0, T]$, since $u, u_t \in H^{r+3}(\Omega)$, using Lemma 2.1, there exist $\widetilde{W}, \widehat{W} \in S_h$ such that

$$\|u - \widetilde{W}\|_\ell \le Ch^{r+1-\ell} \|u\|_{r+1}, \qquad \|u_t - \widehat{W}\|_\ell \le Ch^{r+1-\ell} \|u_t\|_{r+1}, \qquad \ell = 0, 1, 2. \tag{3.17}$$

Hence from Lemma 3.1(c)

$$\|\Delta(u - \widetilde{W})\|_h \le C(u)h^{r-1}, \qquad \|\Delta(u_t - \widehat{W})\|_h \le C(u)h^{r-1}. \tag{3.18}$$

Next we bound the comparison function error η in H^j norm, and $\mathcal{D}_j\eta$ in discrete norm $\|\cdot\|_h$, for $j = 0, 1, 2$, where $\mathcal{D}_j$ is as defined in (3.2) with $\alpha = \gamma = \beta_1 = \beta_2 = 1$. Using (3.1), the comparison function error $\eta = u - W$ is the solution of

$$(\mathcal{A}u\,\Delta\eta, \Delta\Psi)_h = 0, \quad \Psi \in S_h, \quad t \in [0, T].$$

Theorem 3.2 *For each $t \in [0, T]$, (3.1) has a unique solution and*

$$\|\eta\|_j \le C(u)h^{r+1-j}, \qquad \|\mathcal{D}_j\eta\|_h \le C(u)h^{r+1-j}, \qquad j = 0, 1, 2. \tag{3.19}$$

Proof Since the linear system corresponding to (3.1) is square, it is enough to show that the solution of (3.1) is unique. If there exist two solutions W and $Z \in S_h$ of (3.1), then

$$(\mathcal{A}u\,\Delta(W-Z),\,\Delta\Psi)_h = 0, \quad \Psi \in S_h. \tag{3.20}$$

Using (2.2), Lemma 2.3, (1.4), and $\Psi = W - Z$ in (3.20), we get

$$\begin{aligned}\|W-Z\|_2^2 &\le C\|\Delta(W-Z)\|_h^2 \le C\|\sqrt{\mathcal{A}u}\,\Delta(W-Z)\|_h^2\\ &= C(\mathcal{A}u\,\Delta(W-Z),\,\Delta(W-Z))_h = 0,\end{aligned}$$

and hence $W = Z$.

Since $\eta = u - W$, with $u \in H^{r+3}(\Omega)$ and $W \in S_h$, using Lemma 3.1(c), it is sufficient to prove the first inequality in (3.19). The triangle inequality and (3.17) yield

$$\|\eta\|_2 \le \|u - \widetilde{W}\|_2 + \|\widetilde{W} - W\|_2 \le Ch^{r-1}\|u\|_{r+1} + \|\widetilde{W} - W\|_2. \tag{3.21}$$

Using (2.2), Lemma 2.3, (1.4), (3.1), (2.3), (3.18), and Lemma 2.3,

$$\begin{aligned}\|\widetilde{W} - W\|_2^2 &\le C\|\Delta(\widetilde{W} - W)\|_h^2\\ &\le C\|\sqrt{\mathcal{A}u}\,\Delta(\widetilde{W} - W)\|_h^2\\ &= C\left(\mathcal{A}u\,\Delta(\widetilde{W} - W),\,\Delta(\widetilde{W} - W)\right)_h\\ &= C\left(\mathcal{A}u\,\Delta(\widetilde{W} - u),\,\Delta(\widetilde{W} - W)\right)_h\\ &\le C\|\mathcal{A}u\|_\infty\|\Delta(\widetilde{W} - u)\|_h\|\Delta(\widetilde{W} - W)\|_h\\ &\le Ch^{r-1}\|\mathcal{A}u\|_\infty\|u\|_{r+3}\|\widetilde{W} - W\|_2, \quad t \in [0,\,T].\end{aligned} \tag{3.22}$$

Hence from (3.21) and (3.22), we obtain $\|\eta\|_2 \le C(u)h^{r-1}$.

Next we bound η in L^2 and H^1 norms. For each $t \in [0,\,T]$, using Lemma 2.6, with $\widetilde{\phi} = \eta$, and ϕ, ϕ^u as in (2.4), there exists $\Phi \in S_h$ and such that

$$\|\phi^u - \Delta\Phi\|_0 \le C(u)h^2\|\eta\|_0, \qquad \|\Delta\Phi\|_{2,\mathcal{T}_h} \le C(u)\|\eta\|_0. \tag{3.23}$$

Using ϕ, $\eta = 0$ on $\partial\Omega$, integration by parts and (3.1), we obtain

$$\begin{aligned}\|\eta\|_0^2 &= (\Delta\phi,\,\eta) = (\Delta(\mathcal{A}u\phi^u),\,\eta) = (\mathcal{A}u\phi^u,\,\Delta\eta)\\ &= (\phi^u - \Delta\Phi,\,\mathcal{A}u\,\Delta\eta) + \{(\Delta\Phi,\,\mathcal{A}u\,\Delta\eta) - (\Delta\Phi,\,\mathcal{A}u\,\Delta\eta)_h\} \equiv J_1 + J_2.\end{aligned} \tag{3.24}$$

The Cauchy–Schwarz inequality, (3.23), and (3.19) for $\ell = 2$ yield

$$|J_1| \le C\|\phi^u - \Delta\Phi\|_0\|\mathcal{A}u\,\Delta\eta\|_0 \le C(u)h^2\|\eta\|_0\|\eta\|_2 \le C(u)h^{r+1}\|\eta\|_0. \tag{3.25}$$

Since the first inequality in (3.19) holds for $j = 2$, using Lemma 3.1(a) (with $\alpha = 1$ being the only non-zero coefficient in (3.2)), (3.3) and (3.23), we obtain

$$|J_2| \le C(u)h^{r+1}\|\eta\|_0. \tag{3.26}$$

Using (3.25) and (3.26) in (3.24) we get the first inequality in (3.19) for $j = 0$.

Further, the Poincaré inequality, integration by parts, the Cauchy–Schwarz inequality, and the first inequality in (3.19) for $j = 0,\,2$ yield

$$\|\eta\|_1^2 \le -C(\Delta\eta,\,\eta) \le C\,\|\eta\|_2\,\|\eta\|_0 \le C(u)h^{2r},$$

and hence, the first desired result in (3.19) is obtained for $j = 0, 1, 2$. Using (3.2), the triangle inequality and the first bound in (3.19), we obtain $\|\mathcal{D}_j\eta\|_0 \le C(u)h^{r+1-j}$. Using this and (3.5) with η in place of $\widetilde{\phi}$, we obtain the second bound in (3.19). □

Next we show that η_t has optimal error bounds in H^1 and H^2 norms, and hence we obtain a sub-optimal error bound of η_t in discrete norm $\|\cdot\|_h$.

Theorem 3.3 *For each* $t \in [0, T]$,

$$\|\eta_t\|_2 \le C(u)h^{r-1}, \qquad \|\eta_t\|_1 \le C(u)h^r, \qquad \|\eta_t\|_h \le C(u)h^r. \tag{3.27}$$

Proof Differentiating both sides of (3.1) with respect to t, we get that $W, W_t \in S_h$ satisfy

$$(\mathcal{A}_t u\,\Delta W + \mathcal{A}u\,\Delta W_t, \Delta\Psi)_h = (\mathcal{A}_t u\,\Delta u + \mathcal{A}u\,\Delta u_t, \Delta\Psi)_h\,, \quad v \in S_h, \quad t \in [0, T], \tag{3.28}$$

where $[\mathcal{A}_t u(t)] = \frac{\partial}{\partial t}[\mathcal{A}u(t)]$.

Following (3.1), let $W^* : [0, T] \to S_h$ be such that

$$\left(\mathcal{A}u\,\Delta W^*, \Delta\Psi\right)_h = (\mathcal{A}u\,\Delta u_t, \Delta\Psi)_h\,, \quad \Psi \in S_h, \quad t \in [0, T]. \tag{3.29}$$

Using the triangle inequality, (3.17), arguments similar to (3.22) (with W, $\widetilde{W}$ and (3.1) replaced by W^*, $\widehat{W}$ and (3.29) respectively), we obtain

$$\|u_t - W^*\|_2 \le \|u_t - \widehat{W}\|_2 + \|\widehat{W} - W^*\|_2 \le C(u)h^{r-1}. \tag{3.30}$$

The triangle inequality and (3.30) yield

$$\|\eta_t\|_2 \le \|u_t - W^*\|_2 + \|W^* - W_t\|_2 \le C(u)h^{r-1} + \|W^* - W_t\|_2. \tag{3.31}$$

Using (2.2), Lemma 2.3, (1.4), (3.28), (3.29), (2.3), Theorem 3.2 and Lemma 2.3, we get

$$\begin{aligned}\|W^* - W_t\|_2^2 &\le C\|\Delta(W^* - W_t)\|_h^2 \le C\|\sqrt{\mathcal{A}u}\,\Delta(W^* - W_t)\|_h^2 \\ &= C\left(\mathcal{A}u\,\Delta(W_t - W^*), \Delta(W_t - W^*)\right)_h = C\left(\mathcal{A}_t u\,\Delta\eta, \Delta(W_t - W^*)\right)_h \\ &\le C\|\mathcal{A}_t u\|_\infty\|\Delta\eta\|_h\|\Delta(W^* - W_t)\|_h \le C(u)h^{r-1}\|W^* - W_t\|_2. \end{aligned}\tag{3.32}$$

Using (3.32) in (3.30), we obtain $\|\eta_t\|_2 \le C(u)h^{r-1}$. Next, to bound η_t in H^1 norm, we define a function $p \in H^2(\Omega)$ and obtain its bound using the triangle inequality, (3.27) and the first inequality in (3.19) with $j = 2$:

$$p = \eta_t + \frac{\mathcal{A}_t u}{\mathcal{A}u}\eta, \qquad \|p\|_2 \le C(\|\eta_t\|_2 + C(u)\|\eta\|_2) \le C(u)h^{r-1}. \tag{3.33}$$

Since

$$\Delta p = \Delta\eta_t + q + \frac{\mathcal{A}_t u}{\mathcal{A}u}\Delta\eta, \qquad q = \Delta\left(\frac{\mathcal{A}_t u}{\mathcal{A}u}\right)\eta + 2\nabla\left(\frac{\mathcal{A}_t u}{\mathcal{A}u}\right).\nabla\eta, \tag{3.34}$$

using (3.28), it is easy to check that

$$(\mathcal{A}u\,\Delta p, \Delta\Psi)_h = (\mathcal{A}u\,q, \Delta\Psi)_h\,, \quad \Psi \in S_h, \quad t \in [0, T]. \tag{3.35}$$

We approximate p by P, where $P \in S_h$ is the solution of

$$(\mathcal{A}u\,\Delta P, \Delta\Psi)_h = (\mathcal{A}u\,q, \Delta\Psi)_h, \quad \Psi \in S_h, \quad t \in [0, T]. \tag{3.36}$$

Using (2.2), Lemma 2.3, (3.36), the Cauchy–Schwarz inequality, (3.34), (1.4), and Theorem 3.2, we obtain

$$\|P\|_2^2 \le C\|\Delta P\|_0^2 \le C\|\Delta P\|_h^2 \le C(\mathcal{A}u\Delta P, \Delta P)_h \le C(u)h^r\|\Delta P\|_0, \tag{3.37}$$

Further, using (2.2), the triangle inequality, (3.33), and (3.37), we obtain

$$\|p - P\|_2 \le C\|\Delta(p - P)\|_0 \le C\,(\|\Delta p\|_0 + \|\Delta P\|_0) \le C(u)h^{r-1} \tag{3.38}$$

Using Lemma 2.6, with $\widetilde{\phi} = p - P$, and ϕ, ϕ^u as in (2.4), there exists $\Phi \in S_h$ such that

$$\|\phi^u - \Delta\Phi\|_0 \le C(u)h^2\|p - P\|_0, \qquad \|\Delta\Phi\|_{2,\mathcal{T}_h} \le C(u)\|p - P\|_0, \tag{3.39}$$

Using (3.35) and (3.36),

$$(\mathcal{A}u\Delta(p - P), \Delta\Phi)_h = 0. \tag{3.40}$$

Hence $\Delta(\mathcal{A}u\,\phi^u) = p - P$ and (3.40) yield

$$\begin{aligned}\|p - P\|_0^2 &= (\phi^u - \Delta\Phi, \mathcal{A}u\,\Delta(p - P)) \\ &\quad + \{(\Delta\Phi, \mathcal{A}u\,\Delta(p - P)) - (\Delta\Phi, \mathcal{A}u\,\Delta(p - P))_h\} \equiv \widetilde{J}_1 + \widetilde{J}_2.\end{aligned} \tag{3.41}$$

Using arguments similar to the ones used to bound J_1 in (3.25), with η replaced by $p - P$, we have

$$\widetilde{J}_1 \le C(u)h^{r+1}\|p - P\|_0. \tag{3.42}$$

To bound $\widetilde{J}_2$, we use (3.34), the triangle inequality, and write $\widetilde{J}_2 \le \sum_{i=1}^5 \widetilde{J}_2^i$, where,

$$\begin{aligned}
\widetilde{J}_2^1 &= \left|(\Delta\Phi, \mathcal{A}u\,\Delta\eta_t) - (\Delta\Phi, \mathcal{A}u\,\Delta\eta_t)_h\right| \\
\widetilde{J}_2^2 &= \left|\left(\Delta\Phi, \mathcal{A}u\,\Delta\left(\frac{\mathcal{A}_t u}{\mathcal{A}u}\right)\eta\right) - \left(\Delta\Phi, \mathcal{A}u\,\Delta\left(\frac{\mathcal{A}_t u}{\mathcal{A}u}\right)\eta\right)_h\right| \\
\widetilde{J}_2^3 &= 2\left|\left(\Delta\Phi, \mathcal{A}u\,\nabla\left(\frac{\mathcal{A}_t u}{\mathcal{A}u}\right).\nabla\eta\right) - \left(\Delta\Phi, \mathcal{A}u\,\nabla\left(\frac{\mathcal{A}_t u}{\mathcal{A}u}\right).\nabla\eta\right)_h\right| \\
\widetilde{J}_2^4 &= \left|(\Delta\Phi, \mathcal{A}_t u\,\Delta\eta) - (\Delta\Phi, \mathcal{A}_t u\,\Delta\eta)_h\right| \\
\widetilde{J}_2^5 &= \left|\,(\Delta\Phi, \mathcal{A}u\,\Delta P) - (\Delta\Phi, \mathcal{A}u\,\Delta P)_h\,\right|.
\end{aligned}$$

Since the first inequalities in (3.19) and (3.27) hold for $j = 2$, $\|\Delta P\|_0 \le Ch^r$, and using Lemma 3.1 (with coefficients in (3.2) chosen appropriately to be zero or one), (3.3), (3.4) and (3.39), we obtain

$$\widetilde{J}_2 \le \sum_{i=1}^{5} \widetilde{J}_2^i \le C(u)h^{r+1}\|p - P\|_0. \tag{3.43}$$

Therefore, using (3.42) and (3.43) in (3.41), we get

$$\|p - P\|_0 \le C(u)h^{r+1}. \tag{3.44}$$

The Poincaré inequality, integration by parts, the Cauchy–Schwarz inequality, (3.38), and (3.44) yield

$$\|p - P\|_1^2 \le -C(\Delta(p-P), p-P) \le C\,\|\Delta(p-P)\|_0\,\|(p-P)\|_0 \le C(u)h^{2r}. \quad (3.45)$$

Now we are ready to bound $\|\eta_t\|_1$. Using (3.33), the triangle inequality, (3.19), (3.45), and (3.37), we obtain

$$\|\eta_t\|_1 = \|p - \frac{\mathcal{A}_t u}{\mathcal{A}u}\eta\|_1 \le C(u)[\|p\|_1 + \|\eta\|_1] \le C(u)[\|p-P\|_1 + \|P\|_1 + h^r] \le C(u)h^r.$$

Finally, since $u_t \in H^{r+3}(\Omega)$, $W_t \in S_h$, the last inequality in (3.27) follows from the bound $\|\eta_t\|_0 \le \|\eta_t\|_1 \le C(u)h^r$, and Lemma 3.1 (b) (with $j=0,\ \gamma=1,\ k=r,\ \widetilde{\phi}=\eta_t$). □

We complete this section with uniform boundedness of the comparison functions W, W_t in various norms.

Lemma 3.4 *For each* $t \in [0, T]$,

$$\|W\| \le C(u), \quad \|W_t\| \le C(u), \quad \|\cdot\| = \|\cdot\|_{1,\infty},\ \text{or}\ \|\cdot\| = \|\cdot\|_2,\ \text{or}\ \|\cdot\| = \|\cdot\|_{2,\infty,\mathcal{T}_h}.$$

Proof Since $u \in H^{r+3}(\Omega)$, using Lemma 2.2, there exists a spline Ψ of degree $r+2$ such

$$\|u - \Psi\|_k \le Ch^{r+3-k}\,\|u\|_{r+3} \qquad k = 0, \cdots, r+2.$$

Hence using the triangle and inverse inequalities, the Sobolev embedding theorem, and (3.19), we obtain

$$\begin{aligned}\|W\| &\le \|W - \Psi\| + \|u - \Psi\| + \|u\| \\ &\le Ch^{-1}\|W - \Psi\|_2 + C\|u - \Psi\|_{r+2} + C\|u\|_{r+2} \\ &\le Ch^{-1}[\|\eta\|_2 + \|u - \Psi\|_2] + C(u) \le C(u),\end{aligned}$$

and hence, the first desired inequality is obtained. Similarly (using (3.27) instead of (3.19)), we obtain the second inequality. □

4 Convergence analysis

In this section, we prove optimal order convergence of our scheme by analyzing the error $u^n - U^n$, for $n = 0, \ldots, N^t$ in H^1 norm, where U^1 and $\{U^n\}_{n=2}^{N^t}$ are defined by (1.18) and (1.15) respectively, and $U^0 = W^0$ (with W as in (3.1)).

Since $u^n - U^n = \eta^n - \xi^n$, with $\eta^n = u^n - W^n$, $\xi^n = U^n - W^n$, $n = 0, \ldots, N^t$, first we derive bounds on ξ between any two consecutive time levels in Theorem 4.1. Then we bounded the term ξ^1 in Lemma 4.1 using the fact that $\xi^0 = 0$ and assuming that the time discretization parameter τ and the spatial mesh size h satisfy the relation $\tau \le Ch^{2/3}$.

We prove the following two results in Appendix (see Section 6.2).

Theorem 4.1 *If h and τ are sufficiently small, then for $n = 1, \cdots, N^t - 1$, we have*

$$-(\xi^{n+1}, \Delta\xi^{n+1})_h + (\xi^n, \Delta\xi^n)_h \le C(u)\tau[d(n,u,U)]^2 \left[\tau^4 + h^{2r} + \sum_{i=n-1}^{n+1} (\xi^i, -\Delta\xi^i)_h\right], \tag{4.1}$$

where

$$d(n,u,U) = \left[\exp\left(\max_{0\le t\le T}\|u\|_{1,\infty}\right) + \exp(\|\widehat{U}^{n+\frac{1}{2}}\|_{1,\infty})\right]^{s+1}. \tag{4.2}$$

Lemma 4.2 *Assume that h and τ are sufficiently small, $\tau \le Ch^{2/3}$, $V^0 = U^0 = W^0$ and $V^1,\ U^1 \in S_h$ are defined by (1.17) and (1.18). Then*

$$-(\xi^1, \Delta\xi^1)_h \le C(u)\left[\tau^4 + h^{2r}\right]. \tag{4.3}$$

Now we are ready to prove uniform boundedness and optimal order convergence of the approximate solutions to the exact solution u of (1.1), (1.2) and (1.3) in H^1 norm. We need to restrict our analysis to H^1, mainly due to obtaining only h^{2r} error bound in (4.1). For convergence results in L^2 norm, we need h^{2r+2} instead of h^{2r} in (4.1).

In addition to obtaining optimal order error bounds in H^1 norm for the algorithm (1.15), (1.16), (1.17) and (1.18), an important contribution in this paper is to prove uniform boundedness of the approximate solutions in supremum norm (see (4.8)) with mild nonlinear conditions (1.5), (1.6), (1.7) and (1.8) for the model problem (1.1), (1.2) and (1.3).

Theorem 4.3 *Assume that h and τ are sufficiently small, $\tau \le Ch^{2/3}$, $V^0 = U^0 = W^0$ and $V^1, U^1 \in S_h$ are defined by (1.17) and (1.18). Then*

$$\max_{0\le n\le N^t}\|u^n - U^n\|_1 \le C(u)\left[\tau^2 + h^r\right]. \tag{4.4}$$

Proof Using Lemmas 3.4 and 4.2, there exists a constant $C_1(u)$ independent of h and τ such that

$$\max_{0\le t\le T}\|W\|_{1,\infty} \le C_1(u)/4, \qquad (\xi^1, -\Delta\xi^1)_h \le C_1(u)(\tau^4 + h^{2r}). \tag{4.5}$$

The relation $\widehat{U}^{1+\frac{1}{2}} = \left[3(W^1 + \xi^1) - W^0\right]/2$, the triangle and inverse inequalities, Lemma 2.3, (4.5), $\tau^3 \le Ch^2$, and h sufficiently small yield a constant $C_2(u)$ independent of h and τ such that

$$\|\widehat{U}^{1+\frac{1}{2}}\|_{1,\infty} \le C\left(\|W^0\|_{1,\infty} + \|W^1\|_{1,\infty} + h^{-1}\|\xi^1\|_1\right) \le C_2(u). \tag{4.6}$$

Let

$$C_3(u) = \max\{C_1(u), C_2(u)\}. \tag{4.7}$$

Next, we claim using mathematical induction that

$$\|\widehat{U}^{n+\frac{1}{2}}\|_{1,\infty} \le C_3(u), \quad n = 1, \ldots, N^t. \tag{4.8}$$

Using (4.6) and (4.7), clearly (4.8) holds for $n = 1$. Assuming that $\|\widehat{U}^{n+\frac{1}{2}}\|_{1,\infty} \le C_3(u)$, $n = 1, \ldots, \ell$, for some $1 \le \ell \le N^t - 1$, and using (4.2), we have $d(n, u, U) \le C(u)$, $n = 1 \ldots, \ell$. Hence Theorem 4.1 gives

$$-(\xi^{n+1}, \Delta\xi^{n+1})_h + (\xi^n, \Delta\xi^n)_h \le C(u)\tau \left[\tau^4 + h^{2r} + \sum_{i=n-1}^{n+1} (\xi^i, -\Delta\xi^i)_h \right], \quad n = 1, \ldots, \ell. \tag{4.9}$$

Summing both sides of (4.9) for $n = 1, \ldots, k-1$, where $2 \le k \le \ell + 1$, we obtain

$$-(\xi^k, \Delta\xi^k)_h \le C(u) \left[-(\xi^1, \Delta\xi^1)_h + \tau^4 + h^{2r} + \tau \sum_{n=0}^{k} (\xi^n, -\Delta\xi^n)_h \right]. \tag{4.10}$$

Clearly (4.10) holds for $k = 1$, and since $\xi^0 = 0$, it also holds for $k = 0$. Hence, for τ sufficiently small, the discrete analogue of Gronwall's inequality (see Lemma 4.7 in [12]) yields

$$-(\xi^n, \Delta\xi^n)_h \le C(u) \left[-(\xi^1, \Delta\xi^1)_h + \tau^4 + h^{2r} \right], \quad n = 0, \ldots, \ell + 1. \tag{4.11}$$

Using the inverse inequality, Lemma 2.3, (4.11), (4.5), $\tau^4 \le Ch^{\frac{8}{3}}$, and taking h sufficiently small, we get

$$\|\xi^n\|_{1,\infty}^2 \le Ch^{-2}\|\xi^n\|_1^2 \le Ch^{-2}(\xi^n, -\Delta\xi^n)_h \le C(u)h^{-2}\left(\tau^4 + h^{2r}\right) \le \frac{[C_3(u)]^2}{16}, \quad n = \ell, \ell + 1. \tag{4.12}$$

Using the relation $\widehat{U}^{\ell+1+\frac{1}{2}} = \widehat{W}^{\ell+1+\frac{1}{2}} + \widehat{\xi}^{\ell+1+\frac{1}{2}}$, the triangle inequality, (4.5), and (4.12), we obtain

$$\|\widehat{U}^{\ell+1+\frac{1}{2}}\|_{1,\infty} \le \frac{3}{2}\|W^{\ell+1}\|_{1,\infty} + \frac{1}{2}\|W^{\ell}\|_{1,\infty} + \frac{3}{2}\|\xi^{\ell+1}\|_{1,\infty} + \frac{1}{2}\|\xi^{\ell}\|_{1,\infty} \le C_3(u),$$

which completes the proof of (4.8) by induction.

To show (4.4), we use (4.8) and follow the derivation of (4.11) to obtain

$$(\xi^n, -\Delta\xi^n)_h \le C(u) \left[(\xi^1, -\Delta\xi^1)_h + \tau^4 + h^{2r} \right], \quad n = 0, \ldots, N^t.$$

Hence Lemmas 2.3 and 4.2 imply that

$$\|\xi^n\|_1^2 \le C(u)\left(\tau^4 + h^{2r}\right), \quad n = 0, \ldots, N^t. \tag{4.13}$$

Since $u^n - U^n = \eta^n - \xi^n$, (4.4) follows from the triangle inequality, Theorem 3.2, and (4.13). □

5 Numerical experiments

We consider the nonlinear parabolic test problem

$$\frac{\partial u}{\partial t}(\mathbf{x}, t) = \nabla \cdot \left[u^2(\mathbf{x}, t)\nabla u(\mathbf{x}, t)\right] + t^3 \exp(u(\mathbf{x}, t)) + \widetilde{f}(\mathbf{x}, t), \quad (\mathbf{x}, t) \in \Omega \times (0, T], \quad (5.1)$$

$$u(\mathbf{x}, 0) = g(\mathbf{x}), \quad \mathbf{x} \in \overline{\Omega}, \quad (5.2)$$

$$u(\mathbf{x}, t) = 0, \quad (\mathbf{x}, t) \in \partial\Omega \times (0, T], \quad (5.3)$$

where Ω is the unit square, the forcing term $\widetilde{f}(\mathbf{x}, t)$ and initial condition $g(\mathbf{x})$ are chosen so that the exact solution of (5.1), (5.2) and (5.3) is

$$u(\mathbf{x}, t) = [\sin(t) + \cos(t)] \sin(\pi x) \sin(\pi y), \qquad \mathbf{x} = (x, y) \in \Omega, \qquad t \in [0, T].$$

The model problem (5.1), (5.2) and (5.3) with (Frank–Kamenetskii type) exponential reaction, and polynomial advection and diffusion terms can be written in the form (1.1), (1.2) and (1.3), with $a(\mathbf{x}, u(\mathbf{x}, t)) = u^2(\mathbf{x}, t)$ (not satisfying the positivity requirement (1.4) in analysis),

$$f(\mathbf{x}, t, u, \nabla u) = t^3 \exp(u(\mathbf{x}, t)) + \widetilde{f}(\mathbf{x}, t) + 2u(\mathbf{x}, t)\left[\frac{\partial u}{\partial x}(\mathbf{x}, t)\right]^2 + 2u(\mathbf{x}, t)\left[\frac{\partial u}{\partial y}(\mathbf{x}, t)\right]^2.$$

We computed the numerical solutions $U^n \in S_h$, $n = 1, \cdots, N^t$, of (5.1), (5.2) and (5.3) by solving the linear systems (1.15), (1.17) and (1.18) in two finite element spaces S_h, spanned by the cubic ($r = 3$) and quartic ($r = 4$) spline bases. The basis functions were constructed on various triangulations of Ω using rectangular elements with a uniform mesh size parameter h. The uniform time partition parameter $\tau = T/N^t$ was chosen so that $\tau^2 \leq h^{r+1}$ to measure the rate of convergence of the computed solutions.

The initial approximate solution $U^0 \in S_h$, required for (1.17) and (1.18), was computed by solving the linear system (3.1) with $t = 0$, using the initial condition (5.2). We used the quadrature described in Section 6.1. Our analysis and the number of quadrature points prescribed in Section 6.1 is optimal in the sense that reducing the precision of quadrature by even one degree did not yield optimal experimental convergence rates observed in the following tables. We calculated the errors

$$\epsilon_{k,\infty} = \max_{0 \leq n \leq N^t} \|u^n - U^n\|_k, \quad k = 0, 1, 2,$$

using 25600 Gauss quadrature points in Ω, and then calculated the rate of convergence ($R(H^k)$) in H^k norm, for $k = 0, 1, 2$. Results in Tables 1 and 2 confirm the

Table 1 Errors and optimal order convergence rates for the case $r = 3$

h	$\epsilon_{0,\infty}$	$R(H^0)$	$\epsilon_{1,\infty}$	$R(H^1)$	$\epsilon_{2,\infty}$	$R(H^2)$
0.25000	5.5713e-03		6.0496e-02		1.0520e+00	
0.12500	2.3454e-04	4.5701	4.1345e-03	3.8710	1.4383e-01	2.8707
0.06250	1.0904e-05	4.4269	3.0451e-04	3.7631	2.0137e-02	2.8364
0.03125	6.7513e-07	4.0135	3.1648e-05	3.2663	4.3848e-03	2.1993

Table 2 Errors and optimal order convergence rates for the case $r = 4$

h	$\epsilon_{0,\infty}$	$R(H^0)$	$\epsilon_{1,\infty}$	$R(H^1)$	$\epsilon_{2,\infty}$	$R(H^2)$
0.25000	1.1427e-03		1.7491e-02		4.4372e-01	
0.12500	1.4639e-05	6.2865	2.0892e-04	6.3876	1.0837e-02	5.3555
0.06250	3.0518e-07	5.5840	8.9139e-06	4.5507	7.2817e-04	3.8956
0.03125	9.0067e-09	5.0825	4.0232e-07	4.4696	6.6759e-05	3.4472

optimal convergence rates $(r + 1 - k)$ of our scheme in H^k norm, proved in Theorem 4.3 for $k = 1$, and expected for $k = 0, 2$.

6 Appendix

In Section 6.1, we describe a special quadrature treatment for rectangular elements mentioned in Remark 2.5, and in Section 6.2 we prove Theorem 4.1 and Lemma 4.2.

6.1 Rectangular elements and quadrature

As described in Remark 2.5, in case of rectangular elements, we need to be precise about the choice of quadrature rule to obtain results in Section 2. We give details below by assuming in this subsection that Ω is the unit square.

For a given small parameter h, it is convenient to use a quasi-uniform collection $\mathcal{T}_h$ of rectangular elements ρ to partition Ω, with $h = \max(h^x, h^y)$, where $\rho = (x_{k-1}, x_k) \times (y_l, y_{l-1})$, $h^{x,k} = x_k - x_{k-1}$ $k = 1, \cdots, N^x$, $l = 1, \cdots, N^y$, $h^x = \max_{1 \le k \le N^x} h^{x,k}$ and h^y defined similarly.

In this case, $S_h \subset H^2(\Omega) \cap H_0^1(\Omega)$ is a tensor-product space consisting of splines that are polynomials of degree at most r (with respect to *each* variable x, y) on each rectangular element ρ. The approximation properties stated in Section 2 (Lemmas 2.1 and 2.2) hold for this finite dimensional space. (See for example Theorem 4 in [11] and Theorem 3.3 in [12]). Next we consider a concrete quadrature rule.

We start with a $J = r - 1$-point Gauss quadrature rule on $(0, 1)$ (for $r \ge 4$) with weights and nodes $\{w_j\}_{j=1}^J$ and $\{\xi_j\}_{j=1}^J$. This rule induces a quadrature rule with degree of precision $2J - 1$, on $\rho = (x_{k-1}, x_k) \times (y_l, y_{l-1})$, $k = 1, \cdots, N^x$, $l = 1, \cdots, N^y$, defined by

$$\int_\rho v(\mathbf{x}) z(\mathbf{x})\, d\mathbf{x} = (v, z)_\rho \approx (v, z)_{\rho,h} = h^{x,k} h^{y,l} \sum_{m,n=1}^{J} w_m w_n (vz)(x_{k,m}, y_{l,m}), \tag{6.4}$$

with $x_{k,m} = x_{k-1} + h^{x,k}\xi_m$ and $y_{l,n} = y_{l-1} + h^{y,l}\xi_n$. Lemma 2.4 for this rectangular element discrete inner product follows from application of [6, Lemma 2.6], provided that $J \ge (r + 2)/2$. Hence we need $J = r$ for the cubic spline $r = 3$ case. (The proof of [6, Lemma 2.6] is based on arguments used in [9, Chapter 4]).

Below we give a proof of Lemma 2.3 to complete all results in Section 2 for rectangular elements. Let Φ, $\Psi \in S_h$. First we prove that

$$(\Phi_{xx}, \Phi_{yy})_h \ge 0\,. \tag{6.5}$$

For $r=3$, since $\Phi\in S_h$, the chosen quadrature rule has degree of precision $2r-1$, $\Phi_{yy}(\alpha, y)=\Phi_{xx}(x,\alpha)=0$ for $\alpha=0,1,\ x\in[0,1]$, using integration by parts we obtain

$$(\Phi_{xx},\Phi_{yy})_h=(\Phi_{xx},\Phi_{yy})=-(\Phi_x,\Phi_{xyy})=(\Phi_{xy},\Phi_{xy})=\|\Phi_{xy}\|_0^2\geq 0. \tag{6.6}$$

For $r\geq 4$, using Lemma 3.1 in [10] and $\Phi_{xx}(x,\alpha)=0$ for $\alpha=0,1,\ x\in[0,1]$, with $y_{l-\frac{1}{2}}=\frac{y_{l-1}+y_l}{2}$, we obtain

$$\begin{aligned}
(\Phi_{xx},\Phi_{yy})_h&=\sum_{k=1}^{N^x}h^{x,k}\sum_{m=1}^{r-1}w_m\sum_{l=1}^{N^y}h^{y,l}\sum_{n=1}^{r-1}w_n(\Phi_{xx}\Phi_{yy})(x_{k,m},y_{l,n})\\
&=-\sum_{k=1}^{N^x}h^{x,k}\sum_{m=1}^{r-1}w_m\left(\int_0^1(\Phi_{xxy}\Phi_y)(x_{k,m},y)\,dy\right.\\
&\quad\left.+C\sum_{l=1}^{N^y}(h^{y,l})^{2r-1}\frac{\partial^{2r}}{\partial y^{2r}}(\Phi_{xx}\Phi)(x_{k,m},y_{l-\frac{1}{2}})\right)\\
&=-\int_0^1\sum_{k,m=1}^{N^x,r-1}h^{x,k}w_m(\Phi_{xxy}\Phi_y)(x_{k,m},y)\,dy\\
&\quad-C\sum_{l=1}^{N^y}(h^{y,l})^{2r-1}\sum_{k,m=1}^{N^x,r-1}h^{x,k}w_m\frac{\partial^{2r}}{\partial y^{2r}}(\Phi_{xx}\Phi)(x_{k,m},y_{l-\frac{1}{2}}).
\end{aligned}\tag{6.7}$$

Similarly and using $\Phi_y(\alpha,y)=0$ for $\alpha=0,1,\ y\in[0,1]$, we get

$$\begin{aligned}
&-\sum_{k=1}^{N^x}h^{x,k}\sum_{m=1}^{r-1}w_m(\Phi_{xxy}\Phi_y)(x_{k,m},y)\\
&=\int_0^1(\Phi_{xy}\Phi_{xy})(x,y)\,dx\\
&\quad+C\sum_{k=1}^{N^x}(h^{x,k})^{2r-1}\left(\frac{\partial^r}{\partial x^r}\Phi_y\right)^2(x_{k-\frac{1}{2}},y)\geq C\|\Phi_{xy}\|^2,
\end{aligned}\tag{6.8}$$

and for $l=1,\cdots,N^y$,

$$\begin{aligned}
&-\sum_{k=1}^{N^x}h^{x,k}\sum_{m=1}^{r-1}w_m\frac{\partial^{2r}}{\partial y^{2r}}(\Phi_{xx}\Phi)(x_{k,m},y_{l-\frac{1}{2}})\\
&=\int_0^1\left(\frac{\partial^r}{\partial y^r}\Phi_x\right)^2(x,y_{l-\frac{1}{2}})\,dx\\
&\quad+C\sum_{k=1}^{N^x}(h^{x,k})^{2r-1}\left(\frac{\partial^r}{\partial x^r}\Phi^k\right)^2(x,y_{l-\frac{1}{2}})\geq 0.
\end{aligned}\tag{6.9}$$

Therefore, using (6.7), (6.8) and (6.9) we obtain (6.5).

Using (1.6) in [20] and (17) in [13], we obtain

$$\|\Phi_{xx}\|_h \le C\|\Phi_{xx}\|_0 \le C\|\Phi_{xx}\|_h, \quad \|\Phi_{yy}\|_h \le C\|\Phi_{yy}\|_0 \le C\|\Phi_{xx}\|_h. \tag{6.10}$$

Using (2.3), the Cauchy's inequality, (6.10), and [18, p.180, (8.24)], we obtain

$$\begin{aligned}\|\Delta\Phi\|_h^2 &= (\Delta\Phi, \Delta\Phi)_h = \|\Phi_{xx}\|_h^2 + \|\Phi_{yy}\|_h^2 + 2(\Phi_{xx}, \Phi_{yy})_h \\ &\le C\left[\|\Phi_{xx}\|_h^2 + \|\Phi_{xx}\|_h^2\right] \le C\|\Delta\Phi\|_0^2.\end{aligned} \tag{6.11}$$

The triangle inequality, (6.10), and the inequality $(\Phi_{xx}, \Phi_{yy})_h \ge 0$ yield

$$\|\Delta\Phi\|_0^2 \le C\left[\|\Phi_{xx}\|_0^2 + \|\Phi_{xx}\|_0^2\right] \le C\left[\|\Phi_{xx}\|_h^2 + \|\Phi_{xx}\|_h^2\right] \le C\|\Delta\Phi\|_h^2, \tag{6.12}$$

and hence, the first result in Lemma 2.3 follows from (6.11) and (6.12).

Using (2.6) in [20] and Lemma 3.3 in [10], we get

$$\begin{aligned}\|\Phi_x\|_0^2 &= \int_0^1 \left(\int_0^1 \Phi_x^2(x, y)dy\right) dx \le C\sum_{l=1}^{N^y} h^{y,l} \sum_{n=1}^{J} w_n \int_0^1 \Phi_x^2(x, y_{l,n})dx \\ &\le -C\sum_{l=1}^{N^y} h^{y,l} \sum_{n=1}^{J} w_n \left(\sum_{k=1}^{N^x} h^{x,k} \sum_{m=1}^{J} w_m(\Phi\Phi_{xx})(x_{k,m}, y_{l,n})\right) \\ &= -C(\Phi, \Phi_{xx})_h.\end{aligned} \tag{6.13}$$

Similarly, we obtain

$$\|\Phi_y\|_0^2 \le -C(\Phi, \Phi_{yy})_h \tag{6.14}$$

Since $\Phi = 0$ on $\partial\Omega$, the second desired result follows from the Poincaré inequality, (6.13) and (6.14). The last inequality in Lemma 2.3, for $r = 3$, is obtained by using the fact that the degree of precision is $2r - 2$ and using integration by parts . For $r \ge 4$, we use (1.13) and Lemma 3.1 in [10], to obtain

$$\begin{aligned}(\Phi, \Delta\Psi)_h &= \sum_{k=1}^{N^x} h^{x,k} \sum_{m=1}^{r-1} w_m \sum_{l=1}^{N^y} h^{y,l} \sum_{n=1}^{r-1} w_n(\Phi[\Psi_{xx} + \Psi_{yy}])(x_{k,m}, y_{l,n}) \\ &= \sum_{k=1}^{N^x} h^{x,k} \sum_{m=1}^{r-1} w_m \sum_{l=1}^{N^y} h^{y,l} \sum_{n=1}^{r-1} w_n(\Phi_{xx}\Psi + \Phi_{yy}\Psi)(x_{k,m}, y_{l,n}) = (\Delta\Phi, \Psi)_h.\end{aligned}$$

6.2 Proofs of Theorem 4.1 and Lemma 4.2

The first part of a proof of Theorem 4.1 is based on the idea of expanding and splitting first two terms of the method in (1.15) into eight terms, by choosing appropriate test functions and U in (1.15) replaced by the error ξ. The major part of the proof is then to bound each of the eight terms.

Proof of Theorem 4.1 We first consider the first two terms of the method in (1.15) for $n = 1, \ldots, N^t - 1$, with $\Psi \in S_h$ and U replaced by $\xi = U - W$. Using $\eta = u - W$, and (3.1),

$$
\begin{aligned}
&- (\partial_t \xi^n, \Delta\Psi)_h + (A\widehat{U}^{n+\frac{1}{2}} \Delta\xi^{n+\frac{1}{2}}, \Delta\Psi)_h \\
&= -(\partial_t U^n, \Delta\Psi)_h + (\mathcal{A}\widehat{U}^{n+\frac{1}{2}} \Delta U^{n+\frac{1}{2}}, \Delta\Psi)_h \\
&\quad + (\partial_t W^n, \Delta\Psi)_h - (\mathcal{A}\widehat{U}^{n+\frac{1}{2}} \Delta W^{n+\frac{1}{2}}, \Delta\Psi)_h \\
&= -(\mathcal{F}(t_{n+\frac{1}{2}})\widehat{U}^{n+\frac{1}{2}}, \Delta\Psi)_h \\
&\quad + (\partial_t (u^n - \eta^n), \Delta\Psi)_h \\
&\quad + [(\mathcal{A}\widehat{u}^{n+\frac{1}{2}} \Delta W^{n+\frac{1}{2}}, \Delta\Psi)_h - (\mathcal{A}\widehat{U}^{n+\frac{1}{2}} \Delta W^{n+\frac{1}{2}}, \Delta\Psi)_h] \\
&\quad - [(\mathcal{A}\widehat{u}^{n+\frac{1}{2}} \Delta W^{n+\frac{1}{2}}, \Delta\Psi)_h - (\mathcal{A}u^{n+\frac{1}{2}} \Delta W^{n+\frac{1}{2}}, \Delta\Psi)_h] - (\mathcal{A}u^{n+\frac{1}{2}} \Delta W^{n+\frac{1}{2}}, \Delta\Psi)_h.
\end{aligned}
\tag{6.15}
$$

Further using (1.14) we get

$$
\begin{aligned}
&- (\mathcal{A}u^{n+\frac{1}{2}} \Delta W^{n+\frac{1}{2}}, \Delta\Psi)_h \\
&= (1/2)[(\mathcal{A}u^{n+1} \Delta(W^{n+1} - W^n), \Delta\Psi)_h - (\mathcal{A}u^{n+\frac{1}{2}} \Delta(W^{n+1} - W^n), \Delta\Psi)_h] \\
&\quad - ([\mathcal{A}u^{n+\frac{1}{2}} - (\mathcal{A}u)^{n+\frac{1}{2}}] \Delta W^n, \Delta\Psi)_h - ((\mathcal{A}u \Delta W)^{n+\frac{1}{2}}, \Delta\Psi)_h.
\end{aligned}
\tag{6.16}
$$

Adding and subtracting $\mathcal{F}(t_{n+\frac{1}{2}})u(t_{n+\frac{1}{2}})$ in (6.15) and using (6.16), (3.1), and (1.1), we get the following identity for $n = 1, \ldots, N^t - 1$

$$
\begin{aligned}
&- (\partial_t \xi^n, \Delta\Psi)_h + (A\widehat{U}^{n+\frac{1}{2}} \Delta\xi^{n+\frac{1}{2}}, \Delta\Psi)_h = (\partial_t u^n - u_t(t_{n+\frac{1}{2}}), \Delta\Psi)_h - (\partial_t \eta^n, \Delta\Psi)_h \\
&\quad + [(\mathcal{A}\widehat{u}^{n+\frac{1}{2}} \Delta W^{n+\frac{1}{2}}, \Delta\Psi)_h - (\mathcal{A}\widehat{U}^{n+\frac{1}{2}} \Delta W^{n+\frac{1}{2}}, \Delta\Psi)_h] \\
&\quad - [(\mathcal{A}\widehat{u}^{n+\frac{1}{2}} \Delta W^{n+\frac{1}{2}}, \Delta\Psi)_h - (\mathcal{A}u^{n+\frac{1}{2}} \Delta W^{n+\frac{1}{2}}, \Delta\Psi)_h] \\
&\quad + (1/2)[(\mathcal{A}u^{n+1} \Delta(W^{n+1} - W^n), \Delta\Psi)_h - (\mathcal{A}u^{n+\frac{1}{2}} \Delta(W^{n+1} - W^n), \Delta\Psi)_h] \\
&\quad - ([\mathcal{A}u^{n+\frac{1}{2}} - (\mathcal{A}u)^{n+\frac{1}{2}}] \Delta W^n, \Delta\Psi)_h + ((\mathcal{A}u \Delta u)(t_{n+\frac{1}{2}}) - (\mathcal{A}u \Delta u)^{n+\frac{1}{2}}, \Delta\Psi)_h \\
&\quad + (\mathcal{F}(t_{n+\frac{1}{2}})u(t_{n+\frac{1}{2}}) - \mathcal{F}(t_{n+\frac{1}{2}})\widehat{U}^{n+\frac{1}{2}}, \Delta\Psi)_h.
\end{aligned}
\tag{6.17}
$$

With $\Psi = \xi^{n+\frac{1}{2}}$ in (6.17), we obtain

$$
-(\partial_t \xi^n, \Delta\xi^{n+\frac{1}{2}})_h + (\mathcal{A}\widehat{U}^{n+\frac{1}{2}} \Delta\xi^{n+\frac{1}{2}}, \Delta\xi^{n+\frac{1}{2}})_h = \sum_{i=1}^{8} I_i^n, \quad n = 1, \ldots, N^t - 1, \tag{6.18}
$$

where

$$I_1^n = (\partial_t u^n - u_t(t_{n+\frac{1}{2}}), \Delta\xi^{n+\frac{1}{2}})_h, \quad I_2^n = -(\partial_t \eta^n, \Delta\xi^{n+\frac{1}{2}})_h, \tag{6.19}$$

$$I_3^n = (\mathcal{A}\widehat{u}^{n+\frac{1}{2}}\Delta W^{n+\frac{1}{2}}, \Delta\xi^{n+\frac{1}{2}})_h - (\mathcal{A}\widehat{U}^{n+\frac{1}{2}}\Delta W^{n+\frac{1}{2}}, \Delta\xi^{n+\frac{1}{2}})_h, \tag{6.20}$$

$$I_4^n = (\mathcal{A}u^{n+\frac{1}{2}}\Delta W^{n+\frac{1}{2}}, \Delta\xi^{n+\frac{1}{2}})_h - (\mathcal{A}\widehat{u}^{n+\frac{1}{2}}\Delta W^{n+\frac{1}{2}}, \Delta\xi^{n+\frac{1}{2}})_h, \tag{6.21}$$

$$\begin{aligned} I_5^n = (1/2)[&(\mathcal{A}u^{n+1}\Delta(W^{n+1} - W^n), \Delta\xi^{n+\frac{1}{2}})_h \\ &- (\mathcal{A}u^{n+\frac{1}{2}}\Delta(W^{n+1} - W^n), \Delta\xi^{n+\frac{1}{2}})_h], \end{aligned} \tag{6.22}$$

$$I_6^n = -([\mathcal{A}u^{n+\frac{1}{2}} - (\mathcal{A}u)^{n+\frac{1}{2}}]\Delta W^n, \Delta\xi^{n+\frac{1}{2}})_h, \tag{6.23}$$

$$I_7^n = ((\mathcal{A}u\Delta u)(t_{n+\frac{1}{2}}) - (\mathcal{A}u\Delta u)^{n+\frac{1}{2}}, \Delta\xi^{n+\frac{1}{2}})_h, \tag{6.24}$$

$$I_8^n = (\mathcal{F}(t_{n+\frac{1}{2}})u(t_{n+\frac{1}{2}}) - \mathcal{F}(t_{n+\frac{1}{2}})\widehat{U}^{n+\frac{1}{2}}, \Delta\xi^{n+\frac{1}{2}})_h. \tag{6.25}$$

For the first term on the left hand side of (6.18), we use Lemma 2.3 to obtain

$$\begin{aligned} -(\partial_t\xi^n, \Delta\xi^{n+\frac{1}{2}})_h &= -\frac{1}{2\tau}(\xi^{n+1} - \xi^n, \Delta\xi^{n+1} + \Delta\xi^n)_h \\ &= -\frac{1}{2\tau}\{(\xi^{n+1}, \Delta\xi^{n+1})_h - (\xi^n, \Delta\xi^n)_h\}. \end{aligned} \tag{6.26}$$

For the second term on the left hand side of (6.18), we use (1.16), $a \geq a_{\min} > 0$ and Lemma 2.4 to obtain

$$(\mathcal{A}\widehat{U}^{n+\frac{1}{2}}\Delta\xi^{n+\frac{1}{2}}, \Delta\xi^{n+\frac{1}{2}})_h \geq C\|\Delta\xi^{n+\frac{1}{2}}\|_h^2 \geq C\|\Delta\xi^{n+\frac{1}{2}}\|_0^2. \tag{6.27}$$

We bound terms on the right hand side of (6.18) starting with I_1^n. Since

$$\partial_t u^n - u_t(t_{n+\frac{1}{2}}) = \frac{1}{2\tau}\left[\int_{t_n}^{t_{n+\frac{1}{2}}}(s - t_n)^2 u_{ttt}\,ds + \int_{t_{n+\frac{1}{2}}}^{t_{n+1}}(s - t_{n+1})^2 u_{ttt}\,ds\right],$$

using (2.3) and Lemma 2.3, the triangle and ϵ inequalities, we obtain

$$\begin{aligned} |I_1^n| &\leq C\|u_t(t_{n+\frac{1}{2}}) - \partial_t u^n\|_h\|\Delta\xi^{n+\frac{1}{2}}\|_h \leq C\|u_t(t_{n+\frac{1}{2}}) - \partial_t u^n\|_\infty\|\Delta\xi^{n+\frac{1}{2}}\|_0 \\ &\leq C\tau\int_{t_n}^{t_{n+1}}\|u_{ttt}\|_\infty\,ds\|\Delta\xi^{n+\frac{1}{2}}\|_0 \leq \epsilon_1\|\Delta\xi^{n+\frac{1}{2}}\|_0^2 + C(u)\epsilon_1^{-1}\tau^4. \end{aligned} \tag{6.28}$$

Next we bound I_2^n. Since for any $\phi \in \mathcal{C}^1[0, T]$ and any t', $t'' \in [0, T]$,

$$\phi(t'') - \phi(t') = \int_{t'}^{t''} \phi_t \, ds, \tag{6.29}$$

using (3.27), we have

$$\left\| \partial_t \eta^n \right\|_h \leq C\tau^{-1} \int_{t_n}^{t_{n+1}} \|\eta_t\|_h \, ds \leq C(u)h^r. \tag{6.30}$$

Hence, (2.3), (6.30), Lemma 2.3, and the ϵ inequality yield

$$|I_2^n| \leq C\|\partial_t \eta^n\|_h \|\Delta\xi^{n+\frac{1}{2}}\|_h \leq C(u)h^r \|\Delta\xi^{n+\frac{1}{2}}\|_0 \leq \epsilon_2 \|\Delta\xi^{n+\frac{1}{2}}\|_0^2 + C(u)\epsilon_2^{-1}h^{2r}. \tag{6.31}$$

Using (6.28) and (6.31), we get

$$|I_1^n| + |I_2^n| \leq (\epsilon_1 + \epsilon_2)\|\Delta\xi^{n+\frac{1}{2}}\|_0^2 + C(u)\left(\epsilon_1^{-1} + \epsilon_2^{-1}\right)\left(h^{2r} + \tau^4\right). \tag{6.32}$$

Using (6.20), (2.3), Lemma 3.4, (1.5), Lemma 2.3, the triangle inequality, $U^n - u^n = \xi^n - \eta^n$, and Theorem 3.2, we obtain

$$\begin{aligned}
|I_3^n| &\leq C\|\Delta W^{n+\frac{1}{2}}\|_{\infty,\mathcal{T}_h} \|\mathcal{A}\widehat{u}^{n+\frac{1}{2}} - \mathcal{A}\widehat{U}^{n+\frac{1}{2}}\|_h \|\Delta\xi^{n+\frac{1}{2}}\|_h \\
&\leq C(u)d(n, u, U)\|\widehat{u}^{n+\frac{1}{2}} - \widehat{U}^{n+\frac{1}{2}}\|_h \|\Delta\xi^{n+\frac{1}{2}}\|_0 \\
&\leq C(u)d(n, u, U) \sum_{i=n-1}^{n} (\|\xi^i\|_h + \|\eta^i\|_h)\|\Delta\xi^{n+\frac{1}{2}}\|_0 \\
&\leq C(u)d(n, u, U) \left(\sum_{i=n-1}^{n} \|\xi^i\|_0 + h^{r+1} \right) \|\Delta\xi^{n+\frac{1}{2}}\|_0,
\end{aligned}$$

and thus, the ϵ inequality, and Lemma 2.3 yield

$$|I_3^n| \leq \epsilon_3 \|\Delta\xi^{n+\frac{1}{2}}\|_0^2 + C(u)\epsilon_3^{-1}[d(n, u, U)]^2 \left(h^{2r+2} + \sum_{i=n-1}^{n} (\xi^i, -\Delta\xi^i)_h \right). \tag{6.33}$$

Using (6.21), (2.3), Lemma 3.4, (1.5), and Lemma 2.3, we get

$$\begin{aligned}
|I_4^n| &\leq C\|\Delta W^{n+\frac{1}{2}}\|_{\infty,\mathcal{T}_h} \|\mathcal{A}\widehat{u}^{n+\frac{1}{2}} - \mathcal{A}u^{n+\frac{1}{2}}\|_h \|\Delta\xi^{n+\frac{1}{2}}\|_h \\
&\leq C(u)\|\widehat{u}^{n+\frac{1}{2}} - u^{n+\frac{1}{2}}\|_h \|\Delta\xi^{n+\frac{1}{2}}\|_0.
\end{aligned} \tag{6.34}$$

Since, for any $\phi \in \mathcal{C}^2[0, T]$, and t', t'', $t^* \in [0, T]$, where $t^* = \frac{t'+t''}{2}$,

$$\frac{\phi(t') + \phi(t'')}{2} - \phi(t^*) = \frac{1}{2}\left[\int_{t'}^{t^*} (s - t')\phi_{tt} \, ds - \int_{t^*}^{t''} (s - t'')\phi_{tt} \, ds \right], \tag{6.35}$$

using (1.14),

$$\begin{aligned}
\|\widehat{u}^{n+\frac{1}{2}} - u^{n+\frac{1}{2}}\|_h &= \left\| \frac{u^{n+1} + u^{n-1}}{2} - u^n \right\|_h \\
&= \frac{1}{2} \left\| \int_{t_{n-1}}^{t_n} (s - t_{n-1})u_{tt} \, ds - \int_{t_n}^{t_{n+1}} (s - t_{n+1})u_{tt} \, ds \right\|_h \leq C(u)\tau^2.
\end{aligned} \tag{6.36}$$

Similarly using (6.35), we obtain

$$\|u^{n+\frac{1}{2}} - u(t_{n+\frac{1}{2}})\|_h \le C(u)\tau^2. \tag{6.37}$$

Using (6.34), (6.36), and the ϵ inequality, we get

$$|I_4^n| \le \epsilon_4 \|\Delta\xi^{n+\frac{1}{2}}\|_0^2 + C(u)\epsilon_4^{-1}\tau^4. \tag{6.38}$$

Using (6.22), (2.3), (1.5), (6.29) (with ϕ replaced by u and W respectively), Lemmas 2.3 and 3.4, and the ϵ inequality, we have

$$\begin{aligned}
|I_5^n| &\le |([\mathcal{A}u^{n+1} - \mathcal{A}u^{n+\frac{1}{2}}]\Delta(W^{n+1} - W^n), \Delta\xi^{n+\frac{1}{2}})_h| \\
&\le \|\mathcal{A}u^{n+1} - \mathcal{A}u^{n+\frac{1}{2}}\|_\infty \|\Delta W^{n+1} - \Delta W^n\|_h \|\Delta\xi^{n+\frac{1}{2}}\|_h \\
&\le C(u)\|u^{n+1} - u^{n+\frac{1}{2}}\|_\infty \|\Delta W^{n+1} - \Delta W^n\|_h \|\Delta\xi^{n+\frac{1}{2}}\|_0 \\
&\le C(u)\int_{t_n}^{t_{n+1}} \|u_t\|_\infty \, ds \int_{t_n}^{t_{n+1}} \|\Delta W_t\|_h \, ds \|\Delta\xi^{n+\frac{1}{2}}\|_0 \le C(u)\tau^2\|\Delta\xi^{n+\frac{1}{2}}\|_0 \\
&\le \epsilon_5\|\Delta\xi^{n+\frac{1}{2}}\|_0^2 + C(u)\epsilon_5^{-1}\tau^4.
\end{aligned} \tag{6.39}$$

Next (6.23), (2.3), Lemmas 2.3 and 3.4, and the triangle inequality, yield

$$\begin{aligned}
|I_6^n| &\le \|\mathcal{A}u^{n+\frac{1}{2}} - (\mathcal{A}u)^{n+\frac{1}{2}}\|_\infty \|\Delta W^n\|_h \|\Delta\xi^{n+\frac{1}{2}}\|_h \\
&\le C(u)\|\mathcal{A}u^{n+\frac{1}{2}} - (\mathcal{A}u)^{n+\frac{1}{2}}\|_\infty \|\Delta\xi^{n+\frac{1}{2}}\|_0 \\
&\le C(u)[\|\mathcal{A}u^{n+\frac{1}{2}} - \mathcal{A}u(t_{n+\frac{1}{2}})\|_\infty + \|\mathcal{A}u(t_{n+\frac{1}{2}}) - (\mathcal{A}u)^{n+\frac{1}{2}}\|_\infty]\|\Delta\xi^{n+\frac{1}{2}}\|_0.
\end{aligned} \tag{6.40}$$

Using (1.5) and (6.37), we obtain

$$\|\mathcal{A}u^{n+\frac{1}{2}} - \mathcal{A}u(t_{n+\frac{1}{2}})\|_\infty \le C(u)\|u^{n+\frac{1}{2}} - u(t_{n+\frac{1}{2}})\|_\infty \le C(u)\tau^2. \tag{6.41}$$

Using (6.35) with $\mathcal{A}u$ in place of ϕ,

$$\mathcal{A}u(t_{n+\frac{1}{2}}) - (\mathcal{A}u)^{n+\frac{1}{2}} = -\frac{1}{2}\left[\int_{t_n}^{t_{n+\frac{1}{2}}} (s - t_n)\mathcal{A}_{tt}u \, ds - \int_{t_{n+\frac{1}{2}}}^{t_{n+1}} (s - t_{n+1})\mathcal{A}_{tt}u \, ds\right],$$

where $[\mathcal{A}_{tt}u(t)] = \frac{\partial^2}{\partial t^2}[\mathcal{A}u(t)]$, and hence the triangle inequality and (6.29) give

$$\|\mathcal{A}u(t_{n+\frac{1}{2}}) - (\mathcal{A}u)^{n+\frac{1}{2}}\|_\infty \le C\tau \int_{t_n}^{t_{n+1}} \|\mathcal{A}_{tt}u\|_\infty \, ds \le C(u)\tau^2. \tag{6.42}$$

Using (6.40), (6.41) and (6.42), and the ϵ inequality, we obtain

$$|I_6^n| \le CC_4(u)\tau^2\|\Delta\xi^{n+\frac{1}{2}}\|_0 \le \epsilon_6\|\Delta\xi^{n+\frac{1}{2}}\|_0 + C(u)\epsilon_6^{-1}\tau^4. \tag{6.43}$$

Using (2.3), Lemma 2.3, (6.35) with $\mathcal{A}u\Delta u$ in place of ϕ, and the ϵ inequality we get

$$\begin{aligned}|I_7^n| &\le C\|(\mathcal{A}u\Delta u)(t_{n+\frac{1}{2}}) - (\mathcal{A}u\Delta u)^{n+\frac{1}{2}}\|_\infty \|\Delta\xi^{n+\frac{1}{2}}\|_0 \\ &\le C\tau \int_{t_n}^{t_{n+1}} \|(\mathcal{A}u\Delta u)_{tt}\|_\infty \, ds \|\Delta\xi^{n+\frac{1}{2}}\|_0 \le C(u)\tau^2\|\Delta\xi^{n+\frac{1}{2}}\|_0. \\ &\le \epsilon_7\|\Delta\xi^{n+\frac{1}{2}}\|_0^2 + C(u)\epsilon_7^{-1}\tau^4. \end{aligned} \tag{6.44}$$

Using (6.39), (6.43), and (6.44), we obtain

$$\sum_{i=5}^{7} |I_i^n| \le \left(\sum_{i=5}^{7} \epsilon_i\right) \|\Delta\xi^{n+\frac{1}{2}}\|_0^2 + C(u)\left(\sum_{i=5}^{7} \epsilon_i^{-1}\right)\tau^4. \tag{6.45}$$

To bound I_8^n in (6.25). we use (1.16), (2.3), the triangle inequality, and (1.6), (1.7) and (1.8),

$$\begin{aligned}|I_8^n| &\le C\|\mathcal{F}(t_{n+\frac{1}{2}})u(t_{n+\frac{1}{2}}) - \mathcal{F}(t_{n+\frac{1}{2}})\widehat{U}^{n+\frac{1}{2}}\|_h \|\Delta\xi^{n+\frac{1}{2}}\|_h \\ &\le C\Big[\big(\|\mathcal{F}(t_{n+\frac{1}{2}})u(t_{n+\frac{1}{2}}) - f\big(., t_{n+\frac{1}{2}}, u(t_{n+\frac{1}{2}}), u_x(t_{n+\frac{1}{2}}), \widehat{U}_y^{n+\frac{1}{2}}\big)\|_h \\ &\qquad + \|f\big(., t_{n+\frac{1}{2}}, u(t_{n+\frac{1}{2}}), u_x(t_{n+\frac{1}{2}}), \widehat{U}_y^{n+\frac{1}{2}}\big) - f\big(., t_{n+\frac{1}{2}}, u(t_{n+\frac{1}{2}}), \widehat{U}_x^{n+\frac{1}{2}}, \widehat{U}_y^{n+\frac{1}{2}}\big)\|_h \\ &\qquad + \|f\big(., t_{n+\frac{1}{2}}, u(t_{n+\frac{1}{2}}), \widehat{U}_x^{n+\frac{1}{2}}, \widehat{U}_y^{n+\frac{1}{2}}\big) - \mathcal{F}(t_{n+\frac{1}{2}})\widehat{U}^{n+\frac{1}{2}}\|_h\Big] \|\Delta\xi^{n+\frac{1}{2}}\|_h \\ &\le Cd(n,u,U)\Big[\|u(t_{n+\frac{1}{2}}) - \widehat{U}^{n+\frac{1}{2}}\|_h + \|u_x(t_{n+\frac{1}{2}}) \\ &\qquad - \widehat{U}_x^{n+\frac{1}{2}}\|_h + \|u_y(t_{n+\frac{1}{2}}) - \widehat{U}_y^{n+\frac{1}{2}}\|_h\Big] \|\Delta\xi^{n+\frac{1}{2}}\|_h. \end{aligned} \tag{6.46}$$

The relation $u(t_{n+\frac{1}{2}}) - \widehat{U}^{n+\frac{1}{2}} = [u(t_{n+\frac{1}{2}}) - u^{n+\frac{1}{2}}] + [u^{n+\frac{1}{2}} - \widehat{u}^{n+\frac{1}{2}}] + \widehat{\eta}^{n+\frac{1}{2}} - \widehat{\xi}^{n+\frac{1}{2}}$, the triangle inequality, Lemma 2.3, (6.36), (6.37), and Theorem 3.2, give

$$\begin{aligned}\|u(t_{n+\frac{1}{2}}) - \widehat{U}^{n+\frac{1}{2}}\|_h &\le C\Big[\|u(t_{n+\frac{1}{2}}) - u^{n+\frac{1}{2}}\|_h + \|u^{n+\frac{1}{2}} - \widehat{u}^{n+\frac{1}{2}}\|_h + \|\widehat{\eta}^{n+\frac{1}{2}}\|_h + \|\widehat{\xi}^{n+\frac{1}{2}}\|_h\Big] \\ &\le C\left(C(u)[\tau^2 + h^{r+1}] + \|\widehat{\xi}^{n+\frac{1}{2}}\|_0\right). \end{aligned} \tag{6.47}$$

Similarly with u_x (or u_y) and U_x (or U_y) in place of u and U respectively, we obtain

$$\max\left\{\|u_x(t_{n+\frac{1}{2}}) - \widehat{U}_x^{n+\frac{1}{2}}\|_h, \|u_y(t_{n+\frac{1}{2}}) - \widehat{U}_y^{n+\frac{1}{2}}\|_h\right\} \le C\left(C(u)[\tau^2 + h^r] + \|\widehat{\xi}^{n+\frac{1}{2}}\|_1\right). \tag{6.48}$$

Using (6.46), (6.47) and (6.48), the ϵ inequality, and Lemma 2.3, we get

$$|I_8^n| \le \epsilon_8\|\Delta\xi^{n+\frac{1}{2}}\|_0^2 + C(u)\epsilon_8^{-1}[d(n,u,U)]^2\left(\tau^4 + h^{2r} + \sum_{i=n-1}^{n} (\xi^i, -\Delta\xi^i)_h\right). \tag{6.49}$$

Using (6.18), (6.26), (6.27), (6.32), (6.33), (6.38), (6.45), (6.49), and the triangle inequality, we obtain for $n = 1, \ldots, N^t - 1$,

$$-\frac{1}{2\tau}[(\xi^{n+1}, \Delta\xi^{n+1})_h + (\xi^n, \Delta\xi^n)_h] + C\|\Delta\xi^{n+\frac{1}{2}}\|_0^2 \le \left(\sum_{i=1}^{8}\epsilon_i\right)\|\Delta\xi^{n+\frac{1}{2}}\|_0^2$$
$$+ C(u)\left([d(n,u,U)]^2 + \sum_{i=1}^{8}\epsilon_i^{-1}\right)\left[\tau^4 + h^{2r} + \sum_{i=n-1}^{n+1}(\xi^i, -\Delta\xi^i)_h\right],$$

and hence, taking ϵ_i, $i = 1, \cdots, 8$, sufficiently small, and multiplying through by 2τ, we obtain for $n = 1, \ldots, N^t - 1$,

$$-(\xi^{n+1}, \Delta\xi^{n+1})_h + (\xi^n, \Delta\xi^n)_h \le C(u)[d(n,u,U)]^2\tau\left[\tau^4 + h^{2r} + \sum_{i=n-1}^{n+1}(\xi^i, -\Delta\xi^i)_h\right]. \quad \square$$

We follow arguments used in proving Theorem 4.1, to prove Lemma 4.2 for the predictor-corrector scheme (1.17) and (1.18). Main changes are in obtaining concrete error bounds for all terms involved in the splitting first two terms in (1.17) and (1.18) with appropriate test functions and error terms. Below we show that the predictor scheme (1.17) yields only the sub-optimal $\tau^{3/2}$ error in time; this is then corrected by using (1.18) to obtain optimal second-order in time error bound.

Proof of Lemma 4.2 Let $\widetilde{\xi}^n = V^n - W^n$, $n = 0, 1$. Then, as in the first part of proof of Theorem 4.1, using (1.17) and (1.18), we have

$$-\left(\partial_t\widetilde{\xi}^0, \Delta\Psi\right)_h + \left(\mathcal{A}V^0\Delta\widetilde{\xi}^{\frac{1}{2}}, \Delta\Psi\right)_h = \left(\partial_t u^0 - u_t(t_{\frac{1}{2}}), \Delta\Psi\right)_h$$
$$- (\partial_t\eta^0, \Delta\Psi)_h + [(\mathcal{A}u^0\Delta W^{\frac{1}{2}}, \Delta\Psi)_h - (\mathcal{A}V^0\Delta W^{\frac{1}{2}}, \Delta\Psi)_h]$$
$$- \left[\left(\mathcal{A}u^0\Delta W^{\frac{1}{2}}, \Delta\Psi\right)_h - \left(\mathcal{A}u^{\frac{1}{2}}\Delta W^{\frac{1}{2}}, \Delta\Psi\right)_h\right] + (1/2)\left[\left(\mathcal{A}u^1\Delta\left(W^1 - W^0\right), \Delta\Psi\right)_h\right.$$
$$\left. - \left(\mathcal{A}u^{\frac{1}{2}}\Delta\left(W^1 - W^0\right), \Delta\Psi\right)_h\right] - \left(\left[\mathcal{A}u^{\frac{1}{2}} - (\mathcal{A}u)^{\frac{1}{2}}\right]\Delta W^0, \Delta\Psi\right)_h + \left((\mathcal{A}u\Delta u)\left(t_{\frac{1}{2}}\right)\right.$$
$$\left. - \left(\mathcal{A}u\Delta u\right)^{\frac{1}{2}}, \Delta\Psi\right)_h + \left(\mathcal{F}\left(t_{\frac{1}{2}}\right)u\left(t_{\frac{1}{2}}\right) - \mathcal{F}\left(t_{\frac{1}{2}}\right)V^0, \Delta\Psi\right)_h, \quad \Psi \in S_h, \tag{6.50}$$

and

$$-\left(\partial_t\xi^0, \Delta\Psi\right)_h + \left(\mathcal{A}V^{\frac{1}{2}}\Delta\xi^{\frac{1}{2}}, \Delta\Psi\right)_h = \left(\partial_t u^0 - u_t\left(t_{\frac{1}{2}}\right), \Delta\Psi\right)_h - \left(\partial_t\eta^0, \Delta\Psi\right)_h$$
$$+ \left[\left(\mathcal{A}u^{\frac{1}{2}}\Delta W^{\frac{1}{2}}, \Delta\Psi\right)_h - \left(\mathcal{A}V^{\frac{1}{2}}\Delta W^{\frac{1}{2}}, \Delta\Psi\right)_h\right]$$
$$+ (1/2)\left[\left(\mathcal{A}u^1\Delta(W^1 - W^0), \Delta\Psi\right)_h - \left(\mathcal{A}u^{\frac{1}{2}}\Delta(W^1 - W^0), \Delta\Psi\right)_h\right]$$
$$- \left(\left[\mathcal{A}u^{\frac{1}{2}} - (\mathcal{A}u)^{\frac{1}{2}}\right]\Delta W^0, \Delta\Psi\right)_h + \left((\mathcal{A}u\Delta u)\left(t_{\frac{1}{2}}\right) - (\mathcal{A}u\Delta u)^{\frac{1}{2}}, \Delta\Psi\right)_h$$
$$+ \left(\mathcal{F}\left(t_{\frac{1}{2}}\right)u\left(t_{\frac{1}{2}}\right) - \mathcal{F}\left(t_{\frac{1}{2}}\right)V^{\frac{1}{2}}, \Delta\Psi\right)_h, \quad \Psi \in S_h. \tag{6.51}$$

Taking $\Psi = \widetilde{\xi}^{\frac{1}{2}}$ in (6.50), we obtain

$$-(\partial_t \widetilde{\xi}^0, \Delta \widetilde{\xi}^{\frac{1}{2}})_h + (\mathcal{A}V^0 \Delta \widetilde{\xi}^{\frac{1}{2}}, \Delta \widetilde{\xi}^{\frac{1}{2}})_h = \sum_{i=1}^{8} I_i^0, \tag{6.52}$$

where I_i^0, $i = 1, \cdots, 8$ are defined as in (6.19), (6.20), (6.21), (6.22), (6.23), (6.24) and (6.25), with $n = 0$, ξ, $\widehat{u}^{\frac{1}{2}}$ and $\widehat{U}^{\frac{1}{2}}$ replaced by $\widetilde{\xi}$, u^0 and V^0 respectively. Using $V^0 = W^0$ and Lemma 3.4,

$$\|V^0\|_{1,\infty} = \|W^0\|_{1,\infty} \leq C(u). \tag{6.53}$$

Following the derivations of (6.34) and (6.49) with $n = 0$, and u^0, V^0, and $\Delta\widetilde{\xi}^{\frac{1}{2}}$, in place of $\widehat{u}^{\frac{1}{2}}$, $\widehat{U}^{\frac{1}{2}}$, and $\Delta\xi^{\frac{1}{2}}$, respectively, and using the relation $V^0 - u^0 = -\eta^0$, and (6.53), we obtain

$$|I_4^0| + |I_8^0| \leq C(u)\left[\|u^1 - u^0\|_{1,\infty} + \|u^0 - u(t_{\frac{1}{2}})\|_{1,\infty} + h^r\right]\|\Delta\widetilde{\xi}^{\frac{1}{2}}\|_0. \tag{6.54}$$

Using (6.29),

$$\|u^1 - u^0\|_{1,\infty} + \|u^0 - u(t_{\frac{1}{2}})\|_{1,\infty} \leq C(u)\int_0^{t_1} \|u_t\|_{1,\infty}\, ds \leq C(u)\tau. \tag{6.55}$$

Hence, using (6.54), (6.55), and the ϵ inequality, we have

$$|I_4^0| + |I_8^0| \leq \epsilon_4 \|\Delta\widetilde{\xi}^{\frac{1}{2}}\|_0^2 + C(u)\epsilon_4^{-1}\left(h^{2r} + \tau^2\right). \tag{6.56}$$

Using (6.52), and (6.26), (6.27), (6.33), (6.32), (6.45), with $n = 0$ and u^0, $\tilde{\xi}^{\frac{1}{2}}$, and V^0 in place of $\widehat{u}^{\frac{1}{2}}$, $\xi^{\frac{1}{2}}$, and $\widehat{U}^{\frac{1}{2}}$, respectively, we get

$$\begin{aligned} -\frac{1}{2\tau}(\widetilde{\xi}^1, \Delta\tilde{\xi}^1)_h + C\|\Delta\widetilde{\xi}^{\frac{1}{2}}\|_0^2 \leq & \left(\sum_{i=1}^{3}\epsilon_i + \sum_{i=5}^{7}\epsilon_i\right)\|\Delta\widetilde{\xi}^{\frac{1}{2}}\|_0^2 + |I_4^0| + |I_8^0| \\ & + C\left[C(u) + \sum_{i=1}^{3}\epsilon_i^{-1} + \sum_{i=5}^{7}\epsilon_i^{-1}\right] \\ & \times \left[\tau^2 + h^{2r} + (\widetilde{\xi}^1, -\Delta\widetilde{\xi}^1)_h\right], \end{aligned}$$

and hence, (6.56), taking ϵ_i, $i = 1, \cdots, 7$, sufficiently small, and multiplying through by 2τ yield

$$-(\tilde{\xi}^1, \Delta\tilde{\xi}^1)_h \leq C(u)\left[\tau^3 + \tau h^{2r} + \tau(\widetilde{\xi}^1, -\Delta\widetilde{\xi}^1)_h\right],$$

and consequently, for τ sufficiently small, we get

$$-(\widetilde{\xi}^1, \Delta\widetilde{\xi}^1)_h \leq C(u)\left(\tau^3 + h^{2r}\right). \tag{6.57}$$

The relation $V^{\frac{1}{2}} = (W^1 + W^0 + \widetilde{\xi}^1)/2$, the triangle and inverse inequalities, Lemmas 2.3 and 3.4, (6.57), and $\tau^3 \leq Ch^2$ we get

$$\|V^{\frac{1}{2}}\|_{1,\infty} \leq C\left(\|W^0\|_{1,\infty} + \|W^1\|_{1,\infty} + h^{-1}\|\tilde{\xi}^1\|_1\right) \leq C(u). \tag{6.58}$$

Next, we use (6.57) and (6.58) to bound $-(\xi^1, \Delta\xi^1)_h$. Taking $v = \xi^{\frac{1}{2}}$ in (6.51), we obtain

$$-(\partial_t \xi^0, \Delta\xi^{\frac{1}{2}})_h + (\mathcal{A}V^{\frac{1}{2}}\Delta\xi^{\frac{1}{2}}, \Delta\xi^{\frac{1}{2}})_h = \sum_{i=1}^{8} I_i^0, \tag{6.59}$$

where I_i^0, $i = 1, \cdots, 8$ are defined in (6.19), (6.20), (6.21), (6.22), (6.23), (6.24) and (6.25), with $\widehat{u}^{\frac{1}{2}}$ and $\widehat{U}^{\frac{1}{2}}$ replaced by $u^{\frac{1}{2}}$ and $V^{\frac{1}{2}}$ respectively. Following (6.31) (6.32) and (6.33), (6.49) with $n = 0$, $V^{\frac{1}{2}}$ in place of $\widehat{U}^{\frac{1}{2}}$, $u^{\frac{1}{2}}$ in place of $\widehat{u}^{\frac{1}{2}}$, and using the relation $V^{\frac{1}{2}} - u^{\frac{1}{2}} = \tilde{\xi}^{\frac{1}{2}} - \eta^{\frac{1}{2}}$, $\tilde{\xi}^0 = 0$, and (6.58), we obtain

$$|I_3^0| + |I_8^0| \le \epsilon_3 \|\Delta\xi^{\frac{1}{2}}\|_0^2 + C(u)\epsilon_3^{-1}\left[h^{2r} + \tau^4 + (\widetilde{\xi}^1, -\Delta\widetilde{\xi}^1)_h\right]. \tag{6.60}$$

Using (6.59), $\xi^0 = 0$, (6.26), (6.27), (6.32), (6.38), (6.45), with $n = 0$, (6.58), (6.60), taking ϵ_i, $i = 1, \cdots, 7$, sufficiently small, and multiplying through by 2τ, we obtain

$$-(\xi^1, \Delta\xi^1)_h \le C(u)\left\{\tau^5 + \tau h^{2r} + \tau(\xi^1, -\Delta\xi^1)_h + \tau(\tilde{\xi}^1, -\Delta\tilde{\xi}^1)_h\right\},$$

and hence, for τ sufficiently small, the desired result follows from (6.57). □

References

1. Aitbayev, R.: Multilevel preconditioners for a quadrature Galerkin solution of a biharmonic problem. Numer. Methods Partial Differ. Equ. **22**, 847–866 (2006)
2. Aitbayev, R.: Convergence analysis of a quadrature finite element Galerkin scheme for a biharmonic problem. IMA J. Numer. Anal. (to appear)
3. Baker, G.: Simplified proofs of error estimates for the least squares method for Dirichlet's problem. Math. Comput. **327**, 229–235 (1973)
4. Bramble, J., Schatz, A.: Rayleigh–Ritz–Galerkin methods for Dirichlet's problem using subspaces without boundary conditions. Commun. Pure Appl. Math. **23**, 653–675 (1970)
5. Bramble, J., Schatz, A.: Least-squares methods for 2*m*th order elliptic boundary value problems. Math. Comput. **25**, 1–32 (1971)
6. Bialecki, B., Ganesh, M., Mustapha, K.: A Petrov-Galerkin method with quadrature for elliptic boundary value problems. IMA J. Numer. Anal. **24**, 157–177 (2004)
7. Bialecki, B., Fairweather, G.: Orthogonal spline collocation methods for partial differential equations. J. Comput. Appl. Math. **128**, 55–85 (2001)
8. Blum, H., Rannacher, R.:On the boundary value problem of the biharmonic operator on domains with angular corners. Math. Methods Appl. Sci. **2**, 556–581 (1980)
9. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. North-Holland, Amsterdam (1978)
10. Douglas, J., Dupont, T.: Collocation method for parabolic equations in a single variable. In: Lecture Notes in Math, vol. 385. Springer, Berlin Heidelberg New York (1974)
11. Douglas, J., Dupont, T., Rachford, H.H., Wheeler, M.F.: Local H^{-1} Galerkin and adjoint local H^{-1} Galerkin procedures for elliptic equations. RAIRO Anal. Numer. **11**, 3–12 (1977)
12. Fairweather, G.: Finite element Galerkin methods for differential equations. In: Lecture Notes in Pure and Applied Mathematics, vol. 34. Marcel Dekker, New York (1978)
13. Grigorieff, R., Sloan, I.H., Brandts, J.H.: Superapproximation and commutator properties of discrete orthogonal projections for continuous splines. J. Approx. Theory **107**, 244–267 (2000)
14. Gilbarg, D., Trudinger, N.S.: Elliptic Partial Differential Equations of Second Order. Springer, Berlin Heidelberg New York (1977)
15. He, Y., Jin, M., Gu, X., Qin, H.: A $\mathcal{C}^1$ globally interpolatory spline of arbitrary topology. Lect. Notes Comput. Sci. **3752**, 295–306 (2005)
16. He, Y., Gu, X., Qin, H.: Automatic shape control of triangular B-splines of arbitrary topology. J. Comput. Sci. Technol. **21**, 232–237 (2006)

17. Hackbusch, W.: Elliptic Differential Equations, Theory and Numerical Treatment. Springer, Berlin Heidelberg New York (1992)
18. Ladyzhenskaya, O., Uraltseva, N.: Linear and Quasilinear Elliptic Equations. Academic, New York (1968)
19. Ma, N., Lu, T., Yang, D.: Analysis of incompressible miscible displacement in porous media by characteristics collocation method. Numer. Methods Partial Differ. Equ. **22**, 797–814 (2006)
20. Percell, P., Wheeler, M.F.: A $\mathcal{C}^1$ finite element collocation method for elliptic equations. SIAM J. Numer. Anal. **17**, 605–622 (1980)