Analyticity, maximal regularity and maximum-norm stability of semi-discrete finite element solutions of parabolic equations in nonconvex polyhedra *

Buyang Li †

Abstract

In general polygons and polyhedra, possibly nonconvex, the analyticity of the finite element heat semigroup in the L^q norm, $1 \le q \le \infty$, and the maximal L^p-regularity of semi-discrete finite element solutions of parabolic equations are proved. By using these results, the problem of maximum-norm stability of the finite element parabolic projection is reduced to the maximum-norm stability of the Ritz projection, which currently is known to hold for general polygonal domains and convex polyhedral domains.

Key words: analytic semigroup, maximal L^p-regularity, maximum-norm stability, finite element method, parabolic equation, nonconvex polyhedra

1 Introduction

Let Ω be a polygonal or polyhedral domain in $\mathbb{R}^N$, $N = 2, 3$, and consider the heat equation

$$\begin{cases} \dfrac{\partial u(t,x)}{\partial t} - \Delta u(t,x) = f(t,x) & \text{for } (t,x) \in \mathbb{R}_+ \times \Omega, \\ u(t,x) = 0 & \text{for } (t,x) \in \mathbb{R}_+ \times \partial\Omega, \\ u(0,x) = u_0(x) & \text{for } x \in \Omega. \end{cases} \tag{1.1}$$

In the case $f = 0$, it is well known that the solution of (1.1) is given by $u(t,x) = (e^{t\Delta}u_0)(x)$, where $E(t) = e^{t\Delta}$ extends to a bounded analytic semigroup on $C_0(\overline{\Omega})$ and $L^q(\Omega)$ for arbitrary $1 \le q < \infty$ (cf. [38]), satisfying the following estimates:

$$\begin{aligned} &\sup_{t>0} \big(\|E(t)v\|_{L^q(\Omega)} + t\|\partial_t E(t)v\|_{L^q(\Omega)}\big) \le C\|v\|_{L^q(\Omega)}\,, \quad v \in L^q(\Omega), \ \ 1 \le q < \infty, \\ &\sup_{t>0} \big(\|E(t)v\|_{C_0(\overline{\Omega})} + t\|\partial_t E(t)v\|_{C_0(\overline{\Omega})}\big) \le C\|v\|_{C_0(\overline{\Omega})}, \quad v \in C_0(\overline{\Omega}). \end{aligned} \tag{1.2}$$

*This work is partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region (Project No. 15300817).

†Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong. Email: buyang.li@polyu.edu.hk

In the case $u_0 = 0$, the solution of (1.1) possesses the maximal L^p-regularity in the space $L^q(\Omega)$, namely, for all $f \in L^p(\mathbb{R}_+; L^q(\Omega))$,

$$\|\partial_t u\|_{L^p(\mathbb{R}_+;L^q(\Omega))} + \|\Delta u\|_{L^p(\mathbb{R}_+;L^q(\Omega))} \le C_{p,q}\|f\|_{L^p(\mathbb{R}_+;L^q(\Omega))}, \quad \text{if } u_0 = 0,\ 1 < p, q < \infty. \tag{1.3}$$

Such maximal L^p-regularity as (1.3) has important applications in the analysis of nonlinear partial differential equations (PDEs) [4, 7, 35], and has been widely studied in the literature; see [25] and the references therein.

This paper is concerned with the discrete analogues of (1.2)-(1.3), namely,

$$\sup_{t>0}\big(\|E_h(t)v_h\|_{L^q(\Omega)} + t\|\partial_t E_h(t)v_h\|_{L^q(\Omega)}\big) \le C\|v_h\|_{L^q(\Omega)}, \qquad \forall\, v_h \in S_h,\ 1 \le q \le \infty, \tag{1.4}$$

$$\|\partial_t u_h\|_{L^p(\mathbb{R}_+;L^q(\Omega))} + \|\Delta_h u_h\|_{L^p(\mathbb{R}_+;L^q(\Omega))} \le C_{p,q}\|f\|_{L^p(\mathbb{R}_+;L^q(\Omega))}, \quad \text{if } u_{h,0} = 0,\ 1 < p, q < \infty, \tag{1.5}$$

where $E_h(t) = e^{t\Delta_h}$ is the semigroup generated by the discrete Laplacian operator Δ_h (on a finite element subspace $S_h \subset H_0^1(\Omega)$ with mesh size h), defined by

$$(\Delta_h \phi_h, \varphi_h) = -(\nabla\phi_h, \nabla\varphi_h), \quad \forall\, \phi_h, \varphi_h \in S_h, \tag{1.6}$$

and u_h is the finite element solution of (1.1), i.e.

$$\begin{cases} (\partial_t u_h, v_h) + (\nabla u_h, \nabla v_h) = (f, v_h), \quad \forall\, v_h \in S_h,\ \forall\, t \in (0,T), \\ u_h(0) = u_{h,0}. \end{cases} \tag{1.7}$$

The constants C and $C_{p,q}$ in (1.4)-(1.5) should be independent of f and h. For the maximal L^p-regularity (1.5) we require $u_{h,0} = 0$ (as the continuous problem), while for error estimate we choose $u_{h,0}$ to be the L^2 projection of u_0 (see Corollary 2.3).

The discrete analyticity (1.4) and the discrete maximal L^p-regularity (1.5) are important mathematical tools for numerical analysis of parabolic equations. For example, (1.4) can be used to derive error estimates for both semi-discrete and fully discrete finite element methods [18, 37, 39, 46], and (1.5) has been used to study the convergence rates of finite element solutions of semilinear parabolic equations [15] as well as nonlinear parabolic equations with nonsmooth diffusion coefficients [32]. The time-discrete extension of the maximal L^p-regularity (1.3) has been used to study the stability and convergence of time discretization methods for nonlinear parabolic equations with general (possibly degenerate) nonlinearities [2, 3, 24].

Being the foundation for many existing numerical analyses, the discrete analyticity (1.4) and the discrete maximal regularity (1.5) have been studied by many authors in the literature. In the case $q = 2$, (1.4) holds trivially [46, Lemma 3.2] and (1.5) is an immediate consequence of (1.4) due to the Hilbert space structure of $L^2(\Omega)$ (cf. [20]). The discrete analyticity (1.4) for $q \in [1,\infty]\backslash 2$ is a simple consequence of the result in the end-point case $q = \infty$ (via complex interpolation and duality), which was proved in [43] for $N = 2$ and $r = 1$ and was proved in [36] for $N = 1, 2, 3$ and $r \ge 4$, where r is the degree of finite elements. The general case $N \geq 2$ and $r \geq 1$ was proved in [44] and [47] for the Neumann and Dirichlet boundary conditions, respectively, and was extended to parabolic equations with nonsmooth diffusion coefficients in [31]. The analyses presented in these works were all restricted to smooth domains. The discrete analyticity (1.4) was proved in [40] for convex polygons in the case $N = 2$ and $r = 1$ with a logarithmic factor $|\ln h|^{\frac{3}{2}}$, and was proved in [33] for convex polyhedra in the case $N = 2, 3$ and $r \geq 1$. In the presence of an extra logarithmic factor $|\ln h|$, the discrete analyticity (1.4) can be extended to general two-dimensional polygons (cf. [46, Theorems 6.1 and 6.3]). However, the sharp estimate (without logarithmic factor) of (1.4) remains open in nonconvex polygons and polyhedra.

Similarly, (1.5) has been proved in smooth domains and convex polygons/polyhedra [14, 33]. The extension of (1.5) to the fully discrete finite element methods has been considered in [22, section 6] and [30, 34], which rely on the semi-discrete results. For the lumped mass method, both (1.4) and (1.5) have been proved in general polygons by using the maximum principle [8, 21]. However, for the finite element method, sharp estimates of (1.4) and (1.5) remain open in nonconvex polygons and polyhedra. In particular, the techniques used in the existing works rely on the H^2-regularity of elliptic equations, which only holds in smooth or convex domains.

It is worth to mention that the proof of the discrete maximal L^p-regularity (1.5) is closely related to the proof of the maximum-norm stability (best approximation property) of finite element solutions of parabolic equations, namely,

$$\|u - u_h\|_{L^\infty(0,T;L^\infty(\Omega))} \leq C \ln(2 + 1/h) \inf_{\chi_h} \|u - \chi_h\|_{L^\infty(0,T;L^\infty(\Omega))}, \tag{1.8}$$

where the infimum extends over all $\chi_h \in L^\infty(0, T; S_h)$, and the logarithmic factor "$\ln(2 + 1/h)$" in (1.8) is sharp for piecewise linear finite elements (possibly removable for higher order finite elements). Such a priori L^∞-norm best approximation property has been proved in smooth domains [27, 31, 36, 43, 44, 47] and convex polygons (2D) [40], but remains open in convex polyhedra and nonconvex polygons/polyhedra, though the maximum-norm a posteriori error estimates for finite element solutions of parabolic equations have been derived in general polyhedra [11]. The a priori L^∞-norm best approximation property has been proved in the fully discrete settings with discontinuous Galerkin time-stepping methods [29] (the result does not cover the semi-discrete case due to a logarithmic dependence on the time-step size). Related maximum-norm stability for finite element solutions of elliptic equations can be found in [12, 17, 28, 41, 42].

In this paper, we prove (1.4)-(1.5) in general polygons and polyhedra, possibly nonconvex (cf. Theorem 2.1), and we reduce (1.8) to the maximum-norm stability of the Ritz projection (cf. Corollary 2.2). In particular, (1.8) is proved completely in nonconvex polygons and convex polyhedra (cf. Corollary 2.3). The proof of these results relies on a dyadic decomposition of the domain $(0, 1) \times \Omega = \cup_{*,j} Q_j$ together with some local $L^2 H^{1+\alpha}(Q_j)$ and $L^\infty H^{1+\alpha}(Q_j)$ estimates of the Green's function (Lemma 4.1) and a local energy error estimate for finite element solutions of parabolic equations (Lemma 5.1). In contrast to the existing work (cf. [33, 44]), the local energy error estimate used here does not require any superapproximation property of the Ritz projection (which only holds in convex domains). These results help to prove the key lemma (Lemma 4.4) for the proof of our main results. The maximal L^p-regularity (1.5) is first proved for $p = q$ and then extended to $p \neq q$ by using the singular integral operator approach (Sections 4.3–4.4).

2 Main results and their consequences

Let $L^q = L^q(\Omega)$. Let $\Gamma_h(t,x,x_0)$ be the kernel of the operator $E_h(t)$, i.e.

$$(E_h(t)v_h)(x_0) = \int_\Omega \Gamma_h(t,x,x_0)v_h(x)\mathrm{d}x, \quad \forall\, v_h \in S_h, \tag{2.1}$$

and define $|E_h(t)|$ to be the linear operator on L^q with the kernel $|\Gamma_h(t,x,x_0)|$, namely,

$$(|E_h(t)|v)(x_0) := \int_\Omega |\Gamma_h(t,x,x_0)|v(x)\mathrm{d}x, \quad \forall\, v \in L^q. \tag{2.2}$$

The main result of this paper is the following theorem.

Theorem 2.1 *Let Ω be a polygon in $\mathbb{R}^2$ or a polyhedron in $\mathbb{R}^3$ (possibly nonconvex), and let S_h, $0 < h < h_0$, be a family of finite element subspaces of $H^1_0(\Omega)$ consisting of piecewise polynomials of degree $r \geq 1$ subject to a quasi-uniform triangulation of the domain Ω (with mesh size h). Then we have the following analytic semigroup estimate and maximal function estimate:*

$$\sup_{t>0}\left(\|E_h(t)v_h\|_{L^q} + t\|\partial_t E_h(t)v_h\|_{L^q}\right) \leq C\|v_h\|_{L^q}, \quad \forall\ v_h \in S_h,\ \ \forall\ 1 \leq q \leq \infty, \tag{2.3}$$

$$\left\|\sup_{t>0}|E_h(t)|\,|v|\right\|_{L^q} \leq C_q\|v\|_{L^q}, \qquad \forall\, v \in L^q,\ \ \forall\, 1 < q \leq \infty. \tag{2.4}$$

Further, if $u_{h,0} = 0$ and $f \in L^p(0,T;L^q)$, then the finite element solution given by (1.7) possesses the following maximal L^p-regularity:

$$\|\partial_t u_h\|_{L^p(0,T;L^q)} + \|\Delta_h u_h\|_{L^p(0,T;L^q)} \leq \max(p\,,\,(p-1)^{-1})C_q\|f\|_{L^p(0,T;L^q)}, \quad \forall\ 1 < p,q < \infty, \tag{2.5}$$

$$\|\partial_t u_h\|_{L^\infty(0,T;L^q)} + \|\Delta_h u_h\|_{L^\infty(0,T;L^q)} \leq C\ell_h\|f\|_{L^\infty(0,T;L^q)}, \qquad \forall\ 1 \leq q \leq \infty, \tag{2.6}$$

where $\ell_h := \log(2 + 1/h)$.

The constant C in (2.3) and (2.6) is independent of f, h, p, q and T, and the constant C_q in (2.4) and (2.5) is independent of f, h, p and T.

Remark 2.1 By the theory of analytic semigroups [51, page 254], the inequality (2.3) implies the existence of a positive constant $\theta \in (0,\pi/2)$, independent of h and q, such that the semigroup $\{E_h(t)\}_{t>0}$ extends to be a bounded analytic semigroup $\{E_h(z)\}_{z\in\Sigma_\theta}$ in the sector $\Sigma_\theta := \{z \in \mathbb{C}\backslash\{0\} : |\arg(z)| < \theta\}$, i.e.

$$E_h(z_1 + z_2) = E_h(z_1)E_h(z_2), \quad \forall\ z_1, z_2 \in \Sigma_\theta, \tag{2.7}$$

$$\sup_{z\in\Sigma_\theta}\left(\|E_h(z)v_h\|_{L^q} + |z|\|\partial_z E_h(z)v_h\|_{L^q}\right) \leq C\|v_h\|_{L^q}, \quad \forall\ v_h \in S_h,\ \ \forall\ 1 \leq q \leq \infty. \tag{2.8}$$

An immediate consequence of Theorem 2.1 is the following R-boundedness result for the discrete heat semigroup and discrete resolvent operator, which has important application in deriving the time-discrete maximal ℓ^p-regularity of the fully discrete finite element solutions discretized with backward Euler, Crank-Nicolson, second-order BDF and A-stable Runge-Kutta schemes (cf. [22, Section 6]).

Corollary 2.1 (R-boundedness of the discrete resolvent) *Under the assumptions of Theorem 2.1, for any $1 < q < \infty$ there exists $\theta_q > 0$ (independent of h) such that*

(1) *The semigroup of operators $\{E_h(z) : z \in \Sigma_{\theta_q}\}$ is R-bounded in $\mathcal{L}(L^q, L^q)$ (the space of bounded linear operators on L^q), and the R-bound is independent of h.*

(2) *The collection of finite element resolvent operators $\{z(z-\Delta_h)^{-1} : z \in \Sigma_{\frac{\pi}{2}+\theta_q}\}$ is R-bounded in $\mathcal{L}(L^q, L^q)$, and the R-bound is independent of h.*

Proof. It is easy to see that the maximal semigroup estimate (2.4) implies the maximal ergodic estimate

$$\left\| \sup_{t>0} \frac{1}{t} \int_0^t |E_h(s)||v| \mathrm{d}s \right\|_{L^q} \le C_q \|v\|_{L^q}, \qquad \forall\, 1 < q \le \infty. \tag{2.9}$$

According to [50, Lemma 4.c], for $q \in (1, 2]$ the above maximal ergodic estimate implies the R-boundedness of the semigroup of operators $\{E_h(z)\}_{z\in\Sigma_{\theta_q}}$ in $\mathcal{L}(L^q, L^q)$ with $\theta_q = (\theta - \epsilon)q/2$, where ϵ can be arbitrarily small. For $q \in [2, \infty)$, a duality argument shows that the semigroup $\{E_h(z)\}_{z\in\Sigma_{\theta_q}}$ is R-bounded in $\mathcal{L}(L^q, L^q)$ with angle $\theta_q = (\theta - \epsilon)q'/2$ (cf. [50, Proof of Lemma 4.d]).

The second statement in Corollary 2.1 is actually a consequence of the first statement (cf. [49, Theorem 4.2]). □

Recall that the L^2 projection $P_h : L^2(\Omega) \to S_h$ and Ritz projection $R_h : H^1_0(\Omega) \to S_h$ onto the finite element spaces are defined by

$$(P_h\phi, \varphi_h) = (\phi, \varphi_h), \qquad \forall\, \phi \in L^2(\Omega),\ \forall\, \varphi_h \in S_h, \tag{2.10}$$

$$(\nabla R_h\phi, \nabla\varphi_h) = (\nabla\phi, \nabla\varphi_h), \qquad \forall\, \phi \in H^1_0(\Omega),\ \forall\, \varphi_h \in S_h. \tag{2.11}$$

In particular, the L^2 projection actually can be extended to $L^q(\Omega)$, $1 \le q \le \infty$, satisfying the following estimate:

$$\|P_h\phi\|_{L^q} \le C\|\phi\|_{L^q}, \quad \forall\, \phi \in L^q(\Omega),\ 1 \le q \le \infty, \tag{2.12}$$

where the constant C is independent of the mesh size h. The estimate above is a consequence of [46, Lemma 6.1] and the self-adjointness of P_h; also see [19], [48, Lemma 7.2] and the properties of the finite element spaces stated in Section 3.2.

Besides Corollary 2.1, the maximal L^p-regularity results (2.5)-(2.6) also imply the following sharp $L^p(0, T; L^q)$ error estimates for finite element solutions of parabolic equations.

Corollary 2.2 *Let u and u_h be the solutions of (1.1) and (1.7), respectively. Then, under the assumptions of Theorem 2.1, we have*

$$\|u_h - P_h u\|_{L^p(0,T;L^q)} \le C_{p,q}(\|u - R_h u\|_{L^p(0,T;L^q)} + \|P_h u(0) - u_h(0)\|_{L^q}),$$
$$\|u_h - P_h u\|_{L^\infty(0,T;L^\infty)} \le C\ell_h(\|u - R_h u\|_{L^\infty(0,T;L^\infty)} + \|P_h u(0) - u_h(0)\|_{L^\infty}),$$

for $1 < p, q < \infty$, where P_h and R_h denote the L^2-projection and Ritz projection onto the finite element space S_h, respectively, and the constants $C_{p,q}$ and C are independent of u and T.

Proof. Let $\phi_h := P_h u - u_h - e^{-t}(P_h u(0) - u_h(0))$. Then ϕ_h satisfies the following operator equation:

$$\begin{cases} \partial_t \phi_h - \Delta_h \phi_h = \Delta_h\big(R_h u - P_h u + e^{-t}(P_h u(0) - u_h(0))\big) + e^{-t}(P_h u(0) - u_h(0)), \\ \phi_h(0) = 0. \end{cases}$$

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

$$\begin{cases} \partial_t(\Delta_h^{-1}\phi_h) - \Delta_h(\Delta_h^{-1}\phi_h) = R_h u - P_h u + (e^{-t} + e^{-t}\Delta_h^{-1})(P_h u(0) - u_h(0)), \\ \Delta_h^{-1}\phi_h(0) = 0. \end{cases} \tag{2.13}$$

By applying (2.6) to the equation above (with $q = \infty$), we have

$$\begin{aligned} \|\phi_h\|_{L^\infty(0,T;L^\infty)} &= \|\Delta_h(\Delta_h^{-1}\phi_h)\|_{L^\infty(0,T;L^\infty)} \\ &\le C\ell_h\|R_h u - P_h u + (e^{-t} + e^{-t}\Delta_h^{-1})(P_h u(0) - u_h(0))\|_{L^\infty(0,T;L^\infty)} \\ &\le C\ell_h(\|R_h u - P_h u\|_{L^\infty(0,T;L^\infty)} + \|P_h u(0) - u_h(0)\|_{L^\infty}), \end{aligned} \tag{2.14}$$

where we have used the following L^∞ estimate of finite element solutions of the Poisson equation (a proof is given in Appendix C)

$$\|\Delta_h^{-1}(P_h u(0) - u_h(0))\|_{L^\infty} \le C\|P_h u(0) - u_h(0)\|_{L^\infty}. \tag{2.15}$$

By using the L^∞ stability of the L^2 projection (i.e., using (2.12) with $q = \infty$), (2.14) further reduces to

$$\begin{aligned} \|\phi_h\|_{L^\infty(0,T;L^\infty)} &\le C\ell_h(\|P_h(R_h u - u)\|_{L^\infty(0,T;L^\infty)} + \|P_h u(0) - u_h(0)\|_{L^\infty}) \\ &\le C\ell_h(\|R_h u - u\|_{L^\infty(0,T;L^\infty)} + \|P_h u(0) - u_h(0)\|_{L^\infty}). \end{aligned} \tag{2.16}$$

This proves the second statement of Corollary 2.2. The first statement of Corollary 2.2 can be proved similarly by applying (2.5) to (2.13). □

One of the advantages of Corollary 2.2 is that it reduces the L^∞ stability of finite element solutions of parabolic equations to the L^∞ stability of the Ritz projection, which immediately implies the following L^∞ stability results in nonconvex polygons and convex polyhedra.

Corollary 2.3 *Under the assumptions of Theorem 2.1, if Ω is a polygon in $\mathbb{R}^2$ (possibly nonconvex) or a convex polyhedra in $\mathbb{R}^3$, and $u_{h,0} = P_h u_0$ or $u_{h,0} = R_h u_0$, then the solutions of (1.1) and (1.7) satisfy*

$$\|u - u_h\|_{L^\infty(0,T;L^\infty)} \leq C\ell_h^2 \inf_{\chi_h} \|u - \chi_h\|_{L^\infty(0,T;L^\infty)}, \tag{2.17}$$

where the constant C is independent of h and T, and the infimum extends over all $\chi_h \in L^\infty(0,T;S_h)$.

Proof. In a two-dimensional polygon (possibly nonconvex) or a convex polyhedra, both the L^2-projection P_h and the Ritz projection R_h have been proved to be stable in the maximum norm (cf. [46, Lemma 6.1] and [28, 42]), i.e.

$$\|u - P_h u\|_{L^\infty} + \|u - R_h u\|_{L^\infty} \leq C\ell_h \inf_{\chi_h \in S_h} \|u - \chi_h\|_{L^\infty}.$$

Hence, Corollary 2.2 and the inequality above imply (2.17). □

In the next section, we introduce the notations to be used in this paper. The proof of Theorem 2.1 is presented in Section 4.

3 Notations

3.1 Function spaces

We use the conventional notations of Sobolev spaces $W^{s,q}(\Omega)$, $s \geq 0$ and $1 \leq q \leq \infty$ (cf. [1]), with the abbreviations $L^q = W^{0,q}(\Omega)$, $W^{s,q} = W^{s,q}(\Omega)$ and $H^s := W^{s,2}(\Omega)$. The notation $H^{-s}(\Omega)$ denotes the dual space of $H_0^s(\Omega)$, the closure of $C_0^\infty(\Omega)$ in $H^s(\Omega)$.

For any given function $f : (0,T) \to W^{s,q}$ we define the Bochner norm

$$\|f\|_{L^p(0,T;W^{s,q})} = \big\| \|f(\cdot)\|_{W^{s,q}} \big\|_{L^p(0,T)}, \quad \forall\, 1 \leq p, q \leq \infty,\ s \in \mathbb{R}. \tag{3.1}$$

For any subdomain $D \subset \Omega$, we define

$$\|f\|_{W^{s,q}(D)} := \inf_{\widetilde{f}|_D = f} \|\widetilde{f}\|_{W^{s,q}(\Omega)}, \quad \forall\, 1 \leq q \leq \infty,\ s \in \mathbb{R}, \tag{3.2}$$

where the infimum extends over all possible $\widetilde{f}$ defined on Ω such that $\widetilde{f} = f$ in D. Similarly, for any subdomain $Q \subset \mathcal{Q} = (0,1) \times \Omega$, we define

$$\|f\|_{L^p W^{s,q}(Q)} := \inf_{\widetilde{f}|_Q = f} \|\widetilde{f}\|_{L^p(0,T;W^{s,q})}, \quad \forall\, 1 \leq p, q \leq \infty,\ s \in \mathbb{R}, \tag{3.3}$$

where the infimum extends over all possible $\widetilde{f}$ defined on $\mathcal{Q}$ such that $\widetilde{f} = f$ in Q.

We use the abbreviations

$$(\phi, \varphi) := \int_\Omega \phi(x)\varphi(x)\mathrm{d}x, \qquad [u, v] := \int_0^T \int_\Omega u(t,x)v(t,x)\mathrm{d}x\,\mathrm{d}t, \tag{3.4}$$

and denote $w(t) = w(t,\cdot)$ for any function w defined on $\mathcal{Q}$. The notation $1_{0<t<T}$ will denote the characteristic function of the time interval $(0,T)$, i.e. $1_{0<t<T}(t) = 1$ if $t \in (0,T)$ while $1_{0<t<T}(t) = 0$ if $t \notin (0,T)$.

3.2 Properties of the finite element space

For any subdomain $D \subset \Omega$, we denote by $S_h(D)$ the space of functions of S_h restricted to the domain D, and denote by $S_h^0(D)$ the subspace of $S_h(D)$ consisting of functions which equal zero outside D. For any given subset $D \subset \Omega$, we denote $B_d(D) = \{x \in \Omega : \mathrm{dist}(x, D) \le d\}$ for $d > 0$. On a quasi-uniform triangulation of the domain Ω, there exist positive constants K and κ such that the triangulation and the corresponding finite element space S_h possess the following properties (K and κ are independent of the subset D and h).

(P1) Quasi-uniformity:

For all triangles (or tetrahedron) τ_l^h in the partition, the diameter h_l of τ_l^h and the radius ρ_l of its inscribed ball satisfy

$$K^{-1}h \le \rho_l \le h_l \le Kh.$$

(P2) Inverse inequality:

If D is a union of elements in the partition, then

$$\|\chi_h\|_{W^{l,p}(D)} \le Kh^{-(l-k)-(N/q-N/p)}\|\chi_h\|_{W^{k,q}(D)}, \quad \forall\, \chi_h \in S_h,$$

for $0 \le k \le l \le 1$ and $1 \le q \le p \le \infty$.

(P3) Local approximation and superapproximation:

There exists an operator $I_h : H_0^1(\Omega) \to S_h$ with the following properties (cf. Appendix B):

(1)

$$\|v - I_h v\|_{L^2} + h\|\nabla(v - I_h v)\|_{L^2} \le Kh^{1+\alpha}\|v\|_{H^{1+\alpha}}, \quad \forall\, v \in H^{1+\alpha}(\Omega) \cap H_0^1(\Omega), \;\; \alpha \in [0,1],$$

(2) If $d \ge 2h$ then the value of $I_h v$ in D depends only on the value of v in $B_d(D)$. If $d \ge 2h$ and $\mathrm{supp}(v) \subset \overline{D}$, then $I_h v \in S_h^0(B_d(D))$.

(3) If $d \ge 2h$, $\omega = 0$ outside D and $|\partial^\beta \omega| \le Cd^{-|\beta|}$ for all multi-index β, then

$$\begin{aligned}
&\psi_h \in S_h(B_d(D)) \implies I_h(\omega\psi_h) \in S_h^0(B_d(D)),\\
&\|\omega\psi_h - I_h(\omega\psi_h)\|_{L^2} + h\|\omega\psi_h - I_h(\omega\psi_h)\|_{H^1} \le Khd^{-1}\|\psi_h\|_{L^2(B_d(D))}.
\end{aligned}$$

(4) If $d \ge 2h$ and $\omega \equiv 1$ on $B_d(D)$, then $I_h(\omega\psi_h) = \psi_h$ on D.

The properties (P1)-(P3) hold for any quasi-uniform triangulation with the standard finite element spaces consisting of globally continuous piecewise polynomials of degree $r \ge 1$ (cf. [45, Appendix]), and have been used in many works in studying the discrete maximal L^p-regularity and maximum-norm stability of finite element solutions of parabolic equations; see [14, 27, 31, 33, 44, 45, 47]. Property (P3)-(1) and Definition (3.2) imply the following estimate for $\alpha \in [0,1]$:

$$\|v - I_h v\|_{L^2(D)} + h\|v - I_h v\|_{H^1(D)} \le Kh^{1+\alpha}\|v\|_{H^{1+\alpha}(B_d(D))}, \quad \forall\, v \in H^{1+\alpha}(B_d(D)) \cap H_0^1(\Omega). \tag{3.5}$$

3.3 Green's functions

For any $x_0 \in \tau_l^h$ (where τ_l^h is a triangle or a tetrahedron in the triangulation of Ω), there exists a function $\widetilde{\delta}_{x_0} \in C^3(\overline{\Omega})$ with support in τ_l^h such that

$$\chi_h(x_0) = \int_\Omega \chi_h \widetilde{\delta}_{x_0} \mathrm{d}x, \quad \forall\, \chi_h \in S_h,$$

and

$$\|\widetilde{\delta}_{x_0}\|_{W^{l,p}} \le K h^{-l-N(1-1/p)} \quad \text{for } 1 \le p \le \infty,\ \ l = 0,1,2,3, \tag{3.6}$$

$$\sup_{y\in\Omega} \int_\Omega |\widetilde{\delta}_y(x)| \mathrm{d}x + \sup_{x\in\Omega} \int_\Omega |\widetilde{\delta}_y(x)| \mathrm{d}y \le C. \tag{3.7}$$

The construction of $\widetilde{\delta}_{x_0}$ can be found in [47, Lemma 2.2].

Let δ_{x_0} denote the Dirac Delta function centered at x_0. In other words, $\int_\Omega \delta_{x_0}(y)\varphi(y)\mathrm{d}y = \varphi(x_0)$ for arbitrary $\varphi \in C(\overline{\Omega})$. Then the discrete Delta function

$$\delta_{h,x_0} := P_h \delta_{x_0} = P_h \widetilde{\delta}_{x_0}$$

decays exponentially away from x_0 (cf. [48, Lemma 7.2]):

$$|\delta_{h,x_0}(x)| = |P_h \widetilde{\delta}_{x_0}(x)| \le K h^{-N} e^{-\frac{|x-x_0|}{Kh}}, \quad \forall\, x, x_0 \in \Omega. \tag{3.8}$$

Let $G(t,x,x_0)$ denote the Green's function of the parabolic equation, i.e. $G = G(\cdot,\cdot,x_0)$ is the solution of

$$\begin{cases} \partial_t G(\cdot,\cdot,x_0) - \Delta G(\cdot,\cdot,x_0) = 0 & \text{in } (0,T)\times\Omega, \\ G(\cdot,\cdot,x_0) = 0 & \text{on } (0,T)\times\partial\Omega, \\ G(0,\cdot,x_0) = \delta_{x_0} & \text{in } \Omega. \end{cases} \tag{3.9}$$

The Green's function $G(t,x,y)$ is symmetric with respect to x and y. It has an analytic extension to the right half-plane, satisfying the following Gaussian estimate (cf. [10, p. 103]):

$$|G(z,x,y)| \le C_\theta |z|^{-\frac{N}{2}} e^{-\frac{|x-y|^2}{C_\theta |z|}}, \quad \forall\, z \in \Sigma_\theta,\ \forall\, x,y \in \Omega, \quad \forall\, \theta \in (0,\pi/2), \tag{3.10}$$

where the constant C_θ depends only on θ. Then Cauchy's integral formula says that

$$\partial_t^k G(t,x,y) = \frac{k!}{2\pi i} \int_{|z-t|=\frac{t}{2}} \frac{G(z,x,y)}{(z-t)^{k+1}} \mathrm{d}z, \tag{3.11}$$

which further yields the following Gaussian pointwise estimate for the time derivatives of Green's function (cf. [14, Appendix B with $\alpha = \beta = 0$]):

$$|\partial_t^k G(t,x,x_0)| \le \frac{C_k}{t^{k+N/2}} e^{-\frac{|x-x_0|^2}{C_k t}}, \quad \forall\, x, x_0 \in \Omega,\ \forall\, t > 0,\ k = 0,1,2,\dots \tag{3.12}$$

Let $\Gamma = \Gamma(\cdot,\cdot,x_0)$ be the regularized Green's function of the parabolic equation, defined by

$$\begin{cases} \partial_t \Gamma(\cdot,\cdot,x_0) - \Delta \Gamma(\cdot,\cdot,x_0) = 0 & \text{in } (0,T)\times\Omega, \\ \Gamma(\cdot,\cdot,x_0) = 0 & \text{on } (0,T)\times\partial\Omega, \\ \Gamma(0,\cdot,x_0) = \widetilde{\delta}_{x_0} & \text{in } \Omega, \end{cases} \tag{3.13}$$

and let $\Gamma_h = \Gamma_h(\cdot,\cdot,x_0)$ be the finite element approximation of Γ, defined by

$$\begin{cases} (\partial_t \Gamma_h(t,\cdot,x_0), v_h) + (\nabla \Gamma_h(t,\cdot,x_0), \nabla v_h) = 0, \quad \forall\, v_h \in S_h,\ t \in (0,T), \\ \Gamma_h(0,\cdot,x_0) = \delta_{h,x_0}. \end{cases} \tag{3.14}$$

By using the Green's function and discrete Green's function, the solutions of (1.1) and (1.7) can be represented by

$$u(t,x_0) = \int_\Omega G(t,x,x_0)u_0(x)\mathrm{d}x + \int_0^t \int_\Omega G(t-s,x,x_0)f(s,x)\mathrm{d}x\mathrm{d}s, \tag{3.15}$$

$$u_h(t,x_0) = \int_\Omega \Gamma_h(t,x,x_0)u_{0,h}(x)\mathrm{d}x + \int_0^t \int_\Omega \Gamma_h(t-s,x,x_0)f(s,x)\mathrm{d}x\mathrm{d}s, \tag{3.16}$$

and

$$(E(t)v)(x_0) = \int_\Omega G(t,x,x_0)v(x)\mathrm{d}x, \qquad (E_h(t)v_h)(x_0) = \int_\Omega \Gamma_h(t,x,x_0)v_h(x)\mathrm{d}x. \tag{3.17}$$

The regularized Green's function can be represented by

$$\Gamma(t,x,x_0) = \int_\Omega G(t,y,x)\widetilde{\delta}_{x_0}(y)\mathrm{d}y = \int_\Omega G(t,x,y)\widetilde{\delta}_{x_0}(y)\mathrm{d}y. \tag{3.18}$$

From the representation (3.18) one can easily derive that the regularized Green's function Γ also satisfies the Gaussian pointwise estimate:

$$|\partial_t^k \Gamma(t,x,x_0)| \le \frac{C_k}{t^{k+N/2}} e^{-\frac{|x-x_0|^2}{C_k t}}, \quad \forall\, x, x_0 \in \Omega,\ \forall\, t > 0 \text{ such that } \max(|x-x_0|, \sqrt{t}) \ge 2h, \tag{3.19}$$

with $k = 0, 1, 2, \ldots$

3.4 Dyadic decomposition of the domain $\mathcal{Q} = (0,1) \times \Omega$

In the proof of Theorem 2.1, we need to partition the domain $\mathcal{Q} = (0,1) \times \Omega$ into subdomains, and present estimates of the finite element solutions in each subdomain. The following dyadic decomposition of $\mathcal{Q}$ was introduced in [44] and has been used by many authors [14, 27, 31, 33, 47]. The readers may pass this subsection if they are familiar with such dyadic decompositions.

For any integer j, we define $d_j = 2^{-j}$. For a given $x_0 \in \Omega$, we let $J_1 = 1$, $J_0 = 0$ and J_* be an integer satisfying $2^{-J_*} = C_* h$ with $C_* \ge 16$ to be determined later. If

$$h < 1/(4C_*), \tag{3.20}$$

then

$$2 \le J_* = \log_2[1/(C_* h)] \le \log_2(2 + 1/h). \tag{3.21}$$

Let

$$Q_*(x_0) = \{(x,t) \in \Omega_T : \max(|x-x_0|, t^{1/2}) \le d_{J_*}\},$$
$$\Omega_*(x_0) = \{x \in \Omega : |x-x_0| \le d_{J_*}\}.$$

We define

$$\begin{aligned} &Q_j(x_0) = \{(x,t) \in \Omega_T : d_j \le \max(|x-x_0|, t^{1/2}) \le 2d_j\} && \text{for } j \ge 1, \\ &\Omega_j(x_0) = \{x \in \Omega : d_j \le |x-x_0| \le 2d_j\} && \text{for } j \ge 1, \\ &D_j(x_0) = \{x \in \Omega : |x-x_0| \le 2d_j\} && \text{for } j \ge 1, \end{aligned}$$

and

$$Q_0(x_0) = \mathcal{Q} \backslash \big(\cup_{j=1}^{J_*} Q_j(x_0) \cup Q_*(x_0) \big),$$
$$\Omega_0(x_0) = \Omega \backslash \big(\cup_{j=1}^{J_*} \Omega_j(x_0) \cup \Omega_*(x_0) \big).$$

For $j < 0$, we simply define $Q_j(x_0) = \Omega_j(x_0) = \emptyset$. For all integer $j \ge 0$, we define

$$\begin{aligned} &\Omega_j'(x_0) = \Omega_{j-1}(x_0) \cup \Omega_j(x_0) \cup \Omega_{j+1}(x_0), \quad Q_j'(x_0) = Q_{j-1}(x_0) \cup Q_j(x_0) \cup Q_{j+1}(x_0), \\ &\Omega_j''(x_0) = \Omega_{j-2}(x_0) \cup \Omega_j'(x_0) \cup \Omega_{j+2}(x_0), \quad Q_j''(x_0) = Q_{j-2}(x_0) \cup Q_j'(x_0) \cup Q_{j+2}(x_0), \\ &D_j'(x_0) = D_{j-1}(x_0) \cup D_j(x_0), \quad D_j''(x_0) = D_{j-2}(x_0) \cup D_j'(x_0). \end{aligned}$$

Then we have

$$\Omega_T = \bigcup_{j=0}^{J_*} Q_j(x_0) \cup Q_*(x_0) \quad \text{and} \quad \Omega = \bigcup_{j=0}^{J_*} \Omega_j(x_0) \cup \Omega_*(x_0). \tag{3.22}$$

We refer to $Q_*(x_0)$ as the "innermost" set. We shall write $\sum_{*,j}$ when the innermost set is included and $\sum_j$ when it is not. When x_0 is fixed, if there is no ambiguity, we simply write $Q_j = Q_j(x_0)$, $Q_j' = Q_j'(x_0)$, $Q_j'' = Q_j''(x_0)$, $\Omega_j = \Omega_j(x_0)$, $\Omega_j' = \Omega_j'(x_0)$ and $\Omega_j'' = \Omega_j''(x_0)$.

We shall use the notations

$$\|v\|_{k,D} = \left(\int_D \sum_{|\alpha| \le k} |\partial^\alpha v|^2 \mathrm{d}x \right)^{\frac{1}{2}}, \qquad \|\!|v|\!\|_{k,Q} = \left(\int_Q \sum_{|\alpha| \le k} |\partial^\alpha v|^2 \mathrm{d}x\mathrm{d}t \right)^{\frac{1}{2}}, \tag{3.23}$$

for any subdomains $D \subset \Omega$ and $Q \subset (0,1) \times \Omega$. Throughout this paper, we denote by C a generic positive constant that is independent of h, x_0 and C_* (until C_* is determined in Section 5). To simplify the notations, we also denote $d_* = d_{J_*}$.

4 Proof of Theorem 2.1

4.1 Estimates of the Green's function

In this subsection, we prove the following local $L^2H^{1+\alpha}$ and $L^\infty H^{1+\alpha}$ estimates for the Green's function and regularized Green's function. These local estimates are needed in the proof of Theorem 2.1.

Lemma 4.1 *Let Ω be a polygon in $\mathbb{R}^2$ or a polyhedron in $\mathbb{R}^3$ (possibly nonconvex). There exists $\alpha \in (\frac{1}{2}, 1]$ and $C > 0$, independent of h and x_0, such that the Green's function G defined in (3.9) and the regularized Green's function Γ defined in (3.13) satisfy the following estimates:*

$$\begin{aligned}
&d_j^{-4-\alpha+N/2}\|\Gamma(\cdot,\cdot,x_0)\|_{L^\infty(Q_j(x_0))} + d_j^{-4-\alpha}|||\nabla\Gamma(\cdot,\cdot,x_0)|||_{L^2(Q_j(x_0))}\\
&+ d_j^{-4}|||\Gamma(\cdot,\cdot,x_0)|||_{L^2H^{1+\alpha}(Q_j(x_0))} + d_j^{-2}|||\partial_t\Gamma(\cdot,\cdot,x_0)|||_{L^2H^{1+\alpha}(Q_j(x_0))}\\
&+ |||\partial_{tt}\Gamma(\cdot,\cdot,x_0)|||_{L^2H^{1+\alpha}(Q_j(x_0))} \le Cd_j^{-N/2-4-\alpha},
\end{aligned} \tag{4.1}$$

$$\|G(\cdot,\cdot,x_0)\|_{L^\infty H^{1+\alpha}(\cup_{k\le j}Q_k(x_0))} + d_j^2\|\partial_t G(\cdot,\cdot,x_0)\|_{L^\infty H^{1+\alpha}(\cup_{k\le j}Q_k(x_0))} \le Cd_j^{-N/2-1-\alpha}. \tag{4.2}$$

To prove Lemma 4.1, we need to use the following two lemmas.

Lemma 4.2 *Let Ω be a polygon in $\mathbb{R}^2$ or a polyhedron in $\mathbb{R}^3$ (possibly nonconvex). Then there exists a positive constant $\alpha \in (\frac{1}{2}, 1]$ (depending on the domain Ω) such that the solution of the Poisson equation*

$$\begin{cases} \Delta\varphi = f & \text{in } \Omega,\\ \varphi = 0 & \text{on } \partial\Omega, \end{cases}$$

satisfies

$$\|\varphi\|_{H^{1+\alpha}} \le C\|f\|_{H^{-1+\alpha}}.$$

Lemma 4.2 is a consequence of [9, Theorem 18.13], where either "$n_x = 2$ (x is an edge point) and $\mu_1(\Gamma_x) > 1/2$ (first eigenvalue of the Laplacian in a 2D sector)" or "$n_x = 3$ (x is a vertex) and $\mu_1(\Gamma_x) > 0$ (first eigenvalue of the Laplacian in a 3D cone)". Such regularity also holds for the Neumann Laplacian (cf. [9, Corollary 23.5]).

Lemma 4.3 *Let Ω be a polygon in $\mathbb{R}^2$ or a polyhedron in $\mathbb{R}^3$ (possibly nonconvex). Then there exists $\alpha \in (\frac{1}{2}, 1]$ such that*

$$\|u\|_{H^{1+\alpha}} \le C\|\nabla u\|_{L^2}^{1-\alpha}\|\Delta u\|_{L^2}^{\alpha}, \quad \forall\, u \in H_0^1 \text{ such that } \Delta u \in L^2. \tag{4.3}$$

Proof. Let ϕ_j, $j = 1, 2, \dots$, be the orthornormal eigenfunctions of the Dirichlet Laplacian $-\Delta$, corresponding to the eigenvalues $\lambda_j > 0$, $j = 1, 2, \dots$, respectively. With these notations, if $u = \sum_{j=1}^\infty a_j\phi_j$ then

$$(-\Delta)^{\alpha/2}u := \sum_{j=1}^\infty \lambda_j^{\alpha/2}a_j\phi_j, \qquad \alpha \in [0, 2].$$

The norm $\|(-\Delta)^{\alpha/2}u\|_{L^2} := \left(\sum_{j=1}^\infty \lambda_j^\alpha a_j^2\right)^{\frac{1}{2}}$ can be viewed as the a weighted ℓ^2 norm of the sequence $(a_1, a_2, \cdots)$. Since $\|(-\Delta)^0 v\|_{L^2} = \|v\|_{L^2}$ and $C^{-1}\|v\|_{H^1} \le \|(-\Delta)^{1/2}v\|_{L^2} \le C\|v\|_{H^1}$, the complex interpolation method (cf. [5, Theorem 5.4.1]) yields the following equivalence of norms:

$$C^{-1}\|v\|_{H^{1-\alpha}(\Omega)} \le \|(-\Delta)^{(1-\alpha)/2}v\|_{L^2(\Omega)} \le C\|v\|_{H^{1-\alpha}(\Omega)}, \qquad \forall\, v \in H^{1-\alpha}(\Omega),\ \forall\, \alpha \in (\tfrac{1}{2}, 1].$$

Hence, we have

$$\begin{aligned}|(-\Delta u, v)| &= |((-\Delta)^{(1+\alpha)/2}u, (-\Delta)^{(1-\alpha)/2}v)| \\ &\le \|(-\Delta)^{(1+\alpha)/2}u\|_{L^2}\|(-\Delta)^{(1-\alpha)/2}v\|_{L^2} \\ &\le C\|(-\Delta)^{(1+\alpha)/2}u\|_{L^2}\|v\|_{H^{1-\alpha}}, \qquad \forall\, v\in C_0^\infty(\Omega),\end{aligned}$$

which implies (via the duality argument)

$$\begin{aligned}\|\Delta u\|_{H^{-1+\alpha}}^2 &\le C\|(-\Delta)^{(1+\alpha)/2}u\|_{L^2}^2 \\ &= \sum_j \lambda_j^{1+\alpha}a_j^2 \\ &= \sum_j (\lambda_j a_j^2)^{1-\alpha}(\lambda_j^2 a_j^2)^{\alpha} \\ &\le \Big(\sum_j \lambda_j a_j^2\Big)^{1-\alpha}\Big(\sum_j \lambda_j^2 a_j^2\Big)^{\alpha} \qquad \text{(Hölder's inequality)} \\ &= \|(-\Delta)^{1/2}u\|_{L^2}^{2(1-\alpha)}\|\Delta u\|_{L^2}^{2\alpha} \\ &= \|\nabla u\|_{L^2}^{2(1-\alpha)}\|\Delta u\|_{L^2}^{2\alpha}.\end{aligned}$$

Lemma 4.2 implies the existence of $\alpha\in(\frac12,1]$ such that $\|u\|_{H^{1+\alpha}}\le C\|\Delta u\|_{H^{-1+\alpha}}$ for all $u\in H_0^1$ such that $\Delta u\in L^2$. This (together with the inequality above) yields Lemma 4.3. ∎

Proof of Lemma 4.1. To simplify the notations, we denote $Q_j = Q_j(x_0)$. Let $0\le\omega_j(x,t)\le 1$ and $0\le\widetilde\omega_j(x,t)\le 1$ be smooth cut-off functions vanishing outside Q_j' and equals 1 in Q_j, such that $\widetilde\omega_j = 1$ on the support of ω_j, and

$$|\partial_t^{k_1}\nabla^{k_2}\omega_j| + |\partial_t^{k_1}\nabla^{k_2}\widetilde\omega_j| \le Cd_j^{-2k_1-k_2} \tag{4.4}$$

for all nonnegative integers k_1 and k_2. By the definition in (3.3), it suffices to prove the corresponding global estimates for the function $\omega_1 G$, which equals G in Q_j.

Consider $\omega_j G$, which is the solution of

$$\partial_t(\omega_j G) - \Delta(\omega_j G) = \widetilde\omega_j G\partial_t\omega_j + \widetilde\omega_j G\Delta\omega_j - \nabla\cdot(2\widetilde\omega_j G\nabla\omega_j) \tag{4.5}$$

in the domain $(0,\infty)\times\Omega$, with zero boundary and initial conditions. The standard energy estimate yields (cf. [26, Lemma 2.1 of Chapter III], with $q_1 = r_1 = (2N+4)/(N+4)$)

$$\begin{aligned}&\|\omega_j G\|_{L^\infty(0,T;L^2)} + \|\omega_j G\|_{L^2(0,T;H^1)} \\ &\le C\|\widetilde\omega_j G\partial_t\omega_j\|_{L^{(2N+4)/(N+4)}(\mathcal{Q})} + C\|\widetilde\omega_j G\Delta\omega_j\|_{L^{(2N+4)/(N+4)}(\mathcal{Q})} + C\|2\widetilde\omega_j G\nabla\omega_j\|_{L^2(0,T;L^2)} \\ &\le Cd_j^{-2}\|G\|_{L^{(2N+4)/(N+4)}(Q_j')} + Cd_j^{-1}\|G\|_{L^2(Q_j')} \\ &\le Cd_j^{-N/2},\end{aligned} \tag{4.6}$$

where we have used the Gaussian estimate (3.12) in the last step. The last inequality implies

$$\|G\|_{L^\infty L^2(Q_j)} + \|G\|_{L^2H^1(Q_j)} \le Cd_j^{-N/2}, \tag{4.7}$$

and

$$\begin{aligned}\|\nabla\cdot(2\widetilde{\omega}_jG\nabla\omega_j)\|_{L^2(\mathcal{Q})}&\le C\|G\Delta\omega_j\|_{L^2(\mathcal{Q})}+C\|\nabla G\cdot\nabla\omega_j\|_{L^2(\mathcal{Q})}\\&\le Cd_j^{-2}\|G\|_{L^2(Q_j')}+Cd_j^{-1}\|\nabla G\|_{L^2(Q_j')}\\&\le Cd_j^{-1}\|G\|_{L^\infty L^2(Q_j')}+Cd_j^{-1}\|\nabla G\|_{L^2(Q_j')}\\&\le Cd_j^{-1-N/2},\end{aligned}\tag{4.8}$$

where we have used (4.7) in the last inequality (replacing Q_j by Q_j'). By applying the energy estimate to (4.5), we have

$$\begin{aligned}&\|\partial_t(\omega_jG)\|_{L^2(0,T;L^2)}+\|\Delta(\omega_jG)\|_{L^2(0,T;L^2)}\\&\le C\|\widetilde{\omega}_jG\partial_t\omega_j\|_{L^2(\mathcal{Q})}+C\|\widetilde{\omega}_jG\Delta\omega_j\|_{L^2(\mathcal{Q})}+C\|\nabla\cdot(2\widetilde{\omega}_jG\nabla\omega_j)\|_{L^2(\mathcal{Q})}\\&\le Cd_j^{-1-N/2},\end{aligned}\tag{4.9}$$

where we have used (3.12), (4.4) and (4.8) in the last step.

Lemma 4.3 implies the existence of $\alpha\in(\frac{1}{2},1]$ (depending on the domain Ω) such that

$$\begin{aligned}\|\omega_jG\|_{L^2(0,T;H^{1+\alpha})}&\le C\|\nabla(\omega_jG)\|_{L^2(0,T;L^2)}^{1-\alpha}\|\Delta(\omega_jG)\|_{L^2(0,T;L^2)}^{\alpha}\\&\le Cd_j^{-\alpha-N/2},\end{aligned}\tag{4.10}$$

where we have used (4.6) and (4.9) in the last step.

Similarly (replacing G by ∂_tG and $\partial_{tt}G$ in the estimates above), one can prove the following estimates:

$$d_j^{-\alpha}\|\nabla(\omega_j\partial_tG)\|_{L^2(0,T;L^2)}+d_j^{1-\alpha}\|\Delta(\omega_j\partial_tG)\|_{L^2(0,T;L^2)}+\|\omega_j\partial_tG\|_{L^2(0,T;H^{1+\alpha})}\le Cd_j^{-\alpha-2-N/2},\tag{4.11}$$

$$d_j^{-\alpha}\|\nabla(\omega_j\partial_{tt}G)\|_{L^2(0,T;L^2)}+d_j^{1-\alpha}\|\Delta(\omega_j\partial_{tt}G)\|_{L^2(0,T;L^2)}+\|\omega_j\partial_{tt}G\|_{L^2(0,T;H^{1+\alpha})}\le Cd_j^{-\alpha-4-N/2}.\tag{4.12}$$

By using (3.12) and (4.4), the last two inequalities imply

$$\|\partial_t(\omega_jG)\|_{L^2(0,T;H^{1+\alpha})}\le\|\partial_t\omega_jG\|_{L^2(0,T;H^{1+\alpha})}+\|\omega_j\partial_tG\|_{L^2(0,T;H^{1+\alpha})}\le Cd_j^{-\alpha-2-N/2},\tag{4.13}$$

$$\begin{aligned}\|\partial_{tt}(\omega_jG)\|_{L^2(0,T;H^{1+\alpha})}&\le\|\partial_{tt}\omega_jG\|_{L^2(0,T;H^{1+\alpha})}+2\|\partial_t\omega_j\partial_tG\|_{L^2(0,T;H^{1+\alpha})}+\|\omega_j\partial_{tt}G\|_{L^2(0,T;H^{1+\alpha})}\\&\le Cd_j^{-\alpha-4-N/2}.\end{aligned}\tag{4.14}$$

The estimates (4.7) and (4.10)-(4.14) imply

$$\begin{aligned}&d_j^{-4-\alpha+N/2}\|G(\cdot,\cdot,x_0)\|_{L^\infty(Q_j(x_0))}+d_j^{-4-\alpha}|\!|\!|\nabla G(\cdot,\cdot,x_0)|\!|\!|_{L^2(Q_j(x_0))}\\&+d_j^{-4}|\!|\!|G(\cdot,\cdot,x_0)|\!|\!|_{L^2H^{1+\alpha}(Q_j(x_0))}+d_j^{-2}|\!|\!|\partial_tG(\cdot,\cdot,x_0)|\!|\!|_{L^2H^{1+\alpha}(Q_j(x_0))}\\&+|\!|\!|\partial_{tt}G(\cdot,\cdot,x_0)|\!|\!|_{L^2H^{1+\alpha}(Q_j(x_0))}\le Cd_j^{-N/2-4-\alpha}.\end{aligned}\tag{4.15}$$

The estimate (4.15) can also be proved for the regularized Green's function Γ by using the following expression:

$$\Gamma(t,x,x_0) = \int_{\tau_l^h} G(t,y,x)\widetilde{\delta}_{x_0}(y)\mathrm{d}y = \int_{\tau_l^h} G(t,x,y)\widetilde{\delta}_{x_0}(y)\mathrm{d}y, \tag{4.16}$$

where τ_l^h is the triangle/tetrahedron containing x_0 (thus $\widetilde{\delta}_{x_0}$ is supported in τ_l^h). For example, if $y \in \tau_l^h$ then $(t,x) \in Q_j(x_0) \Rightarrow (t,x) \in Q_j'(y)$, which implies

$$\begin{aligned}
\|\Gamma(\cdot,\cdot,x_0)\|_{L^2H^{1+\alpha}(Q_j(x_0))} &= \left\| \int_{\tau_l^h} G(\cdot,\cdot,y)\widetilde{\delta}_{x_0}(y)\mathrm{d}y \right\|_{L^2H^{1+\alpha}(Q_j(x_0))} \\
&\leq \int_{\tau_l^h} \|G(\cdot,\cdot,y)\|_{L^2H^{1+\alpha}(Q_j(x_0))}|\widetilde{\delta}_{x_0}(y)|\mathrm{d}y \\
&\leq \int_{\tau_l^h} \|G(\cdot,\cdot,y)\|_{L^2H^{1+\alpha}(Q_j'(y))}|\widetilde{\delta}_{x_0}(y)|\mathrm{d}y \\
&\leq \int_{\Omega} Cd_j^{-\alpha-N/2}|\widetilde{\delta}_{x_0}(y)|\mathrm{d}y \\
&\leq Cd_j^{-\alpha-N/2}.
\end{aligned}$$

This completes the proof of (4.1).

From (4.6) and (4.9) we see that

$$\|\nabla G\|_{L^2(\cup_{k\leq j}Q_k')} \leq \sum_{k\leq j} \|\nabla G\|_{L^2(Q_k')} \leq C\sum_{k\leq j} d_k^{-N/2} \leq Cd_j^{-N/2}, \tag{4.17}$$

$$\|\Delta G\|_{L^2(\cup_{k\leq j}Q_k')} \leq \sum_{k\leq j} \|\Delta G\|_{L^2(Q_k')} \leq C\sum_{k\leq j} d_k^{-1-N/2} \leq Cd_j^{-1-N/2}, \tag{4.18}$$

Let χ_j be a smooth cut-off function which equals 1 on $\cup_{k\leq j}Q_k$ and equals zero outside $\cup_{k\leq j}Q_k'$, satisfying $|\partial_t^l\nabla^m\chi_j| \leq Cd_j^{-2l-m}$ for all nonnegative integers l and m. Then χ_jG is a function defined on $\mathcal{Q}$ and equals G on $\cup_{k\leq j}Q_k$. The inequalities (3.12) and (4.17)-(4.18) imply

$$\|\nabla(\chi_jG)\|_{L^2(0,T;L^2)} \leq Cd_j^{-N/2}, \qquad \|\Delta(\chi_jG)\|_{L^2(0,T;L^2)} \leq Cd_j^{-1-N/2}. \tag{4.19}$$

Then Lemma 4.3 implies

$$\|\chi_jG\|_{L^2(0,T;H^{1+\alpha})} \leq C\|\nabla(\chi_jG)\|_{L^2(0,T;L^2)}^{1-\alpha}\|\Delta(\chi_jG)\|_{L^2(0,T;L^2)}^{\alpha} \leq Cd_j^{-\alpha-N/2}, \tag{4.20}$$

Similarly one can prove (by using (4.11)-(4.12))

$$\|\partial_t(\chi_jG)\|_{L^2(0,T;H^{1+\alpha})} \leq C\|\nabla\partial_t(\chi_jG)\|_{L^2(0,T;L^2)}^{1-\alpha}\|\Delta\partial_t(\chi_jG)\|_{L^2(0,T;L^2)}^{\alpha} \leq Cd_j^{-\alpha-2-N/2}, \tag{4.21}$$

$$\|\partial_{tt}(\chi_jG)\|_{L^2(0,T;H^{1+\alpha})} \leq C\|\nabla\partial_{tt}(\chi_jG)\|_{L^2(0,T;L^2)}^{1-\alpha}\|\Delta\partial_{tt}(\chi_jG)\|_{L^2(0,T;L^2)}^{\alpha} \leq Cd_j^{-\alpha-4-N/2}. \tag{4.22}$$

Hence, the interpolation between the last two inequalities yield

$$\|\chi_j G\|_{L^\infty(0,T;H^{1+\alpha})} \leq \|\chi_j G\|_{L^2(0,T;H^{1+\alpha})}^{1/2} \|\partial_t(\chi_j G)\|_{L^2(0,T;H^{1+\alpha})}^{1/2} \leq C d_j^{-\alpha-1-N/2}, \tag{4.23}$$

$$\|\partial_t(\chi_j G)\|_{L^\infty(0,T;H^{1+\alpha})} \leq \|\partial_t(\chi_j G)\|_{L^2(0,T;H^{1+\alpha})}^{1/2} \|\partial_{tt}(\chi_j G)\|_{L^2(0,T;H^{1+\alpha})}^{1/2} \leq C d_j^{-\alpha-3-N/2}. \tag{4.24}$$

This completes the proof of (4.2). □

Besides Lemma 4.1, we also need the following lemma in the proof of Theorem 2.1. The proof of this lemma is deferred to Section 5.

Lemma 4.4 *Under the assumptions of Theorem 2.1, the functions* $\Gamma_h(t,x,x_0)$, $\Gamma(t,x,x_0)$ *and* $F(t,x,x_0) := \Gamma_h(t,x,x_0) - \Gamma(t,x,x_0)$ *satisfy*

$$\sup_{t\in(0,\infty)} (\|\Gamma_h(t,\cdot,x_0)\|_{L^1(\Omega)} + t\|\partial_t\Gamma_h(t,\cdot,x_0)\|_{L^1(\Omega)}) \leq C, \tag{4.25}$$

$$\sup_{t\in(0,\infty)} (\|\Gamma(t,\cdot,x_0)\|_{L^1(\Omega)} + t\|\partial_t\Gamma(t,\cdot,x_0)\|_{L^1(\Omega)}) \leq C, \tag{4.26}$$

$$\|\partial_t F(\cdot,\cdot,x_0)\|_{L^1((0,\infty)\times\Omega)} + \|t\partial_{tt}F(\cdot,\cdot,x_0)\|_{L^1((0,\infty)\times\Omega)} \leq C, \tag{4.27}$$

$$\|\partial_t\Gamma_h(t,\cdot,x_0)\|_{L^1} \leq Ce^{-\lambda_0 t}, \qquad \forall\, t \geq 1, \tag{4.28}$$

where the constants C *and* λ_0 *are independent of* h*.*

4.2 Proof of (2.3)-(2.4)

By denoting

$$\Gamma_h(t) = \Gamma_h(t,\cdot,x_0), \qquad \Gamma(t) = \Gamma(t,\cdot,x_0) \quad \text{and} \quad F(t) = \Gamma_h(t) - \Gamma(t),$$

and using the Green's function representation (3.17), we have

$$\begin{aligned}(E_h(t)v_h)(x_0) &= (\Gamma_h(t), v_h) = (F(t), v_h) + (\Gamma(t), v_h)\\ &= \int_0^t (\partial_s F(s), v_h)\mathrm{d}s + (F(0), v_h) + (\Gamma(t), v_h)\end{aligned}$$

and

$$\begin{aligned}(t\partial_t E_h(t)v_h)(x_0) &= (t\partial_t F(t), v_h) + (t\partial_t\Gamma(t), v_h)\\ &= \int_0^t (s\partial_{ss}F(s) + \partial_s F(s), v_h)\mathrm{d}s + (t\partial_t\Gamma(t), v_h),\end{aligned}$$

with (cf. (3.6)-(3.8) and Lemma 4.4)

$$\|F(0)\|_{L^1} + \|\Gamma(t)\|_{L^1} + \|t\partial_t\Gamma(t)\|_{L^1} \leq C(\|\delta_{h,x_0} - \widetilde{\delta}_{x_0}\|_{L^1} + \|\Gamma(t)\|_{L^1} + \|t\partial_t\Gamma(t)\|_{L^1}) \leq C.$$

By applying Lemma 4.4 to the last two equations, we obtain

$$|(E_h(t)v_h)(x_0)| \leq \big(\|\partial_t F\|_{L^1((0,\infty)\times\Omega)} + \|F(0)\|_{L^1} + \|\Gamma(t)\|_{L^1}\big)\|v_h\|_{L^\infty} \leq C\|v_h\|_{L^\infty},$$

$$|(t\partial_t E_h(t)v_h)(x_0)| \le \big(\|t\partial_{tt}F\|_{L^1((0,\infty)\times\Omega)} + \|\partial_t F\|_{L^1((0,\infty)\times\Omega)} + \|t\partial_t\Gamma(t)\|_{L^1}\big)\|v_h\|_{L^\infty} \le C\|v_h\|_{L^\infty}.$$

This proves (2.3) in the case $q=\infty$. The case $2\le q\le\infty$ follows from the two endpoint cases $q=2$ and $q=\infty$ via interpolation, and the case $1\le q\le 2$ follows from the case $2\le q\le\infty$ via duality (the operators $E_h(t)$ and $\partial_t E_h(t)$ are self-adjoint). The proof of (2.3) is complete.

In order to prove (2.4), we need to construct a symmetrically truncated Green's function (since the regularized Green's function $\Gamma(t,x,x_0)$ may not be symmetric with respect to x and x_0). In fact, there exists a truncated Green's function $G^*_{\rm tr}(t,x,y)$ satisfying the following conditions (cf. [31, 33]):

(1) $G^*_{\rm tr}(t,x,y)$ is symmetric with respect to x and y, namely, $G^*_{\rm tr}(t,x,y) = G^*_{\rm tr}(t,y,x)$.

(2) $G^*_{\rm tr}(\cdot,\cdot,y) = 0$ in $Q_*(y) := \{(t,x)\in\mathcal{Q} : \max(|x-y|,\sqrt{t}) < d_*\}$, and $G^*_{\rm tr}(0,\cdot,y)\equiv 0$ in Ω.

(3) $0\le G^*_{\rm tr}(t,x,y)\le G(t,x,y)$ and $G^*_{\rm tr}(t,x,y) = G(t,x,y)$ when $\max(|x-y|,\sqrt{t}) > 2d_*$,

(4) $|\partial_t G^*_{\rm tr}(t,x,y)|\le Cd_*^{-N-2}$ when $\max(|x-y|,t^{1/2})\le 2d_*$.

Note that for the fixed triangle/tetrahedron τ_l^h and the point $x_0\in\tau_l^h$, the function $\widetilde{\delta}_{x_0}$ is supported in $\tau_l^h\subset\Omega_*(x_0)$ with $\int_\Omega\widetilde{\delta}_{x_0}(y){\rm d}y = 1$. By using Lemma 4.1, there exists $\alpha\in(\frac12,1]$ such that (with $Q_{2*}(y) := \{(t,x)\in\mathcal{Q} : \max(|x-y|,\sqrt{t}) < 2d_*\}$)

$$\begin{aligned}
&\iint_{[(0,\infty)\times\Omega]\backslash Q_{2*}(x_0)} |\partial_t\Gamma(t,x,x_0) - \partial_t G^*_{\rm tr}(\tau,x,x_0)|{\rm d}x{\rm d}t\\
&= \iint_{[(0,\infty)\times\Omega]\backslash Q_{2*}(x_0)} |\partial_t\Gamma(t,x,x_0) - \partial_t G(t,x,x_0)|{\rm d}x{\rm d}t\\
&\le \iint_{[(0,1)\times\Omega]\backslash Q_{2*}(x_0)} \bigg|\int_{\tau_l^h}\partial_t G(t,x,y)\widetilde{\delta}_{x_0}(y){\rm d}y - \int_{\tau_l^h}\partial_t G(t,x,x_0)\widetilde{\delta}_{x_0}(y){\rm d}y\bigg|{\rm d}x{\rm d}t\\
&\quad + \iint_{(1,\infty)\times\Omega} |\partial_t\Gamma(t,x,x_0) - \partial_t G(t,x,x_0)|{\rm d}x{\rm d}t\\
&\le C\iint_{\mathcal{Q}\backslash Q_{2*}(x_0)} h^{\alpha-(N-2)/2}|\partial_t G(t,x,\cdot)|_{C^{\alpha-(N-2)/2}(\overline{\tau}_l^h)}{\rm d}x{\rm d}t\\
&\quad + \iint_{(1,\infty)\times\Omega} |\partial_t\Gamma(t,x,x_0) - \partial_t G(t,x,x_0)|{\rm d}x{\rm d}t\\
&=: \mathcal{I}_1 + \mathcal{I}_2.
\end{aligned} \tag{4.29}$$

By using (3.12) and (3.19) we have

$$\mathcal{I}_2 = \iint_{(1,\infty)\times\Omega} |\partial_t\Gamma(t,x,x_0) - \partial_t G(t,x,x_0)|{\rm d}x{\rm d}t \le C\iint_{(1,\infty)\times\Omega}\frac{C}{t^{1+N/2}}e^{-\frac{|x-x_0|^2}{Ct}}{\rm d}x{\rm d}t \le C.$$

For $(t,x)\in Q_j(x_0)$ and $y\in\overline{\tau}_l^h$, we have $(t,y)\in Q_j'(x)$, which implies that

$$\mathcal{I}_1 \le C\sum_j\iint_{Q_j(x_0)} h^{\alpha-(N-2)/2}|\partial_t G(t,x,\cdot)|_{C^{\alpha-(N-2)/2}(\overline{\tau}_l^h)}{\rm d}x{\rm d}t$$

$$
\begin{aligned}
&\le C\sum_j \iint_{Q_j(x_0)} h^{1+\alpha-N/2}|\partial_t G(t,\cdot,x)|_{C^{\alpha-(N-2)/2}(\overline{Q_j'(x)})}\mathrm{d}x\mathrm{d}t\\
&\le C\sum_j d_j^{N+2}h^{1+\alpha-N/2}\sup_{x\in\Omega}\|\partial_t G(\cdot,\cdot,x)\|_{L^\infty C^{\alpha-(N-2)/2}(Q_j'(x))}\\
&\le C\sum_j d_j^{N+2}h^{1+\alpha-N/2}\sup_{x\in\Omega}\|\partial_t G(\cdot,\cdot,x)\|_{L^\infty H^{1+\alpha}(Q_j'(x))} \qquad (\text{because } H^{1+\alpha}(\Omega)\hookrightarrow C^{\alpha-(N-2)/2}(\overline{\Omega}))\\
&\le C\sum_j d_j^{N+2}h^{1+\alpha-N/2}d_j^{-\alpha-3-N/2} \qquad (\text{here we use Lemma 4.1})\\
&\le C\sum_j (h/d_j)^{1+\alpha-N/2}\\
&\le C,
\end{aligned}
$$

where $\sum_j$ indicates summation over $j=0,1,\dots,J_*$ (see the notations at the end of Section 3.3), and the last inequality is due to the fact that $h2^{J_*}\le C$.

Substituting the estimates of $\mathcal{I}_1$ and $\mathcal{I}_2$ into (4.29) yields

$$
\iint_{[(0,\infty)\times\Omega]\backslash Q_{2*}(x_0)} |\partial_t\Gamma(t,x,x_0)-\partial_t G^*_{\mathrm{tr}}(\tau,x,x_0)|\mathrm{d}x\mathrm{d}t\le C. \tag{4.30}
$$

Furthermore, by using the basic energy estimate, we have

$$
\begin{aligned}
\iint_{Q_{2*}(x_0)} |\partial_t\Gamma(t,x,x_0)|\mathrm{d}x\mathrm{d}t &\le C d_*^{N/2+1}\|\partial_t\Gamma(\cdot,\cdot,x_0)\|_{L^2(\mathcal{Q})}\\
&\le C d_*^{N/2+1}\|\widetilde{\delta}_{x_0}\|_{H^1(\Omega)}\\
&\le C d_*^{N/2+1}h^{-N/2-1}\le CC_*^{N/2+1},
\end{aligned} \tag{4.31}
$$

and (cf. Property (4) of the function $G^*_{\mathrm{tr}}(t,x,x_0)$)

$$
\iint_{Q_{2*}(x_0)} |\partial_t G^*_{\mathrm{tr}}(t,x,x_0)|\mathrm{d}x\mathrm{d}t\le C d_*^{-N-2}d_*^{N+2}\le C.
$$

The last three inequalities imply $\iint_{(0,\infty)\times\Omega}|\partial_t\Gamma(t,x,x_0)-\partial_t G_{\mathrm{tr}}(t,x,x_0)|\mathrm{d}x\mathrm{d}t\le C$, which together with Lemma 4.4 further implies

$$
\iint_{(0,\infty)\times\Omega} |\partial_t\Gamma_h(t,x,x_0)-\partial_t G_{\mathrm{tr}}(t,x,x_0)|\mathrm{d}x\mathrm{d}t\le C. \tag{4.32}
$$

Since both $\Gamma_h(t,x,y)$ and $G^*_{\mathrm{tr}}(t,x,y)$ are symmetric with respect to x and y, from the last inequality we see that the kernel

$$
K(x,y):=\int_0^\infty |\partial_t\Gamma_h(t,x,y)-\partial_t G^*_{\mathrm{tr}}(t,x,y)|\mathrm{d}t
$$

is symmetric with respect to x and y, and satisfies

$$
\sup_{y\in\Omega}\int_\Omega K(x,y)\mathrm{d}x+\sup_{x\in\Omega}\int_\Omega K(x,y)\mathrm{d}y\le C.
$$

By Schur's lemma [23, Lemma 1.4.5], the operator M_K defined by

$$M_K v(x) := \int_\Omega K(x,y)v(y)\mathrm{d}y \tag{4.33}$$

is bounded on L^q for all $1 \le q \le \infty$, i.e.

$$\|M_K v\|_{L^q} \le C\|v\|_{L^q}, \quad \forall\, 1 \le q \le \infty. \tag{4.34}$$

Then we have

$$\begin{aligned}
&\sup_{t>0}(|E_h(t)|\,|v|)(x_0)\\
&= \sup_{t>0}\left|\int_\Omega |\Gamma_h(t,x,x_0)|\,|v(x)|\mathrm{d}x\right|\\
&\le \sup_{t>0}\int_\Omega |\Gamma_h(t,x,x_0) - G^*_{\mathrm{tr}}(t,x,x_0)|\,|v(x)|\mathrm{d}x + \sup_{t>0}\int_\Omega |G^*_{\mathrm{tr}}(t,x,x_0)|\,|v(x)|\mathrm{d}x\\
&\le \sup_{t>0}\int_\Omega \bigg(|\Gamma_h(0,x,x_0) - G^*_{\mathrm{tr}}(0,x,x_0)| + \int_0^t |\partial_t(\Gamma_h(s,x,x_0) - G^*_{\mathrm{tr}}(s,x,x_0))|\mathrm{d}s\bigg)|v(x)|\mathrm{d}x\\
&\quad + \sup_{t>0}\int_\Omega |G^*_{\mathrm{tr}}(t,x,x_0)|\,|v(x)|\mathrm{d}x\\
&= \sup_{t>0}\int_\Omega \bigg(|\delta_{h,x_0}(x)| + \int_0^t |\partial_t(\Gamma_h(s,x,x_0) - G^*_{\mathrm{tr}}(s,x,x_0))|\mathrm{d}s\bigg)|v(x)|\mathrm{d}x + \sup_{t>0}\int_\Omega G(t,x,x_0)|u_h(x)|\mathrm{d}x\\
&\le \int_\Omega Kh^{-N}e^{-\frac{|x-x_0|}{Kh}}|v(x)|\mathrm{d}x + (M_K|v|)(x) + \sup_{t>0}(E(t)|v|)(x_0),
\end{aligned} \tag{4.35}$$

where we have used (3.8) and (4.33) in the last step.

From (3.12) we know that $G(t,x,x_0) \le t^{-n/2}\Phi((x-x_0)/\sqrt{t})$ with $\Phi(x) := Ce^{-|x|^2/C}$, which is a radially decreasing and integrable function. Let $\widetilde{u}_h$ denote the zero extension of u_h from Ω to $\mathbb{R}^N$. Then Corollary 2.1.12 of [16] implies

$$\begin{aligned}
\sup_{t>0}(E(t)|v|)(x_0) &= \sup_{t>0}\int_\Omega G(t,x,x_0)|v(x)|\mathrm{d}x\\
&\le \sup_{t>0}\int_{\mathbb{R}^N} t^{-n/2}\Phi((x-x_0)/\sqrt{t})|\widetilde{v}(x)|\mathrm{d}x\\
&\le \|\Phi\|_{L^1(\mathbb{R}^N)}\mathcal{M}|\widetilde{v}|(x_0),
\end{aligned} \tag{4.36}$$

where $\mathcal{M}$ is the Hardy–Littlewood maximal operator. Since the Hardy–Littlewood maximal operator is strong-type (∞,∞) and weak-type $(1,1)$ ([16], Theorem 2.1.6), it follows that (via real interpolation)

$$\|\mathcal{M}|\widetilde{v}|\|_{L^q(\mathbb{R}^N)} \le \frac{Cq}{q-1}\|\widetilde{v}\|_{L^q(\mathbb{R}^N)} = \frac{Cq}{q-1}\|v\|_{L^q(\Omega)}, \quad \forall\, 1 < q \le \infty. \tag{4.37}$$

By substituting (4.34) and (4.36)-(4.37) into (4.35), we obtain (2.4). □

4.3 Proof of (2.5) for $2 \le p = q < \infty$

In this subsection, we prove (2.5) in the simple case $2 \le p = q < \infty$. The general case $1 < p, q < \infty$ will be proved in the next subsection based on the result of this subsection, by using the mathematical tool of singular integral operators.

Let $f_h = P_h f$ and consider the expression

$$\begin{aligned}
&\partial_t u_h(t, x_0) \\
&= \partial_t \int_0^t (E_h(t-s) f_h(s, \cdot))(x_0) \mathrm{d}s \\
&= \int_0^t (\partial_t E_h(t-s) f_h(s, \cdot))(x_0) \mathrm{d}s + f_h(t, x_0) \\
&= \int_0^t \int_\Omega \partial_t F(t-s, x, x_0) f_h(s, x) \mathrm{d}x \mathrm{d}s + \int_0^t \int_\Omega \partial_t \Gamma(t-s, x, x_0) f_h(s, x) \mathrm{d}x \mathrm{d}s + f_h(t, x_0) \\
&=: \mathcal{M}_h f_h + \mathcal{K}_h f_h + f_h, \qquad (4.38)
\end{aligned}$$

where $\mathcal{M}_h$ and $\mathcal{K}_h$ are certain linear operators. By Lemma 4.4 we have

$$\int_0^t \int_\Omega |\partial_t F(t-s, x, x_0)| \mathrm{d}x \mathrm{d}s \le \int_0^\infty \int_\Omega |\partial_t F(t, x, x_0)| \mathrm{d}x \mathrm{d}t \le C, \qquad (4.39)$$

which implies

$$\|\mathcal{M}_h f_h\|_{L^\infty(\mathbb{R}_+; L^\infty)} \le C \|f_h\|_{L^\infty(\mathbb{R}_+; L^\infty)}. \qquad (4.40)$$

Since the classical energy estimate implies

$$\|\mathcal{M}_h f_h\|_{L^2(\mathbb{R}_+; L^2)} \le C \|f_h\|_{L^2(\mathbb{R}_+; L^2)}, \qquad (4.41)$$

the interpolation of the last two inequalities yields

$$\|\mathcal{M}_h f_h\|_{L^q(\mathbb{R}_+; L^q)} \le C \|f_h\|_{L^q(\mathbb{R}_+; L^q)}, \quad \forall\, 2 \le q \le \infty. \qquad (4.42)$$

It remains to prove

$$\|\mathcal{K}_h f_h\|_{L^q(\mathbb{R}_+; L^q)} \le C_q \|f_h\|_{L^q(\mathbb{R}_+; L^q)}, \quad \forall\, 2 \le q < \infty. \qquad (4.43)$$

To this end, we express $\mathcal{K}_h f_h$ as

$$\begin{aligned}
\mathcal{K}_h f_h(t, x_0) &= \int_0^t \int_\Omega \partial_t \Gamma(t-s, x, x_0) f_h(s, x) \mathrm{d}x \mathrm{d}s \\
&= \int_0^t \int_\Omega \int_\Omega \partial_t G(t-s, x, y) \widetilde{\delta}_{x_0}(y) f_h(s, x) \mathrm{d}y \mathrm{d}x \mathrm{d}s \\
&= \int_\Omega \widetilde{\delta}_{x_0}(y) \bigg(\int_0^t \int_\Omega \partial_t G(t-s, x, y) f_h(s, x) \mathrm{d}x \mathrm{d}s \bigg) \mathrm{d}y.
\end{aligned}$$

In view of (3.7), Schur's lemma [23, Lemma 1.4.5] implies

$$\|\mathcal{K}_h f_h(\cdot, t)\|_{L^q} \le C \bigg\| \int_0^t \int_\Omega \partial_t G(t-s, x, \cdot) f_h(x, s) \mathrm{d}x \mathrm{d}s \bigg\|_{L^q}, \quad \forall\, 1 \le q \le \infty,$$

and so

$$
\begin{aligned}
\|\mathcal{K}_h f_h\|_{L^q(\mathbb{R}_+;L^q)} &\leq C\left\| \int_0^t \int_\Omega \partial_t G(t-s,x,y) f_h(x,s)\mathrm{d}x\mathrm{d}s \right\|_{L_t^q(\mathbb{R}_+;L_y^q)} \\
&= C\left\| \partial_t \int_0^t \int_\Omega G(t-s,x,y) f_h(x,s)\mathrm{d}x\mathrm{d}s - f_h(y,t) \right\|_{L_t^q(\mathbb{R}_+;L_y^q)} \\
&= C\|\partial_t W - f_h\|_{L^q(\mathbb{R}_+;L^q)}, \qquad (4.44)
\end{aligned}
$$

where W is the solution of the PDE problem

$$
\begin{cases}
\partial_t W - \Delta W = f_h & \text{in } \mathbb{R}_+ \times \Omega, \\
W = 0 & \text{on } \mathbb{R}_+ \times \partial\Omega, \\
W(0,\cdot) = 0 & \text{in } \Omega,
\end{cases} \qquad (4.45)
$$

which possesses the following maximal L^q-regularity (in view of (1.3)):

$$
\|\partial_t W\|_{L^q(\mathbb{R}_+;L^q)} \leq C_q \|f_h\|_{L^q(\mathbb{R}_+;L^q)}, \quad \forall\, 2 \leq q < \infty. \qquad (4.46)
$$

The last inequality implies (4.43). Then substituting (4.42)-(4.43) into (4.38) yields

$$
\|\partial_t u_h\|_{L^q(\mathbb{R}_+;L^q)} \leq C_q \|f_h\|_{L^q(\mathbb{R}_+;L^q)}, \qquad \forall\ 2 \leq q < \infty, \qquad (4.47)
$$

Since replacing $f_h(t,x)$ by $f_h(t,x)1_{0<t<T}$ does not affect the value of $u_h(t,x)$ for $t \in (0,T)$, the last inequality implies (2.5) for $2 \leq p = q < \infty$.

4.4 Proof of (2.5) for $1 < p, q < \infty$

In the last subsection, we have proved (2.5) for $2 \leq p = q < \infty$ by showing that the operator $\mathcal{E}_h$ defined by

$$
(\mathcal{E}_h f_h)(t,\cdot) := \int_0^t \partial_t E_h(t-s) f_h(s,\cdot)\mathrm{d}s = (\mathcal{M}_h f_h)(t,\cdot) + (\mathcal{K}_h f_h)(t,\cdot) \qquad (4.48)
$$

satisfies

$$
\|\mathcal{E}_h f_h\|_{L^q(\mathbb{R}_+;L^q)} \leq C_q \|f_h\|_{L^q(\mathbb{R}_+;L^q)}, \quad \forall\ 2 \leq q < \infty. \qquad (4.49)
$$

In this subsection, we prove (2.5) for all $1 < p, q < \infty$ via a duality argument and the singular integral operator approach.

In fact, by the same method, one can also prove that the operator $\mathcal{E}_h'$ defined by

$$
(\mathcal{E}_h' f_h)(s,\cdot) = \int_s^\infty \partial_t E_h(t-s) f_h(t,\cdot)\mathrm{d}t \qquad (4.50)
$$

satisfies

$$
\|\mathcal{E}_h' f_h\|_{L^q(\mathbb{R}_+;L^q)} \leq C_q \|f_h\|_{L^q(\mathbb{R}_+;L^q)}, \quad \forall\ 2 \leq q < \infty. \qquad (4.51)
$$

Since $\mathcal{E}_h'$ is the dual of $\mathcal{E}_h$, by duality we have

$$
\|\mathcal{E}_h f_h\|_{L^q(\mathbb{R}_+;L^q)} \leq C_q \|f_h\|_{L^q(\mathbb{R}_+;L^q)}, \quad \forall\ 1 < q \leq 2. \qquad (4.52)
$$

The two inequalities (4.49) and (4.52) can be summarized as

$$\|\mathcal{E}_h f_h\|_{L^q(\mathbb{R}_+;L^q)} \le C_q \|f_h\|_{L^q(\mathbb{R}_+;L^q)}, \quad \forall\ 1<q<\infty. \tag{4.53}$$

Therefore, we have

$$\|\partial_t u_h\|_{L^q(\mathbb{R}_+;L^q)} \le C_q \|f_h\|_{L^q(\mathbb{R}_+;L^q)}, \qquad \forall\ 1<q<\infty. \tag{4.54}$$

Overall, for any fixed $1<q<\infty$ the operator $\mathcal{E}_h$ is bounded on $L^q(\mathbb{R}_+;L^q)$, and $\{E_h(t)\}_{t>0}$ is an analytic semigroup satisfying (see Lemma 4.4):

$$\|\partial_t E_h(t-s)\|_{\mathcal{L}(L^q,L^q)} \le C(t-s)^{-1}, \qquad \forall\ t>s>0, \tag{4.55}$$

$$\|\partial_{tt} E_h(t-s)\|_{\mathcal{L}(L^q,L^q)} \le C(t-s)^{-2}, \qquad \forall\ t>s>0. \tag{4.56}$$

From (4.48) and (4.55)-(4.56) we see that $\mathcal{E}_h$ is an operator-valued singular integral operator whose kernel $\partial_t E_h(t-s)1_{t>s}$ satisfying the Hörmander conditions (cf. [16, condition (4.6.2)]):

$$\sup_{s,s_0\in\mathbb{R}} \int_{|t-s_0|\ge 2|s-s_0|} \|\partial_t E_h(t-s)1_{t>s} - \partial_t E_h(t-s_0)1_{t>s_0}\|_{\mathcal{L}(L^q,L^q)}\,\mathrm{d}t \le C, \tag{4.57}$$

$$\sup_{t,t_0\in\mathbb{R}} \int_{|t_0-s|\ge 2|t-t_0|} \|\partial_t E_h(t-s)1_{t>s} - \partial_t E_h(t_0-s)1_{t_0>s}\|_{\mathcal{L}(L^q,L^q)}\,\mathrm{d}s \le C. \tag{4.58}$$

Under the conditions (4.57)-(4.58), the theory of singular integral operators (cf. [16, Theorem 4.6.1]) says that if $\mathcal{E}_h$ is bounded on $L^q(\mathbb{R}_+;L^q)$ for some $q\in(1,\infty)$ as proved in (4.54), then it is bounded on $L^p(\mathbb{R}_+;L^q)$ for all $p\in(1,\infty)$:

$$\|\mathcal{E}_h f_h\|_{L^p(\mathbb{R}_+;L^q)} \le \max(p\,,\,(p-1)^{-1})\, C_q \|f_h\|_{L^p(\mathbb{R}_+;L^q)}. \tag{4.59}$$

Since replacing $f_h(t,x)$ by $f_h(t,x)1_{0<t<T}$ does not affect the value of $u_h(t,x)$ for $t\in(0,T)$, the last inequality implies (2.5) for all $1<p,q<\infty$.

4.5 Proof of (2.6)

Again, we consider

$$(\mathcal{E}_h f_h)(t,\cdot) = \int_0^t \partial_t E_h(t-s) f_h(s,\cdot)\mathrm{d}s = \int_0^t \int_\Omega \partial_t \Gamma_h(t-s,x,\cdot) f_h(s,x)\mathrm{d}x\mathrm{d}s \tag{4.60}$$

and use the following inequality: for $t\in(0,T)$

$$\begin{aligned}\|\mathcal{E}_h f_h(t,\cdot)\|_{L^\infty} &\le \left(\sup_{x_0\in\Omega}\int_0^t\int_\Omega |\partial_t\Gamma_h(t-s,x,x_0)|\mathrm{d}x\mathrm{d}s\right)\|f_h\|_{L^\infty(0,T;L^\infty)}\\ &\le \left(\sup_{x_0\in\Omega}\int_0^\infty\int_\Omega |\partial_t\Gamma_h(t,x,x_0)|\mathrm{d}x\mathrm{d}t\right)\|f_h\|_{L^\infty(0,T;L^\infty)}\\ &\le \left(\int_0^\infty \sup_{x_0\in\Omega}\int_\Omega |\partial_t\Gamma_h(t,x,x_0)|\mathrm{d}x\mathrm{d}t\right)\|f_h\|_{L^\infty(0,T;L^\infty)},\end{aligned} \tag{4.61}$$

$$
\begin{aligned}
\|\mathcal{E}_h f_h(t,\cdot)\|_{L^1} &\le \left(\int_0^t \sup_{x\in\Omega}\int_\Omega |\partial_t\Gamma_h(t-s,x,x_0)|\mathrm{d}x_0\mathrm{d}s\right)\|f_h\|_{L^\infty(0,T;L^1)}\\
&= \left(\int_0^t \sup_{x_0\in\Omega}\int_\Omega |\partial_t\Gamma_h(t-s,x,x_0)|\mathrm{d}x\mathrm{d}s\right)\|f_h\|_{L^\infty(0,T;L^1)}\\
&\le \left(\int_0^\infty \sup_{x_0\in\Omega}\int_\Omega |\partial_t\Gamma_h(t,x,x_0)|\mathrm{d}x\mathrm{d}t\right)\|f_h\|_{L^\infty(0,T;L^1)}, \qquad (4.62)
\end{aligned}
$$

where we have used the symmetry $\partial_t\Gamma_h(t-s,x,x_0)=\partial_t\Gamma_h(t-s,x_0,x)$, due to the self-adjointness of the operator $E_h(t-s)$. By interpolation between L^∞ and L^1, we get

$$
\|\mathcal{E}_h f_h\|_{L^\infty(0,T;L^q)} \le \left(\int_0^\infty \sup_{x_0\in\Omega}\int_\Omega |\partial_t\Gamma_h(t,x,x_0)|\mathrm{d}x\mathrm{d}t\right)\|f_h\|_{L^\infty(0,T;L^q)}, \quad \forall\, 1\le q\le\infty. \qquad (4.63)
$$

It remains to prove

$$
\int_0^\infty \sup_{x_0\in\Omega}\int_\Omega |\partial_t\Gamma_h(t,x,x_0)|\mathrm{d}x\mathrm{d}t \le C\log(2+1/h). \qquad (4.64)
$$

To this end, we note that $\partial_t\Gamma_h(t,\cdot,x_0)=\Delta_h\Gamma_h(t,\cdot,x_0)=E_h(t)\Delta_h P_h\widetilde{\delta}_{x_0}$. By using (4.25) of Lemma 4.4 and (2.3) (proved in Section 4.2), we have

$$
\|\partial_t\Gamma_h(t,\cdot,x_0)\|_{L^1} \le Ct^{-1}, \qquad (4.65)
$$

$$
\|\partial_t\Gamma_h(t,\cdot,x_0)\|_{L^1} \le C\|\Delta_h P_h\widetilde{\delta}_{x_0}\|_{L^1} \le Ch^{-2}\|P_h\widetilde{\delta}_{x_0}\|_{L^1} \le Ch^{-2}. \qquad (4.66)
$$

The interpolation of the last two inequalities gives $\|\partial_t\Gamma_h(t,\cdot,x_0)\|_{L^1} \le \dfrac{C}{h^{2\theta}t^{1-\theta}}$ for arbitrary $\theta\in(0,1)$, where the constant C is independent of θ. Hence, we have $\displaystyle\int_0^1 \sup_{x_0\in\Omega}\|\partial_t\Gamma(t,\cdot,x_0)\|_{L^1}\mathrm{d}t \le \frac{C}{\theta h^{2\theta}}$ for arbitrary $\theta\in(0,1)$. By choosing $\theta=1/\log(2+1/h)$, we obtain

$$
\int_0^1 \sup_{x_0\in\Omega}\int_\Omega |\partial_t\Gamma(t,x,x_0)|\mathrm{d}x\mathrm{d}t \le C\log(2+1/h). \qquad (4.67)
$$

The estimate (4.28) implies

$$
\int_1^\infty \sup_{x_0\in\Omega}\int_\Omega |\partial_t\Gamma(t,x,x_0)|\mathrm{d}x\mathrm{d}t \le C. \qquad (4.68)
$$

The last two inequalities imply (4.64), and this completes the proof of (2.6).

The proof of Theorem 2.1 is complete (up to the proof of Lemma 4.4). □

5 Proof of Lemma 4.4

In this section we prove Lemma 4.4, which is used in proving Theorem 2.1 in the last section. To this end, we use the following local energy error estimate for finite element solutions of parabolic equations, which extends the existing work [44, Lemma 6.1] and [33, Proposition 3.2] to nonconvex polyhedra without using the superapproximation results of the Ritz projection (cf. [44, Theorem 5.1] and [33, Proposition 3.1], which only hold in convex domains).

Lemma 5.1 *Suppose that* $\phi \in L^2(0,T;H_0^1(\Omega)) \cap H^1(0,T;L^2(\Omega))$ *and* $\phi_h \in H^1(0,T;S_h)$ *satisfy the equation*

$$(\partial_t(\phi-\phi_h),\chi_h) + (\nabla(\phi-\phi_h),\nabla\chi_h) = 0, \quad \forall\, \chi_h \in S_h,\ a.e.\ t>0, \tag{5.1}$$

with $\phi(0)=0$ *in* Ω_j''. *Then, under the assumptions of Theorem 2.1, we have*

$$\begin{aligned}
&|||\partial_t(\phi-\phi_h)|||_{Q_j} + d_j^{-1}|||\phi-\phi_h|||_{1,Q_j} \\
&\le C\epsilon^{-3}\big(I_j(\phi_h(0)) + X_j(I_h\phi-\phi) + d_j^{-2}|||\phi-\phi_h|||_{Q_j'}\big) \\
&\quad + (Ch^{1/2}d_j^{-1/2} + C\epsilon^{-1}hd_j^{-1}+\epsilon)\big(|||\partial_t(\phi-\phi_h)|||_{Q_j'} + d_j^{-1}|||\phi-\phi_h|||_{1,Q_j'}\big),
\end{aligned} \tag{5.2}$$

where

$$\begin{aligned}
I_j(\phi_h(0)) &= \|\phi_h(0)\|_{1,\Omega_j'} + d_j^{-1}\|\phi_h(0)\|_{\Omega_j'}\,, \\
X_j(I_h\phi-\phi) &= d_j|||\partial_t(I_h\phi-\phi)|||_{1,Q_j'} + |||\partial_t(I_h\phi-\phi)|||_{Q_j'} \\
&\quad + d_j^{-1}|||I_h\phi-\phi|||_{1,Q_j'} + d_j^{-2}|||I_h\phi-\phi|||_{Q_j'}\,,
\end{aligned}$$

where $\epsilon \in (0,1)$ *is an arbitrary positive constant, and the positive constant* C *is independent of* h*,* j *and* C_**; the norms* $|||\cdot|||_{k,Q_j'}$ *and* $|||\cdot|||_{k,\Omega_j'}$ *are defined in* (3.23).

The proof of Lemma 5.1 is presented in Appendix A. In the rest of this section, we apply Lemma 5.1 to prove Lemma 4.4 by denoting $\alpha \in (\frac{1}{2},1]$ a fixed constant satisfying Lemma 4.1. The proof consists of three parts. The first part is concerned with estimates for $t\in(0,1)$, where we convert the L^1 estimates on $\mathcal{Q}=(0,1)\times\Omega = Q_*\bigcup\big(\cup_{j=0}^J Q_j\big)$ into weighted L^2 estimates on the subdomains Q_* and Q_j, $j=0,1,\dots,J$. The second part is concerned with estimates for $t\ge 1$, which is a simple consequence of the parabolic regularity. The third part is concerned with the proof of (4.25)-(4.26), which are simple consequences of the results proved in the first two parts.

Part I. First, we present estimates in the domain $\mathcal{Q}=(0,1)\times\Omega$ with the restriction $h<1/(4C_*)$; see (3.20). In this case, the basic energy estimate gives

$$\|\partial_t\Gamma\|_{L^2(\mathcal{Q})} + \|\partial_t\Gamma_h\|_{L^2(\mathcal{Q})} \le C(\|\Gamma(0)\|_{H^1} + \|\Gamma_h(0)\|_{H^1}) \le Ch^{-1-N/2}, \tag{5.3}$$

$$\|\Gamma\|_{L^\infty L^2(\mathcal{Q})} + \|\Gamma_h\|_{L^\infty L^2(\mathcal{Q})} \le C(\|\Gamma(0)\|_{L^2} + \|\Gamma_h(0)\|_{L^2}) \le Ch^{-N/2}, \tag{5.4}$$

$$\|\nabla\Gamma\|_{L^2(\mathcal{Q})} + \|\nabla\Gamma_h\|_{L^2(\mathcal{Q})} \le C(\|\Gamma(0)\|_{L^2} + \|\Gamma_h(0)\|_{L^2}) \le Ch^{-N/2}, \tag{5.5}$$

$$\|\partial_{tt}\Gamma\|_{L^2(\mathcal{Q})} + \|\partial_{tt}\Gamma_h\|_{L^2(\mathcal{Q})} \le C(\|\Delta\Gamma(0)\|_{H^1} + \|\Delta_h\Gamma_h(0)\|_{H^1}) \le Ch^{-3-N/2}, \tag{5.6}$$

$$\|\nabla\partial_t\Gamma\|_{L^2(\mathcal{Q})} + \|\nabla\partial_t\Gamma_h\|_{L^2(\mathcal{Q})} \le C(\|\Delta\Gamma(0)\|_{L^2} + \|\Delta_h\Gamma_h(0)\|_{L^2}) \le Ch^{-2-N/2}, \tag{5.7}$$

where we have used (3.6) and (3.8) to estimate $\Gamma(0)$ and $\Gamma_h(0)$, respectively. Hence, we have

$$|||\Gamma|||_{Q_*} + |||\Gamma_h|||_{Q_*} \le Cd_*\|\Gamma\|_{L^\infty L^2(Q_*)} + Cd_*\|\Gamma_h\|_{L^\infty L^2(Q_*)} \le Cd_* h^{-N/2} \le CC_* h^{1-N/2}. \tag{5.8}$$

Since the volume of Q_j is Cd_j^{2+N}, we can decompose $\|\partial_t F\|_{L^1(\mathcal{Q})} + \|t\partial_{tt}F\|_{L^1(\mathcal{Q})}$ in the following way:

$$\begin{aligned}
&\|\partial_t F\|_{L^1(\mathcal{Q})} + \|t\partial_{tt}F\|_{L^1(\mathcal{Q})} \\
&\le \|\partial_t F\|_{L^1(Q_*)} + \|t\partial_{tt}F\|_{L^1(Q_*)} + \sum_j \big(\|\partial_t F\|_{L^1(Q_j)} + \|t\partial_{tt}F\|_{L^1(Q_j)}\big) \\
&\le Cd_*^{1+N/2}\big(|||\partial_t F|||_{Q_*} + d_*^2|||\partial_{tt}F|||_{Q_*}\big) + \sum_j Cd_j^{1+N/2}\big(|||\partial_t F|||_{Q_j} + d_j^2|||\partial_{tt}F|||_{Q_j}\big) \\
&\le CC_*^{3+N/2} + \mathscr{K},
\end{aligned} \tag{5.9}$$

where we have used (5.3) and (5.6) to estimate $Cd_*^{1+N/2}\big(|||\partial_t F|||_{Q_*} + d_*^2|||\partial_{tt}F|||_{Q_*}\big)$, and introduced the notation

$$\mathscr{K} := \sum_j d_j^{1+N/2}(d_j^{-1}|||F|||_{1,Q_j} + |||\partial_t F|||_{Q_j} + d_j|||\partial_t F|||_{1,Q_j} + d_j^2|||\partial_{tt}F|||_{Q_j}). \tag{5.10}$$

It remains to estimate $\mathscr{K}$. To this end, we set "$\phi_h = \Gamma_h$, $\phi = \Gamma$, $\phi_h(0) = P_h\widetilde{\delta}_{x_0}$ and $\phi(0) = \widetilde{\delta}_{x_0}$" and "$\phi_h = \partial_t\Gamma_h$, $\phi = \partial_t\Gamma$, $\phi_h(0) = \Delta_h P_h\widetilde{\delta}_{x_0}$ and $\phi(0) = \Delta\widetilde{\delta}_{x_0}$" in Lemma 5.1, respectively. Then we obtain

$$\begin{aligned}
d_j^{-1}|||F|||_{1,Q_j} + |||\partial_t F|||_{Q_j} &\le C\epsilon_1^{-3}(\widehat{I_j} + \widehat{X_j} + d_j^{-2}|||F|||_{Q_j'}) \\
&\quad + (Ch^{1/2}d_j^{-1/2} + C\epsilon_1^{-1}hd_j^{-1} + \epsilon_1)\big(d_j^{-1}|||F|||_{1,Q_j'} + |||\partial_t F|||_{Q_j'}\big)
\end{aligned} \tag{5.11}$$

and

$$\begin{aligned}
d_j|||\partial_t F|||_{1,Q_j} + d_j^2|||\partial_{tt}F|||_{Q_j} &\le C\epsilon_2^{-3}(\overline{I_j} + \overline{X_j} + |||\partial_t F|||_{Q_j'}) \\
&\quad + (Ch^{1/2}d_j^{-1/2} + C\epsilon_2^{-1}hd_j^{-1} + \epsilon_2)\big(d_j|||\partial_t F|||_{1,Q_j'} + d_j^2|||\partial_{tt}F|||_{Q_j'}\big),
\end{aligned} \tag{5.12}$$

respectively, where $\epsilon_1, \epsilon_2 \in (0,1)$ are arbitrary positive constants. By using (3.5) (local interpolation error estimate), (3.8) (exponential decay of $P_h\widetilde{\delta}_{x_0}$) and Lemma 4.1 (estimates of regularized Green's function), we have

$$\widehat{I_j} = \|P_h\widetilde{\delta}_{x_0}\|_{1,\Omega_j'} + d_j^{-1}\|P_h\widetilde{\delta}_{x_0}\|_{\Omega_j'} \le Ch^2 d_j^{-3-N/2}, \tag{5.13}$$

$$\begin{aligned}
\widehat{X_j} &= d_j|||\partial_t(I_h\Gamma - \Gamma)|||_{1,Q_j'} + |||\partial_t(I_h\Gamma - \Gamma)|||_{Q_j'} \\
&\quad + d_j^{-1}|||I_h\Gamma - \Gamma|||_{1,Q_j'} + d_j^{-2}|||I_h\Gamma - \Gamma|||_{Q_j'} \\
&\le (d_jh^\alpha + h^{1+\alpha})|||\partial_t\Gamma|||_{L^2H^{1+\alpha}(Q_j'')} + (d_j^{-1}h^\alpha + d_j^{-2}h^{1+\alpha})|||\Gamma|||_{L^2H^{1+\alpha}(Q_j'')}
\end{aligned}$$

$$\leq Ch^{\alpha} d_j^{-1-\alpha-N/2}. \tag{5.14}$$

and

$$\overline{I_j} = d_j^2 \|\Delta_h P_h \widetilde{\delta}_{x_0}\|_{1,\Omega'_j} + d_j \|\Delta_h P_h \widetilde{\delta}_{x_0}\|_{\Omega'_j} \leq Ch^2 d_j^{-3-N/2}, \tag{5.15}$$

$$\begin{aligned}
\overline{X_j} &= d_j^3 |\!|\!| I_h \partial_{tt}\Gamma - \partial_{tt}\Gamma |\!|\!|_{1,Q'_j} + d_j^2 |\!|\!| I_h \partial_{tt}\Gamma - \partial_{tt}\Gamma |\!|\!|_{Q'_j} \\
&\quad + d_j |\!|\!| I_h \partial_t \Gamma - \partial_t \Gamma |\!|\!|_{1,Q'_j} + |\!|\!| I_h \partial_t \Gamma - \partial_t \Gamma |\!|\!|_{Q'_j} \\
&\leq (d_j^3 h^{\alpha} + d_j^2 h^{1+\alpha}) |\!|\!| \partial_{tt}\Gamma |\!|\!|_{L^2 H^{1+\alpha}(Q''_j)} + (d_j h^{\alpha} + h^{1+\alpha}) |\!|\!| \partial_t \Gamma |\!|\!|_{L^2 H^{1+\alpha}(Q''_j)} \\
&\leq Ch^{\alpha} d_j^{-1-\alpha-N/2}.
\end{aligned} \tag{5.16}$$

By choosing $\epsilon_1 = \epsilon^4$ and $\epsilon_2 = \epsilon$ in (5.11)-(5.12), and substituting (5.11)-(5.16) into the expression of $\mathscr{K}$ in (5.10), we have

$$\begin{aligned}
\mathscr{K} &= \sum_j d_j^{1+N/2} (d_j^{-1} |\!|\!| F |\!|\!|_{1,Q_j} + |\!|\!| \partial_t F |\!|\!|_{Q_j} + d_j |\!|\!| \partial_t F |\!|\!|_{1,Q_j} + d_j^2 |\!|\!| \partial_{tt} F |\!|\!|_{Q_j}) \\
&\leq C_{\epsilon} \sum_j d_j^{1+N/2} \big(h^2 d_j^{-3-N/2} + h^{\alpha} d_j^{-1-s-N/2} + d_j^{-2} |\!|\!| F |\!|\!|_{Q'_j} \big) \\
&\quad + (C_{\epsilon} h^{1/2} d_j^{-1/2} + C_{\epsilon} h d_j^{-1} + \epsilon) \sum_j d_j^{1+N/2} (d_j^{-1} |\!|\!| F |\!|\!|_{1,Q'_j} + |\!|\!| \partial_t F |\!|\!|_{Q'_j}) \\
&\quad + (C_{\epsilon} h^{1/2} d_j^{-1/2} + C_{\epsilon} h d_j^{-1} + \epsilon) \sum_j d_j^{1+N/2} (d_j |\!|\!| \partial_t F |\!|\!|_{1,Q'_j} + d_j^2 |\!|\!| \partial_{tt} F |\!|\!|_{Q'_j}) \\
&\leq C_{\epsilon} + C_{\epsilon} \sum_j d_j^{-1+N/2} |\!|\!| F |\!|\!|_{Q'_j} \\
&\quad + (C_{\epsilon} h^{1/2} d_j^{-1/2} + C_{\epsilon} h d_j^{-1} + \epsilon) \sum_j d_j^{1+N/2} (d_j^{-1} |\!|\!| F |\!|\!|_{1,Q'_j} + |\!|\!| \partial_t F |\!|\!|_{Q'_j}) \\
&\quad + (C_{\epsilon} h^{1/2} d_j^{-1/2} + C_{\epsilon} h d_j^{-1} + \epsilon) \sum_j d_j^{1+N/2} (d_j |\!|\!| \partial_t F |\!|\!|_{1,Q'_j} + d_j^2 |\!|\!| \partial_{tt} F |\!|\!|_{Q'_j}).
\end{aligned} \tag{5.17}$$

Since $|\!|\!| F |\!|\!|_{Q'_j} \leq C(|\!|\!| F |\!|\!|_{Q_{j-1}} + |\!|\!| F |\!|\!|_{Q_j} + |\!|\!| F |\!|\!|_{Q_{j+1}})$, we can convert the Q'_j-norm in the inequality above to the Q_j-norm:

$$\begin{aligned}
\mathscr{K} &\leq C_{\epsilon} + C_{\epsilon} \sum_j d_j^{-1+N/2} |\!|\!| F |\!|\!|_{Q_j} + C_{\epsilon} d_*^{-1+N/2} |\!|\!| F |\!|\!|_{Q_*} \\
&\quad + (C_{\epsilon} h^{1/2} d_j^{-1/2} + C_{\epsilon} h d_j^{-1} + \epsilon) \sum_j d_j^{1+N/2} (d_j^{-1} |\!|\!| F |\!|\!|_{1,Q_j} + |\!|\!| \partial_t F |\!|\!|_{Q_j}) \\
&\quad + (C_{\epsilon} h^{1/2} d_j^{-1/2} + C_{\epsilon} h d_j^{-1} + \epsilon) \sum_j d_j^{1+N/2} (d_j |\!|\!| \partial_t F |\!|\!|_{1,Q_j} + d_j^2 |\!|\!| \partial_{tt} F |\!|\!|_{Q_j}) \\
&\quad + (C_{\epsilon} h^{1/2} d_*^{-1/2} + C_{\epsilon} h d_*^{-1} + \epsilon) d_*^{1+N/2} (d_*^{-1} |\!|\!| F |\!|\!|_{1,Q_*} + |\!|\!| \partial_t F |\!|\!|_{Q_*}) \\
&\quad + (C_{\epsilon} h^{1/2} d_*^{-1/2} + C_{\epsilon} h d_*^{-1} + \epsilon) d_*^{1+N/2} (d_* |\!|\!| \partial_t F |\!|\!|_{1,Q_*} + d_*^2 |\!|\!| \partial_{tt} F |\!|\!|_{Q_*}) \\
&\leq C_{\epsilon} + C_{\epsilon} C_*^{3+N/2} + \sum_j C d_j^{-1+N/2} |\!|\!| F |\!|\!|_{Q_j} + C(C_{\epsilon} C_*^{-1/2} + C_{\epsilon} C_*^{-1} + \epsilon) \mathscr{K}.
\end{aligned} \tag{5.18}$$

where we have used $d_j \geq C_* h$ and (5.3)-(5.7) to estimate

$$|||F|||_{1,Q_*}, \;\; |||\partial_t F|||_{Q_*}, \;\; |||\partial_t F|||_{1,Q_*} \;\; \text{and} \;\; |||\partial_{tt} F|||_{Q_*}$$

and used the expression of $\mathscr{K}$ in (5.10) to bound the terms involving Q_j. By choosing ϵ small enough and then choosing C_* large enough (C_* is still to be determined later), the last term on the right-hand side of (5.18) will be absorbed by the left-hand side. Hence, we obtain

$$\mathcal{K} \leq C + CC_*^{3+N/2} + \sum_j C d_j^{-1+N/2} |||F|||_{Q_j}. \tag{5.19}$$

It remains to estimate $|||F|||_{Q_j}$. To this end, we apply a duality argument below. Let w be the solution of the backward parabolic equation

$$-\partial_t w - \Delta w = v \quad \text{with} \;\; w(T) = 0,$$

where v is a function supported on Q_j and $|||v|||_{Q_j} = 1$. Multiplying the above equation by F yields (using integration by parts, with the notations (3.4))

$$[F, v] = (F(0), w(0)) + [F_t, w] + [\nabla F, \nabla w], \tag{5.20}$$

where (since $\widetilde{\delta}_{x_0} = 0$ on Ω_j'')

$$\begin{aligned}
(F(0), w(0)) &= (P_h\widetilde{\delta}_{x_0} - \widetilde{\delta}_{x_0}, w(0)) \\
&= (P_h\widetilde{\delta}_{x_0} - \widetilde{\delta}_{x_0}, w(0) - I_h w(0)) \\
&= (P_h\widetilde{\delta}_{x_0}, w(0) - I_h w(0))_{\Omega_j''} + (P_h\widetilde{\delta}_{x_0} - \widetilde{\delta}_{x_0}, w(0) - I_h w(0))_{(\Omega_j'')^c} \\
&=: \mathcal{I}_1 + \mathcal{I}_2.
\end{aligned}$$

By using Property (P3) and (3.8) (the exponential decay of $P_h\widetilde{\delta}_{x_0}$), we derive that

$$\begin{aligned}
|\mathcal{I}_1| &\leq Ch \|P_h\widetilde{\delta}_{x_0}\|_{L^2(\Omega_j'')} \|w(0)\|_{H^1(\Omega)} \\
&\leq Ch^{1-N/2} e^{-Cd_j/h} |||v|||_{Q_j} \\
&= C(d_j/h)^{1+N/2} e^{-Cd_j/h} h^2 d_j^{-1-N/2} \\
&\leq Ch^2 d_j^{-1-N/2},
\end{aligned} \tag{5.21}$$

$$\begin{aligned}
|\mathcal{I}_2| &\leq C\|\widetilde{\delta}_{x_0}\|_{L^2} \|w(0) - I_h w(0)\|_{L^2((\Omega_j'')^c)} \\
&\leq Ch^{1+\alpha} \|\widetilde{\delta}_{x_0}\|_{L^2} \inf_{\widetilde{w}} \|\widetilde{w}\|_{H^{1+\alpha}(\Omega)} \\
&= Ch^{1+\alpha} \|\widetilde{\delta}_{x_0}\|_{L^2} \|w(0)\|_{H^{1+\alpha}((\Omega_j')^c)} \\
&\leq Ch^{1+\alpha-N/2} \|w(0)\|_{H^{1+\alpha}((\Omega_j')^c)},
\end{aligned} \tag{5.22}$$

where the infimum extends over all possible $\widetilde{w}$ extending $w(0)$ from $(\Omega_j')^c$ to Ω, and we have used (3.6) in the last step.

To estimate $\|w(0)\|_{H^{1+\alpha}((\Omega_j')^c)}$, we let W_j be a set containing $(\Omega_j')^c$ but its distance to Ω_j is larger than $C^{-1}d_j$. Since

$$|x-y| + s^{1/2} \geq C_1^{-1} d_j \quad \text{for } x \in W_j \text{ and } (s,y) \in Q_j$$

for some positive constant C_1, it follows that $(s,x) \in \bigcup_{k\leq j+\log_2 C_1} Q_k(y)$ for $(s,y) \in Q_j$. Now, if we denote $\widetilde{G}(\cdot,\cdot,y)$ as any extension of $G(\cdot,\cdot,y)$ from $\bigcup_{k\leq j+\log_2 C_1} Q_k(y)$ to $\mathcal{Q}$, then for $x \in W_j$ we have

$$w(0,x) = \int_0^T \int_\Omega G(s,y,x)v(s,y)\mathrm{d}y\mathrm{d}s = \iint_{Q_j} G(s,x,y)v(s,y)\mathrm{d}y\mathrm{d}s = \iint_{Q_j} \widetilde{G}(s,x,y)v(s,y)\mathrm{d}y\mathrm{d}s,$$

where we have used the symmetric $G(s,y,x) = G(s,x,y)$ and the compact support of v in Q_j. Hence, we have

$$\begin{aligned}\|w(0,\cdot)\|_{H^{1+\alpha}(W_j)} \leq \|w(0,\cdot)\|_{H^{1+\alpha}(\Omega)} &\leq C\int_0^t \int_\Omega \|\widetilde{G}(s,\cdot,y)\|_{H^{1+\alpha}(\Omega)} |v(s,y)|\mathrm{d}y\mathrm{d}s \\ &\leq C\sup_{y\in\Omega} \|\widetilde{G}(\cdot,\cdot,y)\|_{L^\infty H^{1+\alpha}(\Omega)} \|v\|_{L^1(Q_j)}.\end{aligned} \tag{5.23}$$

Since the last inequality holds for all possible $\widetilde{G}(\cdot,\cdot,y)$ extending $G(\cdot,\cdot,y)$ from $\bigcup_{k\leq j+\log_2 C_1} Q_k(y)$ to $\mathcal{Q}$, it follows that (cf. definition (3.3))

$$\begin{aligned}\|w(0,\cdot)\|_{H^{1+\alpha}(W_j)} &\leq C\sup_{y\in\Omega} \|G(\cdot,\cdot,y)\|_{L^\infty H^{1+\alpha}(\bigcup_{k\leq j+\log_2 C_1} Q_k(y))} \|v\|_{L^1(Q_j)} \\ &\leq C d_j^{-1-\alpha-N/2} \|v\|_{L^1(Q_j)} \qquad \text{(here we use (4.2))} \\ &\leq C d_j^{-\alpha} |||v|||_{Q_j} = C d_j^{-\alpha}.\end{aligned} \tag{5.24}$$

From (5.21)-(5.24), we see that the first term on the right-hand side of (5.20) is bounded by

$$|(F(0), w(0))| \leq Ch^2 d_j^{-N/2-1} + Ch^{1+\alpha-N/2} d_j^{-\alpha} \leq Ch^{1+\alpha-N/2} d_j^{-\alpha}, \tag{5.25}$$

and the rest terms are bounded by (recall that $F = \Gamma_h - \Gamma$, where Γ_h and Γ are solutions of (3.13)-(3.14))

$$\begin{aligned}[F_t, w] + [\nabla F, \nabla w] &= [F_t, w - I_h w] + [\nabla F, \nabla(w - I_h w)] \\ &\leq \sum_{*,i} \Big(C|||F_t|||_{Q_i} |||w - I_h w|||_{Q_i} + C|||F|||_{1,Q_i} |||w - I_h w|||_{1,Q_i} \Big) \\ &\leq \sum_{*,i} (Ch^{1+\alpha} |||F_t|||_{Q_i} + Ch^\alpha |||F|||_{1,Q_i}) \|w\|_{L^2 H^{1+\alpha}(Q_i')}.\end{aligned} \tag{5.26}$$

where we have used Property (P3) of Section 3.2 in the last step.

To estimate $\|w\|_{L^2 H^{1+\alpha}(Q_i')}$ in (5.26), we consider the expression (v is supported in Q_j)

$$w(t,x) = \int_0^T \int_\Omega G(s-t,x,y)v(s,y)1_{s>t}\,\mathrm{d}y\mathrm{d}s = \iint_{Q_j} G(s-t,x,y)v(s,y)1_{s>t}\,\mathrm{d}y\mathrm{d}s. \tag{5.27}$$

For $i \le j-3$ (so that $d_i > d_j$), we have

$$\begin{aligned}
&\text{if } t > 4d_j^2 \text{ then } w(t,x) = 0 \text{ (because } v \text{ is supported in } Q_j);\\
&\text{if } t \le 4d_j^2,\ (t,x) \in Q_i' \text{ and } (s,y) \in Q_j, \text{ then } d_i/4 \le |x-y| \le 4d_i \text{ and } s-t \in (0, d_i^2)\\
&\qquad\qquad\qquad\qquad\qquad \text{thus } (s-t,x) \in Q_i'(y).
\end{aligned} \tag{5.28}$$

Hence, from (5.27) we derive

$$\begin{aligned}
\|w\|_{L^2H^{1+\alpha}(Q_i')} &\le \sup_y \|G(\cdot,\cdot,y)\|_{L^2H^{1+\alpha}(Q_i'(y))}\|v\|_{L^1(Q_j)}\\
&\le Cd_i \sup_y \|G(\cdot,\cdot,y)\|_{L^\infty H^{1+\alpha}(Q_i'(y))}\|v\|_{L^1(Q_j)}\\
&\le Cd_i^{-\alpha-N/2} d_j^{1+N/2} |\!|\!|v|\!|\!|_{Q_j} \qquad \text{(here we use (4.2))}\\
&\le Cd_j^{1-\alpha}\left(\frac{d_j}{d_i}\right)^{\alpha+N/2}.
\end{aligned}$$

For $i \ge j+3$ ($d_i \le d_j$, including the case $i = *$), we have

$$\begin{aligned}
&\text{if } (t,x) \in Q_i' \text{ and } (s,y) \in Q_j, \text{ then } \max(|s-t|^{1/2}, |x-y|) \ge d_{j+3},\\
&\qquad\qquad\qquad \text{thus } (s-t,x) \in \textstyle\bigcup_{k\le j+3} Q_k(y),\\
&\text{if } (t,x) \in Q_* \text{ and } (s,y) \in Q_j \text{ with } j \le J_*-3, \text{ then } \max(|s-t|^{1/2}, |x-y|) \ge d_{j+3},\\
&\qquad\qquad\qquad \text{thus } (s-t,x) \in \textstyle\bigcup_{k\le j+3} Q_k(y).
\end{aligned} \tag{5.29}$$

If $\widetilde{G}(\cdot,\cdot,y)$ is a function satisfying

$$\widetilde{G}(\cdot,\cdot,y) = G(\cdot,\cdot,y) \quad \text{on} \quad \bigcup_{k\le j+3} Q_k(y), \tag{5.30}$$

then for $(t,x) \in Q_i'$ we have

$$w(t,x) = \int_0^T \int_\Omega G(s-t,x,y)v(y,s)1_{s>t}\,\mathrm{d}y\mathrm{d}s = \iint_{Q_j} \widetilde{G}(s-t,x,y)v(s,y)1_{s>t}\,\mathrm{d}y\mathrm{d}s.$$

Hence, we have

$$\begin{aligned}
\|w\|_{L^2H^{1+\alpha}(Q_i')} &\le Cd_i\|w\|_{L^\infty H^{1+\alpha}(Q_i')} \le Cd_i\|w\|_{L^\infty H^{1+\alpha}(\mathcal{Q})}\\
&\le Cd_i \iint_{Q_j} \|\widetilde{G}(s-t,x,y)1_{s>t}\|_{L_t^\infty H_x^{1+\alpha}(\mathcal{Q})}|v(s,y)|\mathrm{d}y\mathrm{d}s\\
&\le Cd_i \iint_{Q_j} \|\widetilde{G}(\cdot,\cdot,y)\|_{L^\infty H^{1+\alpha}(\mathcal{Q})}|v(s,y)|\mathrm{d}y\mathrm{d}s\\
&\le Cd_i \sup_{y\in\Omega} \|\widetilde{G}(\cdot,\cdot,y)\|_{L^\infty H^{1+\alpha}(\mathcal{Q})}\|v\|_{L^1(Q_j)}
\end{aligned}$$

$$\leq Cd_i \sup_{y\in\Omega} \|\widetilde{G}(\cdot,\cdot,y)\|_{L^\infty H^{1+\alpha}(\mathcal{Q})} |||v|||_{Q_j} d_j^{1+N/2}.$$

In view of the definition (3.3), by taking infimum over all the possible choices of $\widetilde{G}(\cdot,\cdot,y)$ satisfying (5.30), we have

$$\begin{aligned}
\|w\|_{L^2H^{1+\alpha}(Q_i')} &\leq Cd_i \sup_{y\in\Omega} \|G(\cdot,\cdot,y)\|_{L^\infty H^{1+\alpha}(\cup_{k\leq j+3}Q_k(y))} |||v|||_{Q_j} d_j^{1+N/2}. \\
&\leq Cd_i d_j^{-1-\alpha-N/2} d_j^{N/2+1} \qquad\qquad \text{(here we use (4.2))} \\
&= Cd_i^{1-\alpha}\left(\frac{d_i}{d_j}\right)^\alpha .
\end{aligned}$$

For $|i-j|\leq 2$, applying the standard energy estimate yields

$$\|w\|_{L^2H^{1+\alpha}(Q_i')} \leq \|w\|_{L^2H^{1+\alpha}(\mathcal{Q})} \leq C\|v\|_{L^2H^{\alpha-1}(\mathcal{Q})} \leq C\|v\|_{L^2L^{\frac{N}{1+N/2-\alpha}}(\mathcal{Q})} \leq Cd_i^{1-\alpha}|||v|||_{\mathcal{Q}} = Cd_i^{1-\alpha},$$

where we have used the Sobolev embedding $L^{\frac{N}{1+N/2-\alpha}} \hookrightarrow H^{\alpha-1}$ and the Hölder's inequality

$$\|v\|_{L^{\frac{N}{1+N/2-\alpha}}} \leq Cd_j^{1-\alpha}\|v\|_{L^2} \quad \text{(this requires the volume of the support of } v \text{ to be bounded by } d_j^N).$$

Combining the three cases above (corresponding to $i\leq j-3$, $i\geq j+3$ and $|i-j|\leq 2$), we have

$$\|w\|_{L^2H^{1+\alpha}(Q_i')} \leq Cd_i^{1-\alpha}\left(\frac{\min(d_i,d_j)}{\max(d_i,d_j)}\right)^\alpha . \tag{5.31}$$

Substituting (5.25)-(5.26) and (5.31) into (5.20) yields

$$|||F|||_{Q_j} \leq Ch^{1+\alpha-N/2}d_j^{-\alpha} + C\sum_{*,i}(h^{1+\alpha}|||F_t|||_{Q_i} + h^\alpha|||F|||_{1,Q_i})d_i^{1-\alpha}\left(\frac{\min(d_i,d_j)}{\max(d_i,d_j)}\right)^\alpha . \tag{5.32}$$

Since $\alpha > 1/2$, it follows that

$$\sum_j d_j^{N/2-1}\left(\frac{\min(d_i,d_j)}{\max(d_i,d_j)}\right)^\alpha \leq Cd_i^{N/2-1}. \tag{5.33}$$

Hence, we have

$$\begin{aligned}
\mathscr{K} &\leq C + CC_*^{3+N/2} + C\sum_j d_j^{-1+N/2}|||F|||_{Q_j} \\
&\leq C + CC_*^{3+N/2} + C\sum_j \left(\frac{h}{d_j}\right)^{1+\alpha-N/2} \qquad\qquad \text{here we substitute (5.32)} \\
&\quad + C\sum_j d_j^{-1+N/2}\sum_{*,i}(h^{1+\alpha}|||F_t|||_{Q_i} + h^\alpha|||F|||_{1,Q_i})d_i^{1-\alpha}\left(\frac{\min(d_i,d_j)}{\max(d_i,d_j)}\right)^\alpha
\end{aligned}$$

$$
\begin{aligned}
&\le C + CC_*^{3+N/2} + C \qquad \text{here we exchange the order of summation}\\
&\quad + C\sum_{*,i}(h^{1+\alpha}|\!|\!|F_t|\!|\!|_{Q_i} + h^\alpha |\!|\!|F|\!|\!|_{1,Q_i})d_i^{1-\alpha}\sum_j d_j^{N/2-1}\left(\frac{\min(d_i,d_j)}{\max(d_i,d_j)}\right)^\alpha\\
&\le C + CC_*^{3+N/2} + C\sum_{*,i}(h^{1+\alpha}|\!|\!|F_t|\!|\!|_{Q_i} + h^\alpha |\!|\!|F|\!|\!|_{1,Q_i})d_i^{N/2-\alpha} \qquad \text{(here we use (5.33))}\\
&= C + CC_*^{3+N/2} + C\sum_{*,i} d_i^{1+N/2}\left(|\!|\!|F_t|\!|\!|_{Q_i}\left(\frac{h}{d_i}\right)^{1+\alpha} + d_i^{-1}|\!|\!|F|\!|\!|_{1,Q_i}\left(\frac{h}{d_i}\right)^\alpha\right)\\
&\le C + CC_*^{3+N/2} + Cd_*^{1+N/2}\left(|\!|\!|F_t|\!|\!|_{Q_*} + d_j^{-1}|\!|\!|F|\!|\!|_{1,Q_*}\right)\\
&\quad + C\sum_i d_i^{1+N/2}\left(|\!|\!|F_t|\!|\!|_{Q_i} + d_j^{-1}|\!|\!|F|\!|\!|_{1,Q_i}\right)\left(\frac{h}{d_i}\right)^\alpha\\
&\le C + CC_*^{3+N/2} + \frac{C\mathscr{K}}{C_*^\alpha}.
\end{aligned}
$$

By choosing C_* to be large enough (C_* is determined now), the term $\dfrac{C\mathscr{K}}{C_*^\alpha}$ will be absorbed by the left-hand side of the inequality above. In this case, the inequality above implies

$$
\mathscr{K} \le C. \tag{5.34}
$$

Substituting the last inequality into (5.9) yields

$$
\|\partial_t F\|_{L^1(\mathcal{Q})} + \|t\partial_{tt}F\|_{L^1(\mathcal{Q})} \le C. \tag{5.35}
$$

Part II. Second, we present estimates for $(t,x) \in (1,\infty)\times\Omega$. For $t > 1$, we differentiate (3.14) with respect to t and integrate the resulting equation against $\partial_t\Gamma_h$. Then we get

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\|\partial_t\Gamma_h(t,\cdot,x_0)\|_{L^2}^2 + \lambda_0\|\partial_t\Gamma_h(t,\cdot,x_0)\|_{L^2}^2\\
&\le \frac{\mathrm{d}}{\mathrm{d}t}\|\partial_t\Gamma_h(t,\cdot,x_0)\|_{L^2}^2 + (\nabla\partial_t\Gamma_h(t,\cdot,x_0), \nabla\partial_t\Gamma_h(t,\cdot,x_0))\\
&= 0,
\end{aligned}
$$

for $t \ge 1$, where $\lambda_0 > 0$ is the smallest eigenvalue of the operator $-\Delta$. From the last inequality we derive the exponential decay of $\partial_t\Gamma_h$ with respect to t

$$
\|\partial_t\Gamma_h(t,\cdot,x_0)\|_{L^2}^2 \le e^{-\lambda_0(t-1)}\|\partial_t\Gamma_h(1,\cdot,x_0)\|_{L^2}^2 \le Ce^{-\lambda_0(t-1)},
$$

where the inequality $\|\partial_t\Gamma_h(1,\cdot,x_0)\|_{L^2} \le C$ can be proved by a simple energy estimate (omitted here). Similarly, we also have

$$
\|\partial_{tt}\Gamma_h(t,\cdot,x_0)\|_{L^2}^2 + \|\partial_t\Gamma(t,\cdot,x_0)\|_{L^2}^2 + \|\partial_{tt}\Gamma(t,\cdot,x_0)\|_{L^2}^2 \le Ce^{-\lambda_0(t-1)} \quad \text{for } t \ge 1.
$$

The estimate (5.35) and the last two inequalities imply (4.27)-(4.28) in the case $h < h_* := 1/(4C_*)$.

For $h \geq h_*$, some basic energy estimates would yield

$$\begin{aligned}
&\|\partial_t \Gamma_h(t,\cdot,x_0)\|_{L^2}^2 + \|\partial_{tt}\Gamma_h(t,\cdot,x_0)\|_{L^2}^2 + \|\partial_t \Gamma(t,\cdot,x_0)\|_{L^2}^2 + \|\partial_{tt}\Gamma(t,\cdot,x_0)\|_{L^2}^2 \\
&\leq Ce^{-\lambda_0 t}(\|\partial_t \Gamma_h(0,\cdot,x_0)\|_{L^2}^2 + \|\partial_{tt}\Gamma_h(0,\cdot,x_0)\|_{L^2}^2 + \|\partial_t \Gamma(0,\cdot,x_0)\|_{L^2}^2 + \|\partial_{tt}\Gamma(0,\cdot,x_0)\|_{L^2}^2) \\
&= Ce^{-\lambda_0 t}(\|\Delta_h P_h \widetilde{\delta}_{x_0}\|_{L^2}^2 + \|\Delta_h^2 P_h \widetilde{\delta}_{x_0}\|_{L^2}^2 + \|\Delta \widetilde{\delta}_{x_0}\|_{L^2}^2 + \|\Delta^2 \widetilde{\delta}_{x_0}\|_{L^2}^2) \\
&\leq Ce^{-\lambda_0 t}(h_*^{-4-N} + h_*^{-8-N})
\end{aligned}$$

for arbitrary $t > 0$. This implies (4.27)-(4.28) in the case $h \geq h_*$.

Part III. Finally, we note that (4.26) is a simple consequence of (3.6), (3.12) and (3.18), while (4.25) is a consequence of (4.26) and the following inequalities:

$$\|\Gamma_h(t,\cdot,x_0)\|_{L^1(\Omega)} \leq \|\Gamma(t,\cdot,x_0)\|_{L^1(\Omega)} + \|F(t,\cdot,x_0)\|_{L^1(\Omega)}, \tag{5.36}$$

$$\|t\partial_t \Gamma_h(t,\cdot,x_0)\|_{L^1(\Omega)} \leq \|t\partial_t \Gamma(t,\cdot,x_0)\|_{L^1(\Omega)} + \|t\partial_t F(t,\cdot,x_0)\|_{L^1(\Omega)}, \tag{5.37}$$

with

$$\begin{aligned}
\|F(t,\cdot,x_0)\|_{L^1(\Omega)} &\leq \|F(0,\cdot,x_0)\|_{L^1(\Omega)} + \left\| \int_0^t \partial_s F(s,\cdot,x_0) \mathrm{d}s \right\|_{L^1(\Omega)} \\
&\leq \|P_h \widetilde{\delta}_{x_0} - \widetilde{\delta}_{x_0}\|_{L^1(\Omega)} + \|\partial_t F\|_{L^1(\mathcal{Q})} \leq C, \qquad (5.38)
\end{aligned}$$

$$\begin{aligned}
\|t\partial_t F(t,\cdot,x_0)\|_{L^1(\Omega)} &\leq \left\| \int_0^t \Big(s\partial_{ss} F(s,\cdot,x_0) + \partial_s F_h(s,\cdot,x_0) \Big) \mathrm{d}s \right\|_{L^1(\Omega)} \\
&\leq \|t\partial_{tt} F\|_{L^1(\mathcal{Q})} + \|\partial_t F\|_{L^1(\mathcal{Q})} \leq C, \qquad (5.39)
\end{aligned}$$

where we have used (4.27) in the last two inequalities, which was proved in Part I and Part II.

The proof of Lemma 4.4 is complete. □

6 Conclusion

The analyticity and maximal L^p-regularity of finite element solutions of the heat equation are proved in general polygons and polyhedra, possibly nonconvex. The L^∞-stability of the finite element parabolic projection has been reduced to the L^∞-stability of the Ritz projection. Such L^∞-stability of the Dirichlet Ritz projection is currently known in general polygons [42] and convex polyhedra [28], but still remains open in nonconvex polyhedra. The L^∞-stability of the Neumann Ritz projection remains open in both nonconvex polygons and nonconvex polyhedra. This article focuses on the Lagrange finite element method. Extension of the results to other numerical methods, such as finite volume methods and discontinuous Galerkin methods, are interesting and nontrivial. Such extension may need more precise $W^{s,p}$-approximation properties of local elliptic projectors onto finite element spaces (e.g., see [13]).

Appendix A: Proof of Lemma 5.1

In this subsection, we prove Lemma 5.1, which is used in the last section in proving Lemma 4.4. Before we prove Lemma 5.1, we present a local energy estimate for finite element solutions of parabolic equations based on the decomposition $Q_j = [(0,d_j^2)\times\Omega_j]\cup[(d_j^2,4d_j^2)\times D_j]$.

Lemma A.1 *Suppose that $\phi_h(t)\in S_h$, $t\in(0,T)$, satisfies*

$$\begin{aligned}&\big(\partial_t\phi_h,\chi_h\big)+\big(\nabla\phi_h,\nabla\chi_h\big)=0, \qquad \textit{for}\ \ \chi_h\in S_h^0(\Omega_j''),\ \ t\in(0,d_j^2),\\ &\big(\partial_t\phi_h,\chi_h\big)+\big(\nabla\phi_h,\nabla\chi_h\big)=0, \qquad \textit{for}\ \ \chi_h\in S_h^0(D_j''),\ \ t\in(d_j^2/4,2d_j^2).\end{aligned}$$

Then we have

$$\begin{aligned}&\|\partial_t\phi_h\|_{L^2(Q_j)}+d_j^{-1}\|\nabla\phi_h\|_{L^2(Q_j)}\\ &\le (Ch^{1/2}d_j^{-1/2}+C\epsilon^{-1}hd_j^{-1}+\epsilon)\big(\|\partial_t\phi_h\|_{L^2(Q_j')}+d_j^{-1}\|\nabla\phi_h\|_{L^2(Q_j')}\big)\\ &\quad +C\epsilon^{-1}\Big(d_j^{-2}\|\phi_h\|_{L^2(Q_j')}+d_j^{-1}\|\phi_h(0)\|_{L^2(\Omega_j')}+\|\nabla\phi_h(0)\|_{L^2(\Omega_j')}\Big), \qquad \text{(A.1)}\end{aligned}$$

where the constant C is independent of h, j and C_*.

Proof of Lemma A.1. We shall present estimates in the two subdomains $(0,d_j^2)\times\Omega_j$ and $(d_j^2,4d_j^2)\times D_j$, separately.

First, we present estimates in $(0,d_j^2)\times\Omega_j$. To this end, we let ω be a smooth cut-off function which equals 1 on Ω_j and equals 0 outside Ω_j', and let $\widetilde\omega$ be a smooth cut-off function which equals 1 on Ω_j' and equals 0 outside Ω_j'', such that

(1) $\mathrm{dist}\big(\mathrm{supp}(\omega)\cap\Omega,\Omega\backslash\Omega_j'\big)\ge d_j/8\ge 2h$ and $\mathrm{dist}\big(\mathrm{supp}(\widetilde\omega)\cap\Omega,\Omega\backslash\Omega_j''\big)\ge d_j/8\ge 2h$,

(2) $|\partial^\alpha\omega|+|\partial^\alpha\widetilde\omega|\le C_\alpha d_j^{-|\alpha|}$ for any multi-index α.

By Property (P3) of Section 3.2, the function $v_h:=I_h(\widetilde\omega\phi_h)\in S_h^0(\Omega_j'')$ satisfies $v_h=\phi_h$ on Ω_j' and

$$\|v_h\|_{L^2(\Omega)}\le C\|\phi_h\|_{L^2(\Omega_j'')}, \qquad \text{(A.2)}$$

$$\|\nabla v_h\|_{L^2(\Omega)}\le C\|\nabla\phi_h\|_{L^2(\Omega_j'')}+Cd_j^{-1}\|\phi_h\|_{L^2(\Omega_j'')}, \qquad \text{(A.3)}$$

$$\big(\partial_t v_h,\chi_h\big)+\big(\nabla v_h,\nabla\chi_h\big)=0,\quad \forall\,\chi_h\in S_h^0(\Omega_j'),\quad \forall\ t\in(0,d_j^2). \qquad \text{(A.4)}$$

Property (P3) of Section 3.2 also implies that $I_h(\omega^2v_h)\in S_h^0(\Omega_j')$ such that

$$\|\omega^2v_h-I_h(\omega^2v_h)\|_{L^2}+h\|\nabla(\omega^2v_h-I_h(\omega^2v_h))\|_{L^2}\le Chd_j^{-1}\|v_h\|_{L^2}.$$

Since ω and $\widetilde\omega$ are time-independent, it follows that

$$\begin{aligned}&\frac12\frac{\mathrm d}{\mathrm dt}\|\omega v_h\|^2+(\omega^2\nabla v_h,\nabla v_h)\\ &=\big[\big(\partial_t v_h,\omega^2v_h\big)+\big(\nabla v_h,\nabla(\omega^2v_h)\big)\big]-\big(2v_h\omega\nabla\omega,\nabla v_h\big)\\ &=\big[\big(\partial_t v_h,\omega^2v_h-I_h(\omega^2v_h)h\big)+\big(\nabla v_h,\nabla(\omega^2v_h-I_h(\omega^2v_h))\big)\big]-\big(2v_h\nabla\omega,\omega\nabla v_h\big)\\ &\le\big[C\|\partial_t v_h\|_{L^2}\|v_h\|_{L^2}hd_j^{-1}+C\|\nabla v_h\|_{L^2}\|v_h\|_{L^2}d_j^{-1}\big]+Cd_j^{-1}\|v_h\|_{L^2}\|\omega\nabla v_h\|_{L^2}\end{aligned}$$

$$\leq C\|\partial_t v_h\|_{L^2}^2 h^2 + \epsilon^4\|\nabla v_h\|_{L^2}^2 + C\epsilon^{-4}\|v_h\|_{L^2}^2 d_j^{-2}, \qquad \forall\, \epsilon \in (0,1).$$

By using (A.2)-(A.3), integrating the last inequality from 0 to d_j^2 yields

$$\begin{aligned}
&\|\phi_h\|_{L^\infty(0,d_j^2;L^2(\Omega_j))} + \|\nabla\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j))} \\
&\leq C\|\phi_h(0)\|_{L^2(\Omega_j'')} + C\|\partial_t\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j''))}h \\
&\quad + \epsilon^2\|\nabla\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j''))} + C\epsilon^{-2}\|\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j''))}d_j^{-1}.
\end{aligned} \tag{A.5}$$

Furthermore, we have

$$\begin{aligned}
&\|\omega^2\partial_t v_h\|_{L^2}^2 + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\big(\omega^4\nabla v_h, \nabla v_h\big) \\
&= \big(\partial_t v_h, \omega^4\partial_t v_h\big) + \big(\nabla v_h, \nabla(\omega^4\partial_t v_h)\big) - \big(4\partial_t v_h\omega^3\nabla\omega, \nabla v_h\big) \\
&= \big(\partial_t v_h, \omega^4\partial_t v_h - I_h(\omega^4\partial_t v_h)\big) + \big(\nabla v_h, \nabla(\omega^4\partial_t v_h - I_h(\omega^4\partial_t v_h))\big) - \big(4\partial_t v_h\omega^3\nabla\omega, \nabla v_h\big) \\
&\leq C\|\partial_t v_h\|_{L^2}^2 h d_j^{-1} + C\|\nabla v_h\|_{L^2}\|\partial_t v_h\|_{L^2}d_j^{-1} + C\|\omega^2\partial_t v_h\|_{L^2}\|\nabla v_h\|_{L^2}d_j^{-1} \\
&\leq (Chd_j^{-1} + \epsilon^2)\|\partial_t v_h\|_{L^2}^2 + \epsilon^2\|\omega^2\partial_t v_h\|_{L^2}^2 + C\epsilon^{-2}\|\nabla v_h\|_{L^2}^2 d_j^{-2}, \qquad \forall\, \epsilon \in (0,1/2),
\end{aligned}$$

which reduces to

$$\|\omega^2\partial_t v_h\|_{L^2(0,d_j^2;L^2)}^2 \leq (Chd_j^{-1} + \epsilon^2)\|\partial_t v_h\|_{L^2(0,d_j^2;L^2)}^2 + C\epsilon^{-2}\|\nabla v_h\|_{L^2(0,d_j^2;L^2)}^2 d_j^{-2} + C\|\nabla v_h(0)\|_{L^2}^2.$$

By using (A.2)-(A.3), the last inequality further implies

$$\begin{aligned}
\|\partial_t\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j))} &\leq C(\|\nabla\phi_h(0)\|_{L^2(\Omega_j'')} + d_j^{-1}\|\phi_h(0)\|_{L^2(\Omega_j'')}) \\
&\quad + (Ch^{1/2}d_j^{-1/2} + \epsilon)\|\partial_t\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j''))} \\
&\quad + C\epsilon^{-1}(\|\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j''))}d_j^{-2} + \|\nabla\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j''))}d_j^{-1}) \\
&\leq C(\|\nabla\phi_h(0)\|_{L^2(\Omega_j'')} + d_j^{-1}\|\phi_h(0)\|_{L^2(\Omega_j'')}) \\
&\quad + (Ch^{1/2}d_j^{-1/2} + \epsilon)\|\partial_t\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j''))} \\
&\quad + C\epsilon^{-1}(\|\phi_h\|_{L^\infty(0,d_j^2;L^2(\Omega_j''))}d_j^{-1} + \|\nabla\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j''))}d_j^{-1}).
\end{aligned} \tag{A.6}$$

With an obvious change of domains (replacing Ω_j by Ω_j' on the left-hand side of (A.5) and replacing Ω_j'' by Ω_j' on the right-hand side of (A.6)), the two estimates (A.5) and (A.6) imply

$$\begin{aligned}
&\|\phi_h\|_{L^\infty(0,d_j^2;L^2(\Omega_j'))} + \|\nabla\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j'))} \\
&\leq \hat{C}\|\phi_h(0)\|_{L^2(\Omega_j'')} + \hat{C}\|\partial_t\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j''))}h \\
&\quad + \epsilon^2\|\nabla\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j''))} + \hat{C}\epsilon^{-2}\|\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j''))}d_j^{-1}.
\end{aligned} \tag{A.7}$$

and

$$\begin{aligned}
\|\partial_t\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j))} &\leq \hat{C}(\|\nabla\phi_h(0)\|_{L^2(\Omega_j')} + d_j^{-1}\|\phi_h(0)\|_{L^2(\Omega_j')}) \\
&\quad + (\hat{C}h^{1/2}d_j^{-1/2} + \epsilon)\|\partial_t\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j'))}
\end{aligned}$$

$$+\hat{C}\epsilon^{-1}d_j^{-1}(\|\phi_h\|_{L^\infty(0,d_j^2;L^2(\Omega_j'))}+\|\nabla\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j'))}), \tag{A.8}$$

where $\hat{C}\geq 1$ is some positive constant and $\epsilon\in(0,1)$ can be arbitrary. Then $2\hat{C}\epsilon^{-1}d_j^{-1}\times$(A.7)+(A.8) yields (the last term in (A.8) can be absorbed by left-hand side of $2\hat{C}\epsilon^{-1}d_j^{-1}\times$(A.7))

$$\begin{aligned}
&\|\partial_t\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j))}+d_j^{-1}\|\nabla\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j))}\\
&\leq (Ch^{1/2}d_j^{-1/2}+C\epsilon^{-1}hd_j^{-1}+\epsilon)(\|\partial_t\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j''))}+d_j^{-1}\|\nabla\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j''))})\\
&\quad +C\epsilon^{-3}\|\phi_h\|_{L^2(0,d_j^2;L^2(\Omega_j''))}d_j^{-2}+C\epsilon^{-1}(\|\nabla\phi_h(0)\|_{L^2(\Omega_j'')}+d_j^{-1}\|\phi_h(0)\|_{L^2(\Omega_j'')}).
\end{aligned} \tag{A.9}$$

Second, we present estimates in $(d_j^2,4d_j^2)\times D_j$. We re-define $\omega(x,t):=\omega_1(x)\omega_2(t)$ and $\widetilde\omega(x,t):=\widetilde\omega_1(x)\widetilde\omega_2(t)$ such that

(1) $\omega_1=1$ in D_j and $\omega_1=0$ outside D_j', $\widetilde\omega_1=1$ in D_j' and $\widetilde\omega_1=0$ outside D_j'';

(2) $\mathrm{dist}(\mathrm{supp}(\omega_1)\cap\Omega,\Omega\backslash D_j')\geq d_j/4\geq\kappa h$ and $\mathrm{dist}(\mathrm{supp}(\widetilde\omega_1)\cap\Omega,\Omega\backslash D_j'')\geq d_j/4\geq\kappa h$;

(3) $\omega_2=1$ for $t\in(d_j^2,4d_j^2)$ and $\omega_2=0$ for $t\in(0,d_j^2/2)$;

(4) $\widetilde\omega_2=1$ for $t\in(d_j^2/4,4d_j^2)$ and $\widetilde\omega_2=0$ for $t\in(0,d_j^2/8)$;

(5) $|\partial^\alpha\omega_1|+|\partial^\alpha\widetilde\omega_1|\leq Cd_j^{-|\alpha|}$ for any multi-index α;

(6) $|\partial_t^k\omega_2|+|\partial_t^k\widetilde\omega_2|\leq Cd_j^{-2k}$ for any nonnegative integer k.

Then the function $v_h:=I_h(\widetilde\omega\phi_h)\in S_h^0(D_j'')$ satisfies $v_h=\widetilde\omega_2\phi_h$ on D_j' and

$$\|v_h\|_{L^2(\Omega)}\leq C\|\widetilde\omega_2\phi_h\|_{L^2(D_j'')}, \tag{A.10}$$

$$\|\nabla v_h\|_{L^2(\Omega)}\leq C\|\widetilde\omega_2\nabla\phi_h\|_{L^2(D_j'')}+Cd_j^{-1}\|\widetilde\omega_2\phi_h\|_{L^2(D_j'')}, \tag{A.11}$$

$$\big(\partial_t v_h,\chi_h\big)+\big(\nabla v_h,\nabla\chi_h\big)=0,\quad \forall\,\chi_h\in S_h^0(D_j')\quad \forall\, t\in(d_j^2/4,4d_j^2). \tag{A.12}$$

According to (P3) of Section 3.2, the function $\chi_h=I_h(\omega^2v_h)\in S_h^0(D_j')$ satisfies

$$\|\omega^2v_h-\chi_h\|_{L^2}+h\|\nabla(\omega^2v_h-\chi_h)\|_{L^2}\leq Chd_j^{-1}\|v_h\|_{L^2}.$$

Therefore, we have

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\omega v_h\|^2+(\omega^2\nabla v_h,\nabla v_h)\\
&=\big[\big(\partial_t v_h,\omega^2v_h\big)+\big(\nabla v_h,\nabla(\omega^2v_h)\big)\big]+\big(\partial_t\omega v_h,v_h\big)-\big(2v_h\omega\nabla\omega,\nabla v_h\big)\\
&=\big[\big(\partial_t v_h,\omega^2v_h-\chi_h\big)+\big(\nabla v_h,\nabla(\omega^2v_h-\chi_h)\big)\big]+\big(\omega\partial_t\omega v_h,v_h\big)-\big(2v_h\nabla\omega,\omega\nabla v_h\big)\\
&\leq\big[C\|\partial_t v_h\|_{L^2(\Omega)}\|v_h\|_{L^2(\Omega)}hd_j^{-1}+C\|\nabla v_h\|_{L^2(\Omega)}\|v_h\|_{L^2(\Omega)}d_j^{-1}\big]\\
&\quad +Cd_j^{-2}\|v_h\|_{L^2(\Omega)}^2+Cd_j^{-1}\|v_h\|_{L^2(\Omega)}\|\omega\nabla v_h\|_{L^2(\Omega)}\\
&\leq C\|\partial_t v_h\|_{L^2(\Omega)}^2h^2+\epsilon^4\|\nabla v_h\|_{L^2(\Omega)}^2+\epsilon^4\|\omega\nabla v_h\|_{L^2(\Omega)}^2+C\epsilon^{-4}\|v_h\|_{L^2(\Omega)}^2d_j^{-2},\quad\forall\,\epsilon\in(0,1/2).
\end{aligned}$$

Integrating the last inequality in time for $t\in(d_j^2/2,4d_j^2)$, we obtain

$$\|\phi_h\|_{L^\infty(d_j^2,4d_j^2;L^2(D_j))}+\|\nabla\phi_h\|_{L^2(d_j^2,4d_j^2;L^2(D_j))} \tag{A.13}$$

$$\leq C\|\partial_t\phi_h\|_{L^2(d_j^2/2,4d_j^2;L^2(D_j''))}h + \epsilon^2\|\nabla\phi_h\|_{L^2(d_j^2/2,4d_j^2;L^2(D_j''))} + C\epsilon^{-2}\|\phi_h\|_{L^2(d_j^2/2,4d_j^2;L^2(D_j''))}d_j^{-1}.$$

Furthermore, we have

$$\begin{aligned}
&\|\omega^2\partial_t v_h\|_{L^2}^2 + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\big(\omega^4\nabla v_h, \nabla v_h\big) \\
&= \big[\big(\partial_t v_h, \omega^4\partial_t v_h - \eta_h\big) + \big(\nabla v_h, \nabla(\omega^4\partial_t v_h - \eta_h)\big)\big] \\
&\quad + \big(2\omega^3\partial_t\omega\nabla v_h, \nabla v_h\big) - \big(4\partial_t v_h\omega^3\nabla\omega, \nabla v_h\big) \\
&\leq [C\|\partial_t v_h\|_{L^2(\Omega)}^2 h d_j^{-1} + C\|\nabla v_h\|_{L^2(\Omega)}\|\partial_t v_h\|_{L^2(\Omega)}d_j^{-1} \\
&\quad + C\|\nabla v_h\|_{L^2(\Omega)}^2 d_j^{-2} + C\|\omega^2\partial_t v_h\|_{L^2(\Omega)}\|\nabla v_h\|_{L^2(\Omega)}d_j^{-1} \\
&\leq (Chd_j^{-1} + \epsilon)\|\partial_t v_h\|_{L^2(\Omega)}^2 + \epsilon^2\|\omega^2\partial_t v_h\|_{L^2(\Omega)}^2 + C\epsilon^{-2}\|\nabla v_h\|_{L^2(\Omega)}^2 d_j^{-2}, \quad \forall\, \epsilon \in (0, 1/2),
\end{aligned}$$

which implies (by integrating the last inequality in time for $t \in (d_j^2/2, 4d_j^2)$)

$$\|\omega^2\partial_t v_h\|_{L^2(d_j^2,4d_j^2;L^2(\Omega))}^2 \leq (Chd_j^{-1} + \epsilon^2)\|\partial_t v_h\|_{L^2(d_j^2/2,4d_j^2;L^2(\Omega))}^2 + C\epsilon^{-2}\|\nabla v_h\|_{L^2(d_j^2/2,4d_j^2;L^2(\Omega))}^2 d_j^{-2}.$$

By using (A.10)-(A.11), the last inequality further implies

$$\begin{aligned}
&\|\partial_t\phi_h\|_{L^2(d_j^2,4d_j^2;L^2(D_j))} \\
&\leq (Ch^{1/2}d_j^{-1/2} + \epsilon)\|\partial_t\phi_h\|_{L^2(d_j^2/2,4d_j^2;L^2(D_j''))} \\
&\quad + C\epsilon^{-1}(\|\phi_h\|_{L^2(d_j^2/2,4d_j^2;L^2(D_j''))}d_j^{-2} + \|\nabla\phi_h\|_{L^2(d_j^2/2,4d_j^2;L^2(D_j''))}d_j^{-1}) \\
&\leq (Ch^{1/2}d_j^{-1/2} + \epsilon)\|\partial_t\phi_h\|_{L^2(d_j^2/2,4d_j^2;L^2(D_j''))} \\
&\quad + C\epsilon^{-1}(\|\phi_h\|_{L^\infty(d_j^2/2,4d_j^2;L^2(D_j''))}d_j^{-1} + \|\nabla\phi_h\|_{L^2(d_j^2/2,4d_j^2;L^2(D_j''))}d_j^{-1}). \qquad \text{(A.14)}
\end{aligned}$$

With an obvious change of domains, (A.13) and (A.14) imply

$$\begin{aligned}
&\|\phi_h\|_{L^\infty(d_j^2/2,4d_j^2;L^2(D_j'))} + \|\nabla\phi_h\|_{L^2(d_j^2/2,4d_j^2;L^2(D_j'))} \qquad \text{(A.15)} \\
&\leq C\|\partial_t\phi_h\|_{L^2(d_j^2/4,4d_j^2;L^2(D_j''))}h + \epsilon^2\|\nabla\phi_h\|_{L^2(d_j^2/4,4d_j^2;L^2(D_j''))} + C\epsilon^{-2}\|\phi_h\|_{L^2(d_j^2/4,4d_j^2;L^2(D_j''))}d_j^{-1}.
\end{aligned}$$

and

$$\begin{aligned}
\|\partial_t\phi_h\|_{L^2(d_j^2,4d_j^2;L^2(D_j))} &\leq (Ch^{1/2}d_j^{-1/2} + \epsilon)\|\partial_t\phi_h\|_{L^2(d_j^2/2,4d_j^2;L^2(D_j'))} \qquad \text{(A.16)} \\
&\quad + C\epsilon^{-1}(\|\phi_h\|_{L^\infty(d_j^2/2,4d_j^2;L^2(D_j'))}d_j^{-1} + \|\nabla\phi_h\|_{L^2(d_j^2/2,4d_j^2;L^2(D_j'))}d_j^{-1}).
\end{aligned}$$

respectively. The last two inequalities further imply

$$\begin{aligned}
&\|\partial_t\phi_h\|_{L^2(d_j^2,4d_j^2;L^2(D_j))} + d_j^{-1}\|\nabla\phi_h\|_{L^2(d_j^2,4d_j^2;L^2(D_j))} \\
&\leq (Ch^{1/2}d_j^{-1/2} + C\epsilon^{-1}hd_j^{-1} + \epsilon)(\|\partial_t\phi_h\|_{L^2(d_j^2/4,4d_j^2;L^2(D_j''))} + d_j^{-1}\|\nabla\phi_h\|_{L^2(d_j^2/4,4d_j^2;L^2(D_j''))}) \\
&\quad + C\epsilon^{-3}\|\phi_h\|_{L^2(d_j^2/4,4d_j^2;L^2(D_j''))}d_j^{-2}. \qquad \text{(A.17)}
\end{aligned}$$

Finally, combining (A.9) and (A.17) yields

$$\|\partial_t\phi_h\|_{L^2(Q_j)} + d_j^{-1}\|\nabla\phi_h\|_{L^2(Q_j)}$$

$$\leq (Ch^{1/2}d_j^{-1/2} + C\epsilon^{-1}hd_j^{-1} + \epsilon)\big(\|\partial_t\phi_h\|_{L^2(Q_j'')} + d_j^{-1}\|\nabla\phi_h\|_{L^2(Q_j'')}\big) + C\epsilon^{-3}d_j^{-2}\|\phi_h\|_{L^2(Q_j'')}$$
$$+ C\epsilon^{-1}(\|\nabla\phi_h(0)\|_{L^2(\Omega_j'')} + d_j^{-1}\|\phi_h(0)\|_{L^2(\Omega_j'')}). \tag{A.18}$$

Replacing Ω_j'' by Ω_j' and replacing Q_j'' by Q_j' in the last inequality, we obtain (A.1) and complete the proof of Lemma A.1. □

Proof of Lemma 5.1. Let $\widetilde{\omega}(t,x)$ be a smooth cut-off function which equals 1 in Q_j' and vanishes outside Q_j'', and let $\widetilde{\phi} = \widetilde{\omega}\phi$. Then $\widetilde{\phi} = \phi$ in Q_j', which implies that

$$\big(\partial_t(\widetilde{\phi} - \phi_h), \chi_h\big) + \big(\nabla(\widetilde{\phi} - \phi_h), \nabla\chi_h\big) = 0, \qquad \text{for } \chi_h \in S_h^0(\Omega_j'),\ t \in (0, d_j^2),$$
$$\big(\partial_t(\widetilde{\phi} - \phi_h), \chi_h\big) + \big(\nabla(\widetilde{\phi} - \phi_h), \nabla\chi_h\big) = 0, \qquad \text{for } \chi_h \in S_h^0(D_j'),\ t \in (d_j^2/4, 4d_j^2).$$

Let $\widetilde{\phi}_h \in S_h$ be the solution of

$$\big(\partial_t(\widetilde{\phi} - \widetilde{\phi}_h), \chi_h\big) + \big(\nabla(\widetilde{\phi} - \widetilde{\phi}_h), \nabla\chi_h\big) = 0, \qquad \forall\, \chi_h \in S_h, \tag{A.19}$$

with $\widetilde{\phi}_h(0) = \widetilde{\phi}(0) = 0$ so that

$$\big(\partial_t(\widetilde{\phi}_h - \phi_h), \chi_h\big) + \big(\nabla(\widetilde{\phi}_h - \phi_h), \nabla\chi_h\big) = 0, \qquad \text{for } \chi_h \in S_h^0(\Omega_j'),\ t \in (0, d_j^2), \tag{A.20}$$
$$\big(\partial_t(\widetilde{\phi}_h - \phi_h), \chi_h\big) + \big(\nabla(\widetilde{\phi}_h - \phi_h), \nabla\chi_h\big) = 0, \qquad \text{for } \chi_h \in S_h^0(D_j'),\ t \in (d_j^2/4, 4d_j^2). \tag{A.21}$$

We shall estimate $\widetilde{\phi} - \widetilde{\phi}_h$ and $\widetilde{\phi}_h - \phi_h$ separately.

The basic global energy estimates of (A.19) are (substituting $\chi_h = P_h\widetilde{\phi} - \widetilde{\phi}_h$ and $\chi_h = \partial_t(P_h\widetilde{\phi} - \widetilde{\phi}_h)$, respectively)

$$\|\nabla(\widetilde{\phi} - \widetilde{\phi}_h)\|_{L^2(\mathcal{Q})}^2 + \|\widetilde{\phi} - \widetilde{\phi}_h\|_{L^\infty(0,T;L^2(\Omega))}^2 \leq C\|\partial_t(\widetilde{\phi} - \widetilde{\phi}_h)\|_{L^2(\mathcal{Q})}\|\widetilde{\phi}\|_{L^2(\mathcal{Q})} + C\|\nabla\widetilde{\phi}\|_{L^2(\mathcal{Q})}^2,$$
$$\|\partial_t(\widetilde{\phi} - \widetilde{\phi}_h)\|_{L^2(\mathcal{Q})}^2 \leq C\|\partial_t\widetilde{\phi}\|_{L^2(\mathcal{Q})}^2 + C\|\nabla(\widetilde{\phi} - \widetilde{\phi}_h)\|_{L^2(\mathcal{Q})}\|\nabla\partial_t\widetilde{\phi}\|_{L^2(\mathcal{Q})},$$

which imply

$$\begin{aligned}
&\|\partial_t(\widetilde{\phi} - \widetilde{\phi}_h)\|_{L^2(\mathcal{Q})}^2 + d_j^{-2}\|\nabla\widetilde{\phi} - \widetilde{\phi}_h)\|_{L^2(\mathcal{Q})}^2 + d_j^{-2}\|\widetilde{\phi} - \widetilde{\phi}_h\|_{L^\infty(0,T;L^2(\Omega))}^2 \\
&\leq Cd_j^{-2}\|\partial_t(\widetilde{\phi} - \widetilde{\phi}_h)\|_{L^2(\mathcal{Q})}\|\widetilde{\phi}\|_{L^2(\mathcal{Q})} + Cd_j^{-2}\|\nabla\widetilde{\phi}\|_{L^2(\mathcal{Q})}^2, \\
&\quad + C\|\partial_t\widetilde{\phi}\|_{L^2(\mathcal{Q})}^2 + C\|\nabla(\widetilde{\phi} - \widetilde{\phi}_h)\|_{L^2(\mathcal{Q})}\|\nabla\partial_t\widetilde{\phi}\|_{L^2(\mathcal{Q})} \\
&\leq \frac{1}{2}\|\partial_t(\widetilde{\phi} - \widetilde{\phi}_h)\|_{L^2(\mathcal{Q})}^2 + Cd_j^{-4}\|\widetilde{\phi}\|_{L^2(\mathcal{Q})}^2 + Cd_j^{-2}\|\nabla\widetilde{\phi}\|_{L^2(\mathcal{Q})}^2, \\
&\quad + C\|\partial_t\widetilde{\phi}\|_{L^2(\mathcal{Q})}^2 + \frac{1}{2}d_j^{-2}\|\nabla(\widetilde{\phi} - \widetilde{\phi}_h)\|_{L^2(\mathcal{Q})}^2 + Cd_j^2\|\nabla\partial_t\widetilde{\phi}\|_{L^2(\mathcal{Q})}^2.
\end{aligned}$$

The first and fifth terms on the right-hand side above can be absorbed by the left-hand side, and the last inequality further reduces to

$$\|\partial_t(\widetilde{\phi} - \widetilde{\phi}_h)\|_{L^2(\mathcal{Q})} + d_j^{-1}\|\nabla(\widetilde{\phi} - \widetilde{\phi}_h)\|_{L^2(\mathcal{Q})} + d_j^{-1}\|\widetilde{\phi} - \widetilde{\phi}_h\|_{L^\infty(0,T;L^2(\Omega))}$$

$$\le Cd_j^{-2}\|\widetilde{\phi}\|_{L^2(\mathcal{Q})} + Cd_j^{-1}\|\nabla\widetilde{\phi}\|_{L^2(\mathcal{Q})} + C\|\partial_t\widetilde{\phi}\|_{L^2(\mathcal{Q})} + Cd_j\|\nabla\partial_t\widetilde{\phi}\|_{L^2(\mathcal{Q})}$$
$$\le Cd_j^{-2}\|\phi\|_{L^2(Q_j'')} + Cd_j^{-1}\|\nabla\phi\|_{L^2(Q_j'')} + C\|\partial_t\phi\|_{L^2(Q_j'')} + Cd_j\|\nabla\partial_t\phi\|_{L^2(Q_j'')}. \tag{A.22}$$

By applying Lemma A.1 to (A.20)-(A.21), we obtain

$$\begin{aligned}
&\|\partial_t(\widetilde{\phi}_h - \phi_h)\|_{L^2(Q_j)} + d_j^{-1}\|\nabla(\widetilde{\phi}_h - \phi_h)\|_{L^2(Q_j)} \\
&\le C\epsilon^{-1}\left(\|\nabla(\widetilde{\phi}_h - \phi_h)(0)\|_{L^2(\Omega_j')} + d_j^{-1}\|(\widetilde{\phi}_h - \phi_h)(0)\|_{L^2(\Omega_j')}\right) + C\epsilon^{-3}d_j^{-2}\|\widetilde{\phi}_h - \phi_h\|_{L^2(Q_j')} \\
&\quad + (Ch^{1/2}d_j^{-1/2} + C\epsilon^{-1}hd_j^{-1} + \epsilon)\big(\|\partial_t(\widetilde{\phi}_h - \phi_h)\|_{L^2(Q_j')} + d_j^{-1}\|\nabla(\widetilde{\phi}_h - \phi_h)\|_{L^2(Q_j')}\big) \\
&= C\epsilon^{-1}\left(\|\nabla\phi_h(0)\|_{L^2(\Omega_j')} + d_j^{-1}\|\phi_h(0)\|_{L^2(\Omega_j')}\right) + C\epsilon^{-3}d_j^{-2}\|\widetilde{\phi}_h - \phi_h\|_{L^2(Q_j')} \\
&\quad + (Ch^{1/2}d_j^{-1/2} + C\epsilon^{-1}hd_j^{-1} + \epsilon)\big(\|\partial_t(\widetilde{\phi}_h - \phi_h)\|_{L^2(Q_j')} + d_j^{-1}\|\nabla(\widetilde{\phi}_h - \phi_h)\|_{L^2(Q_j')}\big)
\end{aligned} \tag{A.23}$$

where we have used the identity $\widetilde{\phi}_h(0) = 0$ in the last step. Splitting $\widetilde{\phi}_h - \phi_h$ into $(\widetilde{\phi} - \phi_h) + (\widetilde{\phi}_h - \widetilde{\phi})$ in the right-hand side of the last inequality yields

$$\begin{aligned}
&\|\partial_t(\widetilde{\phi}_h - \phi_h)\|_{L^2(Q_j)} + d_j^{-1}\|\nabla(\widetilde{\phi}_h - \phi_h)\|_{L^2(Q_j)} \\
&\le C\epsilon^{-1}\left(\|\nabla\phi_h(0)\|_{L^2(\Omega_j')} + d_j^{-1}\|\phi_h(0)\|_{L^2(\Omega_j')}\right) \\
&\quad + C\epsilon^{-3}d_j^{-2}\|\widetilde{\phi} - \phi_h\|_{L^2(Q_j')} \\
&\quad + (Ch^{1/2}d_j^{-1/2} + C\epsilon^{-1}hd_j^{-1} + \epsilon)\big(\|\partial_t(\widetilde{\phi} - \phi_h)\|_{L^2(Q_j')} + d_j^{-1}\|\nabla(\widetilde{\phi} - \phi_h)\|_{L^2(Q_j')}\big) \\
&\quad + C\epsilon^{-3}d_j^{-2}\|\widetilde{\phi}_h - \widetilde{\phi}\|_{L^2(Q_j')} \\
&\quad + (Ch^{1/2}d_j^{-1/2} + C\epsilon^{-1}hd_j^{-1} + \epsilon)\big(\|\partial_t(\widetilde{\phi}_h - \widetilde{\phi})\|_{L^2(Q_j')} + d_j^{-1}\|\nabla(\widetilde{\phi}_h - \widetilde{\phi})\|_{L^2(Q_j')}\big) \\
&\le C\epsilon^{-1}\left(\|\nabla\phi_h(0)\|_{L^2(\Omega_j')} + d_j^{-1}\|\phi_h(0)\|_{L^2(\Omega_j')}\right) + C\epsilon^{-3}d_j^{-2}\|\phi - \phi_h\|_{L^2(Q_j')} \\
&\quad + (Ch^{1/2}d_j^{-1/2} + C\epsilon^{-1}hd_j^{-1} + \epsilon)\big(\|\partial_t(\phi - \phi_h)\|_{L^2(Q_j')} + d_j^{-1}\|\nabla(\phi - \phi_h)\|_{L^2(Q_j')}\big) \\
&\quad + \left(Ch^{1/2}d_j^{-1/2} + C\epsilon^{-1}hd_j^{-1} + C\epsilon + C\epsilon^{-3}\right)d_j^{-2}\|\phi\|_{L^2(Q_j'')} \\
&\quad + \left(Ch^{1/2}d_j^{-1/2} + C\epsilon^{-1}hd_j^{-1} + C\epsilon + C\epsilon^{-3}\right)\left(d_j^{-1}\|\nabla\phi\|_{L^2(Q_j'')} + \|\partial_t\phi\|_{L^2(Q_j'')} + d_j\|\nabla\partial_t\phi\|_{L^2(Q_j'')}\right)
\end{aligned} \tag{A.24}$$

where we have used the identity $\widetilde{\phi} = \phi$ on Q_j' and (A.22) in the last step. Since

$$\left(Ch^{1/2}d_j^{-1/2} + C\epsilon^{-1}hd_j^{-1} + C\epsilon + C\epsilon^{-3}\right) \le C\epsilon^{-3},$$

the last inequality reduces to

$$\begin{aligned}
&\|\partial_t(\widetilde{\phi}_h - \phi_h)\|_{L^2(Q_j)} + d_j^{-1}\|\nabla(\widetilde{\phi}_h - \phi_h)\|_{L^2(Q_j)} \\
&\le C\epsilon^{-1}\left(\|\nabla\phi_h(0)\|_{L^2(\Omega_j')} + d_j^{-1}\|\phi_h(0)\|_{L^2(\Omega_j')}\right) + C\epsilon^{-3}d_j^{-2}\|\phi - \phi_h\|_{L^2(Q_j')}
\end{aligned}$$

$$+ (Ch^{1/2}d_j^{-1/2} + C\epsilon^{-1}hd_j^{-1} + \epsilon)\big(\|\partial_t(\phi-\phi_h)\|_{L^2(Q_j')} + d_j^{-1}\|\nabla(\phi-\phi_h)\|_{L^2(Q_j')}\big) \tag{A.25}$$

$$+ C\epsilon^{-3}\big(d_j^{-2}\|\phi\|_{L^2(Q_j'')} + d_j^{-1}\|\nabla\phi\|_{L^2(Q_j'')} + \|\partial_t\phi\|_{L^2(Q_j'')} + d_j\|\nabla\partial_t\phi\|_{L^2(Q_j'')}\big),$$

The estimates (A.22) and (A.25) imply

$$\begin{aligned}
&\|\partial_t(\phi-\phi_h)\|_{L^2(Q_j)} + d_j^{-1}\|\nabla(\phi-\phi_h)\|_{L^2(Q_j)} \\
&= \|\partial_t(\widetilde\phi-\phi_h)\|_{L^2(Q_j)} + d_j^{-1}\|\nabla(\widetilde\phi-\phi_h)\|_{L^2(Q_j)} \\
&\le \|\partial_t(\widetilde\phi_h-\phi_h)\|_{L^2(Q_j)} + d_j^{-1}\|\nabla(\widetilde\phi_h-\phi_h)\|_{L^2(Q_j)} \qquad \text{(here we use triangle inequality)} \\
&\quad + \|\partial_t(\widetilde\phi-\widetilde\phi_h)\|_{L^2(Q_j)} + d_j^{-1}\|\nabla(\widetilde\phi-\widetilde\phi_h)\|_{L^2(Q_j)} \\
&\le C\epsilon^{-1}\left(\|\nabla\phi_h(0)\|_{L^2(\Omega_j'')} + d_j^{-1}\|\phi_h(0)\|_{L^2(\Omega_j'')}\right) + C\epsilon^{-3}d_j^{-2}\|\phi-\phi_h\|_{L^2(Q_j'')} \\
&\quad + (Ch^{1/2}d_j^{-1/2} + C\epsilon^{-1}hd_j^{-1} + \epsilon)\big(\|\partial_t(\phi-\phi_h)\|_{L^2(Q_j'')} + d_j^{-1}\|\nabla(\phi-\phi_h)\|_{L^2(Q_j'')}\big) \\
&\quad + C\epsilon^{-3}\big(d_j^{-2}\|\phi\|_{L^2(Q_j'')} + d_j^{-1}\|\nabla\phi\|_{L^2(Q_j'')} + \|\partial_t\phi\|_{L^2(Q_j'')} + d_j\|\nabla\partial_t\phi\|_{L^2(Q_j'')}\big)
\end{aligned}$$

Replacing ϕ by $\phi - I_h\phi$, replacing Ω_j'' by Ω_j', and replacing Q_j'' by Q_j' in the last inequality, we obtain (5.2) and complete the proof of Lemma 5.1. □

Appendix B: Property (P3) and the operator I_h

Let Φ_i be the basis function of the finite element space S_h corresponding to the finite element nodes $x_i \in \Omega$, $i = 1\dots, M$. In other words, we have $\Phi_j(x_i) = \delta_{ij}$ (the Kronecker symbol). Let τ_i denote the union of triangles (or tetrahedra in $\mathbb{R}^3$) whose closure contain the node x_i. For any function $v \in L^2(\Omega)$, we denote by $P_h^{(i)}v$ the local L^2 projection onto $S_h(\tau_i)$ (the space of finite element functions defined on the region τ_i). The operator $I_h : L^2(\Omega) \to S_h$ is defined as (in the spirit of Clément's interpolation operator, cf. [6])

$$(I_h v)(x) = \sum_{i=1}^{M} (P_h^{(i)}v)(x_i)\Phi_i(x), \quad \text{for } x \in \Omega, \tag{B.1}$$

which equals zero on the boundary $\partial\Omega$ (as every Φ_i equals zero on $\partial\Omega$).

Now we prove that the operator I_h defined in (B.1) satisfies property (P3) of Section 3. To this end, we let S_h' be the finite element space subject to the same mesh as S_h, with the same order of finite elements, but not necessarily zero on the boundary $\partial\Omega$. We denote by x_j', $j = 1, \dots, m$, the finite element nodes on the boundary $\partial\Omega$, and we denote by Φ_j' the basis function corresponding to the node x_j'. The notation τ_j' will denote the union of triangles (or tetrahedra in $\mathbb{R}^3$) whose closure contain x_j'. With these notations, the space S_h' is spanned by the basis functions Φ_i, $i = 1, \dots, M$, and Φ_j', $j = 1, \dots, m$. We define an auxiliary operator $\widetilde I_h : H^1(\Omega) \to S_h'$ by setting

$$(\widetilde I_h\phi)(x) = (I_h\phi)(x) + \sum_{j=1}^{m} (\widetilde P_h^{(j)}\phi)(x_j')\Phi_j'(x), \quad \text{for } x \in \Omega, \tag{B.2}$$

where $\widetilde{P}_h^{(j)}\phi$ is the L^2 projection of $\phi|_{\partial\Omega}$ (trace of ϕ on the boundary) onto $S_h(\partial\Omega \cap \overline{\tau}'_j)$ (the space of finite element functions on $\partial\Omega \cap \overline{\tau}'_i$, a piece of the boundary). The definition (B.2) implies

$$I_h \phi = \widetilde{I}_h \phi, \qquad \forall\, \phi \in H_0^1(\Omega). \tag{B.3}$$

Hence, in order to prove property (P3)-(1), we only need to prove the corresponding error estimate for the operator $\widetilde{I}_h$.

In fact, the definition (B.2) guarantees the following local stability:

$$\begin{aligned}
&\|\widetilde{I}_h\phi\|_{L^2(\tau_j')} + h^{\frac{1}{2}}\|\widetilde{I}_h\phi\|_{L^2(\partial\Omega\cap\tau_j')} + h\|\nabla\widetilde{I}_h\phi\|_{L^2(\tau_j')} \\
&\qquad \le C(\|\phi\|_{L^2(\widetilde{\tau}_j')} + h^{\frac{1}{2}}\|\phi\|_{L^2(\partial\Omega\cap\widetilde{\tau}_j')}), \\
&\|\widetilde{I}_h\phi\|_{L^2(\tau_i)} + h\|\nabla\widetilde{I}_h\phi\|_{L^2(\tau_j)} \le C\|\phi\|_{L^2(\widetilde{\tau}_i)},
\end{aligned}$$

where $\widetilde{\tau}_j$ is the union of triangles (or tetrahedra in $\mathbb{R}^3$) whose closure intersect the closure of τ_j' (boundary triangle/tetrahedron), and $\widetilde{\tau}_i$ is the union of triangles (or tetrahedra in $\mathbb{R}^3$) whose closure intersect the closure of τ_i (interior triangle/tetrahedron). Let P_h' denote the L^2 projection from $L^2(\Omega)$ onto the finite element space S_h' . Then substituting $\phi = v - P_h'v$ into the two inequalities above yields

$$\begin{aligned}
&\|v-\widetilde{I}_hv\|_{L^2(\tau_j')} + h^{\frac{1}{2}}\|v-\widetilde{I}_hv\|_{L^2(\partial\Omega\cap\tau_j')} + h\|\nabla(v-\widetilde{I}_hv)\|_{L^2(\tau_j')} \\
&\le C(\|v-P_h'v\|_{L^2(\widetilde{\tau}_j')} + h^{\frac{1}{2}}\|v-P_h'v\|_{L^2(\partial\Omega\cap\widetilde{\tau}_j')} + h\|\nabla(v-P_h'v)\|_{L^2(\widetilde{\tau}_j')})
\end{aligned}$$

and

$$\|v-\widetilde{I}_hv\|_{L^2(\tau_i)} + h\|\nabla(v-\widetilde{I}_hv)\|_{L^2(\tau_j)} \le C\|v-P_h'v\|_{L^2(\widetilde{\tau}_i)} + h\|\nabla(v-P_h'v)\|_{L^2(\tau_j)}.$$

Summing up the two inequalities above for $i = 1, \dots, M$ and $j = 1, \dots, m$ yields

$$\begin{aligned}
&\|v-\widetilde{I}_hv\|_{L^2(\Omega)} + h^{\frac{1}{2}}\|v-\widetilde{I}_hv\|_{L^2(\partial\Omega)} + h\|\nabla(v-\widetilde{I}_hv)\|_{L^2(\Omega)} \\
&\le C(\|v-P_h'v\|_{L^2(\Omega)} + h^{\frac{1}{2}}\|v-P_h'v\|_{L^2(\partial\Omega)} + h\|\nabla(v-P_h'v)\|_{L^2(\Omega)}) \\
&\le C(\|v-P_h'v\|_{L^2(\Omega)} + h^{\frac{1}{2}}\|v-P_h'v\|_{L^2(\Omega)}^{\frac{1}{2}}\|v-P_h'v\|_{H^1(\Omega)}^{\frac{1}{2}} + h\|\nabla(v-P_h'v)\|_{L^2(\Omega)}) \\
&\hspace{20em} \text{(interpolation inequality)} \\
&\le C(\|v-P_h'v\|_{L^2(\Omega)} + h\|v-P_h'v\|_{H^1(\Omega)}) \\
&\hspace{20em} \text{(Hölder's inequality)} \\
&\le \begin{cases} Ch\|v\|_{H^1(\Omega)} & \text{if } v \in H^1(\Omega), \\ Ch^2\|v\|_{H^2(\Omega)} & \text{if } v \in H^2(\Omega), \end{cases}
\end{aligned}$$

where the last inequality is the basic estimate of the L^2 projection $P_h' : L^2(\Omega) \to S_h'$ (without imposing boundary condition). By the complex interpolation method, we have

$$\|v-\widetilde{I}_hv\|_{L^2(\Omega)} + h^{\frac{1}{2}}\|v-\widetilde{I}_hv\|_{L^2(\partial\Omega)} + h\|\nabla(v-\widetilde{I}_hv)\|_{L^2(\Omega)} \le Ch^{1+\alpha}\|v\|_{H^{1+\alpha}(\Omega)}.$$

Hence, for $v \in H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$, we have $I_hv = \widetilde{I}_hv = 0$ on $\partial\Omega$ and

$$\begin{aligned}
&\|v-I_hv\|_{L^2(\Omega)} + h\|\nabla(v-I_hv)\|_{L^2(\Omega)} \\
&= \|v-\widetilde{I}_hv\|_{L^2(\Omega)} + h^{\frac{1}{2}}\|v-\widetilde{I}_hv\|_{L^2(\partial\Omega)} + h\|\nabla(v-\widetilde{I}_hv)\|_{L^2(\Omega)} \\
&\le Ch^{1+\alpha}\|v\|_{H^{1+\alpha}(\Omega)}.
\end{aligned} \tag{B.4}$$

This proves property (P3)-(1) in Section 3. The other properties in (P3) are simple consequences of the definition of the operator I_h. □

Appendix C: Proof of (2.15)

The proof of (2.15) requires some properties of the finite element space described in Section 3.2.

It suffices to prove that the solution $w_h \in S_h$ of the finite element equation $\Delta_h w_h = f_h$ satisfies

$$\|w_h\|_{L^\infty} \leq C\|f_h\|_{L^\infty}. \tag{C.1}$$

To this end, we define $w \in H_0^1(\Omega)$ as the solution of the following PDE problem:

$$\begin{cases} \Delta w = f_h & \text{in } \Omega, \\ w = 0 & \text{on } \partial\Omega. \end{cases} \tag{C.2}$$

Then w_h is the Ritz projection of w, and the following standard H^1-norm error estimate holds for some $\alpha \in \left(\frac{1}{2}, 1\right)$:

$$\|I_h w - w_h\|_{H^1} \leq C\|I_h w - w\|_{H^1} \leq Ch^\alpha \|w\|_{H^{1+\alpha}} \leq Ch^\alpha \|f_h\|_{H^{-1+\alpha}}, \tag{C.3}$$

where the first inequality above is due to H^1-stability of the Ritz projection, the second inequality due to (B.4), and the last inequality due to Lemma 4.2. Consequently, we have

$$\begin{aligned}
\|I_h w - w_h\|_{L^\infty} &\leq Ch^{-\frac{N}{6}}\|I_h w - w_h\|_{L^6} && \text{(inverse inequality)} \\
&\leq Ch^{-\frac{1}{2}}\|I_h w - w_h\|_{H^1} && \text{(Sobolev embedding, } N = 2, 3) \\
&\leq Ch^{\alpha-\frac{1}{2}}\|f_h\|_{H^{-1+\alpha}} && \text{(use (C.3))} \\
&\leq Ch^{\alpha-\frac{1}{2}}\|f_h\|_{L^2} && (L^2 \hookrightarrow H^{-1+\alpha}) \\
&\leq Ch^{\alpha-\frac{1}{2}}\|f_h\|_{L^\infty}, && \text{(C.4)} \\
\|w - I_h w\|_{L^\infty} &\leq \|w - P_h w\|_{L^\infty} + \|P_h w - I_h w\|_{L^\infty} && \text{(triangle inequality)} \\
&\leq Ch^{1+\alpha-\frac{N}{2}}\|w\|_{C^{1+\alpha-\frac{N}{2}}} + Ch^{-\frac{N}{2}}\|P_h w - I_h w\|_{L^2} && \text{(inverse inequality)} \\
&\leq Ch^{1+\alpha-\frac{N}{2}}\|w\|_{C^{1+\alpha-\frac{N}{2}}} + Ch^{-\frac{N}{2}}\|w - I_h w\|_{L^2} && (L^2\text{-stability of } P_h) \\
&\leq Ch^{1+\alpha-\frac{N}{2}}\|w\|_{H^{1+\alpha}} + Ch^{1+\alpha-\frac{N}{2}}\|w\|_{H^{1+\alpha}} && (H^{1+\alpha} \hookrightarrow C^{1+\alpha-\frac{N}{2}} \text{ and (B.4)}) \\
&\leq Ch^{\alpha-\frac{1}{2}}\|f_h\|_{H^{-1+\alpha}} && \text{(use Lemma 4.2 and } N = 2, 3) \\
&\leq Ch^{\alpha-\frac{1}{2}}\|f_h\|_{L^2} && (L^2 \hookrightarrow H^{-1+\alpha} \text{ for } \alpha \in (\tfrac{1}{2}, 1)) \\
&\leq Ch^{\alpha-\frac{1}{2}}\|f_h\|_{L^\infty}. && \text{(C.5)}
\end{aligned}$$

Consequently, the triangle inequality implies

$$\begin{aligned}
\|w_h\|_{L^\infty} &\leq \|w\|_{L^\infty} + \|w - I_h w\|_{L^\infty} + \|I_h w - w_h\|_{L^\infty} \\
&\leq \|w\|_{L^\infty} + Ch^{\alpha-\frac{1}{2}}\|f_h\|_{L^\infty} + Ch^{\alpha-\frac{1}{2}}\|f_h\|_{L^\infty} && \text{(use (C.4) and (C.5))} \\
&\leq C\|w\|_{H^{1+\alpha}} + Ch^{\alpha-\frac{1}{2}}\|f_h\|_{L^\infty} + Ch^{\alpha-\frac{1}{2}}\|f_h\|_{L^\infty} && (H^{1+\alpha} \hookrightarrow L^\infty \text{ for } \alpha \in (\tfrac{1}{2}, 1),\ N = 2, 3) \\
&\leq C\|f_h\|_{H^{-1+\alpha}} + Ch^{\alpha-\frac{1}{2}}\|f_h\|_{L^\infty} && \text{(use Lemma 4.2)}
\end{aligned}$$

$$\leq C\|f_h\|_{L^2} + Ch^{\alpha-\frac{1}{2}}\|f_h\|_{L^\infty} \qquad (L^2 \hookrightarrow H^{-1+\alpha} \text{ for } \alpha \in (\tfrac{1}{2}, 1))$$
$$\leq C\|f_h\|_{L^\infty}. \tag{C.6}$$

This proves (C.1). The proof of (2.15) is complete. □

Acknolwedgement

I thank the anonymous referees for the valuable comments and corrections.

References

[1] R. A. Adams and J. J. F. Fournier. *Sobolev Spaces.* Academic Press, Amsterdam, second edition, 2003.

[2] 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.

[3] G. Akrivis, B. Li, and C. Lubich. Combining maximal regularity and energy estimates for time discretizations of quasilinear parabolic equations. *Math. Comp.*, (86):1527–1552, 2017.

[4] H. Amann. *Linear and Quasilinear Parabolic Problems. Volume I: Abstract linear theory.* Birkhäuser Boston Inc., Boston, MA, 1995.

[5] J. Bergh and J. Löfström. *Interpolation Spaces: An Introduction.* Springer-Verlag Berlin Heidelberg, Printed in Germany, 1976.

[6] P. Clément. Approximation by finite element functions using local regularization. *Revue française d'automatique, informatique, recherche opérationnelle. Analyse numérique*, 9:77–84, 1975.

[7] P. Clément and J. Prüss. Global existence for a semilinear prabolic Volterra equation. *Mathematische Zeitschrift*, 209:17–26, 1992.

[8] M. Crouzeix. Contractivity and analyticity in l^p of some approximation of the heat equation. *Numerical Algorithms*, 33:193–201, 2003.

[9] M. Dauge. *Elliptic Boundary Value Problems in Corner Domains.* Springer-Verlag Berlin Heidelberg, 1988.

[10] E. B. Davies. *Heat Kernels and Spectral Theory.* Cambridge University Press, Cambridge, 1989.

[11] A. Demlow, O. Lakkis, and C. Makridakis. A posteriori error estimates in the maximum norm for parabolic problem. *SIAM J. Numer. Anal.*, 47:2157–2176,, 2009.

[12] A. Demlow, D. Leykekhman, A. H. Schatz, and L. B. Wahlbin. Best approximation property in the norm for finite element methods on graded meshes. *Math. Comp.*, 81:743–764, 2012.

[13] D. A. Di Pietro and J. Droniou. $W^{s,p}$-approximation properties of elliptic projectors on polynomial spaces, with application to the error analysis of a Hybrid High-Order discretisation of Leray–Lions problems. *Math. Models Methods Appl. Sci.*, 27:879–908, 2017.

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

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

[16] L. Grafakos. *Classical Fourier Analysis*. Springer Science + Business Media, LLC, second edition, 2008.

[17] J. Guzmán, D. Leykekhman, J. Rossmann, and A. H. Schatz. Hölder estimates for Green's functions on convex polyhedral domains and their applications to finite element methods. *Numer. Math.*, 112:221–243,, 2009.

[18] A. Hansbo. Strong stability and non-smooth data error estimates for discretizations of linear parabolic problems. *BIT Numer. Math.*, 42:351–379, 2002.

[19] T. D. J. Douglas, Jr. and L. Wahlbin. The stability in L^q of the L^2-projection into finite element function spaces. *Numer. Math.*, 23:193–197, 1975.

[20] N. J. Kalton and G. Lancien. A solution to the problem of L^p-maximal regularity. *Math. Z.*, 235:559–568, 2000.

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

[22] B. Kovács, B. Li, and C. Lubich. A-stable time discretizations preserve maximal parabolic regularity. *SIAM J. Numer. Anal.*, 54:3600–3624, 2016.

[23] S. Krantz. *A Panorama of Harmonic Analysis*. American History Series. Mathematical Association of America, 1999.

[24] P. C. Kunstmann, B. Li, and C. Lubich. Runge-Kutta time discretization of nonlinear parabolic equations studied via discrete maximal parabolic regularity. *Found. Comput. Math.*, to appear.

[25] P. C. Kunstmann and L. Weis. *Maximal L^p-regularity for parabolic equations, Fourier multiplier theorems and H^∞-functional calculus*, pages 65–311. Functional Analytic Methods for Evolution Equations, edited by M. Lannelli, R. Nagel, and S. Piazzera. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

[26] O. Ladyzhenskaia, V. Solonnikov, and N. Ural'tseva. *Linear and Quasi-linear Equations of Parabolic Type*. American Mathematical Society, Translations of Mathematical Monographs.

[27] D. Leykekhman. Pointwise localized error estimates for parabolic finite element equations. *Numer. Math.*, 96:583–600, 2004.

[28] D. Leykekhman and B. Vexler. Finite element pointwise results on convex polyhedral domains. *SIAM J. Numer. Anal.*, 54:561–587, 2016.

[29] D. Leykekhman and B. Vexler. Pointwise best approximation results for Galerkin finite element solutions of parabolic problems. *SIAM J. Numer. Anal.*, 54:1365–1384, 2016.

[30] D. Leykekhman and B. Vexler. Discrete maximal parabolic regularity for Galerkin finite element methods. *Numer. Math.*, 135:923–952, 2017.

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

[32] 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.*, 53:1418–1437, 2015.

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

[34] B. Li and W. Sun. Maximal regularity of fully discrete finite element solutions of parabolic equations. *SIAM J. Numer. Anal.*, 55:521–542, 2017.

[35] P. L. Lions. *Mathematical Topics in Fluid Mechanics. Volume 1. Incompressible Models.* Clarendon Press, Oxford, 1996.

[36] J. A. Nitsche and M. F. Wheeler. L_∞-boundedness of the finite element Galerkin operator for parabolic problems. *Numer. Funct. Anal. Optim.*, 4:325–353, 1982.

[37] J. T. Oden and J. N. Reddy. *An Introduction to the Mathematical Theory of Finite Elements*. Dover Books on Engineering. Dover Publications, 2012.

[38] E.-M. Ouhabaz. Gaussian estimates and holomorphy of semigroups. *Proceedings of the American Mathematical Society*, 123:1465–1474, 1995.

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

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

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

[42] A. H. Schatz. A weak discrete maximum principle and stability of the finite element method in L_∞ on plane polygonal domains. I. *Math. Comp.*, 31:77–91,, 1980.

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

[44] 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.*, 51:1349–1385, 1998.

[45] A. H. Schatz and L. B. Wahlbin. Interior maximum-norm estimates for finite element methods II. *Math. Comp.*, 64:907–928, 1995.

[46] V. Thomée. *Galerkin Finite Element Methods for Parabolic Problems.* Springer-Verlag, New York, second edition, 2006.

[47] 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.*, 87:373–389, 2000.

[48] L. B. Wahlbin. *Local behavior in finite element methods.* "Handbook of Numerical Analysis II: Finite Element Methods", pp. 353-522, edited by P. G. Ciarlet and J. L. Lions. North-Holland, Amsterdam, 1991.

[49] L. Weis. Operator-valued Fourier multiplier theorems and maximal L_p-regularity. *Math. Ann.*, 319.

[50] L. Weis. *A new approach to maximal L_p-regularity.* Lecture Notes in Pure and Applied Mathematics 215: Evolution Equations and Their Applications in Physical and Life Sciences, pp. 195–214, edited by G. Lumer. Marcel Dekker, 2001.

[51] K. Yosida. *Functional Analysis.* Springer-Verlag, New York, sixth edition, 1980.