SIAM J. NUMER. ANAL.
Vol. 55, No. 2, pp. 521–542

MAXIMAL REGULARITY OF FULLY DISCRETE FINITE ELEMENT SOLUTIONS OF PARABOLIC EQUATIONS*

BUYANG LI[†] AND WEIWEI SUN[‡]

Abstract. We establish the maximal ℓ^p-regularity for fully discrete finite element solutions of parabolic equations with time-dependent Lipschitz continuous coefficients. The analysis is based on a discrete $\ell^p(W^{1,q})$ estimate together with a duality argument and a perturbation method. Optimal-order error estimates of fully discrete finite element solutions in the norm of $\ell^p(L^q)$ follows immediately.

Key words. nonlinear parabolic equations, BDF methods, discrete maximal parabolic regularity, maximum-norm error analysis, energy technique, time-dependent norms

AMS subject classifications. Primary, 65M12, 65M60; Secondary, 65L06

DOI. 10.1137/16M1071912

1. Introduction. We study fully discrete finite element approximations of the parabolic problem

$$(1.1)\qquad \begin{cases} \partial_t u - \displaystyle\sum_{i,j=1}^N \partial_i\big(a_{ij}\partial_j u\big) + cu = f - \sum_{j=1}^N \partial_j g_j & \text{in } \Omega\times(0,T),\\ \displaystyle\sum_{i,j=1}^N a_{ij}\partial_j u\, n_i = \sum_{j=1}^N g_j n_j & \text{on } \partial\Omega\times(0,T),\\ u(0)=u^0 & \text{in } \Omega, \end{cases}$$

in a bounded smooth domain $\Omega\subset\mathbb{R}^N$, $N=1,2,3$, with the coefficients $a_{ij}=a_{ji}$ satisfying the strong ellipticity condition

$$(1.2)\qquad \begin{aligned} &K^{-1}|\xi|^2 \le \sum_{i,j=1}^N a_{ij}(x,t)\xi_i\xi_j \le K|\xi|^2 && \text{for } \xi\in\mathbb{R}^N \text{ and } (x,t)\in\Omega\times(0,T],\\ &c(x,t)\ge c_0 && \text{for } (x,t)\in\Omega\times(0,T], \end{aligned}$$

where K and c_0 are some positive constants.

Define the elliptic operator $A(t): H^1(\Omega)\to H^1(\Omega)'$ and its semidiscrete finite element approximation $A_h(t): S_h\to S_h$ by

$$(1.3)\qquad (A(t)w,v) := \sum_{i,j=1}^N (a_{ij}(\cdot,t)\partial_j w, \partial_i v) + (c(\cdot,t)w,v) \qquad \forall\, w,v\in H^1(\Omega),$$

$$(1.4)\qquad (A_h(t)w_h,v_h) := \sum_{i,j=1}^N (a_{ij}(\cdot,t)\partial_j w_h, \partial_i v_h) + (c(\cdot,t)w_h,v_h)\ \forall\, w_h,v_h\in S_h,$$

*Received by the editors April 22, 2016; accepted for publication (in revised form) December 9, 2016; published electronically March 7, 2017.
http://www.siam.org/journals/sinum/55-2/M107191.html

Funding: The work of the first author was partially supported by The Hong Kong Polytechnic University (A/C code: 1-ZE6L), and was partially carried during a research stay at Universität Tübingen, funded by the Alexander von Humboldt foundation and NSFC 11301262. The work of the second author was supported in part by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project CityU 11302915).

[†]Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong (buyang.li@polyu.edu.hk).

[‡]Department of Mathematics, City University of Hong Kong, Kowloon, Hong Kong (maweiw@cityu.edu.hk).

where S_h denotes a finite element subspace of $H^1(\Omega)$ consisting of continuous piecewise polynomials of degree $r \geq 1$ subject to a quasi-uniform triangulation which fit the domain Ω exactly, as assumed in [30, 32]. Then (1.1) can be written in the abstract form

$$\frac{\partial u}{\partial t} + Au = f - \overline{\nabla} \cdot \mathbf{g}, \tag{1.5}$$

and its semidiscrete finite element approximation can be written as

$$\frac{\partial u_h}{\partial t} + A_h u_h = f_h - \overline{\nabla}_h \cdot \mathbf{g}, \tag{1.6}$$

where $f_h = P_h f$ (the L^2 projection of f onto the finite element space) and $\mathbf{g} = (g_1, \ldots, g_N)$; the divergence operators $\overline{\nabla}\cdot : H^1(\Omega)^N \to H^1(\Omega)'$ and $\overline{\nabla}_h\cdot : H^1(\Omega)^N \to S_h$ will be defined in section 2. Analysis of (1.6) has been done for a variety of finite element methods and many other numerical methods, particularly in the spaces $L^\infty(0,T;L^2)$ and $L^2(0,T;H^1)$.

If the coefficients a_{ij} are time independent and $a_{ij} \in W^{1,\infty}(\Omega)$, it is well known that the solution of the parabolic problem (1.1) possesses the maximal parabolic regularity [35, 36]

$$\|\partial_t u\|_{L^p(0,T;L^q)} + \|Au\|_{L^p(0,T;L^q)} \leq C\|f\|_{L^p(0,T;L^q)} \quad \text{if } u(0) = 0 \text{ and } \mathbf{g} = 0, \tag{1.7}$$

$$\|u\|_{L^p(0,T;W^{1,q})} \leq C\|\mathbf{g}\|_{L^p(0,T;L^q)} \quad \text{if } u(0) = 0 \text{ and } f = 0, \tag{1.8}$$

for $1 < p, q < \infty$. These results also have been extended to time-dependent coefficients under various conditions [1, 9]. The extension of these estimates to discrete settings is of significant importance as they provide more precise error estimates and a new tool for the analysis of numerical methods for nonlinear parabolic problems. Numerous efforts have been made in the last several decades. For parabolic equations with time-independent smooth coefficients $a_{ij} = a_{ij}(x) \in C^{2+\alpha}(\overline{\Omega})$, Geissert [10, 11] proved that the finite element solution of the semidiscrete equation (1.6) satisfies the spatially discrete maximal L^p-regularity

$$\|\partial_t u_h\|_{L^p(0,T;L^q)} + \|A_h u_h\|_{L^p(0,T;L^q)} \leq C\|f\|_{L^p(0,T;L^q)} \quad \text{when } u_h(0) = 0 \text{ and } \mathbf{g} = 0. \tag{1.9}$$

The proof is based on the maximum-norm stability analysis established by Schatz, Thomée, and Wahlbin [32]. Recently, the first author [20] proved (1.9) together with

$$\|u_h\|_{L^p(0,T;W^{1,q})} \leq C\|\mathbf{g}\|_{L^p(0,T;L^q)} \quad \text{when } u_h(0) = 0 \text{ and } f = 0 \tag{1.10}$$

for Lipschitz continuous coefficients $a_{ij} \in W^{1,\infty}(\Omega)$. Under the Dirichlet boundary condition, the estimates (1.9)–(1.10) were established in [22] for the problem in convex polyhedra with classical finite element approximations, and the estimate (1.9) was proved by Kemmochi and Saito [14, Theorem I] for the problem in general polyhedra with a lumped mass method by using the discrete maximum principle.

A straightforward application of (1.9)–(1.10) is the error estimate

$$\|P_h u - u_h\|_{L^p(0,T;L^q)} \leq C(\|P_h u^0 - u_h^0\|_{L^q} + \|P_h u - R_h u\|_{L^p(0,T;L^q)}), \tag{1.11}$$

where $R_h(t)$ is the Ritz projection operator associated with the elliptic operator $A(t)$ for $t \in [0,T]$. The error estimate (1.11) is optimal with respect to the regularity of the solution. On the other hand, the space-time maximum-norm error estimate

$$\|P_h u - u_h\|_{L^\infty(0,T;L^\infty)} \leq C\|P_h u_0 - u_h^0\|_{L^\infty} + C\ell_h\|P_h u - u\|_{L^\infty(0,T;L^\infty)}, \tag{1.12}$$

has also been studied by many people [5, 8, 13, 17, 19, 20, 25, 26, 28, 31, 32, 33] for linear autonomous parabolic equations with smooth coefficients, where $\ell_h = |\ln h|^m$ for some nonnegative constant m. The extension of (1.9)–(1.12) to nonautonomous equations with Lipschitz continuous coefficients ($a_{ij} = a_{ij}(x,t)$) is not straightforward, but it plays a key role in studying finite element solutions of nonlinear parabolic problems. In our recent work [21], we proved (1.9)–(1.11) for linear parabolic equations with time-dependent Lipschitz continuous coefficients and then applied them to analyze a special nonlinear parabolic model arising from a porous medium flow, where the diffusion-dispersion coefficient is a nonlinear function of the solution. The analysis was based on a perturbation argument, which converts the problem locally (in time) to a parabolic equation with time-independent coefficients. All these works focused on the semidiscrete finite element solutions of (1.6).

The stability and regularity of fully discrete finite element solutions seem to be more complicated. Less work has been done in this direction. Some early work on fully discrete schemes only focused on one-dimensional models. For a simple two-dimensional parabolic equation with $a_{ij} = \delta_{ij}$ (the Kronecker symbol), Schatz, Thomée, and Wahlbin [31] studied a fully discrete approximation, in which a strongly $A(\theta)$-stable scheme was used in time direction and a linear finite element method was used for spatial discretization. They proved the stability of the discrete semigroup $\{E^n_{h,\tau} = (1+\tau A_h)^{-n}\}_{n=0}^\infty$ in the maximum norm:

$$\|E^n_{h,\tau} v_h\|_{L^\infty} \le C\|v_h\|_{L^\infty} \quad \forall\, v_h \in S_h,\ n = 0,1,2,\ldots. \tag{1.13}$$

Later, Palencia [27] proved the estimate for $r \ge 1$ and $N = 1,2,3$, and Hansbo [13] proved a related $L^s \to L^p$ estimate with $1 \le s \le 2$. In these works, the stability of the fully discrete semigroup was established for some strongly $A(\theta)$-stable schemes, excluding the Crank–Nicolson scheme. The inequality (1.13) can be viewed as a stability estimate with respect to the initial value in general.

The corresponding stability estimates with respect to f and $\mathbf{g}$ are more important since they imply directly error estimates of finite element solutions for linear problems and also play a key role for nonlinear problems. In the case $\mathbf{g} \equiv 0$, the maximal ℓ^p-regularity for time-discrete (spatial continuous) systems of (1.5) was studied by several authors. Ashyralyev, Piskarev, and Weis [2, Remark 5.2] studied the semidiscrete backward Euler and Crank–Nicolson time discretizations of (1.5) (time discrete and space continuous), respectively, and proved the corresponding time-discrete maximal ℓ^p-regularity

$$\|D_\tau \vec{U}\|_{\ell^p(L^q)} + \|\vec{U}\|_{\ell^p(W^{2,q})} \le C\|\vec{f}\|_{\ell^p(L^q)} \qquad \text{if } \vec{g}_j \equiv \vec{0}, \tag{1.14}$$

for $U^0 \equiv 0$ and $1 < p,q < \infty$, where D_τ denotes the backward Euler difference operator and $\vec{U} = (U^1,\ldots,U^M)$ is the solution of a time-discrete system. Leykekhman and Vexler [18] studied a class of discontinuous Galerkin time discretization, and proved the discrete maximal L^p-regularity

$$\begin{aligned} &\|\partial_t U_\tau\|_{L^p(0,T;L^q)} + \left(\sum \tau \left\|\frac{[U_\tau]}{\tau}\right\|^p_{L^q}\right)^{\frac{1}{p}} + \|\Delta U_\tau\|_{L^p(0,T;L^q)} \\ &\quad \le C \ln\left(\frac{T}{\tau}\right)\|f\|_{L^p(0,T;L^q)} \qquad \forall\, 1 \le p,q \le \infty. \end{aligned} \tag{1.15}$$

Recently, Kovács, Li, and Lubich [15] proved the maximal ℓ^p-regularity (1.14) for $1 < p,q < \infty$, for the backward differentiation formula (BDF) and A-stable Runge–Kutta time discretization methods. By noting the R-boundedness of the finite element resolvent operators,

$$z(z + A_h)^{-1}, \quad z \in \Sigma_\vartheta := \{z \in \mathbb{C} : |\arg(z)| \le \vartheta\}, \tag{1.16}$$

which is a consequence of [22, estimate (2.13)] and [35, Lemma 4.c] (also see [20]), the maximal ℓ^p-regularity (1.14) can be extended immediately to fully discrete solutions with classical finite element approximations in space to (cf. [15, Theorem 6.1 and Remark 6.1])

$$\|D_\tau \vec{U}_h\|_{\ell^p(L^q)} + \|\vec{A}_h \vec{U}_h\|_{\ell^p(L^q)} \le C\|\vec{f}\|_{\ell^p(L^q)} \qquad \text{if } \vec{g}_j \equiv \vec{0}, \tag{1.17}$$

where $\vec{U}_h = (U_h^1, \ldots, U_h^M)$ denotes the solution of a fully discrete system and $\vec{A}_h \vec{U}_h = (A_h(t_1)U_h^1, \ldots, A_h(t_N)U_h^M)$. More recently, Kemmochi and Saito [14] investigated both the maximal ℓ^p-regularity for the time-discrete solutions given by the θ-scheme, as well as the maximal ℓ^p-regularity of fully discrete solutions with a lumped mass method for the spatial discretization.

However, all the works mentioned above for fully discrete solutions only considered parabolic equations with time-independent coefficients, $a_{ij} = a_{ij}(x)$ since the semigroup approach used in previous works is applicable only for the problem with time-independent coefficient. In this paper, we establish the maximal ℓ^p-regularity (1.17), together with the $\ell^p(W^{1,q})$ estimate

$$\|D_\tau \vec{U}_h\|_{\ell^p(\widetilde{W}^{-1,q})} + \|\vec{U}_h\|_{\ell^p(W^{1,q})} \le C\left(\|\vec{f}\|_{\ell^p(L^q)} + \sum_{j=1}^{N} \|\vec{g}_j\|_{\ell^p(L^q)}\right), \tag{1.18}$$

for the problem with the Neumann boundary condition, the time-dependent Lipschitz continuous coefficients $a_{ij}(x,t)$, and the bounded measurable coefficient $c(x,t)$, where $\widetilde{W}^{-1,q}$ denotes the dual space of $W^{1,q'}$ with $1/q + 1/q' = 1$. Analysis for the maximal ℓ^p-regularity of fully discrete solutions is based on a perturbation technique with a duality argument and a more precise estimate of time-discrete solutions. For simplicity, here we focus our attention on the backward Euler scheme for the time discretization. Our techniques can be extended to the Crank–Nicolson and BDF methods once the R-boundedness of (1.16) can be proved for the angle ϑ required by [15, Theorems 4.1–4.2]. Although we only consider linear problems in this paper, the estimates derived in this paper are useful for analyzing fully discrete finite element solutions of nonlinear problems in a similar way to [21].

2. Main results. Let $0 = t_0 < t_1 < \cdots < t_M = T$ be a uniform partition of the interval $[0, T]$ for some positive integer M and with $\tau = t_n - t_{n-1}$ and let $(0, T]_\tau := \{t_1, \ldots, t_M\}$ be a measure space equipped with the measure

$$|\{t_{k_1}, t_{k_2}, \ldots, t_{k_m}\}| = m\tau.$$

For any Banach space X and a given function $\vec{f} : (0, T]_\tau \to X$, we define the seminorm

$$\|\vec{f}\|_{\ell^p(t_n, t_m; X)} := \left(\sum_{k=n+1}^{m} \tau \|f\|_X^p\right)^{\frac{1}{p}}$$

and the norm

$$\|\vec{f}\|_{\ell^p(X)} := \|\vec{f}\|_{\ell^p(0,T;X)}.$$

We write L^q and $W^{k,q}$ as the abbreviations of the classical function spaces $L^q(\Omega)$ and $W^{k,q}(\Omega)$, respectively, and define $\widetilde{W}^{-1,q} := W^{1,q'}(\Omega)'$ to be the dual space of $W^{1,q'}(\Omega)$. The dual space of $W_0^{1,q'}(\Omega)$ is denoted by $W^{-1,q}$. The following L^2 inner product

$$(w,v) = \int_\Omega wv \,\mathrm{d}x\,,$$

will be used to simplify the notations, and we denote by L_h^q the finite element space S_h equipped with the norm of L^q.

A fully discrete finite element solution with the backward Euler scheme is defined by

$$\begin{aligned}(2.1)\qquad &\big(D_\tau U_h^n, v_h\big) + \sum_{i,j=1}^N \big(a_{ij}(\cdot,t_n)\partial_j U_h^n, \partial_i v_h\big) + \big(c(\cdot,t_n)U_h^n, v_h\big)\\ &= \big(f^n, v_h\big) + \sum_{j=1}^N \big(g_j^n, \partial_j v_h\big) \qquad \forall\, v_h \in S_h,\end{aligned}$$

where $D_\tau U_h^n = (U_h^n - U_h^{n-1})/\tau$ denotes the backward difference operator, and $U_h^0 \in S_h$ is a given approximation of the initial data U^0. Equation (2.1) can be viewed as the spatial discretization of the time-discrete PDEs [23, 24]

$$(2.2)\qquad \begin{cases} D_\tau U^n - \displaystyle\sum_{i,j=1}^N \partial_i\big(a_{ij}(\cdot,t_n)\partial_j U^n\big) + c(\cdot,t_n)U^n = f^n - \sum_{j=1}^N \partial_j g_j^n & \text{in } \Omega,\\ \displaystyle\sum_{i,j=1}^N a_{ij}(\cdot,t_n)\partial_j U^n n_i = \sum_{j=1}^N g_j^n n_j & \text{on } \partial\Omega,\\ U^0 = u^0, \end{cases}$$

where $f^n(x) = f(x,t_n)$ and $g^n(x) = g(x,t_n)$.

The main result of this paper is the following theorem.

THEOREM 2.1. *Let $\Omega \subset \mathbb{R}^N$, $N = 1,2,3$, be a bounded smooth domain. If the coefficients $a_{ij}(x,t) = a_{ji}(x,t)$ and $c(x,t)$ satisfy condition* (1.2) *and*

$$a_{ij} \in W^{1,\infty}(\Omega\times(0,T)) \quad \textit{and} \quad c \in C([0,T]; L^\infty(\Omega)),$$
$$f, g_j \in C([0,T]; L^q),$$

then there exists a positive constant τ_0, independent of τ and h, such that when $\tau < \tau_0$ the solutions of (2.1) *and* (2.2) *satisfy* (1.14) *and* (1.17)–(1.18) *for $1 < p, q < \infty$, and*

$$(2.3)\qquad \|P_h\vec{U} - \vec{U}_h\|_{\ell^p(L^q)} \le C(\|P_h\vec{U} - \vec{R}_h\vec{U}\|_{\ell^p(L^q)} + \|P_h U^0 - U_h^0\|_{L^q}).$$

Moreover, if $\partial_{tt}u \in L^p(0,T;\widetilde{W}^{-1,q})$ then

$$\begin{aligned}(2.4)\qquad \|P_h\vec{u} - \vec{U}_h\|_{\ell^p(L^q)} &\le C\|\partial_{tt}u\|_{L^p(0,T;\widetilde{W}^{-1,q})}\tau\\ &\quad + C(\|P_h\vec{u} - \vec{R}_h\vec{U}\|_{\ell^p(L^q)} + \|P_h u^0 - U_h^0\|_{L^q}).\end{aligned}$$

Remark 2.1. (i) Since we have only assumed $f, g_j \in C([0,T]; L^q)$, the PDEs (1.1) and (2.2) should be viewed in the variational sense to avoid defining the traces of $\sum_{i,j=1}^N a_{ij}\partial_j u\, n_i$ and g_j in the boundary conditions.

(ii) Since the generic constants C in the estimates (1.14), (1.17)–(1.18), and (2.3)–(2.4) are independent of f and g_j, we only need to prove Theorem 2.1 for smooth $f, g_j \in C([0,T];C_0^\infty(\Omega))$. For general $f, g_j \in C([0,T];L^q)$, we can approximate f and g_j by a sequence of smooth functions $f_m \in C([0,T];C_0^\infty(\Omega))$ and $g_{m,j} \in C([0,T];C_0^\infty(\Omega))$, respectively, and by taking the limit $m \to \infty$ in the corresponding estimates with f_m and $g_{m,j}$, we see that these estimates also hold for f and g_j. Hence, without loss of generality, we assume that f and g_j are smooth functions in $C([0,T];C_0^\infty(\Omega))$ in the rest of this paper.

(iii) Theorem 2.1 is based on a commonly used approximation with classical finite element methods in the spatial direction and the backward Euler scheme in the time direction. Due to the nature of the finite difference scheme, the solution and the source term f should be well defined at each time step t_n and $\partial_{tt}u \in L^p(0,T;\widetilde{W}^{-1,q})$ as usual. Of course, if the coefficients a_{ij}, c and the data f, $\mathbf{g}$, and u^0 are sufficiently smooth, then $\partial_{tt}u \in L^p(0,T;\widetilde{W}^{-1,q})$. It is noted that similar stability estimates to (2.4) have been proved by Chrysafinos and Walkington [7] in the energy norms for discontinuous Galerkin time stepping schemes, with nonsymmetric and time-dependent coefficients a_{ij}. Also a class of discontinuous Galerkin time stepping schemes has been studied by Leykekhman and Vexler [18], where the stability estimate in the $L^p(0,T;L^q)$ norm has been proved for the general case $1 \le p, q \le \infty$, which has important applications in optimal control problems. It is noteworthy that for time-dependent coefficients $a_{ij}(x,t)$ the piecewise constant discontinuous Galerkin time discretization is not equivalent to the backward Euler method.

Before we present our proof, some further notations and a lemma are introduced below. We denote by $a(x,t_n) = [a_{ij}(x,t_n)]_{N\times N}$ the coefficient matrix at the time level $t = t_n$, and define the operators

$$\begin{aligned} &A(t_n) : H^1 \to (H^1)', && A_h(t_n) : S_h \to S_h, \\ &R_h(t_n) : H^1 \to S_h, && P_h : L^2 \to S_h, \\ &\overline{\nabla}\cdot : (H^1)^N \to (H^1)', && \overline{\nabla}_h\cdot : (H^1)^N \to S_h, \end{aligned}$$

by

$$\begin{aligned} \big(A(t_n)\phi, v\big) &= \big(a(\cdot,t_n)\nabla\phi, \nabla v\big) + \big(c(\cdot,t_n)\phi, v\big) \text{ for all } \phi, v \in H^1, \\ \big(A_h(t_n)\phi, v\big) &= \big(a(\cdot,t_n)\nabla\phi, \nabla v\big) + \big(c(\cdot,t_n)\phi, v\big) \text{ for all } \phi, v \in S_h, \\ \big(A_h(t_n)R_h(t_n)w, v\big) &= \big(A(t_n)w, v\big) \text{ for all } w \in H^1 \text{ and } v \in S_h, \\ \big(P_h\phi, v\big) &= \big(\phi, v\big) \qquad \text{for all } \phi \in L^2 \text{ and } v \in S_h, \\ \big(\overline{\nabla}\cdot\vec{w}, v\big) &= -\big(\vec{w}, \nabla v\big) \qquad \text{for all } \vec{w} \in (H^1)^N \text{ and } v \in H^1, \\ \big(\overline{\nabla}_h\cdot\vec{w}, v\big) &= -\big(\vec{w}, \nabla v\big) \qquad \text{for all } \vec{w} \in (H^1)^N \text{ and } v \in S_h. \end{aligned}$$

It is well known (cf. [12, Theorem 2.4.2.7]) that for any $\varphi \in L^q$, $1 < q < \infty$, there exists $v \in W^{2,q}$ such that $A(t)v = \varphi$ and

$$\|v\|_{W^{2,q}} \le C\|\varphi\|_{L^q} = C\|A(t)v\|_{L^q}, \tag{2.5}$$

where the constant C is independent of $t \in [0,T]$.

Clearly, P_h is the L^2 projection operator onto S_h, and $R_h(t_n)$ is the Ritz projection operator associated with the elliptic operator $A(t_n)$, satisfying the following approximation properties for $l \le k$:

$$\|P_h\phi - \phi\|_{W^{l,q}} \le C\|\phi\|_{W^{k,q}} h^{k-l}, \qquad 1 \le q \le \infty,\ l = 0,1,\ k = 0,1,2, \tag{2.6a}$$

$$\|R_h(t_n)\phi - \phi\|_{W^{l,q}} \le C_p\|\phi\|_{W^{k,q}} h^{k-l}, \qquad 1 < q < \infty,\ l = 0,1,\ k = 1,2, \tag{2.6b}$$

where (2.6a) is a consequence of [34, Lemma 7.2]. In the case of Dirichlet boundary condition, a proof of (2.6b) can be found in [6, equations (8.5.4)–(8.5.5), with $\mu = \infty$]. In the case of Neumann boundary condition, (2.6b) can be proved similarly by using [11, Theorem A.3]. Besides the two inequalities above, the following estimates are useful in our proof.

LEMMA 2.1. *Let* $D_q(t_n) := \{v \in W^{2,q} : a(\cdot,t_n)\nabla v \cdot \mathbf{n} = 0$ *on* $\partial\Omega\}$, $1 < q < \infty$. *Then we have*

$$\|A_h(t_n)v_h\|_{L^q} \le Ch^{-1}\|v_h\|_{W^{1,q}} \qquad \forall\, v_h \in S_h, \tag{2.7}$$

$$\|A_h(t_n)P_h\psi\|_{L^q} \le C\|\psi\|_{W^{2,q}} \qquad \forall\, \psi \in D_q(t_n), \tag{2.8}$$

$$\|v_h\|_{W^{1,q}} \le C\|A_h(t_n)v_h\|_{L^q} \qquad \forall\, v_h \in S_h, \tag{2.9}$$

$$\|A_h(t_n)^{-1}P_h\varphi\|_{L^q} \le C\|\varphi\|_{\widetilde{W}^{-1,q}} \qquad \forall\, \varphi \in \widetilde{W}^{-1,q}. \tag{2.10}$$

Proof. In fact, (2.7) follows from the following estimate via duality: for $\eta_h \in S_h$,

$$\begin{aligned}
|(A_h(t_n)v_h,\eta_h)| &= |(a(\cdot,t_n)\nabla v_h,\nabla\eta_h) + (c(\cdot,t_n)v_h,\eta_h)| \\
&\le C\|v_h\|_{W^{1,q}}\|\eta_h\|_{W^{1,q'}} \\
&\le Ch^{-1}\|v_h\|_{W^{1,q}}\|\eta_h\|_{L^{q'}} \quad \text{by the inverse inequality.}
\end{aligned}$$

Similarly, (2.8) follows from the following estimate: for $\eta_h \in S_h$,

$$\begin{aligned}
|(A_h(t_n)P_h\psi,\eta_h)| &= |(a(\cdot,t_n)\nabla P_h\psi,\nabla\eta_h) + (c(\cdot,t_n)P_h\psi,\eta_h)| \\
&\le |(a(\cdot,t_n)\nabla(P_h\psi-\psi),\nabla\eta_h) + (c(\cdot,t_n)(P_h\psi-\psi),\eta_h)| \\
&\quad + |(a(\cdot,t_n)\nabla\psi,\nabla\eta_h) + (c(\cdot,t_n)\psi,\eta_h)| \\
&= |(a(\cdot,t_n)\nabla(P_h\psi-\psi),\nabla\eta_h) + (c(\cdot,t_n)(P_h\psi-\psi),\eta_h)| \\
&\quad + |(A(t_n)\psi,\eta_h)| \\
&\le C\|P_h\psi-\psi\|_{W^{1,q}}\|\eta_h\|_{W^{1,q'}} + C\|\psi\|_{W^{2,q}}\|\eta_h\|_{L^{q'}} \\
&\le C\|\psi\|_{W^{2,q}}h\|\eta_h\|_{W^{1,q'}} + C\|\psi\|_{W^{2,q}}\|\eta_h\|_{L^{q'}} \\
&\le C\|\psi\|_{W^{2,q}}\|\eta_h\|_{L^{q'}} \quad \text{by the inverse inequality.}
\end{aligned}$$

To prove (2.9), we define $v \in W^{2,q}$ as the solution of $A(t_n)v = A_h(t_n)v_h$, which satisfies the following estimate in view of (2.5):

$$\|v\|_{W^{2,q}} \le C\|A(t_n)v\|_{L^q} = C\|A_h(t_n)v_h\|_{L^q}. \tag{2.11}$$

Then $(A(t_n)v,\eta_h) = (A_h(t_n)v_h,\eta_h)$ yields

$$(a(\cdot,t_n)\nabla(v-v_h),\nabla\eta_h) + (c(\cdot,t_n)(v-v_h),\eta_h) = 0 \quad \forall\,\eta_h \in S_h,$$

which means that v_h is the Ritz projection of v onto S_h. Then (2.6b) implies

$$\begin{aligned}\|I_h v - v_h\|_{L^q} &\le \|I_h v - v\|_{L^q} + \|v - v_h\|_{L^q}\\ &\le Ch^2\|v\|_{W^{2,q}}\\ &\le Ch^2\|A_h(t_n)v_h\|_{L^q}.\end{aligned}$$

By using the inverse inequality, we obtain

$$\|I_h v - v_h\|_{W^{1,q}} \le Ch^{-1}\|I_h v - v_h\|_{L^q} \le Ch\|A_h(t_n)v_h\|_{L^q},$$

which further implies

$$\begin{aligned}\|v - v_h\|_{W^{1,q}} &\le \|v - I_h v\|_{W^{1,q}} + \|I_h v - v_h\|_{W^{1,q}}\\ &\le Ch\|v\|_{W^{2,q}} + Ch\|A_h(t_n)v_h\|_{L^q}\\ &\le Ch\|A_h(t_n)v_h\|_{L^q}.\end{aligned}$$

The last inequality and (2.11) yield (2.9).

To prove (2.10), we simply note that for $\eta \in S_h$

$$\begin{aligned}&|(A_h(t_n)^{-1}P_h\varphi, \eta)|\\ &\quad = |(A_h(t_n)^{-1}P_h\varphi, P_h\eta)| && \text{(by the definition of the } L^2 \text{ projection } P_h)\\ &\quad = |(\varphi, A_h(t_n)^{-1}P_h\eta)| && \text{(by the self-adjointness of } A_h(t_n)^{-1})\\ &\quad \le \|\varphi\|_{\widetilde{W}^{-1,q}}\|A_h(t_n)^{-1}P_h\eta\|_{W^{1,q'}} && \text{(by the duality between } \widetilde{W}^{-1,q} \text{ and } W^{1,q'})\\ &\quad \le \|\varphi\|_{W^{-1,q}}\|P_h\eta\|_{L^{q'}} && \text{(by using (2.9))}\\ &\quad \le \|\varphi\|_{W^{-1,q}}\|\eta\|_{L^{q'}}, && \text{(by using (2.6a) with } k = l = 0)\end{aligned}$$

which implies (2.10) via duality. ∎

For any $\vec{w} = (w^1, \dots, w^M)$, $\vec{v} = (v^1, \dots, v^M)$, $\vec{w}_h = (w_h^1, \dots, w_h^M)$ with $w^n \in H^1$, $v^n \in L^2$, and $w_h^n \in S_h$, we define

$$\begin{aligned}&\vec{A}\vec{w} = (A(t_1)w^1, \dots, A(t_M)w^M), && \vec{A}_h\vec{w}_h = (A_h(t_1)w_h^1, \dots, A_h(t_M)w_h^M),\\ &P_h\vec{v} = (P_h v^1, \dots, P_h v^M), && \vec{R}_h\vec{w}_h = (R_h(t_1)w_h^1, \dots, R_h(t_M)w_h^M).\end{aligned}$$

In the rest of this paper, we let $C_{p_1,\dots,p_k}$ denote a generic positive constant which may depend on the parameters $p_1, \dots, p_k$, as well as the domain Ω and the quantities T, K, $\|a_{ij}\|_{W^{1,\infty}(\Omega\times(0,T))}$, and $\|c\|_{L^\infty(\Omega\times(0,T))}$, but will be independent of the time-step size τ and the spatial mesh size h.

3. Equations with time-independent coefficients. In this section, we study the ℓ^p-stability estimates for the time-discrete system (2.2) and the fully discrete system (2.1) by assuming that

$$a_{ij} = a_{ij}(x) \in W^{1,\infty} \quad \text{and} \quad c = 1\ . \tag{3.1}$$

In this autonomous case, we simply denote $A :\equiv A(t) = -\nabla \cdot (a(x)\nabla) + 1$ and $D_q :\equiv D_q(t)$ (see Lemma 2.1). We need the following lemma, which was proved in [2, Theorem 5.5 and Remark 5.2] (we refer to [16] for the notation of R-boundedness of a family of operators); also see [15, Theorem 3.1].

LEMMA 3.1 (abstract maximal ℓ^p-regularity). *Let $1 < q < \infty$ and let X be either L^q or L^q_h. Let $B : X \to X$ be a linear operator such that the resolvent operators*

$$z(z+B)^{-1} \tag{3.2}$$

are R-bounded for $\mathrm{Re}(z) \geq 0$, *and denote its R-bound by C_R. Then the solution of the equation*

$$\begin{cases} D_\tau W^n + BW^n = f^n \quad \text{for } n = 1, 2, \ldots, \\ W^0 = 0, \end{cases} \tag{3.3}$$

possesses the discrete maximal ℓ^p-regularity:

$$\|D_\tau \vec{W}\|_{\ell^p(X)} + \|B\vec{W}\|_{\ell^p(X)} \leq CC_R\|\vec{f}\|_{\ell^p(X)}, \ 1 < p < \infty, \tag{3.4}$$

where C is independent of τ and h.

Let the solution of

$$\begin{cases} \partial_t u + Au = 0, \\ u(0) = u^0, \end{cases} \tag{3.5}$$

be denoted by $u(t) = E(t)u^0$, and let $-\overline{A}$ be the generator of the semigroups $\{E(t)\}_{t>0}$. Then the domain of $\overline{A}$ is D_q (see Lemma 2.1). Moreover, $\overline{A}w = Aw$ for any $w \in D_q$ and $\overline{A}^{-1}v = A^{-1}v$ for any $v \in L^q$. In other words, $A : W^{1,q} \to \widetilde{W}^{-1,q}$ is an extension of the operator $\overline{A} : D_q \to L^q$. In the following, we simply use the notation A to represent $\overline{A}$. Similarly, the solution of

$$\begin{cases} \partial_t u_h + A_h u_h = 0, \\ u_h(0) = u_h^0, \end{cases} \tag{3.6}$$

can be expressed as $u_h(t) = E_h(t)u_h^0$, and the generator of the semigroups $\{E_h(t)\}_{t>0}$ coincides with A_h. In order to prove the $\widetilde{W}^{-1,q}$ estimate of numerical solutions, we need the following lemma on boundedness of the Riesz transform.

LEMMA 3.2 (boundedness of the Riesz transform). *The Riesz transform $\nabla A^{-1/2}$ and its dual operator $A^{-1/2}\overline{\nabla}\cdot$ are both bounded on L^q for $1 < q < \infty$.*

Proof. Since $C_0^\infty(\Omega)$ is dense in L^q, we only need to consider the action of the operators $\nabla A^{-1/2}$ and $A^{-1/2}\overline{\nabla}\cdot$ on the functions of $C_0^\infty(\Omega)$.

Let $A_0 = -\nabla \cdot (a(x)\nabla)$ so that $A = 1 + A_0$. Then the following estimate holds for all $1 < q < \infty$ (cf. [4, Theorem 4, condition 1]):

$$\|A_0^{1/2}v\|_{L^q} \leq C_q\|\nabla v\|_{L^q} \quad \forall\, v \in \mathring{W}^{1,q}(\Omega) := \left\{ w \in W^{1,q}(\Omega) : \int_\Omega w(x)\mathrm{d}x = 0 \right\}.$$

Since the operator $\nabla A_0^{-1}\overline{\nabla}\cdot$ is bounded on L^q for all $1 < q < \infty$ (cf. [3, Theorem 1]), it follows that, for $\mathbf{g} \in C_0^\infty(\Omega)^d$,

$$\|A_0^{-1/2}\overline{\nabla} \cdot \mathbf{g}\|_{L^q} = \|A_0^{1/2}A_0^{-1}\overline{\nabla} \cdot \mathbf{g}\|_{L^q} \leq C\|\nabla A_0^{-1}\overline{\nabla} \cdot \mathbf{g}\|_{L^q} \leq C\|\mathbf{g}\|_{L^q}.$$

Then we have

$$\|A^{-1/2}\overline{\nabla} \cdot \mathbf{g}\|_{L^q} = \|[(1+A_0)^{-1}A_0]^{1/2}A_0^{-1/2}\overline{\nabla} \cdot \mathbf{g}\|_{L^q} \leq C\|A_0^{-1/2}\overline{\nabla} \cdot \mathbf{g}\|_{L^q} \leq C\|\mathbf{g}\|_{L^q},$$

where we have used the boundedness of the symmetric positive operator $(1+A_0)^{-1}A_0$ in the last inequality. This last inequality shows the boundedness of the operator $A^{-1/2}\overline{\nabla}\cdot$ on L^q for all $1 < q < \infty$. Since the Riesz transform $\nabla A^{-1/2}$ is the dual of $A^{-1/2}\overline{\nabla}\cdot$, it follows that $\nabla A^{-1/2}$ is also bounded on L^q. ∎

Lemma 3.3 (time-discrete PDEs with the Neumann boundary condition). *Under the assumptions* (1.2) *and* (3.1), *the solution of*

$$\begin{cases} D_\tau U^n + AU^n = f^n - \overline{\nabla}\cdot \mathbf{g}^n, & n=1,\dots,M,\\ U^0 = 0, \end{cases} \tag{3.7}$$

satisfies

$$\|D_\tau \vec{U}\|_{\ell^p(\widetilde{W}^{-1,q})} + \|\vec{U}\|_{\ell^p(W^{1,q})} \le C\|\vec{\mathbf{g}}\|_{\ell^p(L^q)}, \quad 1<p,q<\infty, \quad \text{if } f\equiv 0, \tag{3.8}$$

$$\|D_\tau \vec{U}\|_{\ell^p(L^q)} + \|\vec{U}\|_{\ell^p(W^{2,q})} \le C\|\vec{f}\|_{\ell^p(L^q)}, \quad 1<p,q<\infty, \quad \text{if } \mathbf{g}\equiv 0. \tag{3.9}$$

Proof. It is well known that the continuous problem

$$\begin{cases} \partial_t u + Au = f,\\ u(0) = 0, \end{cases} \tag{3.10}$$

possesses the maximal-L^p regularity [35, Corollary 4.d], which implies that the collection $\{z(z+A)^{-1} : \mathrm{Re}(z)\ge 0\}$ is R-bounded on L^q [36, Theorem 4.2], for any given $1<q<\infty$. When $\mathbf{g}^n\equiv 0$ we simply apply Lemma 3.1 with the elliptic regularity estimate

$$\|\vec{U}\|_{\ell^p(W^{2,q})} \le C_q\|A\vec{U}\|_{\ell^p(L^q)} \quad \forall\, 1<q<\infty. \tag{3.11}$$

This completes the proof of (3.9).

To prove (3.8), we denote $E_n = (1+\tau A)^{-n}$. The inequality (3.9) implies that the map from $\vec{f}$ to $A\vec{U}$ given by the formula

$$AU^n = \sum_{m=1}^{n} \tau A E_{n-m+1} f^n$$

is bounded in $\ell^p(L^q)$. Since the fractional power operator $A^{-1/2}$ commutes with A, when $f\equiv 0$ and $\mathbf{g}\neq 0$ the solution of (3.7) is given by

$$\nabla U^n = \nabla A^{-1/2} W^n,$$

where

$$W^n = -\sum_{m=1}^{n} \tau A E_{n-m+1} A^{-1/2}\overline{\nabla}\cdot \mathbf{g}^m.$$

The inequality (3.9) implies

$$\|\vec{W}\|_{\ell^p(L^q)} \le C\|A^{-1/2}\overline{\nabla}\cdot\mathbf{g}\|_{\ell^p(L^q)}. \tag{3.12}$$

Since the Riesz transform $\nabla A^{-1/2}$ and its dual operator $A^{-1/2}\overline{\nabla}\cdot$ are bounded on $L^q(\Omega)$ for all $1<q<\infty$ (cf. Lemma 3.2), it follows that

$$\|\nabla\vec{U}\|_{\ell^p(L^q)} \le C\|\vec{W}\|_{\ell^p(L^q)} \le C\|A^{-1/2}\overline{\nabla}\cdot\mathbf{g}\|_{\ell^p(L^q)} \le C\|\vec{\mathbf{g}}\|_{\ell^p(L^q)}. \tag{3.13}$$

By using the last inequality, from (3.7) we can further derive

$$\|D_\tau\vec{U}\|_{\ell^p(\widetilde{W}^{-1,q})} \le C\|\vec{\mathbf{g}}\|_{\ell^p(L^q)}, \quad 1<p,q<\infty, \quad \text{if } f\equiv 0. \tag{3.14}$$

The proof of Lemma 3.3 is completed. ∎

Similar results also hold for time-discrete parabolic equations with the Dirichlet boundary conditions, and the proof is similar (thus omitted).

LEMMA 3.4 (time-discrete PDEs with the Dirichlet boundary condition). *Under the assumptions* (1.2) *and* (3.1), *the solution of the equation*

$$\begin{cases} D_\tau U^n - \sum_{i,j=1}^{N} \partial_i\big(a_{ij}\partial_j U^n\big) + cU^n = f^n - \sum_{j=1}^{N} \partial_j g_j^n & \text{in } \Omega, \\ U^n = 0 & \text{on } \partial\Omega \\ U^0 = u^0, \end{cases} \tag{3.15}$$

satisfies

$$\|D_\tau \vec{U}\|_{\ell^p(W^{-1,q})} + \|\vec{U}\|_{\ell^p(W^{1,q})} \le C\|\vec{\mathbf{g}}\|_{\ell^p(L^q)}, \quad 1 < p, q < \infty, \quad \text{if } f \equiv 0, \tag{3.16}$$

$$\|D_\tau \vec{U}\|_{\ell^p(L^q)} + \|\vec{U}\|_{\ell^p(W^{2,q})} \le C\|\vec{f}\|_{\ell^p(L^q)}, \quad 1 < p, q < \infty, \quad \text{if } \mathbf{g} \equiv 0. \tag{3.17}$$

For fully discrete finite element solutions of parabolic equations, we prove the following result. Here we drop the assumption $c = 1$.

LEMMA 3.5 (fully discrete finite element solutions). *Under the assumptions* (1.2) *and* (3.1), *the solution* $\vec{U}_h = (U_h^n)_{n=1}^N$ *of the equation*

$$\begin{cases} D_\tau U_h^n + A_h U_h^n = P_h f^n - \overline{\nabla}_h \cdot \mathbf{g}^n, \quad n = 1, \ldots, M, \\ U_h^0 = 0\,, \end{cases} \tag{3.18}$$

satisfies that, for $1 < p, q < \infty$,

$$\|\vec{U}_h\|_{\ell^p(\widetilde{W}^{-1,q})} + \|\vec{U}_h\|_{\ell^p(W^{1,q})} \le C\|\vec{\mathbf{g}}\|_{\ell^p(L^q)} \qquad \text{if } \vec{f} \equiv 0, \tag{3.19}$$

$$\|D_\tau \vec{U}_h\|_{\ell^p(L^q)} + \|A_h \vec{U}_h\|_{\ell^p(L^q)} \le C\|\vec{f}_h\|_{\ell^p(L^q)} \qquad \text{if } \vec{\mathbf{g}} \equiv 0. \tag{3.20}$$

Proof. [20, text between (4.10) and (4.11)] implies that the operators $\{z(z + A_h)^{-1} : \mathrm{Re}(z) \ge 0\}$ is R-bounded, and therefore Lemma 3.1 implies (3.20).

To prove (3.19), we introduce the time-discrete PDEs

$$\begin{cases} D_\tau U^n + AU^n = f^n - \overline{\nabla} \cdot \mathbf{g}^n, \quad n = 1, \ldots, M, \\ U^0 = 0, \end{cases} \tag{3.21}$$

and we define $e_h^n = P_h U^n - U_h^n$. Integrating (3.21) against a finite element function v_h gives

$$\begin{cases} (D_\tau P_h U^n, v_h) + (a\nabla U^n, \nabla v_h) = (f^n, v_h) + (\mathbf{g}^n, \nabla v_h), \quad n = 1, \ldots, M, \\ U^0 = 0, \end{cases}$$

and integrating (3.18) against v_h gives

$$\begin{cases} (D_\tau U_h^n, v_h) + (a\nabla U_h^n, \nabla v_h) = (f^n, v_h) + (\mathbf{g}^n, \nabla v_h), \quad n = 1, \ldots, M, \\ U_h^0 = 0\,. \end{cases}$$

The difference of the two equations above yields

$$\begin{cases}(D_\tau e_h^n, v_h) + (a\nabla e_h^n, \nabla v_h) = (a\nabla(P_hU^n - R_hU^n), \nabla v_h) \\ \qquad\qquad\qquad\qquad\qquad\qquad\quad = (A_h(P_hU^n - R_hU^n), v_h), \\ e_h^0 = 0\,. \end{cases}$$

Multiplying the last equation by the operator A_h^{-1}, we obtain

$$\begin{cases} D_\tau(A_h^{-1}e_h^n) + A_h(A_h^{-1}e_h^n) = P_hU^n - R_hU^n, \quad n = 1, \ldots, M, \\ A_h^{-1}e_h^0 = 0. \end{cases} \tag{3.22}$$

By applying (3.20) we derive that

$$\|\vec{e}_h\|_{\ell^p(L^q)} = \|A_h(A_h^{-1}e_h^n)\|_{\ell^p(L^q)} \le C\|P_h\vec{U} - R_h\vec{U}\|_{\ell^p(L^q)} \le C\|\vec{U}\|_{\ell^p(W^{1,q})}h,$$

which together with the inverse inequality gives

$$\|\vec{e}_h\|_{\ell^p(W^{1,q})} \le Ch^{-1}\|\vec{e}_h\|_{\ell^p(L^q)} \le C\|\vec{U}\|_{\ell^p(W^{1,q})}.$$

Therefore, we have

$$\|\vec{U}_h\|_{\ell^p(W^{1,q})} \le \|P_h\vec{U}\|_{\ell^p(W^{1,q})} + \|\vec{e}_h\|_{\ell^p(W^{1,q})} \le C\|\vec{U}\|_{\ell^p(W^{1,q})} \le C\|\vec{\mathbf{g}}\|_{\ell^p(W^{1,q})},$$

where (3.8) is used for deriving the last inequality. ∎

4. Proof of Theorem 2.1. In this section, we consider the parabolic system with time-dependent coefficients $a_{ij} = a_{ij}(x,t)$ and $c = c(x,t)$ by applying Lemmas 3.3–3.5 with a perturbation argument.

4.1. Time-discrete PDEs. First, we assume $U^0 = 0$ and prove (1.14) and the following $\ell^p(W^{1,q})$ estimate

$$\|D_\tau\vec{U}\|_{\ell^p(\widetilde{W}^{-1,q})} + \|\vec{U}\|_{\ell^p(W^{1,q})} \le C\left(\|\vec{f}\|_{\ell^p(L^q)} + \sum_{j=1}^N \|\vec{g}_j\|_{\ell^p(L^q)}\right), \tag{4.1}$$

for the solution of (2.2) by applying Lemma 3.3, under the extra assumption $c = 1$. These extra conditions will be dropped later.

For $1 \le n \le m \le M$, we rewrite (2.2) as

$$\begin{cases} D_\tau U^n - \sum_{i,j=1}^N \partial_i\big(a_{ij}(\cdot,t_m)\partial_jU^n\big) + U^n \\ = f^n - \sum_{j=1}^N \partial_j g_j^n - \sum_{i,j=1}^N \partial_i\big((a_{ij}(\cdot,t_m) - a_{ij}(\cdot,t_n))\partial_jU^n\big) \quad \text{in } \Omega, \\ -\sum_{i,j=1}^N a_{ij}(\cdot,t_m)\partial_jU^n n_i \\ = -\sum_{j=1}^N g_j^n n_j - \sum_{i,j=1}^N (a_{ij}(\cdot,t_m) - a_{ij}(\cdot,t_n))\partial_jU^n n_i \qquad \text{on } \partial\Omega, \\ U^0 = u^0. \end{cases} \tag{4.2}$$

Since the operator on the left-hand side of (4.2) is independent of n, we can apply (3.8)–(3.9) in the time interval $[0, t_m]$ and obtain

$$\begin{aligned} &\|D_\tau \vec{U}\|_{\ell^p(0,t_m;\widetilde{W}^{-1,q})} + \|\vec{U}\|_{\ell^p(0,t_m;W^{1,q})} \\ &\quad \le C \left(\|\vec{f}\|_{\ell^p(0,t_m;L^q)} + \sum_{j=1}^{N} \|\vec{g}_j\|_{\ell^p(0,t_m;L^q)} \right) \\ &\qquad + C\|((a(\cdot,t_m) - a(\cdot,t_n))\nabla U^n)_{n=1}^m\|_{\ell^p(0,t_m;L^q)} \\ &\quad \le C \left(\|\vec{f}\|_{\ell^p(0,t_m;L^q)} + \sum_{j=1}^{N} \|\vec{g}_j\|_{\ell^p(0,t_m;L^q)} \right) \\ &\qquad + C\|(|t_m - t_n|\vec{U}^n)_{n=1}^m\|_{\ell^p(0,t_m;W^{1,q})}. \end{aligned} \tag{4.3}$$

If we denote

$$\begin{aligned} E^m &:= \|\vec{U}\|^p_{\ell^p(0,t_m;W^{1,q})} = \tau \sum_{n=1}^{m} \|U^n\|^p_{W^{1,q}}, \quad m = 1, \dots, M, \\ E^0 &:= 0, \\ F^m &:= \left(\|\vec{f}\|_{\ell^p(0,t_m;W^{-1,q})} + \sum_{j=1}^{N} \|\vec{g}_j\|_{\ell^p(0,t_m;L^q)} \right)^p, \end{aligned}$$

then (4.3) can be rewritten as

$$\begin{aligned} E^m &\le CF^m + C\|(|t_m - t_n|U^n)_{n=1}^m\|^p_{\ell^p(0,t_m;W^{1,q})} \\ &= CF^m + C\tau \sum_{n=1}^{m} |t_m - t_n|^p \|U^n\|^p_{W^{1,q}} \\ &= CF^m + C\tau \sum_{n=1}^{m} |t_m - t_n|^p \frac{E^n - E^{n-1}}{\tau} \\ &= CF^m + C\tau \sum_{n=1}^{m-1} \frac{|t_m - t_{n+1}|^p - |t_m - t_n|^p}{\tau} E^n \quad \text{(discrete integration by parts)} \\ &\le CF^m + C\tau \sum_{n=1}^{m-1} E^n, \end{aligned}$$

which holds for all $1 \le m \le M$. Applying Gronwall's inequality yields

$$E^M \le CF^M,$$

and by substituting this inequality into (4.3) we obtain

$$\|D_\tau \vec{U}\|_{\ell^p(\widetilde{W}^{-1,q})} + \|\vec{U}\|_{\ell^p(W^{1,q})} \leq C\left(\|\vec{f}\|_{\ell^p(L^q)} + \sum_{j=1}^N \|\vec{g}_j\|_{\ell^p(L^q)}\right). \tag{4.4}$$

This proves (4.1) under the extra condition $c = 1$.

Second, we drop the assumption $c = 1$ by rewriting (2.2) as

$$\begin{cases} D_\tau U^n - \sum_{i,j=1}^N \partial_i\big(a_{ij}(\cdot,t_n)\partial_j U^n\big) + U^n = (1-c(\cdot,t_n))U^n + f^n - \sum_{j=1}^N \partial_j g_j^n & \text{in } \Omega, \\ \sum_{i,j=1}^N a_{ij}(\cdot,t_n)\partial_j U^n n_i = \sum_{j=1}^N g_j^n n_j & \text{on } \partial\Omega, \\ U^0 = u^0. \end{cases}$$

By applying (4.4) in the subinterval $[0, t_m]$, we have

$$\begin{aligned} &\|D_\tau \vec{U}\|_{\ell^p(0,t_m;\widetilde{W}^{-1,q})} + \|\vec{U}\|_{\ell^p(0,t_m;W^{1,q})} \\ &\quad\leq C\left(\|\vec{f}\|_{\ell^p(0,t_m;L^q)} + \sum_{j=1}^N \|\vec{g}_j\|_{\ell^p(0,t_m;L^q)}\right) + C\|\vec{U}\|_{\ell^p(0,t_m;L^q)} \\ &\quad\leq C\left(\|\vec{f}\|_{\ell^p(0,t_m;L^q)} + \sum_{j=1}^N \|\vec{g}_j\|_{\ell^p(0,t_m;L^q)}\right) \\ &\qquad + C\|\vec{U}\|^{\frac{1}{2}}_{\ell^p(0,t_m;\widetilde{W}^{-1,q})}\|\vec{U}\|^{\frac{1}{2}}_{\ell^p(0,t_m;W^{1,q})} \quad \text{(the interpolation inequality)} \\ &\quad\leq C\left(\|\vec{f}\|_{\ell^p(0,t_m;L^q)} + \sum_{j=1}^N \|\vec{g}_j\|_{\ell^p(0,t_m;L^q)}\right) \\ &\qquad + C\|\vec{U}\|_{\ell^p(0,t_m;\widetilde{W}^{-1,q})} + \frac{1}{2}\|\vec{U}\|_{\ell^p(0,t_m;W^{1,q})}, \end{aligned} \tag{4.5}$$

which reduces to

$$\begin{aligned} &\|D_\tau \vec{U}\|_{\ell^p(0,t_m;\widetilde{W}^{-1,q})} + \|\vec{U}\|_{\ell^p(0,t_m;W^{1,q})} \\ &\quad\leq C\left(\|\vec{f}\|_{\ell^p(0,t_m;L^q)} + \sum_{j=1}^N \|\vec{g}_j\|_{\ell^p(0,t_m;L^q)}\right) + C\|\vec{U}\|_{\ell^p(0,t_m;\widetilde{W}^{-1,q})}. \end{aligned} \tag{4.6}$$

Since

$$\|U^m\|_{\widetilde{W}^{-1,q}} = \left\|U^0 + \tau\sum_{n=1}^m D_\tau U^n\right\|_{\widetilde{W}^{-1,q}} \leq \|U^0\|_{\widetilde{W}^{-1,q}} + \sum_{n=1}^m \tau\|D_\tau U^n\|_{\widetilde{W}^{-1,q}},$$

it follows that

$$
\begin{aligned}
&\|\vec{U}\|_{\ell^\infty(0,t_m;\widetilde{W}^{-1,q})} \\
&\quad\le \|U^0\|_{\widetilde{W}^{-1,q}} + \sum_{n=1}^{m} \tau \|D_\tau U^n\|_{\widetilde{W}^{-1,q}} \\
&\quad= \|U^0\|_{\widetilde{W}^{-1,q}} + \|D_\tau \vec{U}\|_{\ell^1(0,t_m;\widetilde{W}^{-1,q})} \\
&\quad\le \|U^0\|_{\widetilde{W}^{-1,q}} + T^{\frac{1}{p'}} \|D_\tau \vec{U}\|_{\ell^p(0,t_m;\widetilde{W}^{-1,q})} \\
&\quad\le \|U^0\|_{\widetilde{W}^{-1,q}} + C\|\vec{U}\|_{\ell^p(0,t_m;\widetilde{W}^{-1,q})} \\
&\qquad\quad + C\left(\|\vec{f}\|_{\ell^p(0,t_m;L^q)} + \sum_{j=1}^{N} \|\vec{g}_j\|_{\ell^p(0,t_m;L^q)} \right) \\
&\quad\le \|U^0\|_{\widetilde{W}^{-1,q}} + C\|\vec{U}\|^{\frac{1}{p}}_{\ell^1(0,t_m;\widetilde{W}^{-1,q})} \|\vec{U}\|^{1-\frac{1}{p}}_{\ell^\infty(0,t_m;\widetilde{W}^{-1,q})} \\
&\qquad\quad + C\left(\|\vec{f}\|_{\ell^p(0,t_m;L^q)} + \sum_{j=1}^{N} \|\vec{g}_j\|_{\ell^p(0,t_m;L^q)} \right) \\
&\quad\le \|U^0\|_{\widetilde{W}^{-1,q}} + C\|\vec{U}\|_{\ell^1(0,t_m;\widetilde{W}^{-1,q})} + \frac{1}{2}\|\vec{U}\|_{\ell^\infty(0,t_m;\widetilde{W}^{-1,q})} \\
&\qquad\quad + C\left(\|\vec{f}\|_{\ell^p(0,t_m;L^q)} + \sum_{j=1}^{N} \|\vec{g}_j\|_{\ell^p(0,t_m;L^q)} \right),
\end{aligned} \tag{4.7}
$$

which reduces to

$$
\begin{aligned}
\|\vec{U}\|_{\ell^\infty(0,t_m;\widetilde{W}^{-1,q})} &\le \|U^0\|_{\widetilde{W}^{-1,q}} + C\|\vec{U}\|_{\ell^1(0,t_m;\widetilde{W}^{-1,q})} \\
&\qquad + C\left(\|\vec{f}\|_{\ell^p(0,t_m;L^q)} + \sum_{j=1}^{N} \|\vec{g}_j\|_{\ell^p(0,t_m;L^q)} \right),
\end{aligned} \tag{4.8}
$$

which holds for all $1 \le m \le M$. Then Gronwall's inequality implies

$$
\|\vec{U}\|_{\ell^\infty(\widetilde{W}^{-1,q})} \le C\|U^0\|_{\widetilde{W}^{-1,q}} + C\left(\|\vec{f}\|_{\ell^p(L^q)} + \sum_{j=1}^{N} \|\vec{g}_j\|_{\ell^p(L^q)} \right). \tag{4.9}
$$

Substituting the inequality above into the last term of (4.6) and setting $m = M$, we get (4.1) for the Neumann boundary condition.

By using Lemma 3.4, one can also prove (4.1) under the Dirichlet boundary condition in a similar way (the proof is omitted). The corresponding result for the Dirichlet boundary condition is presented below, which is needed to prove (1.14) for the Neumann boundary condition.

LEMMA 4.1 (Dirichlet boundary condition). *The solution of*

$$
\begin{cases}
D_\tau U^n - \displaystyle\sum_{i,j=1}^{N} \partial_i\big(a_{ij}(\cdot,t_n)\partial_j U^n\big) + c(\cdot,t_n)U^n = f^n - \sum_{j=1}^{N} \partial_j g_j^n & \text{in } \Omega, \\
U^n = 0 & \text{on } \partial\Omega, \\
U^0 = u^0
\end{cases} \tag{4.10}
$$

satisfies (4.1) *for* $1 < p, q < \infty$.

Third, we prove (1.14) for the special case that $\vec{U}$ has a compact support in the unit ball and Ω is the upper half-space, i.e., $\Omega = \{(x_1, \ldots, x_N) \in \mathbb{R}^N : x_N > 0\}$. In this case, we have $n_N = -1$ and $n_k = 0$ for $1 \le k \le N-1$. To present estimates for the gradient of the solution, we consider the equation which governs $w_k^n := \partial_k U^n$ with $1 \le k \le N-1$. By differentiating (2.2) with respect to x_k, we obtain (note that $g_j = 0$)

$$\begin{cases} D_\tau w_k^n - \displaystyle\sum_{i,j=1}^N \partial_i\big(a_{ij}(x,t_n)\partial_j w_k^n\big) + w_k^n \\ \qquad = \partial_k(f^n - cU^n + U^n) + \displaystyle\sum_{i,j=1}^N \partial_i\big(\partial_k a_{ij}(x,t_n)\partial_j U^n\big) \quad \text{in } \Omega, \\ -\displaystyle\sum_{i,j=1}^N a_{ij}(x,t_n)\partial_j w_k^n n_i \\ \qquad = (f^n - cU^n + U^n)n_k + \displaystyle\sum_{i,j=1}^N \partial_k a_{ij}(x,t_n)\partial_j U^n n_i \qquad \text{on } \partial\Omega, \\ w_k^0 = 0. \end{cases}$$

Applying (4.1) to the equation above, we get

$$\begin{aligned} \|D_\tau \vec{w}_k\|_{\ell^p(\widetilde{W}^{-1,q})} + \|\vec{w}_k\|_{\ell^p(W^{1,q})} &\le C\|\vec{f} - c\vec{U}^n + U^n\|_{\ell^p(L^q)} + C\sum_{j=1}^N \|\partial_j \vec{U}\|_{\ell^p(L^q)} \\ &\le C\|\vec{f}\|_{\ell^p(L^q)} \qquad \text{(this step uses (4.1) again)}, \end{aligned}$$

which implies

$$\sum_{k=1}^{N-1} \|D_\tau \partial_k \vec{U}\|_{\ell^p(\widetilde{W}^{-1,q})} + \sum_{i=1}^N \sum_{k=1}^{N-1} \|\partial_i \partial_k \vec{U}\|_{\ell^p(L^q)} \le C\|\vec{f}\|_{\ell^p(L^q)}\,. \tag{4.11}$$

Then we define

$$\varphi^n = \sum_{k=1}^{N-1} \frac{a_{Nk}(x,t_n)}{a_{NN}(x,t_n)} \partial_k U^n$$

with

$$D_\tau \varphi^n = \sum_{k=1}^{N-1} \frac{a_{Nk}(x,t_{n-1})}{a_{NN}(x,t_{n-1})} D_\tau \partial_k U^n + \sum_{k=1}^{N-1} D_\tau\left(\frac{a_{Nk}(x,t_n)}{a_{NN}(x,t_n)}\right)\partial_k U^n.$$

The strong ellipticity condition (1.2) implies $a_{NN}(x,t) \ge K^{-1}$ and $a_{Nk}(x,t) \le K$, which together with the Lipschitz continuity $a_{ij} \in W^{1,\infty}(\Omega \times (0,T))$ yields

$$\left|\frac{a_{Nk}(x,t_{n-1})}{a_{NN}(x,t_{n-1})}\right| \le C, \quad \left|\nabla\left(\frac{a_{Nk}(x,t_{n-1})}{a_{NN}(x,t_{n-1})}\right)\right| \le C, \quad \text{and} \quad \left|D_\tau\left(\frac{a_{Nk}(x,t_n)}{a_{NN}(x,t_n)}\right)\right| \le C.$$

Hence we have

$$\begin{aligned} &\|\vec{\varphi}\|_{\ell^p(W^{1,q})} + \|D_\tau \vec{\varphi}\|_{\ell^p(\widetilde{W}^{-1,q})} \\ &\quad \le C\sum_{k=1}^{N-1}\left(\|\partial_k \vec{U}\|_{\ell^p(W^{1,q})} + \|D_\tau \partial_k \vec{U}\|_{\ell^p(\widetilde{W}^{-1,q})}\right) \\ &\quad \le C\|\vec{f}\|_{\ell^p(L^q)} \qquad \text{(here we have used (4.11))}. \end{aligned} \tag{4.12}$$

We consider $w_N^n = \partial_N U^n + \varphi^n$, which is the solution of

$$\begin{cases} D_\tau w_N^n - \displaystyle\sum_{i,j=1}^N \partial_i\big(a_{ij}(x,t_n)\partial_j w_N^n\big) + w_N^n \\ \qquad = \partial_N(f^n - cU^n + U^n) + \displaystyle\sum_{i,j=1}^N \partial_i\big(\partial_N a_{ij}(x,t_n)\partial_j U^n\big) \\ \qquad\quad + D_\tau\varphi^n - \displaystyle\sum_{i,j=1}^N \partial_i\big(a_{ij}(x,t_n)\partial_j\varphi^n\big) & \text{in } \Omega, \\ w_N^n = 0 & \text{on } \partial\Omega, \\ w_N^0 = 0 & \text{in } \Omega, \end{cases}$$

where the Dirichlet boundary condition of w_N^n is a consequence of the Neumann boundary condition satisfied by U^n, i.e.,

$$a_{NN}(x,t_n)\partial_N U^n + \sum_{j=1}^{N-1} a_{Nj}(x,t_n)\partial_j U^n = \sum_{i,j=1}^N a_{ij}(x,t_n)\partial_j U^n n_i = 0 \quad \text{on } \partial\Omega.$$

By applying Lemma 4.1 we derive

$$\begin{aligned} \|\vec{w}_N\|_{\ell^p(W^{1,q})} &\le C\|f - cU^n\|_{\ell^p(L^q)} + C\|\vec{U}\|_{\ell^p(W^{1,q})} \\ &\quad + C\|D_\tau\vec{\varphi}\|_{\ell^p(W^{-1,q})} + C\|\vec{\varphi}\|_{\ell^p(W^{1,q})} \\ &\le C\|f\|_{\ell^p(L^q)} \qquad \text{(here we have used (4.11) and (4.12))}, \end{aligned}$$

which further implies

$$\|\partial_N\vec{U}\|_{\ell^p(W^{1,q})} \le \|\vec{w}_N\|_{\ell^p(W^{1,q})} + \|\vec{\varphi}\|_{\ell^p(W^{1,q})} \le C\|f\|_{\ell^p(L^q)}.$$

This proves (1.14) for the special case that $\Omega = \{(x_1,\ldots,x_N) \in \mathbb{R}^N : x_N > 0\}$ and $\vec{U}$ has compact support in the unit ball. For a general bounded smooth domain, in terms of a partition of unity and a coordinate transform, the problem can always be transformed into the domain $\Omega = \{(x_1,\ldots,x_N) \in \mathbb{R}^N : x_N > 0\}$.

The proof of (1.14) and (4.1) is complete.

4.2. Fully discrete finite element solutions. To prove (1.18), we let $U_h^0 = P_h U^0 = 0$, denote $e_h^n = P_h U^n - U_h^n$, and let $R_h^n = R_h(t_n)$ be the Ritz projection associated with the elliptic operator $A(t_n)$ (defined in section 2). From (2.1)–(2.2) we see that e_h^n satisfies the equation

$$\begin{aligned} &\big(D_\tau e_h^n, v_h\big) + \sum_{i,j=1}^N \big(a_{ij}(\cdot,t_n)\partial_i e_h^n, \partial_j v_h\big) + \big(c(\cdot,t_n)e_h^n, v_h\big) \\ &\quad = \sum_{i,j=1}^N \big(a_{ij}(\cdot,t_n)\partial_i\xi_h^n, \partial_j v_h\big) + \big(c(\cdot,t_n)\xi_h^n, v_h\big) \qquad \forall\, v_h \in S_h, \end{aligned} \tag{4.13}$$

where $\xi_h^n := P_h U^n - R_h^n U^n$. Equivalently, the equation above can be written in the following operator form:

$$\begin{cases} D_\tau e_h^n + A_h(t_n)e_h^n = A_h(t_n)\xi_h^n, \\ e_h^0 = 0. \end{cases} \tag{4.14}$$

Similar to section 4.1, we rewrite the equation above in the perturbed form

$$\begin{cases} D_\tau e_h^n + A_h(t_m)e_h^n = A_h(t_n)\xi_h^n + (A_h(t_m) - A_h(t_n))e_h^n, \\ e_h^0 = 0, \end{cases} \tag{4.15}$$

where the operator on the left-hand side does not depend on n, so we can apply (3.19) of Lemma 3.5 in the subinterval $[0, t_m]$. Then we have

$$\begin{aligned} &\|D_\tau \vec{e}_h\|_{\ell^p(0,t_m;\widetilde{W}^{-1,q})} + \|\vec{e}_h\|_{\ell^p(0,t_m;W^{1,q})} \\ &\quad \leq C\big(\|\vec{\xi}_h\|_{\ell^p(0,t_m;W^{1,q})} + \|((a(\cdot,t_m) - a(\cdot,t_n))\nabla e_h^n)_{n=1}^m\|_{\ell^p(0,t_m;L^q)}\big) \\ &\quad \leq C\big(\|\vec{\xi}_h\|_{\ell^p(0,t_m;W^{1,q})} + \|(|t_m - t_n| e_h^n)_{n=1}^m\|_{\ell^p(0,t_m;W^{1,q})}\big). \end{aligned} \tag{4.16}$$

If we define

$$\begin{aligned} &E_h^m := \|\vec{e}_h\|^p_{\ell^p(0,t_m;W^{1,q})} = \tau \sum_{n=1}^m \|e_h^n\|^p_{W^{1,q}}, \quad m = 1, \ldots, M, \\ &E_h^0 := 0, \end{aligned}$$

then (4.16) can be written as

$$\begin{aligned} E_h^m &\leq C\|\vec{\xi}_h\|^p_{\ell^p(0,t_m;W^{1,q})} + C\tau \sum_{n=1}^m |t_m - t_n|^p \|e_h^n\|^p_{W^{1,q}} \\ &= C\|\vec{\xi}_h\|^p_{\ell^p(0,t_m;W^{1,q})} + C\tau \sum_{n=1}^m |t_m - t_n|^p \frac{E_h^n - E_h^{n-1}}{\tau} \\ &= C\|\vec{\xi}_h\|^p_{\ell^p(0,t_m;W^{1,q})} + C\tau \sum_{n=1}^{m-1} \frac{|t_m - t_{n+1}|^p - |t_m - t_n|^p}{\tau} E_h^n \\ &\leq C\|\vec{\xi}_h\|^p_{\ell^p(0,t_m;W^{1,q})} + C\tau \sum_{n=1}^{m-1} E_h^n, \end{aligned} \tag{4.17}$$

which yields (via using the discrete Gronwall's inequality)

$$E_h^M \leq C\|\vec{\xi}_h\|^p_{\ell^p(0,T;W^{1,q})} = C\|\vec{\xi}_h\|^p_{\ell^p(W^{1,q})}.$$

Substituting the estimate above into (4.16), we get

$$\|D_\tau \vec{e}_h\|_{\ell^p(\widetilde{W}^{-1,q})} + \|\vec{e}_h\|_{\ell^p(W^{1,q})} \leq C\|\vec{\xi}_h\|_{\ell^p(W^{1,q})}, \tag{4.18}$$

which further implies

(4.19)

$$\begin{aligned} &\|D_\tau \vec{U}_h\|_{\ell^p(\widetilde{W}^{-1,q})} + \|\vec{U}_h\|_{\ell^p(W^{1,q})} \\ &\quad \leq (\|D_\tau \vec{e}_h\|_{\ell^p(\widetilde{W}^{-1,q})} + \|\vec{e}_h\|_{\ell^p(W^{1,q})}) + (\|D_\tau P_h \vec{U}\|_{\ell^p(\widetilde{W}^{-1,q})} + \|P_h \vec{U}\|_{\ell^p(W^{1,q})}) \\ &\quad \leq C\|\vec{\xi}_h\|_{\ell^p(W^{1,q})} + C(\|D_\tau \vec{U}\|_{\ell^p(\widetilde{W}^{-1,q})} + \|\vec{U}\|_{\ell^p(W^{1,q})}) \qquad \text{(use (4.18) here)} \\ &\quad = C\|P_h \vec{U}^n - \vec{R}_h \vec{U}^n\|_{\ell^p(W^{1,q})} + C(\|D_\tau \vec{U}\|_{\ell^p(\widetilde{W}^{-1,q})} + \|\vec{U}\|_{\ell^p(W^{1,q})}) \\ &\quad \leq C(\|D_\tau \vec{U}\|_{\ell^p(\widetilde{W}^{-1,q})} + \|\vec{U}\|_{\ell^p(W^{1,q})}) \qquad \text{(use (2.6b) with } l = k = 1) \\ &\quad \leq C\left(\|\vec{f}\|_{\ell^p(L^q)} + \sum_{j=1}^N \|\vec{g}_j\|_{\ell^p(L^q)}\right) \qquad \text{(use (4.1) here).} \end{aligned}$$

This proves (1.18).

Second, we let $U_h^0 = P_h U^0 = 0$ and prove (1.17). By using (4.14) and the inverse inequality, we have

$$
\begin{aligned}
&\|D_\tau \vec{e}_h\|_{\ell^p(L^q)} + \|\vec{A}_h \vec{e}_h\|_{\ell^p(L^q)} \\
&\quad \le C\|\vec{A}_h \vec{e}_h\|_{\ell^p(L^q)} + \|\vec{A}_h(P_h\vec{U} - \vec{R}_h\vec{U})\|_{\ell^p(L^q)} \quad \text{(use (4.14) here)} \\
&\quad \le Ch^{-1}\|\vec{e}_h\|_{\ell^p(W^{1,q})} + Ch^{-1}\|P_h\vec{U} - \vec{R}_h\vec{U}\|_{\ell^p(W^{1,q})} \quad \text{(use (2.7) here)} \\
&\quad \le Ch^{-1}\|P_h\vec{U} - \vec{R}_h\vec{U}\|_{\ell^p(W^{1,q})} \quad \text{(use (4.18) here)} \\
&\quad \le C\|\vec{U}\|_{\ell^p(W^{2,q})} \quad \text{(use (2.6a)–(2.6b) here)},
\end{aligned} \tag{4.20}
$$

which implies that

$$
\begin{aligned}
&\|D_\tau \vec{U}_h\|_{\ell^p(L^q)} + \|\vec{A}_h \vec{U}_h\|_{\ell^p(L^q)} \\
&\quad \le \|D_\tau \vec{e}_h\|_{\ell^p(L^q)} + \|\vec{A}_h \vec{e}_h\|_{\ell^p(L^q)} + \|D_\tau P_h\vec{U}\|_{\ell^p(L^q)} + \|\vec{A}_h P_h\vec{U}\|_{\ell^p(L^q)} \\
&\quad \le C(\|D_\tau \vec{U}\|_{\ell^p(L^q)} + \|\vec{U}\|_{\ell^p(W^{2,q})}) \quad \text{(use (4.20) and (2.8) here)} \\
&\quad \le C\|\vec{f}\|_{\ell^p(L^q)} \quad \text{(use (1.14) here)}.
\end{aligned} \tag{4.21}
$$

This proves (1.17). Note that the $W^{2,q}$ regularity of U^n required in the inequalities above is obtained from (1.14), which has already been proved in section 4.1.

Finally, we drop the assumption $U_h^0 = P_h U^0 = 0$ and prove (2.4). The proof of (2.3) is similar (by dropping the truncation-error term $\mathcal{E}^n$ in the proof below). From (1.1) and (2.1) we see that the error function $\theta_h^n = P_h u^n - U_h^n - (P_h u^0 - U_h^0)$ satisfies the equation

$$
\begin{cases}
D_\tau \theta_h^n + A_h(t_n)\theta_h^n = A_h(t_n)\eta_h^n + P_h\mathcal{E}^n, \\
\theta_h^0 = 0,
\end{cases} \tag{4.22}
$$

where $\eta_h^n = P_h u^n - R_h^n u^n - (P_h u^0 - U_h^0)$ and

$$
\mathcal{E}^n = D_\tau u^n - \partial_t u(\cdot, t_n) = \int_{t_{n-1}}^{t_n} \frac{s - t_{n-1}}{\tau} \partial_{tt} u(s) \mathrm{d}s
$$

denotes the truncation error of the backward Euler scheme, which satisfies

$$
\begin{aligned}
\|\vec{\mathcal{E}}\|_{\ell^p(\widetilde{W}^{-1,q})} &\le \left(\sum_{n=1}^N \tau \left(\int_{t_{n-1}}^{t_n} \|\partial_{tt} u(s)\|_{\widetilde{W}^{-1,q}} \mathrm{d}s \right)^p \right)^{\frac{1}{p}} \\
&\le \left(\sum_{n=1}^N \tau^{1+\frac{p}{p'}} \int_{t_{n-1}}^{t_n} \|\partial_{tt} u(s)\|_{\widetilde{W}^{-1,q}}^p \mathrm{d}s \right)^{\frac{1}{p}} \quad \text{(use Hölder's inequality)} \\
&= \tau \|\partial_{tt} u\|_{L^p(0,T;\widetilde{W}^{-1,q})}.
\end{aligned} \tag{4.23}
$$

Applying (1.18) to (4.22), we obtain

$$
\begin{aligned}
\|\vec{\theta}_h\|_{\ell^p(W^{1,q})} &\le C\|\vec{\eta}_h\|_{\ell^p(W^{1,q})} + C\|P_h\vec{\mathcal{E}}\|_{\ell^p(L^q)} \\
&\le Ch^{-1}(\|\vec{\eta}_h\|_{\ell^p(L^q)} + \|P_h\vec{\mathcal{E}}\|_{\ell^p(\widetilde{W}^{-1,q})}),
\end{aligned} \tag{4.24}
$$

where we have used the inverse inequalities

$$\|\vec{\eta}_h\|_{\ell^p(W^{1,q})} \le Ch^{-1}\|\vec{\eta}_h\|_{\ell^p(L^q)}, \tag{4.25}$$

$$\|P_h\vec{\mathcal{E}}\|_{\ell^p(L^q)} \le Ch^{-1}\|P_h\vec{\mathcal{E}}\|_{\ell^p(\widetilde{W}^{-1,q})}. \tag{4.26}$$

For any given $\vec{\phi} = (\phi^n)_{n=0}^{M-1} \in \ell^p(L^q)$, we let $\vec{\psi} = (\psi^n)_{n=0}^{M-1}$ be the solution of the backward equation

$$
\begin{cases}
D_\tau\psi^n - A(t_n)\psi^{n-1} = -\phi^n, \\
\psi^M = 0.
\end{cases}
$$

According to (1.14) (which has already been proved), we have $\psi^{n-1} \in D_q(t_n)$, $n = 1, \dots, M$ (see the definition of $D_q(t_n)$ in Lemma 2.1), and

$$\|D_\tau\vec{\psi}\|_{\ell^p(L^q)} + \|\vec{\psi}\|_{\ell^p(W^{2,q})} \le C\|\vec{\phi}\|_{\ell^p(L^q)}, \tag{4.27}$$

where $\vec{\psi} := (\psi^0, \dots, \psi^{M-1})$. Using integration by parts and (2.8)–(2.10), we have

$$
\begin{aligned}
&\sum_{n=1}^M \tau(\theta_h^n, \phi^n) \\
&\quad= \sum_{n=1}^M \tau(\theta_h^n, -D_\tau\psi^n + A(t_n)\psi^{n-1}) \\
&\quad= \sum_{n=1}^N \tau(D_\tau\theta_h^n + A(t_n)\theta_h^n, \psi^{n-1}) \\
&\quad= \sum_{n=1}^M \Big(\tau\big(D_\tau\theta_h^n + A(t_n)\theta_h^n, \psi^{n-1} - P_h\psi^{n-1}\big) + \tau\big(D_\tau\theta_h^n + A_h(t_n)\theta_h^n, P_h\psi^{n-1}\big) \\
&\quad= \sum_{n=1}^M \Big(\tau\big(D_\tau\theta_h^n + A(t_n)\theta_h^n, \psi^{n-1} - P_h\psi^{n-1}\big) \\
&\qquad\qquad + \tau\big(A_h(t_n)\eta_h^n, P_h\psi^{n-1}\big) + \tau(P_h\mathcal{E}^n, P_h\psi^{n-1})\Big) \\
&\qquad\qquad\qquad\qquad \text{(here we have used (4.22))} \\
&\quad= \sum_{n=1}^M \Big(\tau\big(D_\tau\theta_h^n + A(t_n)\theta_h^n, \psi^{n-1} - P_h\psi^{n-1}\big) \\
&\qquad\qquad + \tau\big(A(t_n)\eta_h^n, P_h\psi^{n-1} - \psi^{n-1}\big) + \tau\big(A(t_n)\eta_h^n, \psi^{n-1}\big) + \tau(P_h\mathcal{E}^n, P_h\psi^{n-1})\Big) \\
&\quad= \sum_{n=1}^M \Big(\tau\big(\theta_h^n, A(t_n)(\psi^{n-1} - P_h\psi^{n-1})\big) \\
&\qquad\qquad + \tau\big(\eta_h^n, A(t_n)(P_h\psi^{n-1} - \psi^{n-1})\big) + \tau\big(\eta_h^n, A(t_n)\psi^{n-1}\big) + \tau(P_h\mathcal{E}^n, \psi^{n-1})\Big)
\end{aligned}
$$

$$
\begin{aligned}
&\le C(\|\vec{\theta}_h\|_{\ell^p(W^{1,q})} + \|\vec{\eta}_h\|_{\ell^p(W^{1,q})})\|\vec{\psi} - P_h\vec{\psi}\|_{l^{p'}(W^{1,q'})} \\
&\quad + C\|\vec{\eta}_h\|_{\ell^p(L^q)}\|\vec{A}\vec{\psi}\|_{l^{p'}(L^{q'})} + C\|P_h\vec{\mathcal{E}}\|_{\ell^p(\widetilde{W}^{-1,q})}\|\vec{\psi}\|_{l^{p'}(W^{1,q'})} \\
&\le C(\|\vec{\theta}_h\|_{\ell^p(W^{1,q})}h + \|\vec{\eta}_h\|_{\ell^p(L^q)} + \|P_h\vec{\mathcal{E}}\|_{\ell^p(\widetilde{W}^{-1,q})})\|\vec{\psi}\|_{l^{p'}(W^{2,q'})} \\
&\le C(\|\vec{\eta}_h\|_{\ell^p(L^q)} + \|P_h\vec{\mathcal{E}}\|_{\ell^p(\widetilde{W}^{-1,q})})\|\vec{\psi}\|_{l^{p'}(W^{2,q'})} \qquad \text{(here we have used (4.24))} \\
&\le C(\|\vec{\eta}_h\|_{\ell^p(L^q)} + \|\vec{\mathcal{E}}\|_{\ell^p(\widetilde{W}^{-1,q})})\|\vec{\phi}\|_{l^{p'}(L^{q'})}. \qquad \text{(here we have used (4.27)).}
\end{aligned}
$$

By duality, we have

$$
\begin{aligned}
\|\vec{\theta}_h\|_{\ell^p(L^q)} &\le C(\|\vec{\eta}_h\|_{\ell^p(L^q)} + \|\vec{\mathcal{E}}\|_{\ell^p(\widetilde{W}^{-1,q})}) \\
&\le C(\|P_h\vec{u} - \vec{R}_h\vec{u}\|_{\ell^p(L^q)} + \|P_hu^0 - U_h^0\|_{L^q}) + C\|\vec{\mathcal{E}}\|_{\ell^p(\widetilde{W}^{-1,q})} \\
&\le C(\|P_h\vec{u} - \vec{R}_h\vec{u}\|_{\ell^p(L^q)} + \|P_hu^0 - U_h^0\|_{L^q}) + C\|\partial_{tt}u\|_{L^p(0,T;\widetilde{W}^{-1,q})}\tau.
\end{aligned}
$$

This proves (2.4), and the proof of Theorem 2.1 is complete. ∎

Acknowledgment. We would like to thank the anonymous referees for many valuable comments and suggestions, which were very helpful in improving both the quality and presentation of this paper.

REFERENCES

[1] W. Arendt, R. Chill, S. Fornaro, and C. Poupaud, L^p-maximal regularity for nonautonomous evolution equations, J. Differential Equations, 237 (2007), pp. 1–26.
[2] A. Ashyralyev, S. Piskarev, and L. Weis, *On well-posedness of difference schemes for abstract parabolic equations in $L_p([0,T];E)$ spaces*, Numer. Funct. Anal. Optim., 23 (2002), pp. 669–693.
[3] P. Auscher and M. Qafsaoui, *Observations on $W^{1,p}$ estimates for divergence elliptic equations with VMO coefficients*, Boll. Unione Mat. Ital. (8), 5-B (2002), pp. 487–509.
[4] P. Auscher and Ph. Tchamitchian, *Square roots of elliptic second order divergence operators on strongly Lipschitz domains: L^p theory*, Math. Ann., 320 (2001), pp. 577–623.
[5] N. Bakaev, *Maximum norm resolvent estimates for elliptic finite element operators*, BIT, 41 (2001), pp. 215–239.
[6] S.C. Brenner and L.R. Scott, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Texts Appl. Math. 15, Springer, New York, 2008.
[7] K. Chrysafinos and N.J. Walkington, *Error estimates for the discontinuous Galerkin methods for parabolic equations*, SIAM J. Numer. Anal., 44 (2006), pp. 349–366.
[8] M. Crouzeix, *Contractivity and analyticity in ℓ^p of some approximation of the heat equation*, Numer. Algorithms, 33 (2003), pp. 193–201.
[9] C. Gallarati and M. Veraar, *Maximal regularity for non-autonomous equations with measurable dependence on time*, Potential Anal., https://doi.org/10.1007/s11118-016-9593-7 (2016).
[10] M. Geissert, *Discrete maximal L^p regularity for finite element operators*, SIAM J. Numer. Anal., 44 (2006), pp. 677–698.
[11] M. Geissert, *Applications of discrete maximal L^p regularity for finite element operators*, Numer. Math., 108 (2007), pp. 121–149.
[12] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Classics Appl. Math. 69, SIAM, Philadelphia, 2011.
[13] A. Hansbo, *Strong stability and non-smooth data error estimates for discretizations of linear parabolic problems*, BIT, 42 (2002), pp. 351–379.
[14] T. Kemmochi and N. Saito, *Discrete Maximal Regularity and the Finite Element Method for Parabolic Equations*, preprint, arXiv:1602.06864, 2016.
[15] B. Kovács, B. Li, and C. Lubich, *A-stability implies discrete maximal parabolic regularity*, SIAM J. Numer. Anal., 54 (2016), pp. 3600–3624.
[16] P.C. Kunstmann and L. Weis, *Maximal L^p-regularity for parabolic equations, Fourier multiplier theorems and H^∞-functional calculus*, in Functional Analytic Methods for Evolution Equations, Lecture Notes in Math. 1855, Springer, Berlin, 2004, pp. 65–311.

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

[18] D. Leykekhman and B. Vexler, *Discrete maximal parabolic regularity for Galerkin Finite Element Methods*, Numer. Math., 135 (2017), pp. 923–952.

[19] D. Leykekhman and B. Vexler, *Pointwise Best Approximation Results for Galerkin Finite Element Soutions of Parabolic Problems*, preprint, http://arxiv.org/abs/1508.01165.

[20] 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.

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

[22] B. Li and W. Sun, *Maximal L^p analysis of finite element solutions for parabolic equations with nonsmooth coefficients in convex polyhedra*, Math. Comp., https://doi.org/10.1090/mcom/3133 (2016).

[23] 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. 1959–1977.

[24] B. Li and W. Sun, *Error analysis of linearized semi-implicit Galerkin finite element methods for nonlinear parabolic equations*, Int. J. Numer. Anal. Model., 10 (2013), pp. 622–633.

[25] 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.

[26] Y. P. 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.

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

[28] R. Rannacher, *L^∞-stability estimates and asymptotic error expansion for parabolic finite element equations*, in Extrapolation and Defect Correction, 1990, Bonner Math. Schriften 228, University of Bonn, Bonn, 1991, pp. 74–94.

[29] R. Rannacher and R. Scott, *Some optimal error estimates for piecewise linear finite element approximations*, Math. Comp., 38 (1982), pp. 437–445.

[30] A.H. Schatz, *Pointwise error estimates and asymptotic error expansion inequalities for the finite element method on irregular grids* I: *Global estimates*, Math. Comp., 67 (1998), pp. 877–899.

[31] 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.

[32] 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.

[33] 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.

[34] L. Wahlbin, *Local behavior in finite element methods*, in Handbook of Numerical Analysis II: Finite Element Methods (Part 1), P.G. Ciarlet and J.L. Lions, eds., North-Holland, Amsterdam, 1991, pp. 353–522.

[35] L. Weis, *A new approach to maximal L_p-regularity*, in Evolution Equations and their Applcations in Physical and Life Sciences, Lecture Notes in Pure Appl. Math. 215, Marcel Dekker, New York, 2001, pp. 195–214.

[36] L. Weis, *Operator-valued Fourier multiplier theorems and maximal L_p-regularity*, Math. Ann., 319 (2001), pp. 735–758.