INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 14, Number 4-5, Pages 670–687

MAXIMAL L^p ERROR ANALYSIS OF FEMS FOR NONLINEAR PARABOLIC EQUATIONS WITH NONSMOOTH COEFFICIENTS

BUYANG LI AND WEIWEI SUN

Abstract. The paper is concerned with L^p error analysis of semi-discrete Galerkin FEMs for nonlinear parabolic equations. The classical energy approach relies heavily on the strong regularity assumption of the diffusion coefficient, which may not be satisfied in many physical applications. Here we focus our attention on a general nonlinear parabolic equation (or system) in a convex polygon or polyhedron with a nonlinear and Lipschitz continuous diffusion coefficient. We first establish the discrete maximal L^p-regularity for a linear parabolic equation with time-dependent diffusion coefficients in $L^\infty(0,T;W^{1,N+\epsilon})\cap C(\overline{\Omega}\times[0,T])$ for some $\epsilon>0$, where N denotes the dimension of the domain, while previous analyses were restricted to the problem with certain stronger regularity assumption. With the proved discrete maximal L^p-regularity, we then establish an optimal L^p error estimate and an almost optimal L^∞ error estimate of the finite element solution for the nonlinear parabolic equation.

Key words. Finite element method, nonlinear parabolic equation, polyhedron, nonsmooth coefficients, maximal L^p-regularity, optimal error estimate.

1. Introduction

The paper is to present a general framework for numerical analysis of optimal errors of finite element methods for nonlinear parabolic equations with nonsmooth coefficients. To illustrate our idea, we consider the equation

$$\begin{cases} \partial_t u - \displaystyle\sum_{i,j=1}^{N} \partial_i\big(\sigma_{ij}(u,x)\partial_j u\big) = g(u,\nabla u,x) & \text{in } \Omega\times(0,\infty), \\ u = 0 & \text{on } \partial\Omega\times(0,\infty), \\ u(\cdot,0) = u^0 & \text{in } \Omega, \end{cases} \tag{1}$$

in a polyhedral domain in $\mathbb{R}^N$, $N=2,3$, and its semi-discrete finite element approximation

$$\begin{cases} \big(\partial_t u_h, v_h\big) + \displaystyle\sum_{i,j=1}^{N} \big(\sigma_{ij}(u_h,x)\partial_j u_h, \partial_i v_h\big) = (g(u_h,\nabla u_h,x),v_h),\ \forall\ v_h\in S_h, \\ u_h(0) = u_h^0, \end{cases} \tag{2}$$

where S_h denotes a finite element subspace of $H_0^1(\Omega)$ consisting of continuous piecewise polynomials of degree $r\geq 1$ subject to a quasi-uniform triangulation of Ω with a mesh size h, and $u_h^0 = I_h u^0$ denotes the Lagrange interpolation of the initial data u^0. We only impose certain local conditions on the coefficients $\sigma_{ij}(u,x)=\sigma_{ji}(u,x)$ and the right-hand side $g(u,\nabla u,x)$, *i.e.* we assume that for $|u|\leq M$, $|\eta|\leq M$,

Received by the editors on November 14, 2016, accepted on March 24, 2017.
1991 *Mathematics Subject Classification.* 65M12, 65M15, 65M60.

$x \in \Omega$ and $t \in [0,T]$

$$
\begin{aligned}
&|\sigma_{ij}| + |\partial_u \sigma_{ij}| + |\partial_{x_l} \sigma_{ij}| \le K_M, \\
&K_M^{-1}|\xi|^2 \le \sum_{i,j=1}^N \sigma_{ij}(u,x)\xi_i\xi_j \le K_M|\xi|^2, \\
&|g(u,\eta,x)| + |\partial_u g(u,\eta,x)| + |\partial_{\eta_j} g(u,\eta,x)| + |\partial_{x_j} g(u,\eta,x)| \\
&\quad + |\partial_u^2 g(u,\eta,x)| + |\partial_{u\eta_j}^2 g(u,\eta,x)| + |\partial_{\eta_j\eta_l}^2 g(u,\eta,x)| \le K_M,
\end{aligned} \tag{3}
$$

for some positive constant K_M which may depend on M, where ∂_u, ∂_{x_l} and ∂_{η_j} denote the partial derivatives with respect to u, x and η_j, and ∂^2_{u,η_j} and $\partial^2_{\eta_j\eta_l}$ denote the mixed second-order partial derivatives.

The key to the optimal error estimate for the nonlinear problem (2) is more precise L^p estimates of the finite element solution, defined by

$$
\begin{cases}
\big(\partial_t\phi_h, v_h\big) + \displaystyle\sum_{i,j=1}^N \big(a_{ij}\partial_j\phi_h, \partial_i v_h\big) = (f, v_h), \ \forall\, v_h \in S_h, \\
\phi_h(0) = \phi_h^0
\end{cases} \tag{4}
$$

for the linear parabolic equation

$$
\begin{cases}
\partial_t\phi - \displaystyle\sum_{i,j=1}^N \partial_i\big(a_{ij}\partial_j\phi\big) = f & \text{in } \Omega\times(0,\infty), \\
\phi = 0 & \text{on } \partial\Omega\times(0,\infty), \\
\phi(\cdot,0) = \phi^0 & \text{in } \Omega
\end{cases} \tag{5}
$$

where $a_{ij} = a_{ji}$. Namely, the discrete maximal L^p-regularity

$$
\|\partial_t\phi_h\|_{L^p(0,T;L^q)} + \|A_h\phi_h\|_{L^p(0,T;L^q)} \le C_{p,q}\|f\|_{L^p(0,T;L^q)}, \qquad \text{if } \phi^0 = 0, \tag{6}
$$

$$
\|\partial_t\phi_h\|_{L^p(0,T;W^{-1,q})} + \|\phi_h\|_{L^p(0,T;W^{1,q})} \le C_{p,q}\|f\|_{L^p(0,T;W^{-1,q})}, \quad \text{if } \phi^0 = 0, \tag{7}
$$

and the optimal-order error estimate

$$
\|P_h\phi - \phi_h\|_{L^p(0,T;L^q)} \le C_{p,q}\|P_h\phi^0 - \phi_h^0\|_{L^q} + C_{p,q}\|P_h\phi - R_h\phi\|_{L^p(0,T;L^q)}\,. \tag{8}
$$

Here $A_h(t): S_h \to S_h$ is the discrete counterpart of the differential operator $A(t)u = -\partial_j(a_{ij}(\cdot,t)\partial_i u)$, defined by

$$
(A_h w_h, v_h) := \sum_{i,j=1}^N (a_{ij}\partial_j w_h, \partial_i v_h), \ \forall\, w_h, v_h \in S_h, \tag{9}
$$

$R_h(t)$ is the Ritz projection operator onto the finite element space, defined by

$$
\sum_{i,j=1}^N (a_{ij}\partial_j(w - R_h w), \partial_i v_h) = 0, \ \forall\, w \in H_0^1(\Omega), \ v_h \in S_h, \tag{10}
$$

and P_h is the L^2 projection operator onto the finite element space. It is noted that (6) is a discrete analogue of the continuous maximal L^p-regularity [14, 32] (also see [20, Lemma 2.1])

$$
\|\partial_t\phi\|_{L^p(0,T;L^q)} + \|A\phi\|_{L^p(0,T;L^q)} \le C_{p,q}\|f\|_{L^p(0,T;L^q)}, \qquad 1 < p,q < \infty, \tag{11}
$$

$$
\|\partial_t\phi\|_{L^p(0,T;W^{-1,q})} + \|\phi\|_{L^p(0,T;W^{1,q})} \le C_{p,q}\|f\|_{L^p(0,T;W^{-1,q})}, \quad 1 < p,q < \infty. \tag{12}
$$

In the last several decades, many efforts have been devoted to the maximal estimates of finite element solutions for linear parabolic equations. Most were based on the interior estimate/Green's function approach developed in [25, 26]. Among these, the maximal L^p-regularity (6) was proved in [7] for the linear parabolic equation (5) with time-independent coefficients in smooth domains. The corresponding L^∞ stability estimates have also been studied by many authors, see [4, 15, 21, 22, 24, 25, 26, 27, 28, 29] and the references therein. All these works focused only on linear autonomous parabolic equations in a smooth domain with the diffusion coefficient being smooth enough, *i.e.* $a_{ij} = a_{ij}(x) \in C^{2+\alpha}(\overline{\Omega})$. The discrete maximal L^p-regularity (6)-(7) in nonsmooth domains relies on more precise analysis of the discrete Green's functions. The extension of maximal L^p-regularity to time discretization were made in [16, 11, 12, 13]. For linear equations with time-independent coefficients, the discrete maximal L^p-regularity was presented in our recent work [20] for classical finite element solution in convex polyhedra and by Kemmochi and Saito [12] for a lumped mass method in general polyhedra. The extension of L^p error estimate to a semilinear problem in a smooth domain was made in [8], in which $\sigma_{ij} = \sigma_{ij}(x)$ is assumed to be $C^{2+\alpha}(\overline{\Omega})$ and time-independent. To extend the approach to general nonlinear problems, one of important issues to be considered is the regularity assumption of the diffusion coefficient. Recently, the regularity condition on the coefficients was weakened in [17], and (6)-(8) were proved in [17, 19] for the linear problem (5) with time-independent and time-dependent Lipschitz continuous coefficients, respectively, with the Neumann boundary condition in smooth domains. A specific weakly nonlinear elliptic-parabolic system from the model of incompressible miscible flow in porous media with the Neumann boundary condition in a smooth domain was also investigated [19]. The application of maximal L^p-regularity to time discretization of nonlinear parabolic equations can be found in [1, 2]. No analysis has been provided for finite element solutions of the nonlinear system (1) with locally Lipschitz continuous coefficients in convex polyhedra.

In this paper, we focus our attention on the finite element solutions of (1) in convex polygons and polyhedra. To deal with the nonlinear problem, we first extend the discrete maximal L^p estimates (6)-(8) for the autonomous case of the linear parabolic PDE (5), which was proved in [20], to the non-autonomous case with $a_{ij} = a_{ij}(x,t) \in L^\infty(0,T;W^{1,N+\epsilon_0}(\Omega)) \cap C(\overline{\Omega} \times [0,T])$. Due to the nonsmoothness of the polyhedra, the extension is made with certain restrictions on q. By utilizing the proved L^p estimates of the linear non-autonomous problem, we establish an optimal L^p error estimate and an almost optimal L^∞ error estimate of the finite element solution for the nonlinear parabolic equation. Our theoretical analysis provides a fundamental tool in establishing optimal error estimates of Galerkin FEMs for general nonlinear parabolic equations with coefficients of weak regularity.

We present our main results in the following two theorems.

Theorem 1.1. *Let Ω be either a convex polygon in $\mathbb{R}^2$ or a convex polyhedron in $\mathbb{R}^3$, and assume that the coefficients $a_{ij}(x,t) \in L^\infty(0,T;W^{1,N+\epsilon_0}(\Omega)) \cap C(\overline{\Omega} \times [0,T])$, $i,j = 1,\cdots,N$, satisfy the ellipticity condition*

$$K_0^{-1}|\xi|^2 \le \sum_{i,j=1}^{N} a_{ij}(x,t)\xi_i\xi_j \le K_0|\xi|^2, \quad \textit{for } x \in \Omega \ \textit{ and } \ t \in (0,T),$$

for some positive constant $\epsilon_0, K_0 > 0$. *Then there exists* $q_0 > 2$ *such that the solutions of* (4)-(5) *satisfy*

(6) *for* $1 < p < \infty$ *and* $1 < q < q_0$,
(7) *for* $1 < p < \infty$ *and* $1 < q < \infty$,
(8) *for* $1 < p < \infty$ *and* $q_0' < q < \infty$

where q_0' *is a positive integer satisfying* $1/q_0' + 1/q_0 = 1$. *The constant* q_0 *depends on* ϵ_0 *and the interior angle of the corners/edges of the polygon/polyhedron.*

Remark 1.1 The restriction on the index q is due to the smoothness of the domain and the time-dependency of the coefficients. Fortunately, both p and q can be arbitrarily close to infinity in the error estimate (8), which allows one to control the strong nonlinear terms involved in error analysis. If either the domain Ω is smooth or the coefficients a_{ij} are independent of t, then (6) and (8) hold for all $1 < p, q < \infty$.

Remark 1.2 For a two-dimensional convex polygon Ω with $a_{ij} \in L^\infty(0, T; W^{1,\infty}(\Omega)) \cap C(\overline{\Omega} \times [0, T])$, the constant q_0 can be chosen as $q_0 = 2/(2 - \min(\pi/\omega, 2))$, where ω denotes the maximal interior angle of the convex polygon (if $\omega \in (0, \pi/2]$ then $q_0 = \infty$).

Remark 1.3 The results are presented for a nonlinear problem with very general assumptions on σ_{ij} and g in (3). For example, the function $g = \pm e^u$ and even $g = \pm e^{|\nabla u|^2}$ satisfy (3). With such strongly nonlinearities, the problem (1) may not have a globally smooth solution. Nevertheless, (1) always has a smooth solution for some short time interval $[0, T]$ (local existence and uniqueness), provided that the initial data is smooth enough. In this article, we assume that the solution exists and sufficiently smooth (thus unique) in the time interval $[0, T]$, and investigate the stability and convergence of the semi-discrete finite element solutions.

Theorem 1.2. *Let* Ω *be either a convex polygon in* $\mathbb{R}^2$ *or a convex polyhedron in* $\mathbb{R}^3$. *Assume that the condition* (3) *is satisfied and the solution of* (1) *satisfies* $u \in L^\infty(0, T; W^{k+1,q})$ *for some fixed* $q > N$ *and integer* $k \in [1, r]$. *Then there exists a positive constant* h_0 *such that when* $h < h_0$ *the solutions of* (1)-(2) *satisfy*

$$\|u - u_h\|_{L^p(0,T;L^q)} \leq C_p h^{k+1}, \quad \forall\ 2 < p < \infty, \tag{13}$$

$$\|u - u_h\|_{L^\infty(0,T;L^q)} \leq C h^{k+1-\varepsilon_h}, \tag{14}$$

where $\varepsilon_h \to 0$ *as* $h \to 0$.

A traditional way to the optimal error estimate of Galerkin FEMs is based on an estimate in an energy-norm, *i.e.*

$$\|u - u_h\|_{L^\infty(0,T;L^2)} + \int_0^T \|u - u_h\|_{H^1}^2 dt \leq C h^{2r+2}.$$

The main difficulty in such an approach for the general nonlinear equation (1) is the low regularity of $\sigma_{ij}(u, x)$ and the strong nonlinearities in $g(u, \nabla u, x)$. A well-known technique in the approach is to use the elliptic Ritz projection $R_h(t) : H_0^1(\Omega) \to S_h$ [6, 33] defined by (10). This approach requires the a priori estimate

$$\|\partial_t(u - R_h u)\|_{L^2(\Omega\times(0,T))} \leq C h^{r+1}. \tag{15}$$

The above estimate was established in [33] under the regularity assumption

$$\|\nabla_x \partial_t \sigma_{ij}(u(x,t),x)\|_{L^\infty(\Omega\times(0,T))} \le C \tag{16}$$

for a general nonlinear parabolic equation. The condition (16) was required when Nitsche's trick (duality) was used in establishing (15). However, in some physical applications, the coefficients σ_{ij} may not satisfy the regularity condition (16). One of examples is the incompressible miscible flow in porous media [5, 18, 30], where $[\sigma_{ij}]_{i,j=1}^N$ denotes the diffusion-dispersion tensor which is locally Lipschitz continuous in many cases. For σ_{ij} being Lipschitz continuous, with a more precise analysis the above approach may yield a suboptimal error estimate

$$\|u-u_h\|_{L^\infty(0,T;L^2)} \le Ch^{r+1/2}, \tag{17}$$

which is half-order lower than the optimal order. Instead of the elliptic Ritz projection, one may use the corresponding interpolation. However, the error estimate obtained is usually one order lower than the optimal one, except some special case [31].

On the other hand, to deal with the strongly nonlinear term $g(u,\nabla u,x)$, the boundedness of $\|\nabla u_h\|_{L^\infty}$ is often needed in the error analysis since here, g may not satisfy a Lipschitz condition. This boundedness is usually proved by using certain inverse inequality and the error estimate to be proved (in terms of mathematics induction or a truncation approach), *i.e.*

$$\|P_hu-u_h\|_{L^\infty(0,T;W^{1,\infty})} \le Ch^{-1-N/2}\|P_hu-u_h\|_{L^\infty(0,T;L^2)} \le Ch^{r-1/2-N/2}. \tag{18}$$

Since the above condition requires $r-1/2-N/2>0$ to control $\|P_hu-u_h\|_{L^\infty(0,T;W^{1,\infty})}$, the frequent-used linear and quadratic FEMs are excluded and H^3 regularity of the solution is required for higher-order methods.

The rest part of this paper is organized as follows. In Section 2, we introduce some notations to be used in this paper. We prove Theorem 1.1 and Theorem 1.2 in Section 3 and Section 4, respectively.

2. Notations and lemma

Let $W^{k,p}(\Omega)$ be the standard Sobolev space of functions defined in Ω, where k is any nonnegative integer and $1\le p\le\infty$. Let $W_0^{1,p}(\Omega)$ be the subspace of $W^{1,p}(\Omega)$ consisting of functions whose traces vanish on $\partial\Omega$, and denote the dual space of $W_0^{1,p}(\Omega)$ by $W^{-1,p'}(\Omega)$ for $1\le p<\infty$. As conventions, we also use the notations $H^k(\Omega):=W^{k,2}(\Omega)$ and $L^p(\Omega):=W^{0,p}(\Omega)$.

For any Banach space X and a given $T>0$, $L^p(0,T;X)$ denotes the Bochner space equipped with the norm

$$\|f\|_{L^p(0,T;X)} = \begin{cases} \left(\int_0^T \|f(t)\|_X^p dt\right)^{\frac{1}{p}}, & 1\le p<\infty, \\ \operatorname*{ess\ sup}_{t\in(0,T)} \|f(t)\|_X. & p=\infty, \end{cases}$$

To simplify notations, we write L^p, H^k and $W^{k,p}$ as the abbreviations of $L^p(\Omega)$, $H^k(\Omega)$ and $W^{k,p}(\Omega)$, respectively, and define

$$(\phi,\varphi) := \int_\Omega \phi(x)\varphi(x)\mathrm{d}x.$$

For a given $t \in (0,T)$, we use the notation $w(t)$ to denote the function $w(\cdot,t)$ defined on Ω.

Let $a(x,t) = \big(a_{ij}(x,t)\big)_{N\times N}$ be the coefficient matrix and define the operators

$$A(t) : H_0^1 \to H^{-1}, \;\; A_h(t) : S_h \to S_h,$$
$$R_h(t) : H_0^1 \to S_h, \;\; P_h : L^2 \to S_h,$$

by

$$\big(A(t)\phi, v\big) = \big(a(\cdot,t)\nabla\phi, \nabla v\big), \qquad \text{for all } \phi, v \in H_0^1, \tag{19}$$

$$\big(A_h(t)\phi_h, v_h\big) = \big(a(\cdot,t)\nabla\phi_h, \nabla v_h\big), \qquad \text{for all } \phi_h \in S_h \text{ and } v_h \in S_h, \tag{20}$$

$$\big(A_h(t)R_h(t)\phi, v_h\big) = \big(A(t)\phi, v_h\big), \qquad \text{for all } \phi \in H_0^1 \text{ and } v_h \in S_h, \tag{21}$$

$$\big(P_h\phi, v_h\big) = \big(\phi, v_h\big), \qquad \text{for all } \phi \in L^2 \text{ and } v_h \in S_h. \tag{22}$$

Clearly, $R_h(t)$ is the conventional Ritz projection operator associated to the elliptic operator $A(t)$ and P_h is the L^2 projection operator onto the finite element space.

Under the assumptions of Theorem 1.1, there exists $q_0 > 2$ such that the following regularity estimates hold [9, 10]:

$$\|w\|_{W^{2,q}} \le C_q \|\nabla\cdot(a\nabla w)\|_{L^q}, \quad \text{for any } 1 < q < q_0, \;\; \forall\, w \in H_0^1, \tag{23}$$

$$\|\nabla w\|_{L^\infty} \le C_q \|\nabla\cdot(a\nabla w)\|_{L^q}, \quad \text{for any } q > N, \;\; \forall\, w \in H_0^1. \tag{24}$$

Moreover, we have the following projection error estimates:

$$\|\phi - P_h\phi\|_{W^{k,q}} \le Ch^{l-k}\|\phi\|_{W^{l,q}}, \quad 1 \le q \le \infty, \;\; 0 \le k \le 1, \;\; k \le l \le r+1, \tag{25}$$

$$\|\phi - R_h\phi\|_{W^{1,q}} \le Ch^{l-1}\|\phi\|_{W^{l,q}}, \quad 1 < q \le \infty, \;\; 1 \le l \le r+1, \tag{26}$$

$$\|\phi - R_h\phi\|_{L^q} \le Ch^{l}\|\phi\|_{W^{l,q}}, \quad q_0' < q < \infty, \;\; 1 \le l \le r+1, \tag{27}$$

where (25)-(26) are standard error estimates of FEMs in convex polygons and polyhedra [3], and (27) follows a duality argument by using the regularity estimate (23). Finally, we denote by $C = C_{p,q}$ a generic positive constant which is independent of h. To simplify the notations, we omit the subscripts in the generic constant when there is no ambiguity.

We present a generalized Gronwall's inequality in the following lemma, which was proved in [19].

Lemma 2.1. *Let $1 < p < \infty$ and let $Y = Y(t)$ be a continuous function defined on $[0,T]$. If the function $Y(t)$ satisfies*

$$\|Y\|_{L^p(\tau_1,\tau_2)} \le \alpha\|Y\|_{L^1(\tau_1,\tau_2)} + \alpha Y(\tau_1) + \beta$$

for any $0 \le \tau_1 < \tau_2 \le s$ and $s \in (0,T]$, with some positive constants α and β, then we have

$$\|Y\|_{L^p(0,s)} \le C_{T,\alpha,p}(Y(0) + \beta),$$

where the constant $C_{T,\alpha,p}$ is independent of $s \in (0,T]$.

3. Proof of Theorem 1.1

In this section, we study the linear parabolic equation (5) and the corresponding FE solution and prove Theorem 1.1. The following lemma was presented in [20] and will be used in our proof.

Lemma 3.1. *Under the assumptions of Theorem 1.1, if the coefficients $a_{ij} = a_{ij}(x)$, $i, j = 1, \cdots, N$, are independent of the time t, then Theorem 1.1 holds.*

To prove Theorem 1.1, firstly we assume that $\phi_h^0 = \phi^0 = 0$ and prove (6)-(7). For this purpose, we partition the time interval $[0, T]$ uniformly into $0 = t_0 < t_1 < \cdots < t_{N_0} = T$, with $t_n - t_{n-1} = \Delta t = T/N_0$, for $1 \leq n \leq N_0$, and rewrite (4) as

$$\begin{cases} \partial_t \phi_h(t) + A_h(t)\phi_h(t) = P_h f(t), \\ \phi_h(0) = 0. \end{cases} \tag{28}$$

From the equation above we can see that the function $\psi_h(t) = P_h\phi(t) - \phi_h(t)$ satisfies

$$\begin{cases} \partial_t \psi_h(t) + A_h(t_n)\psi_h(t) = (A_h(t_n) - A_h(t))\psi_h(t) + A_h(t)(P_h\phi(t) - R_h(t)\phi(t)), \\ \psi_h(0) = 0. \end{cases}$$

We extend ψ_h to be zero for $t \leq 0$. If we let $\varphi_h^n(t) = \psi_h(t) - \psi_h(2t_n - t)$ for $t \in [t_n, t_{n+1}]$, then φ_h^n is the solution of the equation

$$\begin{aligned}
&\partial_t \varphi_h^n(t) + A_h(t_n)\varphi_h^n(t) \\
=&(A_h(t_n) - A_h(t))\psi_h(t) + A_h(t)(P_h\phi(t) - R_h\phi(t)) \\
&- A_h(t_n)\psi_h(2t_n - t) + \partial_t\psi_h(2t_n - t), \\
=&(A_h(t_n) - A_h(t))\psi_h(t) + A_h(t)(P_h\phi(t) - R_h(t)\phi(t)) - A_h(t_n)\psi_h(2t_n - t) \\
&- A_h(2t_n - t)\psi_h(2t_n - t) + A_h(2t_n - t)(P_h\phi(2t_n - t) - R_h(2t_n - t)\phi(2t_n - t))
\end{aligned}$$

in the time interval $[t_n, t_{n+1}]$, with $\varphi_h^n(t_n) = 0$. This is a parabolic equation with time-independent coefficients. In view of Lemma 3.1, we can apply (7) to the equation above in the time interval $[t_n, t_{n+1}]$ to get

$$\begin{aligned}
\|\varphi_h^n\|_{L^p(t_n,t_{n+1};W^{1,q})} \leq C^*\Bigg(&\sup_{t\in[t_n,t_{n+1}]} \|a_{ij}(t_n) - a_{ij}(t)\|_{L^\infty} \|\psi_h\|_{L^p(t_n,t_{n+1};W^{1,q})} \\
&+ \|P_h\phi - R_h\phi\|_{L^p(t_{n-1},t_{n+1};W^{1,q})} + \|\psi_h\|_{L^p(t_{n-1},t_n;W^{1,q})}\Bigg)
\end{aligned}$$

where the constant C^* is independent of Δt. Since $a_{ij} \in C(\overline{\Omega} \times [0, T])$, there exists a positive constant τ_0 such that $C^* \sup_{t\in[t_n,t_{n+1}]} \|a_{ij}(t_n) - a_{ij}(t)\|_{L^\infty} < 1/2$ when $\Delta t < \tau_0$ and therefore, the last inequality reduces to

$$\begin{aligned}
&\|\psi_h\|_{L^p(t_n,t_{n+1};W^{1,q})} \leq C\|\psi_h\|_{L^p(t_{n-1},t_n;W^{1,q})} + C\|P_h\phi - R_h\phi\|_{L^p(t_{n-1},t_{n+1};W^{1,q})}, \\
&\|\psi_h\|_{L^p(t_0,t_1;W^{1,q})} \leq C\|P_h\phi - R_h\phi\|_{L^p(t_0,t_1;W^{1,q})} \quad (\text{since } \psi_h = 0 \text{ for } t \leq 0),
\end{aligned}$$

for $n = 1, 2, \ldots, N_0 - 1$. Note that the number of iterations is bounded by the constant $[T/\tau_0] + 1$, where $[T/\tau_0]$ denotes the largest integer which does not exceed T/τ_0. Iterating the above inequality leads to

$$\|\psi_h\|_{L^p(0,T;W^{1,q})} \leq C\|P_h\phi - R_h\phi\|_{L^p(0,T;W^{1,q})}, \tag{29}$$

which further implies (by the definition of ψ_h)

$$\begin{aligned}
\|\phi_h\|_{L^p(0,T;W^{1,q})} &\leq C\|\psi_h\|_{L^p(0,T;W^{1,q})} + C\|P_h\phi\|_{L^p(0,T;W^{1,q})} \\
&\leq C\|P_h\phi - R_h\phi\|_{L^p(0,T;W^{1,q})} + C\|P_h\phi\|_{L^p(0,T;W^{1,q})} \\
&\leq C\|\phi\|_{L^p(0,T;W^{1,q})} \\
&\leq C\|f\|_{L^p(0,T;W^{-1,q})}, \qquad (30)
\end{aligned}$$

where we have used (12) and (26) with $l = 1$. By using the above inequality and the equation (28), we obtain

$$\begin{aligned}(\partial_t\phi_h, v) &= (\partial_t\phi_h, P_h v)\\ &= (f, P_h v) - (a\nabla\phi_h, \nabla P_h v)\\ &\le \|f\|_{W^{-1,q}}\|P_h v\|_{W^{1,q'}} + \|\phi_h\|_{W^{1,q}}\|P_h v\|_{W^{1,q'}}\\ &\le C(\|f\|_{W^{-1,q}} + \|\phi_h\|_{W^{1,q}})\|v\|_{W^{1,q'}}, \qquad \forall\, v \in W_0^{1,q'},\end{aligned}$$

which implies (via duality)

$$\|\partial_t\phi_h\|_{W^{-1,q}} \le C(\|f\|_{W^{-1,q}} + \|\phi_h\|_{W^{1,q}}),$$

and so

$$\begin{aligned}\|\partial_t\phi_h\|_{L^p(0,T;W^{-1,q})} &\le C(\|f\|_{L^p(0,T;W^{-1,q})} + \|\phi_h\|_{L^p(0,T;W^{1,q})})\\ &\le C\|f\|_{L^p(0,T;W^{-1,q})}.\end{aligned} \tag{31}$$

This completes the proof of (7).

To prove (6), we note that

$$\begin{aligned}|\big(A_h(t)\psi_h(t), v\big)| &= |\big(A_h(t)\psi_h(t), P_h v\big)|\\ &= |\big(a(\cdot,t)\nabla\psi_h(t), \nabla P_h v\big)|\\ &\le C\|\psi_h(t)\|_{W^{1,q}}\|P_h v\|_{W^{1,q'}}\\ &\le Ch^{-1}\|\psi_h(t)\|_{W^{1,q}}\|P_h v\|_{L^{q'}} \qquad \text{(by the inverse inequality)}\\ &\le Ch^{-1}\|\psi_h(t)\|_{W^{1,q}}\|v\|_{L^{q'}}, \qquad \text{(by the inequality (25))}\end{aligned}$$

which gives (via duality)

$$\|A_h(t)\psi_h(t)\|_{L^q} \le Ch^{-1}\|\psi_h(t)\|_{W^{1,q}}\,. \tag{32}$$

As a consequence of (29)-(32), we see that

$$\begin{aligned}&\|A_h\phi_h\|_{L^p(0,T;L^q)}\\ \le&\|A_hP_h\phi\|_{L^p(0,T;L^q)} + \|A_h\psi_h\|_{L^p(0,T;L^q)}\\ \le&\|A_hP_h\phi\|_{L^p(0,T;L^q)} + Ch^{-1}\|\psi_h\|_{L^p(0,T;W^{1,q})} \qquad \text{(by using (32))}\\ \le&\|A_hP_h\phi\|_{L^p(0,T;L^q)} + Ch^{-1}\|R_h\phi - P_h\phi\|_{L^p(0,T;W^{1,q})} \quad \text{(by using (29))}\\ \le&\|A_hP_h\phi\|_{L^p(0,T;L^q)} + C\|\phi\|_{L^p(0,T;W^{2,q})}. \qquad \text{(by using (26))}\end{aligned}$$

By using (23), which requires $1 < q < q_0$, the last inequality further reduces to

$$\|A_h\phi_h\|_{L^p(0,T;L^q)} \le \|A_hP_h\phi\|_{L^p(0,T;L^q)} + C\|f\|_{L^p(0,T;L^q)}. \tag{33}$$

To estimate $\|A_hP_h\phi\|_{L^p(0,T;L^q)}$, we consider

$$\begin{aligned}&|\big(A_h(t)P_h\phi(t), v\big)|\\ =&|\big(A_h(t)P_h\phi(t), P_h v\big)|\\ \le&|\big(A_h(t)(P_h\phi(t) - R_h(t)\phi(t)), P_h v\big)| + |\big(A(t)\phi(t), P_h v\big)| \qquad \text{(by using (21))}\\ \le&C\|P_h\phi(t) - R_h(t)\phi(t)\|_{W^{1,q}}\|P_h v\|_{W^{1,q'}} + \|A(t)\phi(t)\|_{L^q}\|P_h v\|_{L^{q'}}\\ \le&Ch\|\phi(t)\|_{W^{2,q}}\|P_h v\|_{W^{1,q'}} + C\|A(t)\phi(t)\|_{L^q}\|P_h v\|_{L^{q'}}\\ \le&C(\|\phi(t)\|_{W^{2,q}} + \|A(t)\phi(t)\|_{L^q})\|v\|_{L^{q'}}. \qquad \text{(by using (25))}\end{aligned}$$

By duality, the last inequality implies

$$\begin{aligned}\|A_h(t)P_h\phi(t)\|_{L^p(0,T;L^q)} &\leq C(\|\phi(t)\|_{L^p(0,T;W^{2,q})} + \|A(t)\phi(t)\|_{L^p(0,T;L^q)})\\ &\leq C\|f\|_{L^p(0,T;L^q)},\end{aligned}$$

where we have used (11) and (23), which requires $1 < q < q_0$. From the above inequality and (33), we get

$$\|A_h\phi_h\|_{L^p(0,T;L^q)} \leq C\|f\|_{L^p(0,T;L^q)}, \quad \forall\, 1 < p < \infty,\ 1 < q < q_0. \tag{34}$$

By using the above inequality, we see from (28) that

$$\|\partial_t\phi_h\|_{L^p(0,T;L^q)} \leq \|A_h\phi_h\|_{L^p(0,T;L^q)} + \|P_hf\|_{L^p(0,T;L^q)} \leq C\|f\|_{L^p(0,T;L^q)} \tag{35}$$

for $1 < p < \infty$ and $1 < q < q_0$. We complete the proof of (6).

Secondly, we drop the assumption $\phi_h^0 = \phi^0 = 0$ and prove (8). For this purpose, we consider $e_h(t) = P_h\phi(t) - \phi_h(t) + \phi_h^0 - P_h\phi_0$, which is the solution of the equation

$$\begin{cases} \partial_t e_h(t) + A_h(t)e_h(t) = A_h(t)g_h(t), \\ \phi_h(0) = 0. \end{cases} \tag{36}$$

with

$$g_h(t) = \phi_h^0 - P_h\phi_0 + P_h\phi(t) - R_h(t)\phi(t). \tag{37}$$

By applying (7) to the above equation, we derive

$$\|\partial_t e_h\|_{L^p(0,T;W^{-1,q})} + \|e_h\|_{L^p(0,T;W^{1,q})} \leq C\|g_h\|_{L^p(0,T;W^{1,q})}. \tag{38}$$

Moreover, let v be the solution of the backward parabolic equation

$$\begin{cases} \partial_t v + \nabla\cdot\big(a\nabla v\big) = -\varphi & \text{in } \Omega, \\ v = 0 & \text{on } \partial\Omega, \\ v(T) = 0, \end{cases} \tag{39}$$

which obeys the basic estimate (the same as the forward parabolic equation)

$$\|v\|_{L^{p'}(0,T;W^{2,q'})} \leq C\|\varphi\|_{L^{p'}(0,T;L^{q'})}. \tag{40}$$

The last inequality is a consequence of (11) and (23), which requires $q_0' < q < \infty$. From (39), we see that

$$\begin{aligned}
&\int_0^T \big(e_h, \varphi\big)\mathrm{d}t \\
&= \int_0^T \big(e_h, -\partial_t v - \nabla\cdot(a\nabla v)\big)\mathrm{d}t \\
&= \int_0^T \big[\big(\partial_t e_h, v\big) + \big(a\nabla e_h, \nabla v\big)\big]\mathrm{d}t \\
&= \int_0^T [(\partial_t e_h(t), v(t) - P_h v(t)) + (A_h(t)e_h, v(t) - P_h v(t))]\mathrm{d}t \\
&\quad + \int_0^T (A_h(t)g_h(t), P_h v(t))\mathrm{d}t \\
&= \int_0^T \big(A_h(t)e_h(t), R_h(t)v(t) - P_h v(t)\big)\mathrm{d}t + \int_0^T \big(g_h(t), A(t)v(t)\big)\mathrm{d}t \\
&\quad + \int_0^T \big(g_h(t), A_h(t)(P_h v(t) - R_h(t)v(t))\big)\mathrm{d}t \qquad \text{(here we have used (21)-(22))} \\
&\le C\|e_h\|_{L^p(0,T;W^{1,q})}\|R_h v - P_h v\|_{L^{p'}(0,T;W^{1,q'})} + \|g_h\|_{L^p(0,T;L^q)}\|Av\|_{L^{p'}(0,T;L^{q'})} \\
&\quad + C\|g_h\|_{L^p(0,T;W^{1,q})}\|R_h v - P_h v\|_{L^{p'}(0,T;W^{1,q'})} \\
&\le Ch\|e_h\|_{L^p(0,T;W^{1,q})}\|v\|_{L^{p'}(0,T;W^{2,q'})} \\
&\quad + C\|g_h\|_{L^p(0,T;L^q)}(\|Av\|_{L^{p'}(0,T;L^{q'})} + \|v\|_{L^{p'}(0,T;W^{2,q'})}) \\
&\le C(h\|e_h\|_{L^p(0,T;W^{1,q})} + \|g_h\|_{L^p(0,T;L^q)})\|\varphi\|_{L^{p'}(0,T;L^{q'})}. \quad \text{(here we have used (40))}
\end{aligned}$$

By duality and using (38), we derive that

$$\begin{aligned}
&\|e_h\|_{L^p(0,T;L^q)} \\
\le& C(h\|e_h\|_{L^p(0,T;W^{1,q})} + \|g_h\|_{L^p(0,T;L^q)}) \\
\le& C(h\|g_h\|_{L^p(0,T;W^{1,q})} + h\|P_h\phi^0 - \phi_h^0\|_{W^{1,q}} \\
&\quad + \|g_h\|_{L^p(0,T;L^q)}) \qquad \text{(here we have used (38))} \\
\le& C(\|g_h\|_{L^p(0,T;L^q)} + \|P_h\phi^0 - \phi_h^0\|_{L^q}) \qquad \text{(by the inverse inequality)} \\
\le& C(\|P_h\phi - R_h\phi\|_{L^p(0,T;L^q)} + \|P_h\phi^0 - \phi_h^0\|_{L^q}). \qquad \text{(by the definition of } g_h)
\end{aligned}$$

This proves (8).

The proof of Theorem 1.1 is completed.

4. Proof of Theorem 1.2

In this section, we study the nonlinear parabolic equation (1) and the corresponding FE solution and prove Theorem 1.2 by applying Theorem 1.1.

Let $M = \|u\|_{L^\infty(\Omega_T)} + \|\nabla u\|_{L^\infty(\Omega_T)} + \|P_h u\|_{L^\infty(\Omega_T)} + \|\nabla(P_h u)\|_{L^\infty(\Omega_T)} + 2$, and let $I_c = \{x \in \mathbb{R} : |x| \le c\}$ for any $c > 0$. Let $\zeta(u, \mathbf{w})$ be a smooth function defined on $\mathbb{R}\times\mathbb{R}^N$ such that $\zeta(u, \mathbf{w}) = 1$ for $(u, \mathbf{w}) \in I_M \times I_M^N$ and $\zeta(u, \mathbf{w}) = 0$ for

$(u, \mathbf{w}) \notin I_{2M} \times I_{2M}^N$. We define the truncated coefficients and right-hand side by

$$\overline{\sigma}_{ij}(u,x) = \begin{cases} \sigma_{ij}(u,x) & \text{if } -2M \le u \le 2M, \\ \sigma_{ij}(2M,x) & \text{if } u \ge 2M, \\ \sigma_{ij}(-2M,x) & \text{if } u \le -2M, \end{cases}$$
$$\overline{g}(u,\mathbf{w},x) = g(u,\mathbf{w},x)\zeta(u,\mathbf{w}),$$

and define a cut-off function by

$$\chi(s) = \begin{cases} s, & \text{if } |s| \le 2, \\ 2, & \text{if } s \ge 2, \\ -2, & \text{if } s \le -2. \end{cases}$$

Based on the assumption (3), the truncated functions satisfy

$$\begin{aligned} &|\overline{\sigma}_{ij}(u,x)| + |\partial_u \overline{\sigma}_{ij}(u,x)| + |\partial_x \overline{\sigma}_{ij}(u,x)| \le K_{2M}, \\ &K_{2M}^{-1}|\xi|^2 \le \textstyle\sum_{i,j=1}^N \overline{\sigma}_{ij}(u,x,t)\xi_i\xi_j \le K_{2M}|\xi|^2, \\ &|\overline{g}(u,\eta,x)| + |\partial_u \overline{g}(u,\eta,x)| + |\partial_{\eta_j}\overline{g}(u,\eta,x)| \\ &+ |\partial_u^2 \overline{g}(u,\eta,x)| + |\partial_{u,\eta_j}^2 \overline{g}(u,\eta,x)| + |\partial_{\eta_j\eta_l}^2 \overline{g}(u,\eta,x)| \le CK_{2M} \end{aligned} \tag{41}$$

for any $u \in \mathbb{R}$ and $\eta \in \mathbb{R}^N$.

Let $\theta_h = u - P_h u$ be the projection error and let $e_h = u_h - P_h u$. By comparing (1) and (2), it is easy to see that e_h is the solution of the following finite element equation

$$\begin{aligned} &\big(\partial_t e_h, v\big) + \sum_{i,j=1}^N \big(\sigma_{ij}(u,x)\partial_i e_h, \partial_j v\big) \\ &= \sum_{i,j=1}^N \big(\sigma_{ij}(u,x)\partial_i \theta_h, \partial_j v\big) \\ &\quad + \sum_{i,j=1}^N \big((\sigma_{ij}(u,x) - \sigma_{ij}(e_h + P_h u, x))\partial_i(e_h + P_h u), \partial_j v\big) \\ &\quad - \big(g(P_h u, \nabla P_h u, x) - g(e_h + P_h u, \nabla(e_h + P_h u), x), v\big) \\ &\quad + \big(g(P_h u, \nabla P_h u, x) - g(u, \nabla u, x), v\big), \qquad \forall\, v \in S_h. \end{aligned} \tag{42}$$

Also we let $\overline{e}_h \in S_h$ be the solution of the following nonlinear finite element equation:

$$
\begin{aligned}
&\big(\partial_t \overline{e}_h, v\big) + \sum_{i,j=1}^{N} \big(\overline{\sigma}_{ij}(u,x)\partial_i \overline{e}_h, \partial_j v\big) \\
&= \sum_{i,j=1}^{N} \big(\overline{\sigma}_{ij}(u,x)\partial_i \theta_h, \partial_j v\big) \\
&\quad + \sum_{i,j=1}^{N} \big((\overline{\sigma}_{ij}(u,x) - \overline{\sigma}_{ij}(\overline{e}_h + P_h u, x))(\partial_i u + \chi(\partial_i(\overline{e}_h + P_h u - u))), \partial_j v\big) \\
&\quad - \big(\overline{g}(P_h u, \nabla P_h u, x) - \overline{g}(\overline{e}_h + P_h u, \nabla(\overline{e}_h + P_h u), x), v\big) \\
&\quad + \big(\overline{g}(P_h u, \nabla P_h u, x) - \overline{g}(u, \nabla u, x), v\big), \quad \forall\, v \in S_h,
\end{aligned} \tag{43}
$$

with the initial condition $\overline{e}_h(0) = I_h u(0) - P_h u(0)$. With the spatial discretization, (43) is essentially a system of ordinary differential equations. Thus existence and uniqueness of solutions of (43) are obvious and the solution is continuous with respect to $t \in [0,T]$. Moreover, if

$$\|\overline{e}_h + P_h u - u\|_{L^\infty} \le 2, \qquad \|\partial_i(\overline{e}_h + P_h u - u)\|_{L^\infty} \le 2,$$

then $\|\overline{e}_h + P_h u\|_{L^\infty} \le M$ and $\|\nabla(\overline{e}_h + P_h u)\|_{L^\infty} \le M$, which imply

$$
\begin{aligned}
&\overline{\sigma}_{ij}(\overline{e}_h + P_h u, x) = \sigma_{ij}(\overline{e}_h + P_h u, x), \\
&\overline{g}(\overline{e}_h + P_h u, \nabla(\overline{e}_h + P_h u), x) = g(\overline{e}_h + P_h u, \nabla(\overline{e}_h + P_h u), x).
\end{aligned}
$$

In this case, (43) reduces to (42) and so $\overline{e}_h = e_h$.

Now we proceed with a mathematical induction on

$$|\overline{e}_h| < 1/2, \qquad |\nabla \overline{e}_h| < 1/2. \tag{44}$$

Since

$$|\overline{e}_h(0)| + |\nabla \overline{e}_h(0)| = |I_h u(0) - P_h u(0)| + |\nabla(I_h u(0) - P_h u(0))| \le C h^{1-N/q} \|u\|_{W^{2,q}}$$

for $q > N$, there exists a positive constant h_1 such that when $h < h_1$ the inequality (44) holds at the initial time $t = 0$ and

$$\|P_h u - u\|_{L^\infty(0,T;L^\infty)} < 1/2, \qquad \|\nabla(P_h u - u)\|_{L^\infty(0,T;L^\infty)} < 1/2. \tag{45}$$

By continuity, we can assume that (44) holds for $t \in [0, s_0]$ for some $0 < s_0 \le T$. Moreover, if the inequality holds for $t \in [0,s]$ with some $0 < s \le T$, then there exists a positive constant δ_h such that

$$|\overline{e}_h| \le 1, \quad |\nabla \overline{e}_h| \le 1, \quad t \in [0, s+\delta_h] \cap [0,T], \tag{46}$$

$$\|\overline{e}_h + P_h u - u\|_{L^\infty(0,s+\delta_h;L^\infty)}, \quad \|\partial_i(\overline{e}_h + P_h u - u)\|_{L^\infty(0,s+\delta_h;L^\infty)} \le 3/2, \tag{47}$$

which implies

$$\overline{e}_h = e_h = u_h - P_h u, \qquad \text{for } t \in [0, s+\delta_h].$$

In the following, we shall prove that (44) also holds for $t \in [0, s+\delta_h]$. Then by induction, (44) holds in $[0,T]$ or an open subset of $[0,T]$, and by continuity, the latter can be extended to the closure of this subset. This implies that (44) holds for all $t \in [0,T]$.

To prove (44) for $t \in [0, s+\delta_h]$, we need to estimate those terms in the right-hand side of (43). To make use of Lemma 2.1, we consider (43) in the time interval

(τ_1,τ_2), with $0\le\tau_1<\tau_2\le s+\delta_h$. It should be kept in mind that the generic constant C below should be independent of s and δ_h (but may depend on T). By noting the fact

$$\big(\overline{g}(\overline{e}_h+P_hu,\nabla(\overline{e}_h+P_hu),x),v\big),=\big(\overline{g}(u_h,\nabla u_h,x),v\big),$$

we have

$$\big(\overline{g}(P_hu,\nabla P_hu,x)-\overline{g}(u_h,\nabla u_h,x),v\big)=\big(b_1e_h+\mathbf{b}_2\cdot\nabla e_h,v\big),$$

$$\begin{aligned}
&\big(\overline{g}(P_hu,\nabla P_hu,x)-\overline{g}(u,\nabla u,x),v\big)\\
&=\big(\overline{g}(P_hu,\nabla P_hu,x)-\overline{g}(u,\nabla P_hu,x)+\overline{g}(u,\nabla P_hu,x)-\overline{g}(u,\nabla u,x),v\big)\\
&=\big(b_3\theta_h,v\big)-\left(\nabla\theta_h\cdot\int_0^1\nabla_{\mathbf{w}}\overline{g}(u,(1-s)\nabla P_hu+s\nabla u,x)ds,v\right)\\
&=\big(b_3\theta_h,v\big)-\left(\nabla\theta_h\cdot\int_0^1\left[\nabla_{\mathbf{w}}\overline{g}(u,(1-s)\nabla P_hu+s\nabla u,x)-\nabla_{\mathbf{w}}\overline{g}(u,\nabla u,x)\right]ds,v\right)\\
&\quad-\left(\nabla\cdot(\theta_h\nabla_{\mathbf{w}}\overline{g}(u,\nabla u,x))-\theta_h\nabla\cdot(\nabla_{\mathbf{w}}\overline{g}(u,\nabla u,x)),v\right)\\
&=\big(b_3\theta_h,v\big)+\big(B\nabla\theta_h\cdot\nabla\theta_h+\theta_h\nabla\cdot\mathbf{b}_4,v\big)+\big(\mathbf{b}_4\theta_h,\nabla v\big),
\end{aligned}$$

and

$$\sum_{i,j=1}^N\big((\overline{\sigma}_{ij}(u,x)-\overline{\sigma}_{ij}(u_h,x))(\partial_iu+\chi(\partial_i(u_h-u))),\partial_jv\big)=\big((\theta_h-e_h),\mathbf{b}_5\cdot\nabla v\big),$$

where

$$\begin{aligned}
b_1&=-\int_0^1\partial_u\overline{g}((1-s)P_hu+su_h,(1-s)\nabla P_hu+s\nabla u_h,x)\mathrm{d}s,\\
\mathbf{b}_2&=\int_0^1\partial_{\mathbf{w}}\overline{g}((1-s)P_hu+su_h,(1-s)\nabla P_hu+s\nabla u_h,x)\mathrm{d}s,\\
b_3&=-\int_0^1\partial_u\overline{g}((1-s)P_hu+su,\nabla P_hu,x)\mathrm{d}s,\\
\mathbf{b}_4&=\nabla_{\mathbf{w}}\overline{g}(u,\nabla u,x),\\
B&=-\int_0^1\int_0^1\nabla^2_{\mathbf{w}}\overline{g}\big(u,(1-s')((1-s)\nabla u+s\nabla P_hu)+s'\nabla u,x\big)\mathrm{d}s\mathrm{d}s',\\
\mathbf{b}_5\cdot\nabla v&=\sum_{i,j=1}^N\partial_jv(\partial_iu+\chi(\partial_i(u_h-u)))\int_0^1\partial_u\overline{\sigma}_{ij}((1-s)u_h+su,x)\mathrm{d}s.
\end{aligned}$$

With above formulas, (43) reduces to

$$\begin{aligned}
&\big(\partial_t(e_h-\theta_h),v\big)+\sum_{i,j=1}^N\big(\overline{\sigma}_{ij}(u,x)\partial_i(e_h-\theta_h),\partial_jv\big)\\
&=\big(\theta_h,(\mathbf{b}_4+\mathbf{b}_5)\cdot\nabla v\big)+\big(b_3\theta_h+B\nabla\theta_h\cdot\nabla\theta_h+\theta_h\nabla\cdot\mathbf{b}_4,v\big)\\
&\quad-\big(\mathbf{b}_2\cdot\nabla e_h+b_1\overline{e}_h,v\big)-\big(e_h,\mathbf{b}_5\cdot\nabla v\big),\qquad\forall\ v\in S_h,
\end{aligned}\tag{48}$$

where we have noted $(\partial_t\theta_h, v) = 0$. Moreover, we can see

$$\begin{aligned}&\big(\partial_t(e_h + P_h u - w_h - u), v\big)\\ &\quad+\sum_{i,j=1}^N \big(\overline{\sigma}_{ij}(u,x)\partial_i(e_h + P_h u - w_h - u), \partial_j v\big) = 0, \qquad \forall\, v \in S_h,\end{aligned} \tag{49}$$

where w_h is the finite element solution of the equation

$$\begin{aligned}&\big(\partial_t w_h, v\big) + \sum_{i,j=1}^N \big(\overline{\sigma}_{ij}(u,x)\partial_i w_h, \partial_j v\big)\\ &= \big(\theta_h, (\mathbf{b}_4 + \mathbf{b}_5)\cdot\nabla v\big) + \big(b_3\theta_h + B\nabla\theta_h\cdot\nabla\theta_h + \theta_h\nabla\cdot\mathbf{b}_4, v\big)\\ &\quad - \big(\mathbf{b}_2\cdot\nabla e_h + b_1 e_h, v\big) - \big(e_h, \mathbf{b}_5\cdot\nabla v\big)\\ &= \big(\theta_h, (\mathbf{b}_4 + \mathbf{b}_5)\cdot\nabla v\big) + \big(b_3\theta_h + B\nabla\theta_h\cdot\nabla\theta_h + \theta_h\nabla\cdot\mathbf{b}_4, v\big)\\ &\quad - \big((b_1 - \nabla\cdot\mathbf{b}_2)e_h, v\big) - \big(e_h, (\mathbf{b}_5 - \mathbf{b}_2)\cdot\nabla v\big), \qquad \forall\, v \in S_h,\end{aligned} \tag{50}$$

with the initial condition $w_h(\tau_1) = 0$.

By (46), it is easy to see that

$$\begin{aligned}&\|b_3\|_{L^\infty(0,s+\delta_h;L^\infty)} + \|\mathbf{b}_4\|_{L^\infty(0,s+\delta_h;L^\infty)} + \|\mathbf{b}_5\|_{L^\infty(0,s+\delta_h;L^\infty)}\\ &\quad+ \|\nabla\cdot\mathbf{b}_2\|_{L^\infty(0,s+\delta_h;L^\infty)} + \|\nabla\cdot\mathbf{b}_4\|_{L^\infty(0,s+\delta_h;L^\infty)} + \|B\|_{L^\infty(0,s+\delta_h;L^\infty)} \le C\end{aligned}$$

for $t \in [0, s+\delta_h]$ and also we have

$$\|\theta_h\|_{L^p(0,T;L^q)} \le Ch^m\|u\|_{L^p(0,T;W^{m,q})}, \quad 0 \le m \le r+1, \quad 1 \le p,q \le \infty\,.$$

Substituting $v = w_h$ into (50) gives

$$\begin{aligned}&\|w_h\|_{L^\infty(\tau_1,\tau_2;L^2)}\\ \le&C\|\theta_h\|_{L^2(\tau_1,\tau_2;L^2)} + C\||\nabla\theta_h|^2\|_{L^2(\tau_1,\tau_2;L^2)} + C\|e_h\|_{L^2(\tau_1,\tau_2;L^2)}\\ \le&C\|\theta_h\|_{L^2(\tau_1,\tau_2;L^2)} + C\|\nabla\theta_h\|^2_{L^4(\tau_1,\tau_2;L^4)} + C\|e_h\|_{L^2(\tau_1,\tau_2;L^2)}\\ \le&Ch^{k+1}\|u\|_{L^2(\tau_1,\tau_2;H^{k+1})} + Ch^{2k}\|u\|^2_{L^4(\tau_1,\tau_2;W^{k+1,4})} + C\|e_h\|_{L^2(\tau_1,\tau_2;L^2)}\\ \le&Ch^{k+1} + C\|e_h\|_{L^2(\tau_1,\tau_2;L^2)}.\end{aligned}$$

By applying Theorem 1.1 to (50), we derive

$$\begin{aligned}&\|w_h\|_{L^p(\tau_1,\tau_2;W^{1,q})}\\ \le&C\|\theta_h\|_{L^p(\tau_1,\tau_2;L^q)} + C\||\nabla\theta_h|^2\|_{L^p(\tau_1,\tau_2;L^q)} + C\|e_h\|_{L^p(\tau_1,\tau_2;L^q)}\\ \le&C\|\theta_h\|_{L^p(\tau_1,\tau_2;L^q)} + C\|\nabla\theta_h\|^2_{L^{2p}(\tau_1,\tau_2;L^{2q})} + C\|e_h\|_{L^p(\tau_1,\tau_2;L^q)}\\ \le&Ch^{k+1}\|u\|_{L^p(\tau_1,\tau_2;W^{k+1,q})} + Ch^{2k}\|u\|^2_{L^{2p}(\tau_1,\tau_2;W^{k+1,2q})} + C\|e_h\|_{L^p(\tau_1,\tau_2;L^q)}\\ \le&Ch^{k+1} + C\|e_h\|_{L^p(\tau_1,\tau_2;L^q)},\end{aligned}$$

and by the Sobolev interpolation inequality we have (for $q > N$)

$$\begin{aligned}\|w_h\|_{L^p(\tau_1,\tau_2;L^q)} &\le C\|w_h\|^{\frac{2}{q}}_{L^\infty(\tau_1,\tau_2;L^2)}\|w_h\|^{1-\frac{2}{q}}_{L^{2p/q}(\tau_1,\tau_2;L^\infty)}\\ &\le C\|w_h\|^{\frac{2}{q}}_{L^\infty(\tau_1,\tau_2;L^2)}\|w_h\|^{1-\frac{2}{q}}_{L^{2p/q}(\tau_1,\tau_2;W^{1,q})}\\ &\le C(Ch^{k+1} + C\|e_h\|_{L^2(\tau_1,\tau_2;L^2)})^{\frac{2}{q}}\big(Ch^{k+1} + C\|e_h\|_{L^{2p/q}(\tau_1,\tau_2;L^q)}\big)^{1-\frac{2}{q}}\\ &\le Ch^{k+1} + C\|e_h\|_{L^2(\tau_1,\tau_2;L^2)} + C\|e_h\|_{L^{2p/q}(\tau_1,\tau_2;L^q)}.\end{aligned}$$

Again, applying Theorem 1.1 to (49), we obtain

$$\begin{aligned}
&\|e_h + P_h u - w_h - u\|_{L^p(\tau_1,\tau_2;L^q)} \\
\le & C\|e_h(\tau_1) + P_h u(\tau_1) - u(\tau_1)\|_{L^q} + C\|u - P_h u\|_{L^p(\tau_1,\tau_2;L^q)} \\
& + C\|u - R_h u\|_{L^p(\tau_1,\tau_2;L^q)} \\
\le & C\|e_h(\tau_1)\|_{L^q} + Ch^{k+1}\|u\|_{L^\infty(\tau_1,\tau_2;W^{k+1,q})},
\end{aligned} \tag{51}$$

and therefore,

$$\begin{aligned}
\|e_h\|_{L^p(\tau_1,\tau_2;L^q)} &\le \|e_h + P_h u - w_h - u\|_{L^p(\tau_1,\tau_2;L^q)} \\
&\quad + \|u - P_h u\|_{L^p(\tau_1,\tau_2;L^q)} + \|w_h\|_{L^p(\tau_1,\tau_2;L^q)} \\
&\le C\|e_h(\tau_1)\|_{L^q} + Ch^{k+1} + \|w_h\|_{L^p(\tau_1,\tau_2;L^q)} \\
&\le C\|e_h(\tau_1)\|_{L^q} + Ch^{k+1} + C\|e_h\|_{L^2(\tau_1,\tau_2;L^q)} + C\|e_h\|_{L^{2p/q}(\tau_1,\tau_2;L^q)} \\
&\le C\|e_h(\tau_1)\|_{L^q} + Ch^{k+1} + \epsilon\|e_h\|_{L^p(\tau_1,\tau_2;L^q)} + C_\epsilon\|e_h\|_{L^1(\tau_1,\tau_2;L^q)}
\end{aligned} \tag{52}$$

for any $\epsilon \in (0,1)$, which further leads to

$$\|e_h\|_{L^p(\tau_1,\tau_2;L^q)} \le C\|e_h(\tau_1)\|_{L^q} + Ch^{k+1} + C\|e_h\|_{L^1(\tau_1,\tau_2;L^q)}. \tag{53}$$

By Lemma 2.1 together with the above inequality, we arrive at the following estimate

$$\|e_h\|_{L^p(0,s+\delta_h;L^q)} \le C\|e_h(0)\|_{L^q} + Ch^{k+1} \le Ch^{k+1}, \tag{54}$$

where the constant C is independent of s and δ_h.

Note that e_h is the solution of the equation

$$\partial_t e_h + \overline{A}_h e_h = f_h,$$

where the operator $\overline{A}_h$ and the function f_h are defined by

$$\begin{aligned}
&\big(\overline{A}_h(t)\phi_h, v\big) = \big(\overline{\sigma}(u,x,t)\nabla\phi_h, \nabla v\big) \qquad \text{for all } \phi_h \in S_h,\ v \in S_h, \\
&f_h = \overline{A}_h\theta_h - \nabla_h \cdot [(\mathbf{b}_4 + \mathbf{b}_5)\theta_h + (\mathbf{b}_2 - \mathbf{b}_5)e_h] \\
&\qquad + P_h[b_3\theta_h + B\nabla\theta_h\cdot\nabla\theta_h + \theta_h\nabla\cdot\mathbf{b}_4 + [(\nabla\cdot\mathbf{b}_2) - b_1]\overline{e}_h].
\end{aligned}$$

Since

$$\begin{aligned}
&\|\overline{A}_h\theta_h\|_{L^p(0,s+\delta_h;W^{-1,q})} \le C\|\theta_h\|_{L^p(0,s+\delta_h;W^{1,q})} \le Ch^k, \\
&\|\nabla_h \cdot [(\mathbf{b}_4 + \mathbf{b}_5)\theta_h + (\mathbf{b}_2 - \mathbf{b}_5)e_h]\|_{L^p(0,s+\delta_h;W^{-1,q})} \\
&\qquad\qquad \le C\|(\mathbf{b}_4 + \mathbf{b}_5)\theta_h + (\mathbf{b}_2 - \mathbf{b}_5)e_h\|_{L^p(0,s+\delta_h;L^q)} \le Ch^{k+1}, \\
&\|B\nabla\theta_h\cdot\nabla\theta_h\|_{L^p(0,s+\delta_h;W^{-1,q})} \le C\|B\nabla\theta_h\cdot\nabla\theta_h\|_{L^p(0,s+\delta_h;L^q)} \\
&\qquad\qquad \le C\||\nabla\theta_h|^2\|_{L^p(0,s+\delta_h;L^q)} \le Ch^{k+1},
\end{aligned}$$

it follows that $\|f_h\|_{L^p(0,s+\delta_h;W^{-1,q})} \le Ch^k$. By Theorem 1.1, we have

$$\|\partial_t e_h\|_{L^p(0,s+\delta_h;W^{-1,q})} + \|e_h\|_{L^p(0,s+\delta_h;W^{1,q})} \le C\|f_h\|_{L^p(0,s+\delta_h;W^{-1,q})} \le Ch^k. \tag{55}$$

Since $k \ge 1$, with an inverse inequality we have

$$\|\partial_t e_h\|_{L^p(0,s+\delta_h;L^q)} \le Ch^{-1}\|\partial_t e_h\|_{L^p(0,s+\delta_h;W^{-1,q})} \le C \tag{56}$$

and by the Sobolev interpolation inequality,

$$\begin{aligned}\|e_h\|_{L^\infty(0,s+\delta_h;L^q)} &\le C\|e_h\|^{1-1/p}_{L^p(0,s+\delta_h;L^q)}\|e_h\|^{1/p}_{W^{1,p}(0,s+\delta_h;L^q)}\\ &\le Ch^{(k+1)(1-1/p)} \le Ch^{2-2/p}.\end{aligned}$$

Using the inverse inequality again, we see

$$\|e_h\|_{L^\infty(0,s+\delta_h;W^{1,\infty})} \le Ch^{-1-N/q}\|e_h\|_{L^\infty(0,s+\delta_h;L^q)} \le Ch^{2-2/p-N/q},$$

Since $p > 2$ and $q > N$, there exists $h_2 > 0$ such that when $h < h_2$

$$\|e_h\|_{L^\infty(0,s+\delta_h;W^{1,\infty})} < 1/2, \tag{57}$$

The mathematical induction on (44) is completed, which implies that (54)-(57) hold for any $s \in (0, T]$. In particular, (54) implies (13).

Finally, we see that (14) is a simple consequence of (13). In fact, for any fixed $q > N$, we have

$$\|e_h\|_{L^\infty(0,T;L^q)} \le C\|e_h\|^{1-1/p}_{L^p(0,T;L^q)}\|e_h\|^{1/p}_{W^{1,p}(0,T;L^q)} \le C_p h^{k+1-(k+1)/p},$$

where the Sobolev interpolation inequality and (56) have been used. In the above inequality, C_p is chosen as an increasing function of p. Let $h_0 = \min(h_1, h_2, C_3^{-3/(k+1)})$. Since the inequality above holds for any $2 < p < \infty$ and $h < h_0$, by choosing p such that $C_p = h^{-(k+1)/p}$, we get

$$\|e_h\|_{L^\infty(0,T;L^q)} \le h^{k+1-\varepsilon_h},$$

where

$$\varepsilon_h = 2(k+1)/\gamma((k+1)\ln(1/h))$$

and $p = \gamma(\zeta)$ is the inverse function of $\zeta = p\log C_p$. Clearly,

$$\lim_{h\to 0} \varepsilon_h = 0.$$

The proof of Theorem 1.2 is completed.

5. Conclusion

In this paper, we have presented a framework for optimal L^p-norm and almost optimal L^∞-norm error estimates of finite element solutions for general nonlinear parabolic equations in polygons and polyhedra, with nonsmooth diffusion coefficients. This approach is based on the discrete maximal L^p-regularity of linear parabolic finite element equations. Most previous analyses on maximal L^p estimates were restricted to the problem in a smooth domain with stronger regularity assumptions on the diffusion coefficients.

Acknowledgments

The authors would like to thank these two anonymous referees for their helpful suggestions. The work of B. Li was partially supported by a grant from the Germany/Hong Kong Joint Research Scheme sponsored by the Research Grants Council of Hong Kong and the German Academic Exchange Service of Germany (Ref. No. G-PolyU502/16). The work of W. Sun was supported in part by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China. (Project No. CityU 11302915).

References

[1] G. Akrivis and B. Li. Maximum norm analysis of implicit–explicit backward difference formulas for nonlinear parabolic equations. IMA J. Numer. Anal., 2017, DOI: 10.1093/imanum/drx008.

[2] G. Akrivis, B. Li, and C. Lubich. Combining maximal regularity and energy estimates for time discretizations of quasilinear parabolic equations. Math. Comp., 2016, DOI: http://dx.doi.org/10.1090/mcom/3228.

[3] S.C. Brenner and L.R. Scott: The Mathematical Theory of Finite Element Methods, Third Edition, Springer Science+Business Media, LLC, 2008.

[4] H. Chen: An L^2 and L^∞-error analysis for parabolic finite element equations with applications by superconvergence and error expansions, Doctoral Dissertation, Heidelberg University, 1993.

[5] J. Douglas, JR.: The numerical simulation of miscible displacement, Computational Methods in nonlinear Mechanics (J.T. Oden Ed.), North Holland, Amsterdam, 1980.

[6] T. Dupont, G. Fairweather and J.P. Johnson: Three-level Galerkin methods for parabolic equations, SIAM J. Numer. Anal., 11 (1974), pp. 392–410.

[7] M. Geissert: Discrete maximal L^p regularity for finite element operators, SIAM J. Numer. Anal., 44 (2006), pp. 677–698.

[8] M. Geissert: Applications of discrete maximal L^p regularity for finite element operators, Numer. Math., 108 (2007), pp. 121–149.

[9] P. Grisvard, Elliptic Problems in Nonsmooth Domains, SIAM 2011.

[10] M. Grüter and K. Widman: The Green function for uniformly elliptic equations, Manuscripta Math., 37 (1982), pp. 303–342.

[11] T. Kemmochi. Discrete maximal regularity for abstract Cauchy problems. Studia Math., 234 (2016), pp. 241–263.

[12] T. Kemmochi and N. Saito: Discrete maximal regularity and the finite element method for parabolic equations. Preprint. http://arXiv.org/abs/1602.06864

[13] B. Kovács, B. Li, and C. Lubich. A-stable time discretization preserves maximal parabolic regularity. SIAM J. Numer. Anal., 54 (2016), pp. 3600–3624.

[14] P.C. Kunstmann and L. Weis: Maximal L^p-regularity for parabolic equations, Fourier multiplier theorems and H^∞-functional calculus. Functional Analytic Methods for Evolution Equations, Lecture Notes in Mathematics, 1855 (2004), pp. 65–311.

[15] D. Leykekhman: Pointwise localized error estimates for parabolic finite element equations, Numer. Math., 96 (2004), pp. 583–600.

[16] D. Leykekhman and B. Vexler. Discrete maximal parabolic regularity for Galerkin finite element methods. Numer. Math., 135 (2017), pp. 923–952.

[17] B. Li: Maximum-norm stability and maximal L^p regularity of FEMs for parabolic equations with Lipschitz continuous coefficients. Numer. Math., 131 (2015), pp. 489–516.

[18] B. Li and W. Sun: Unconditional convergence and optimal error estimates of a Galerkin-mixed FEM for incompressible miscible flow in porous media, SIAM J. Numer. Anal., 51 (2013), pp. 1949–1977.

[19] B. Li and W. Sun: Regularity of the diffusion-dispersion tensor and error analysis of FEMs for a porous media flow. SIAM J. Numer. Anal., 53 (2015), pp. 1418–1437.

[20] B. Li and W. Sun: Maximal L^p analysis of finite element solutions for parabolic equations with nonsmooth coefficients in convex polyhedra, Math. Comp., 86 (2017), pp. 1071–1102.

[21] Y. Lin: On maximum norm estimates for Ritz-Volterra projection with applications to some time dependent problems, J. Comput. Math., 15 (1997), pp. 159–178.

[22] Y. Lin, V. Thomée and L.B. Wahlbin: Ritz-Volterra projections to finite-element spaces and applications to integrodifferential and related equations, SIAM J. Numer. Anal., 28 (1991), pp. 1047–1070.

[23] T.A. Lucas: Maximum-norm estimates for an immunology model using reaction-diffusion equations with stochastic source terms, SIAM J. Numer. Anal., 49 (2011), pp. 2256–2276.

[24] C. Palencia: Maximum norm analysis of completely discrete finite element methods for parabolic problems, SIAM J. Numer. Anal., 33 (1996), pp. 1654–1668.

[25] A.H. Schatz, V. Thomée and L.B. Wahlbin: Maximum norm stability and error estimates in parabolic finite element equations, Comm. Pure Appl. Math., 33 (1980), pp. 265–304.

[26] A.H. Schatz, V. Thomée and L.B. Wahlbin: Stability, analyticity, and almost best approximation in maximum norm for parabolic finite element equations, Comm. Pure Appl. Math., 51(1998), pp. 1349–1385.
[27] A. Solo: Sharp estimates for finite element approximations to parabolic problems with Neumann boundary data of low regularity, BIT Numer. Math., 48 (2008), pp. 117–137.
[28] V. Thomée and L. B. Wahlbin: Maximum norm stability and error estimates in Galerkin methods for parabolic equations in one space variable, Numer. Math., 41 (1983), pp. 345–371.
[29] V. Thomée and L.B. Wahlbin: Stability and analyticity in maximum-norm for simplicial Lagrange finite element semidiscretizations of parabolic equations with Dirichlet boundary conditions, Numer. Math., 87 (2000), pp. 373–389.
[30] J. Wang, Z. Si and W. Sun: A new error analysis on characteristic methods for incompressible miscible flow in porous media, SIAM J. Numer. Anal., 52 (2014), pp. 3200–3223.
[31] K. Wang, H. Wang, S. Sun and M.F. Wheeler: An optimal-order L^2-error estimate for nonsymmetric discontinuous Galerkin methods for a parabolic equation in multiple space dimensions, Comput. Methods Appl. Mech. Engrg., 198 (2009), pp. 2190–2197.
[32] L. Weis: Operator-valued Fourier multiplier theorems and maximal L_p-regularity, Math. Ann., 319 (2001), pp. 735–758.
[33] M.F. Wheeler: A priori L^2 error estimates for Galerkin approximations to parabolic partial differential equations, SIAM J. Numer. Anal., 10 (1973), pp. 723–759.

Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong.
E-mail: `buyang.li@polyu.edu.hk`
URL: `http://www.polyu.edu.hk/ama/people/detail/739/`

Department of Mathematics, City University of Hong Kong, Hong Kong.
E-mail: `maweiw@math.cityu.edu.hk`
URL: `http://www6.cityu.edu.hk/ma/people/profile/sunww.htm`