

Weak maximum principle of finite element methods for parabolic equations in polygonal domains

Genming Bai · Dmitriy Leykekhman · Buyang Li

Received: date / Accepted: date

Abstract The weak maximum principle of finite element methods for parabolic equations is proved for both semi-discretization in space and fully discrete methods with k -step backward differentiation formulae for $k = 1, \dots, 6$, on a two-dimensional general polygonal domain or a three-dimensional convex polyhedral domain. The semi-discrete result is established via a dyadic decomposition argument and local energy estimates in which the nonsmoothness of the domain can be handled. The fully discrete result for multistep backward differentiation formulae is proved by utilizing the solution representation via the discrete Laplace transform and the resolvent estimates, which are inspired by the analysis of convolutional quadrature for parabolic and fractional-order partial differential equations.

Keywords Parabolic equation, finite element method, weak maximum principle, full discretization, backward differentiation formulae, Laplace transform, analytic semigroup, local energy estimate, nonsmooth domain.

1 Introduction

Maximum principle serves as a fundamental mathematical tool for the study of elliptic and parabolic partial differential equations (PDEs). The discrete counterpart of maximum principle associated to finite element methods (FEMs) has a well-established history of research and remains an active field. However, unlike its continuous counterpart, the discrete maximum principle is not an inherent property and is significantly influenced by the triangulation of the physical domain [32, §5].

In two dimensions, the applicability of the discrete maximum principle for elliptic equations is primarily confined to piecewise linear elements, as quadratic elements necessitate uniform equilateral triangulation [11]. The scenario becomes more intricate in three dimensions [2, 14, 15, 34]. Particularly, it becomes challenging to assure

G. Bai and B. Li

Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Hong Kong. Email address: genming.bai@connect.polyu.hk, buyang.li@polyu.edu.hk

D. Leykekhman

Department of Mathematics, University of Connecticut, Storrs, CT 06269, USA. Email address: dmitriy.leykekhman@uconn.edu

the discrete maximum principle for even piecewise linear elements. Nonetheless, a large number of applications do not require a strong discrete maximum principle. Schatz demonstrated in [26] that a weak maximum principle (also known as the Agmon-Miranda principle) is applicable to a broad spectrum of finite elements on general quasi-uniform triangulation in any two-dimensional polygonal domains. This weak maximum principle has recently been extended to three-dimensional convex polyhedral domains [19] and the Neumann problem in two-dimensional polygonal domains [18, 23]. The situation for parabolic equations is more complex, depending not only on space discretization but also on time discretization.

This paper focuses on the weak maximum principle of semi-discrete and fully discrete FEMs for the following parabolic equation:

$$\begin{cases} \partial_t u - \Delta u = 0 & \text{in } (0, \infty) \times \Omega, \\ u = g & \text{on } [0, \infty) \times \partial\Omega, \\ u|_{t=0} = u(0) & \text{in } \Omega. \end{cases} \quad (1.1)$$

Specifically, we are interested in determining whether the semi-discrete and fully discrete finite element solutions of (1.1) are uniformly bounded by constant multiples of $\|g\|_{L^\infty(\partial\Omega)}$ and $\|u(0)\|_{L^\infty(\Omega)}$.

The discrete maximum principle of fully discrete FEMs for parabolic equations was initially addressed by Fujii in [7], demonstrating that the backward Euler method may satisfy a maximum principle for specific families of acute triangulations. However, his results require a lower bound for the time step size, and therefore do not imply the discrete maximum principle for the semi-discrete FEM (which would necessitate the step size tending to zero). In [31], it was established that without the application of the mass lumping method, the strong discrete maximum principle for the semi-discrete FEM does not hold, even for acute triangulation, thereby making the mass lumping method a necessary requirement. These results were extended to more general conditions as well as more general single-step time-stepping methods (again, with a lower bound for the step size) and numerically illustrated in [3]. To the best of our knowledge, the weak discrete maximum principle of the semi-discrete FEM for parabolic equations has not been addressed, a gap this paper aims to fill.

The weak maximum principle of the semi-discrete FEM with respect to the initial value is equivalent to the uniform boundedness of the finite element heat semigroup in $L^\infty(\Omega)$. The latter was proved in [27] for smooth domains by assuming that the partition of the domain contains curved simplices which fit the geometry of the boundary exactly. The extension to nonsmooth polygonal and polyhedral domains, thus a standard triangulation with flat simplices can fit the boundary exactly, introduces more technical challenges where the elliptic regularity theory and finite element error estimates degenerate. A comprehensive solution to these issues, which emerge due to the domain's nonsmoothness, was presented in a recent paper [21]. However, the weak maximum principle of the semi-discrete FEM with respect to the boundary value g remains unaddressed, and this issue can't be transformed into the uniform boundedness of the semigroup in $L^\infty(\Omega)$. This is proved in the current paper by the representation formula via discrete Laplace transform, discrete resolvent estimates on the contour and the consistency of discrete and continuous Laplace transform together with the elliptic weak maximum principle developed in [19].

Regarding time discretizations, the maximal $L^p L^q$ regularity for multistep BDF and Runge-Kutta time-stepping methods, with $1 < p, q < \infty$, was shown in [16] and [22] for semi-discretization in time and fully discrete FEMs, respectively. The authors utilized the R -boundedness concept to reduce maximal $L^p L^q$ regularity to set inclusion [17, Theorem 1.11], which is further characterized by the concept of $A(\alpha)$ -

stability. In addition to BDF, the discontinuous Galerkin time-stepping method was studied in [20], and the θ -scheme was considered in [13]. These findings are closely tied to the weak maximum principle of time discretizations with respect to the initial value. However, the weak maximum principle of time discretizations with respect to the boundary value has not been addressed. This question is answered affirmatively in the current paper for k -step BDF methods with $k = 1, \dots, 6$. The analysis was particularly challenging due to the existence of boundary data g , the domain's nonsmoothness, and the lack of A -stability for k -step BDF with $3 \leq k \leq 6$. The two main points of analysis in this paper are:

- The weak maximum principle of the semi-discrete FEM requires proving the following estimate:

$$\|\Gamma_h - \Gamma\|_{L^1(0, \infty; L^1(D_h))} \leq h, \quad (1.2)$$

where $D_h = \{x \in \Omega : \text{dist}(x, \partial\Omega) \leq h\}$, Γ and Γ_h are the parabolic Green's functions with respect to the regularized delta function $\tilde{\delta}$ (see [30, Lemma 2.2]) and the discrete delta function δ_h (to be defined in Section 3.3), respectively. This error estimate is not straightforward due to the lack of an approximation property for the initial value $\Gamma_h(0) - \Gamma(0) = \delta_h - \tilde{\delta} = (P_h - 1)\tilde{\delta}$ in the error equation (where P_h denotes the L^2 orthogonal projection onto the finite element space), as the initial value $\tilde{\delta}$ is only bounded by a universal constant C in the (very weak) L^1 norm. The standard nonsmooth data error estimate [29, Chapter 3] and the negative norm error estimate [29, Chapter 5] are both hindered in the case of a nonsmooth domain. Note that the error estimate (1.2) is applicable when the underlying domain is sufficiently smooth [27, Proposition 3.2]. However, for a nonconvex polygon, the argument of [27, Proposition 3.2] is no longer valid due to the degeneracy of the consistency error and the finite element error estimates. The solution is to leverage the local Hölder's inequality, the smoothing property of the discrete heat semigroup, and the lower bound $\alpha > \frac{1}{2}$, where α is the indicator of the domain's smoothness [21, Lemma 4.2]. The lower bound $\frac{1}{2}$ is sharp in our proof in the sense that our proof does not work if $\alpha < \frac{1}{2}$.

- Weak maximum principle of time discretizations is accomplished by studying the representation formula of u_h^n in terms of the generating functions $\tilde{u}(\zeta) = \sum_{n=k}^{\infty} u_h^n \zeta^n$, which may alternatively be understood as the discrete Laplace transform or z -transform of the fully discrete finite element solution. For the representation of $u_{h,2}^n$ the second part of the fully discrete solution readers are referred to equations (5.17), (5.32), and (5.33).

However, it is crucial to note that a direct application of the L^∞ norm to the representation formula of u_h^n introduces a logarithmic factor. This factor emerges from the summation process, which is a consequence of the discrete convolution as in convolutional quadrature [25], [12, Chapter 3]:

$$\sum_{j=1}^{\lfloor T/\tau \rfloor} \frac{1}{j} \approx \log \frac{T}{\tau}. \quad (1.3)$$

In order to eliminate this logarithmic factor, a comparison is conducted between the representation formulas of the discrete and continuous Laplace transforms, as outlined in equation (5.34). This comparison necessitates the development of consistency between these two types of Laplace transform.

We draw attention to a central correspondence relation concerning the boundary data, as stated in Lemma 5.4:

$$\hat{g}_h(z) = \frac{e^{-\tau z} + e^{\tau z} - 2}{z^2 \tau} \tilde{g}_h(e^{-\tau z}),$$

Here, $\tilde{g}_h(\zeta)$ signifies the generating function (the discrete Laplace transform or z -transform) of g_h^n , while $\hat{g}_h(z)$ represents the continuous Laplace transform of $g_h(t)$. The latter transform corresponds to the piecewise linear reconstruction of the nodal values g_h^n . The removal of the logarithmic factor is exactly accomplished by utilizing this consistency, which improves (1.3) to the following result:

$$\sum_{j=1}^{\infty} \frac{1}{j^2} \leq C. \quad (1.4)$$

Furthermore, it should be noted that BDF- k is not A -stable for $3 \leq k \leq 6$. This instability necessitates the adherence to the BDF-1 symbol within the resolvent operator, i.e., $(\frac{1-e^{-\tau z}}{\tau} - \Delta_h)^{-1}$. Consequently, to eliminate the logarithmic factor for BDF- k , consistency of the solution maps associated with BDF-1 and BDF- k is required, as outlined in Lemma 5.5.

The rest of this article is organized as follow. In Section 2 we state the two main theorems regarding the weak maximum principle of semi-discrete and fully discrete FEMs for parabolic equations. In Section 4, we introduce the notation and establish the weak discrete maximum principle of the semi-discrete FEM. In Section 5, we prove the fully discrete weak maximum principle of fully discrete FEMs with multistep BDF time-stepping methods. Finally, in the conclusion section, we briefly discuss the extension to discontinuous Galerkin time-stepping methods with possibly highly variable time step sizes, as well as the limitation of current approach and the possibility to address it based on novel approaches developed recently.

2 The main results

Let S_h , $0 < h < h_0$, be a family of Lagrange finite element subspaces of $H^1(\Omega)$ consisting of all piecewise polynomials of degree $r \geq 1$ subject to a quasi-uniform triangulation of a polygonal/polyhedral domain Ω . We denote by \dot{S}_h the subspace of S_h with zero boundary conditions, and denote by $I_h : C(\bar{\Omega}) \rightarrow S_h$ the Lagrangian interpolation operator onto S_h .

The semi-discrete FEM (FEM) for (1.1) is to find $u_h \in C^1([0, \infty); S_h)$ satisfying the following weak formulation:

$$\begin{cases} (\partial_t u_h(t), v_h) + (\nabla u_h(t), \nabla v_h) = 0 & \forall v_h \in \dot{S}_h, \quad \forall t \in (0, \infty), \\ u_h(t) = g_h(t) & \text{on } \partial\Omega, \quad \forall t \in [0, \infty) \\ u_h(0) = u_h^0 & \text{in } \Omega, \end{cases} \quad (2.1)$$

where $u_h^0 = I_h u(0) \in S_h(\Omega)$ and $g_h \in S_h(\partial\Omega)$ is the Lagrangian interpolation of g on the boundary $\partial\Omega$. If \tilde{g} is any extension of g from $C(\partial\Omega)$ to $C(\bar{\Omega})$, then $g_h = I_h \tilde{g}$ on $\partial\Omega$.

The first main result of this article is the following theorem.

Theorem 2.1 (Weak maximum principle of semi-discrete FEM) *If Ω is a polygon (possibly nonconvex) in \mathbb{R}^2 or a convex polyhedron in \mathbb{R}^3 , then the solution of the semi-discrete FEM in (2.1) satisfies the following weak maximum principle for $r \geq 1$ in \mathbb{R}^2 and $r \geq 2$ in \mathbb{R}^3 :*

$$\|u_h\|_{L^\infty(0,T;L^\infty(\Omega))} \leq C \|u_h^0\|_{L^\infty(\Omega)} + C \|g_h\|_{L^\infty(0,T;L^\infty(\partial\Omega))},$$

where C is a constant independent of h and T .

We denote by $\delta_j, j = 0, \dots, k$, the coefficients of the generating polynomial of the BDF- k method (see [12, equation (3.5)]), i.e.,

$$\delta(\zeta) = \sum_{j=1}^k \frac{1}{j} (1 - \zeta)^j = \sum_{j=0}^k \delta_j \zeta^j,$$

and use the notation $\delta_\tau(\zeta) = \tau^{-1} \delta(\zeta)$, where τ denotes the step size of time discretization. The BDF- k method is known to be $A(\theta_k)$ -stable, i.e., the set $\{z \in \mathbb{C} \setminus \{0\} : |\arg(z)| < \theta_k\}$ is contained in the stability region of the method (see [33, p. 251]), where $\theta_k = 90^\circ, 90^\circ, 86.03^\circ, 73.35^\circ, 51.84^\circ, 17.84^\circ$ for $k = 1, 2, 3, 4, 5, 6$, respectively.

For the BDF- k method, we assume that the starting values $u_h^n, n = 0, \dots, k-1$, are given or computed from some other time-stepping methods (such as the Taylor expansion method or the Runge–Kutta method).

Given any final time $T \geq (k-1)\tau$, we set $N = \lfloor T/\tau \rfloor$. For $n = k, \dots, N$, the fully discrete FEM with BDF- k time-stepping method seeks $u_h^n \in S_h$ such that

$$\begin{cases} \left(\frac{1}{\tau} \sum_{j=0}^k \delta_j u_h^{n-j}, v_h \right) + (\nabla u_h^n, \nabla v_h) = 0 \quad \forall v_h \in \mathring{S}_h, & n = k, \dots, N, \\ u_h^n = g_h^n & \text{on } \partial\Omega, \quad n = k, \dots, N, \end{cases} \quad (2.2)$$

where $g_h^n = I_h g(t_n) \in S_h(\partial\Omega)$.

The second main result of this article is the following theorem.

Theorem 2.2 (Weak maximum principle of fully discrete FEM) *If Ω is a polygon (possibly nonconvex) in \mathbb{R}^2 or a convex polyhedron in \mathbb{R}^3 , then the fully discrete solution given by (2.2) satisfies the following weak maximum principle for $r \geq 1$ in \mathbb{R}^2 and $r \geq 2$ in \mathbb{R}^3 :*

$$\max_{k \leq n \leq N} \|u_h^n\|_{L^\infty(\Omega)} \leq C \max_{0 \leq n \leq k-1} \|u_h^n\|_{L^\infty(\Omega)} + C \max_{k \leq n \leq N} \|g_h^n\|_{L^\infty(\partial\Omega)},$$

where C is a constant independent of τ, h and N .

Remark 2.1 From the proof of Theorems 2.1 and 2.2 in the following sections we can see that, for a convex polyhedron in \mathbb{R}^3 and finite elements of degree $r = 1$, the semi-discrete and fully discrete FEMs satisfy the following weak maximum principle with an additional logarithmic factor, i.e.,

$$\|u_h\|_{L^\infty(0,T;L^\infty(\Omega))} \leq C \ln(2 + 1/h) \|g_h\|_{L^\infty(0,T;L^\infty(\partial\Omega))}, \quad (2.3)$$

$$\max_{k \leq n \leq N} \|u_h^n\|_{L^\infty(\Omega)} \leq C \max_{0 \leq n \leq k-1} \|u_h^n\|_{L^\infty(\Omega)} + C \ln(2 + 1/h) \max_{k \leq n \leq N} \|g_h^n\|_{L^\infty(\partial\Omega)}. \quad (2.4)$$

3 Notation, finite element discretization, and regularized Green's function

3.1 Function spaces and notation

We use the conventional notations of Sobolev spaces $W^{s,q}(\Omega)$, $s \geq 0$ and $1 \leq q \leq \infty$, with 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 latter is defined as the closure of $C_0^\infty(\Omega)$ in $H^s(\Omega)$.

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

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

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

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

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

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

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

We adopt the following notations for the L^2 inner products on Ω and the space-time L^2 inner products on $\mathcal{Q}_T = (0, T) \times \Omega$:

$$(\phi, \varphi) := \int_{\Omega} \phi(x)\varphi(x)dx, \quad [u, v] := \int_0^T \int_{\Omega} u(t, x)v(t, x)dx dt. \quad (3.4)$$

Moreover, we denote $w(t) = w(t, \cdot)$ for any function w defined on \mathcal{Q}_T . 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 spaces

For any subdomain $D \subset \Omega$, we denote by $\mathring{S}_h(D)$ the space of functions of \mathring{S}_h restricted to the domain D , and denote by $\mathring{S}_h^0(D)$ the subspace of $\mathring{S}_h(D)$ consisting of functions which equal zero outside D . For any given subset $D \subset \Omega$, we denote $D_d = \{x \in \Omega : \text{dist}(x, D) \leq d\}$ 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 \mathring{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

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

(P2) Inverse inequality:

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

$$\|\chi\|_{W^{l,p}(D)} \leq Kh^{-(l-k)-(d/q-d/p)} \|\chi\|_{W^{k,q}(D)}, \quad \forall \chi \in \mathring{S}_h,$$

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

(P3) Local approximation and superapproximation:

There exists an operator $I_h : H_0^1(\Omega) \rightarrow \mathring{S}_h$ with the following properties:

(1) For $v \in H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$ the following estimate holds:

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

(2) If $d \geq 2h$ then the value of $I_h v$ in D depends only on the value of v in D_d . If $d \geq 2h$ and $\text{supp}(v) \subset \overline{D}$, then $I_h v \in \mathring{S}_h^0(D_d)$.

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

$$\psi_h \in \mathring{S}_h(D_d) \implies I_h(\omega\psi_h) \in \mathring{S}_h^0(D_d),$$

$$\|\omega\psi_h - I_h(\omega\psi_h)\|_{L^2} + h \|\omega\psi_h - I_h(\omega\psi_h)\|_{H^1} \leq Kh d^{-1} \|\psi_h\|_{L^2(D_d)}.$$

(4) If $d \geq 2h$ and $\omega \equiv 1$ on 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 $r \geq 1$ (cf. [28, Appendix]). Property (P3)-(1) and the definition (3.2) imply the following local estimate for $\alpha \in [0, 1]$ and $v \in H^{1+\alpha}(D_d) \cap H_0^1(\Omega)$:

$$\|v - I_h v\|_{L^2(D)} + h\|v - I_h v\|_{H^1(D)} \leq Kh^{1+\alpha}\|v\|_{H^{1+\alpha}(D_d)}. \quad (3.5)$$

In addition to I_h , we will also need the orthogonal L^2 projection $P_h: L^2(\Omega) \rightarrow \dot{S}_h(\Omega)$ and Ritz-Projection $R_h: H_0^1(\Omega) \rightarrow \dot{S}_h(\Omega)$, which are defined by

$$\begin{aligned} (P_h v, \chi)_\Omega &= (v, \chi)_\Omega, & \forall \chi \in \dot{S}_h(\Omega), \\ (\nabla R_h v, \nabla \chi)_\Omega &= (\nabla v, \nabla \chi)_\Omega, & \forall \chi \in \dot{S}_h(\Omega). \end{aligned}$$

For $v \in H^{1+\alpha}(\Omega) \cap H_0^1(\Omega)$ with $\alpha \in [0, 1]$, the following global estimate follows from choosing $D = \Omega$ in (3.5):

$$\|v - P_h v\|_{L^2(\Omega)} + h\|v - P_h v\|_{H^1(\Omega)} \leq Kh^{1+\alpha}\|v\|_{H^{1+\alpha}(\Omega)}. \quad (3.6)$$

Using (3.5) and a duality argument, we also have the following global estimate for the Ritz projection:

$$\|v - R_h v\|_{L^2(\Omega)} + h^\alpha\|v - R_h v\|_{H^1(\Omega)} \leq Kh^{2\alpha}\|v\|_{H^{1+\alpha}(\Omega)}. \quad (3.7)$$

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 $\tilde{\delta}_{x_0} \in C^3(\bar{\Omega})$ with support in τ_l^h such that

$$\chi(x_0) = \int_{\Omega} \chi \tilde{\delta}_{x_0} dx, \quad \forall \chi \in \dot{S}_h,$$

and

$$\|\tilde{\delta}_{x_0}\|_{W^{l,p}} \leq Kh^{-l-d(1-1/p)} \quad \text{for } 1 \leq p \leq \infty, \quad l = 0, 1, 2, 3, \quad d \geq 1, \quad (3.8)$$

$$\sup_{y \in \Omega} \int_{\Omega} |\tilde{\delta}_y(x)| dx + \sup_{x \in \Omega} \int_{\Omega} |\tilde{\delta}_y(x)| dy \leq C. \quad (3.9)$$

The construction of $\tilde{\delta}_{x_0}$ can be found in [30, Lemma 2.2].

Let δ_{x_0} denote the Dirac Delta function centered at x_0 . In other words, the following relation holds:

$$\int_{\Omega} \delta_{x_0}(y) \varphi(y) dy = \varphi(x_0) \quad \forall \varphi \in C(\bar{\Omega}).$$

Then the discrete Delta function

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

decays exponentially away from x_0 (cf. [30, Lemma 2.3]):

$$|\delta_{h,x_0}(x)| = |P_h \tilde{\delta}_{x_0}(x)| \leq Kh^{-d} e^{-\frac{|x-x_0|}{Kh}}, \quad \forall x, x_0 \in \Omega. \quad (3.10)$$

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

$$\begin{cases} \partial_t G(\cdot, \cdot, x_0) - \Delta G(\cdot, \cdot, x_0) = 0 & \text{in } (0, T] \times \Omega, \\ G(\cdot, \cdot, x_0) = 0 & \text{on } (0, T] \times \partial\Omega, \\ G(0, \cdot, x_0) = \delta_{x_0} & \text{in } \Omega. \end{cases} \quad (3.11)$$

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. [6,

Lemma 2] and [5, Theorem 3.4.8]):

$$|G(z, x, y)| \leq C_\theta |z|^{-\frac{d}{2}} e^{-\frac{|x-y|^2}{C_\theta |z|}}, \quad \forall z \in \Sigma_\theta, \forall x, y \in \Omega, \quad \forall \theta \in (0, \pi/2), \quad (3.12)$$

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

$$\partial_t^k G(t, x, y) = \frac{k!}{2\pi i} \int_{|z-t|=\frac{t}{2}} \frac{G(z, x, y)}{(z-t)^{k+1}} dz, \quad (3.13)$$

which further yields the following Gaussian pointwise estimate for the time derivatives of Green's function (cf. [6, Corollary 5] and [8, Appendix B with $\alpha = \beta = 0$]):

$$|\partial_t^k G(t, x, x_0)| \leq \frac{C_k}{t^{k+d/2}} e^{-\frac{|x-x_0|^2}{C_k t}}, \quad \forall x, x_0 \in \Omega, \quad \forall t > 0, \quad k = 0, 1, 2, \dots \quad (3.14)$$

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

$$\begin{cases} \partial_t \Gamma(\cdot, \cdot, x_0) - \Delta \Gamma(\cdot, \cdot, x_0) = 0 & \text{in } (0, T] \times \Omega, \\ \Gamma(\cdot, \cdot, x_0) = 0 & \text{on } (0, T] \times \partial\Omega, \\ \Gamma(0, \cdot, x_0) = \tilde{\delta}_{x_0} & \text{in } \Omega, \end{cases} \quad (3.15)$$

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

$$\begin{cases} (\partial_t \Gamma_h(t, \cdot, x_0), v_h) + (\nabla \Gamma_h(t, \cdot, x_0), \nabla v_h) = 0, \quad \forall v_h \in \mathring{S}_h, \quad t \in (0, T), \\ \Gamma_h(0, \cdot, x_0) = \delta_{h, x_0}. \end{cases} \quad (3.16)$$

The regularized Green's function can be expressed in terms of the Green's function as follows:

$$\Gamma(t, x, x_0) = \int_{\Omega} G(t, y, x) \tilde{\delta}_{x_0}(y) dy = \int_{\Omega} G(t, x, y) \tilde{\delta}_{x_0}(y) dy. \quad (3.17)$$

From the expression in (3.17) one can easily derive that the regularized Green's function Γ also satisfies the Gaussian pointwise estimate for $k \geq 0$:

$$|\partial_t^k \Gamma(t, x, x_0)| \leq \frac{C_k}{t^{k+d/2}} e^{-\frac{|x-x_0|^2}{C_k t}}, \quad \forall x, x_0 \in \Omega, \quad \forall t > 0 \text{ for } \max(|x-x_0|, \sqrt{t}) \geq 2h. \quad (3.18)$$

4 Weak maximum principle of semi-discrete FEM

Since the problem is linear, the first part of the result, i.e., $\|u_h\|_{L^\infty(0, T; L^\infty(\Omega))} \leq C \|u_h^0\|_{L^\infty(\Omega)}$, follows from the pointwise stability of the semi-group established in [21]. To establish the second part of the result, i.e., $\|u_h\|_{L^\infty(0, T; L^\infty(\Omega))} \leq C \|u_h^0\|_{L^\infty(\Omega)} + C \|g_h\|_{L^\infty(0, T; L^\infty(\partial\Omega))}$, we will use the technique based on the dyadic decomposition and local energy estimates.

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

In the proof of Theorem 2.1, we need to partition the domain $\mathcal{Q} = \Omega_T$ with $T = 1$, i.e. $\mathcal{Q} = (0, 1) \times \Omega$, into subdomains, and present estimates of the finite element solutions in each subdomain. The following dyadic decomposition of \mathcal{Q} was introduced in [27].

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 $C_* \geq 16$ to be determined later. If

$$h < 1/(4C_*), \quad (4.1)$$

then

$$2 \leq J_* = \log_2[1/(C_* h)] \leq \log_2(2 + 1/h). \quad (4.2)$$

Let

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

We define

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

and

$$\begin{aligned} Q_0(x_0) &= \mathcal{Q} \setminus \left(\cup_{j=1}^{J_*} Q_j(x_0) \cup Q_*(x_0) \right), \\ \Omega_0(x_0) &= \Omega \setminus \left(\cup_{j=1}^{J_*} \Omega_j(x_0) \cup \Omega_*(x_0) \right). \end{aligned}$$

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

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

Then we have

$$\Omega_T = \bigcup_{j=0}^{J_*} Q_j(x_0) \cup Q_*(x_0) \quad \text{and} \quad \Omega = \bigcup_{j=0}^{J_*} \Omega_j(x_0) \cup \Omega_*(x_0). \quad (4.3)$$

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

$$\|v\|_{k,D} = \left(\int_D \sum_{|\alpha| \leq k} |\partial^\alpha v|^2 dx \right)^{\frac{1}{2}}, \quad \|v\|_{k,Q} = \left(\int_Q \sum_{|\alpha| \leq k} |\partial^\alpha v|^2 dx dt \right)^{\frac{1}{2}}, \quad (4.4)$$

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 4.4). To simplify the notations, we also denote $d_* = d_{J_*}$.

4.2 Technical lemmas

Let Ω be a polygon in \mathbb{R}^2 or a polyhedron in \mathbb{R}^3 (possibly nonconvex), and let \mathring{S}_h , $0 < h < h_0$, be a family of finite element subspaces of $H_0^1(\Omega)$ consisting of piecewise polynomials of degree $r \geq 1$ subject to a quasi-uniform triangulation of the domain Ω (with mesh size h).

The first two of the following three technical lemmas were proved in [21, Lemma 4.1, Lemma 4.4] for general polyhedron that may be nonconvex. In the case Ω is a convex polyhedron, α can be slightly bigger than 1. We include this slightly different result for the convex case and omit the proof, which is almost the same as the proof for general polyhedra except some minor difference.

Lemma 4.1 *There exist $\alpha \in (\frac{1}{2}, 2]$ and $C > 0$, independent of h and x_0 , such that the Green's function G defined in (3.11) and the regularized Green's function Γ defined in (3.15) satisfy the following estimates:*

$$d_j^{-4-\alpha+d/2} \|\Gamma(\cdot, \cdot, x_0)\|_{L^\infty(Q_j(x_0))} + d_j^{-4-\alpha} \|\nabla \Gamma(\cdot, \cdot, x_0)\|_{L^2(Q_j(x_0))}$$

$$\begin{aligned}
& + d_j^{-4} \|\Gamma(\cdot, \cdot, x_0)\|_{L^2 H^{1+\alpha}(Q_j(x_0))} + d_j^{-2} \|\partial_t \Gamma(\cdot, \cdot, x_0)\|_{L^2 H^{1+\alpha}(Q_j(x_0))} \\
& + \|\partial_{tt} \Gamma(\cdot, \cdot, x_0)\|_{L^2 H^{1+\alpha}(Q_j(x_0))} \leq C d_j^{-d/2-4-\alpha}, \tag{4.5}
\end{aligned}$$

$$\|\Gamma(\cdot, \cdot, x_0)\|_{L^\infty H^{1+\alpha}(\cup_{k \leq j} Q_k(x_0))} + d_j^2 \|\partial_t \Gamma(\cdot, \cdot, x_0)\|_{L^\infty H^{1+\alpha}(\cup_{k \leq j} Q_k(x_0))} \leq C d_j^{-d/2-1-\alpha}. \tag{4.6}$$

In the case $d = 2$, $\alpha \in (\frac{1}{2}, 1]$. In the case $d = 3$ and Ω is convex, $\alpha \in (1, 2]$.

In the rest of this paper, we denote by α a number depending on which case we are working on, i.e., $\alpha \in (\frac{1}{2}, 1]$ in the case $d = 2$, and $\alpha \in (1, 2]$ in the case $d = 3$ and Ω is convex.

Lemma 4.2 *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*

$$\sup_{t \in (0, \infty)} \left(\|\Gamma_h(t, \cdot, x_0)\|_{L^1(\Omega)} + t \|\partial_t \Gamma_h(t, \cdot, x_0)\|_{L^1(\Omega)} \right) \leq C, \tag{4.7}$$

$$\sup_{t \in (0, \infty)} \left(\|\Gamma(t, \cdot, x_0)\|_{L^1(\Omega)} + t \|\partial_t \Gamma(t, \cdot, x_0)\|_{L^1(\Omega)} \right) \leq C, \tag{4.8}$$

$$\|\partial_t F(\cdot, \cdot, x_0)\|_{L^1((0, \infty) \times \Omega)} + \|t \partial_{tt} F(\cdot, \cdot, x_0)\|_{L^1((0, \infty) \times \Omega)} \leq C, \tag{4.9}$$

$$\|\partial_t \Gamma_h(t, \cdot, x_0)\|_{L^1} \leq C e^{-\lambda_0 t}, \quad \forall t \geq 1, \tag{4.10}$$

where the constants C and λ_0 are independent of h .

Lemma 4.3 *The following results hold:*

1. $\|\Gamma(t)\|_{L^2(\frac{1}{4}, \infty; H^{1+\alpha}(\Omega))} + \|\partial_t \Gamma(t)\|_{L^2(\frac{1}{4}, \infty; H^{1+\alpha}(\Omega))} \leq C$.
2. $(-\Delta_h)^{-1} : (\dot{S}_h(\Omega), \|\cdot\|_{L^1(\Omega)}) \rightarrow (\dot{S}_h(\Omega), \|\cdot\|_{L^2(\Omega)})$ is continuous.
3. $\|\nabla(\Gamma_h - \Gamma)\|_{L^1(0,1; L^1(\Omega))} \leq C h^{\frac{1}{2}}$.

Proof To show the first result, we recall the estimate in [21, p. 27] which can be proved by the standard energy method. Here we additionally allow $t \geq \frac{1}{4}$ instead of $t \geq 1$ in [21, p. 27]:

$$\begin{aligned}
& \|\partial_t \Gamma(t, \cdot, x_0)\|_{L^2}^2 + \|\partial_t \Gamma_h(t, \cdot, x_0)\|_{L^2}^2 + \|\partial_{tt} \Gamma(t, \cdot, x_0)\|_{L^2}^2 + \|\partial_{tt} \Gamma_h(t, \cdot, x_0)\|_{L^2}^2 \\
& \leq C e^{-\lambda_0(t-\frac{1}{4})}.
\end{aligned}$$

Therefore from elliptic regularity theory, we get the following boundedness results

$$\|\Gamma(t)\|_{L^2(\frac{1}{4}, \infty; H^{1+\alpha})} \leq C \|\Delta \Gamma(t)\|_{L^2(\frac{1}{4}, \infty; L^2)} = C \|\partial_t \Gamma(t)\|_{L^2(\frac{1}{4}, \infty; L^2)} \leq C,$$

$$\|\partial_t \Gamma(t)\|_{L^2(\frac{1}{4}, \infty; H^{1+\alpha})} \leq C \|\Delta \partial_t \Gamma(t)\|_{L^2(\frac{1}{4}, +\infty; L^2)} = C \|\partial_{tt} \Gamma(t)\|_{L^2(\frac{1}{4}, +\infty; L^2)} \leq C.$$

To prove the second one, for any given $f_h \in \dot{S}_h(\Omega)$, we define $u_h = (-\Delta_h)^{-1} f_h$ and $u \in H_0^1(\Omega)$ to be the solution of the elliptic equation

$$-\Delta u = f_h.$$

By the definition we have the relation $u_h = R_h u$. For $d = 2, 3$, fix $p = 2^* = \frac{2d}{d-2}$ whose Hölder conjugate is $p' = \frac{2d}{d-2}$, and we apply the Sobolev embedding $W^{1,p} \hookrightarrow L^2$, the stability of Ritz projection R_h (see [19, equation (2.5)]), the elliptic regularity theory on corner domain (see [4, Theorem 3.2, Corollary 3.10, Corollary 3.12]) and the embedding $W^{1,p'} \hookrightarrow L^\infty$ successively to derive

$$\|(-\Delta_h)^{-1} f_h\|_{L^2} = \|u_h\|_{L^2} \leq C \|u_h\|_{W^{1,p}} \leq C \|u\|_{W^{1,p}} \leq C \|f_h\|_{W^{-1,p}} \leq C \|f_h\|_{L^1}.$$

For the third result, we modify the quantity \mathcal{K} in the proof of [21, Lemma 4.4] to be

$$\mathcal{K} := \sum_j d_j^{1+d/2} \left(d_j^{-1/2} h^{-1/2} \|F\|_{1, Q_j} + \|\partial_t F\|_{Q_j} + d_j \|\partial_t F\|_{1, Q_j} + d_j^2 \|\partial_{tt} F\|_{Q_j} \right),$$

and then the same proof will show $\mathcal{K} \leq C$. Sketch of the proof: The lowest powers of h in the consistency part in the proof of [21, Lemma 4.4] are h^α (see [21, equation (5.14) and (5.15)]) and $h^{1+\alpha-d/2}$ (see [21, equation (5.32)]). Therefore if we multiply $h^{-1/2}$ in the front, we get $h^{-1/2+\alpha} + h^{1/2+\alpha-d/2} \leq C$. For the stability part, as [21, equation (5.32)], we will get an additional power of $h^{-1/2}h^{1+\alpha} = h^{1/2+\alpha}$ in front of $\|F_t\|_{Q_i}$ and an additional power of h^α in front of $\|F\|_{1,Q_i}$ (note that according to the definition of \mathcal{K} , there already exists $h^{-1/2}$ in front of $\|F\|_{1,Q_i}$ on the left hand side). Such additional positive powers of h ensure the stability of \mathcal{K} . \square

4.3 Reduction of the problem

To simplify the notation, we relax the dependence of the Green's functions on x_0 and denote

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

If u_h is the finite element solution of

$$\begin{cases} (\partial_t u_h, v_h) + (\nabla u_h, \nabla v_h) = 0 & \forall v_h \in \dot{S}_h, \text{ for } t \in (0, T), \\ u_h = g_h & \text{on } [0, T] \times \partial\Omega, \\ u_h(0) = 0 & \text{in } \Omega, \end{cases} \quad (4.11)$$

then substituting $v_h = \Gamma_h(s-t)$, with $0 < t < s \leq T$, and integrating the result for $t \in (0, s)$ yield

$$\begin{aligned} u_h(s, x_0) &= - \int_0^s \left[(u_h, \partial_t \Gamma_h(s-t)) + (\nabla u_h, \nabla \Gamma_h(s-t)) \right] dt \\ &= - \int_0^s \left[(u_h - w_h, \partial_t \Gamma_h(s-t)) + (\nabla(u_h - w_h), \nabla \Gamma_h(s-t)) \right] dt, \end{aligned} \quad (4.12)$$

which holds for all $w_h \in L^2(0, T; \dot{S}_h)$, where the last inequality is a consequence of (3.16). Let $w_h \in \dot{S}_h$ be the finite element function which equals u_h at interior nodes and equals 0 on $\partial\Omega$. Then

$$u_h(s, x_0) = - \int_0^s \left[(\tilde{g}_h, \partial_t \Gamma_h(s-t)) + (\nabla \tilde{g}_h, \nabla \Gamma_h(s-t)) \right] dt, \quad (4.13)$$

where \tilde{g}_h is the finite element solution which equals g_h on $\partial\Omega$ and equals zero at the interior nodes of Ω . Let $D_h = \{x \in \Omega : \text{dist}(x, \partial\Omega) \leq h\}$. Then

$$u_h(s, x_0) = - \int_0^s \left[(\tilde{g}_h, \partial_t \Gamma_h(s-t))_{D_h} + (\nabla \tilde{g}_h, \nabla \Gamma_h(s-t))_{D_h} \right] dt. \quad (4.14)$$

Let \tilde{u} be the solution of the PDE problem

$$\begin{cases} \partial_t \tilde{u} - \Delta \tilde{u} = 0 & \text{for } t \in (0, T], \\ \tilde{u} = g_h & \text{on } [0, T] \times \partial\Omega, \\ \tilde{u}(0) = 0 & \text{in } \Omega. \end{cases} \quad (4.15)$$

By the maximum principle of parabolic PDEs (see [1, Theorem 6.2.6] and [24, Chapter II]), we have

$$\|\tilde{u}\|_{L^\infty(0, T; L^\infty)} \leq \|g_h\|_{L^\infty(0, T; L^\infty(\partial\Omega))}. \quad (4.16)$$

Testing the first equation of (4.15) by $\Gamma(s-t)$ and integrating the result for $t \in (0, s)$, we obtain

$$(\tilde{u}(s, \cdot), \tilde{\delta}_{x_0}) = - \int_0^s \left[(\tilde{u}, \partial_t \Gamma(s-t)) + (\nabla \tilde{u}, \nabla \Gamma(s-t)) \right] dt$$

$$= - \int_0^s \left[(\tilde{u} - w, \partial_t \Gamma(s-t)) + (\nabla(\tilde{u} - w), \nabla \Gamma(s-t)) \right] dt, \quad (4.17)$$

which holds for all $w \in L^2(0, T; H_0^1(\Omega))$. By choosing $w = \tilde{u} - \tilde{g}_h$ we obtain

$$\begin{aligned} (\tilde{u}(s, \cdot), \tilde{\delta}_{x_0}) &= - \int_0^s \left[(\tilde{g}_h, \partial_t \Gamma(s-t)) + (\nabla \tilde{g}_h, \nabla \Gamma(s-t)) \right] dt \\ &= - \int_0^s \left[(\tilde{g}_h, \partial_t \Gamma(s-t))_{D_h} + (\nabla \tilde{g}_h, \nabla \Gamma(s-t))_{D_h} \right] dt. \end{aligned} \quad (4.18)$$

Subtracting (4.18) from (4.14), we obtain

$$\begin{aligned} u_h(s, x_0) &= (\tilde{u}(s, \cdot), \tilde{\delta}_{x_0}) - \int_0^s (\tilde{g}_h, \partial_t \Gamma_h(s-t) - \partial_t \Gamma(s-t))_{D_h} dt \\ &\quad - \int_0^s (\nabla \tilde{g}_h, \nabla(\Gamma_h(s-t) - \Gamma(s-t)))_{D_h} dt. \end{aligned} \quad (4.19)$$

From (4.9) we know that

$$\int_0^s \int_{\Omega} |\partial_t \Gamma_h(s-t) - \partial_t \Gamma(s-t)| dx dt \leq C,$$

which implies that

$$\begin{aligned} &|u_h(s, x_0)| \\ &\leq \|\tilde{u}(s)\|_{L^\infty} \|\tilde{\delta}_{x_0}\|_{L^1} + C \|\tilde{g}_h\|_{L^\infty(0, s; L^\infty)} + \left| \int_0^s (\nabla \tilde{g}_h, \nabla(\Gamma_h(s-t) - \Gamma(s-t)))_{D_h} dt \right| \\ &\leq C \|g_h\|_{L^\infty(0, s; L^\infty(\partial\Omega))} + \|\tilde{g}_h\|_{L^\infty(0, s; L^\infty(\Omega))} Ch^{-1} \|\nabla F\|_{L^1(D_h^s)} \\ &\leq C \|g_h\|_{L^\infty(0, s; L^\infty(\partial\Omega))} (1 + h^{-1} \|\nabla F\|_{L^1(D_h^s)}), \end{aligned} \quad (4.20)$$

where $D_h^s = (0, s) \times D_h$. If the following estimate can be proved:

$$h^{-1} \|\nabla F\|_{L^1((0, \infty) \times D_h)} \leq C, \quad (4.21)$$

then substituting (4.21) into (4.20) immediately yields

$$|u_h(s, x_0)| \leq C \|g_h\|_{L^\infty(0, s; L^\infty(\partial\Omega))}. \quad (4.22)$$

Since s and x_0 can be arbitrary, this proves the desired result of Theorem 2.1.

It remains to prove (4.21).

4.4 Proof of (4.21)

We need to use the following local energy error estimate for finite element solutions of parabolic equations, which was proved in [21, Lemma 5.1].

Lemma 4.4 *Suppose that $\phi \in L^2(0, T; H_0^1(\Omega)) \cap H^1(0, T; L^2(\Omega))$ and $\phi_h \in H^1(0, T; \dot{S}_h)$ satisfy the equation*

$$(\partial_t(\phi - \phi_h), \chi) + (\nabla(\phi - \phi_h), \nabla \chi) = 0, \quad \forall \chi \in \dot{S}_h, \quad a.e. t > 0, \quad (4.23)$$

with $\phi(0) = 0$ in Ω_j'' . Then

$$\begin{aligned} &\|\partial_t(\phi - \phi_h)\|_{Q_j} + d_j^{-1} \|\phi - \phi_h\|_{1, Q_j} \\ &\leq C \epsilon^{-3} (I_j(\phi_h(0)) + X_j(I_h \phi - \phi) + d_j^{-2} \|\phi - \phi_h\|_{Q_j}) \\ &\quad + \left[C \left(\frac{h}{d_j} \right)^{\frac{1}{2}} + \epsilon_* \right] (\|\partial_t(\phi - \phi_h)\|_{Q_j'} + d_j^{-1} \|\phi - \phi_h\|_{1, Q_j'}), \end{aligned} \quad (4.24)$$

where $\epsilon_* = \epsilon + \epsilon^{-1}/C_*$ and

$$I_j(\phi_h(0)) = \|\phi_h(0)\|_{1, \Omega_j'} + d_j^{-1} \|\phi_h(0)\|_{\Omega_j'},$$

$$\begin{aligned} X_j(I_h\phi - \phi) &= d_j \|\partial_t(I_h\phi - \phi)\|_{1, Q'_j} + \|\partial_t(I_h\phi - \phi)\|_{Q'_j} \\ &\quad + d_j^{-1} \|I_h\phi - \phi\|_{1, Q'_j} + d_j^{-2} \|I_h\phi - \phi\|_{Q'_j}. \end{aligned}$$

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 (4.4).

In the rest of this section, we apply Lemma 4.4 to estimate $h^{-1}\|\nabla F\|_{L^1((0, \infty) \times D_h)}$. The estimation consists of two parts: The first part concerns estimates for $t \in (0, 1)$, and the second part concerns estimates for $t \geq 1$, which is a consequence of the smoothing property of parabolic equations.

Part I. First, we present estimates in the domain $\mathcal{Q} = (0, 1) \times \Omega$ with the restriction $h < 1/(4C_*)$; see (4.1). In this case, the basic energy estimate gives

$$\|\partial_t \Gamma\|_{L^2(\mathcal{Q})} + \|\partial_t \Gamma_h\|_{L^2(\mathcal{Q})} \leq C(\|\Gamma(0)\|_{H^1} + \|\Gamma_h(0)\|_{H^1}) \leq Ch^{-1-d/2}, \quad (4.25)$$

$$\|\Gamma\|_{L^\infty L^2(\mathcal{Q})} + \|\Gamma_h\|_{L^\infty L^2(\mathcal{Q})} \leq C(\|\Gamma(0)\|_{L^2} + \|\Gamma_h(0)\|_{L^2}) \leq Ch^{-d/2}, \quad (4.26)$$

$$\|\nabla \Gamma\|_{L^2(\mathcal{Q})} + \|\nabla \Gamma_h\|_{L^2(\mathcal{Q})} \leq C(\|\Gamma(0)\|_{L^2} + \|\Gamma_h(0)\|_{L^2}) \leq Ch^{-d/2}, \quad (4.27)$$

$$\|\partial_{tt} \Gamma\|_{L^2(\mathcal{Q})} + \|\partial_{tt} \Gamma_h\|_{L^2(\mathcal{Q})} \leq C(\|\Delta \Gamma(0)\|_{H^1} + \|\Delta_h \Gamma_h(0)\|_{H^1}) \leq Ch^{-3-d/2}, \quad (4.28)$$

$$\|\nabla \partial_t \Gamma\|_{L^2(\mathcal{Q})} + \|\nabla \partial_t \Gamma_h\|_{L^2(\mathcal{Q})} \leq C(\|\Delta \Gamma(0)\|_{L^2} + \|\Delta_h \Gamma_h(0)\|_{L^2}) \leq Ch^{-2-d/2}, \quad (4.29)$$

where we have used (3.8) and (3.10) to estimate $\Gamma(0)$ and $\Gamma_h(0)$, respectively. Hence, we have

$$\|\Gamma\|_{Q_*} + \|\Gamma_h\|_{Q_*} \leq Cd_* \|\Gamma\|_{L^\infty L^2(Q_*)} + Cd_* \|\Gamma_h\|_{L^\infty L^2(Q_*)} \leq Cd_* h^{-d/2} \leq CC_* h^{1-d/2}. \quad (4.30)$$

Since the volume of Q_j is Cd_j^{2+d} , we can decompose

$$\|\partial_t F\|_{L^1(D_h^T)} + \|t\partial_{tt} F\|_{L^1(D_h^T)} + h^{-1}\|\nabla F\|_{L^1(D_h^T)}$$

in the following way:

$$\begin{aligned} &\|\partial_t F\|_{L^1(D_h^T)} + \|t\partial_{tt} F\|_{L^1(D_h^T)} + h^{-1}\|\nabla F\|_{L^1(D_h^T)} \\ &\leq \|\partial_t F\|_{L^1(Q_* \cap D_h^T)} + \|t\partial_{tt} F\|_{L^1(Q_* \cap D_h^T)} + h^{-1}\|\nabla F\|_{L^1(Q_* \cap D_h^T)} \\ &\quad + \sum_j (\|\partial_t F\|_{L^1(Q_j \cap D_h^T)} + \|t\partial_{tt} F\|_{L^1(Q_j \cap D_h^T)} + h^{-1}\|\nabla F\|_{L^1(Q_j \cap D_h^T)}) \\ &\leq Cd_*^{1+\frac{d-1}{2}} h^{\frac{1}{2}} (\|\partial_t F\|_{Q_* \cap D_h^T} + d_*^2 \|\partial_{tt} F\|_{Q_* \cap D_h^T} + h^{-1}\|F\|_{1, Q_* \cap D_h^T}) \\ &\quad + \sum_j Cd_j^{1+\frac{d-1}{2}} h^{\frac{1}{2}} (\|\partial_t F\|_{Q_j \cap D_h^T} + d_j^2 \|\partial_{tt} F\|_{Q_j \cap D_h^T} + h^{-1}\|F\|_{1, Q_j \cap D_h^T}) \\ &\leq CC_*^{\frac{5+d}{2}} + \mathcal{K}, \end{aligned} \quad (4.31)$$

where we have used (4.25), (4.27) and (4.28) to estimate

$$Cd_*^{1+\frac{d-1}{2}} h^{\frac{1}{2}} (\|\partial_t F\|_{Q_* \cap D_h^T} + d_*^2 \|\partial_{tt} F\|_{Q_* \cap D_h^T} + h^{-1}\|F\|_{1, Q_* \cap D_h^T}),$$

and introduced the notation

$$\mathcal{K} := \sum_j \left(\frac{h}{d_j}\right)^{-\frac{1}{2}} d_j^{1+\frac{d}{2}} (d_j^{-1}\|F\|_{1, Q_j} + \|\partial_t F\|_{Q_j} + d_j \|\partial_t F\|_{1, Q_j} + d_j^2 \|\partial_{tt} F\|_{Q_j}). \quad (4.32)$$

It remains to estimate \mathcal{K} . To this end, we set “ $\phi_h = \Gamma_h, \phi = \Gamma, \phi_h(0) = P_h \tilde{\delta}_{x_0}$ and $\phi(0) = \tilde{\delta}_{x_0}$ ” and “ $\phi_h = \partial_t \Gamma_h, \phi = \partial_t \Gamma, \phi_h(0) = \Delta_h P_h \tilde{\delta}_{x_0}$ and $\phi(0) = \Delta \tilde{\delta}_{x_0}$ ” in

Lemma 4.4, respectively. Then we obtain

$$\begin{aligned} d_j^{-1} \|F\|_{1,Q_j} + \|\partial_t F\|_{Q_j} &\leq C\epsilon^{-3}(\widehat{I}_j + \widehat{X}_j + d_j^{-2} \|F\|_{Q'_j}) \\ &\quad + \left[C \left(\frac{h}{d_j} \right)^{\frac{1}{2}} + \epsilon_* \right] (d_j^{-1} \|F\|_{1,Q'_j} + \|\partial_t F\|_{Q'_j}) \end{aligned} \quad (4.33)$$

and

$$\begin{aligned} d_j \|\partial_t F\|_{1,Q_j} + d_j^2 \|\partial_{tt} F\|_{Q_j} &\leq C\epsilon^{-3}(\overline{I}_j + \overline{X}_j + \|\partial_t F\|_{Q'_j}) \\ &\quad + \left[C \left(\frac{h}{d_j} \right)^{\frac{1}{2}} + \epsilon_* \right] (d_j \|\partial_t F\|_{1,Q'_j} + d_j^2 \|\partial_{tt} F\|_{Q'_j}), \end{aligned} \quad (4.34)$$

respectively. By using (3.5) (local interpolation error estimate), (3.10) (exponential decay of $P_h \tilde{\delta}_{x_0}$) and Lemma 4.1 (estimates of regularized Green's function), we have

$$\widehat{I}_j = \|P_h \tilde{\delta}_{x_0}\|_{1,\Omega'_j} + h^{-1} \|P_h \tilde{\delta}_{x_0}\|_{\Omega'_j} \leq Ch^2 d_j^{-3-d/2}, \quad (4.35)$$

$$\begin{aligned} \widehat{X}_j &= d_j \|\partial_t(I_h \Gamma - \Gamma)\|_{1,Q'_j} + \|\partial_t(I_h \Gamma - \Gamma)\|_{Q'_j} \\ &\quad + d_j^{-1} \|I_h \Gamma - \Gamma\|_{1,Q'_j} + d_j^{-2} \|I_h \Gamma - \Gamma\|_{Q'_j} \\ &\leq Cd_j h^\alpha \|\partial_t \Gamma\|_{L^2 H^{1+\alpha}(Q'_j)} + d_j^{-1} h^\alpha \|\Gamma\|_{L^2 H^{1+\alpha}(Q'_j)} \\ &\leq Ch^\alpha d_j^{-\alpha-1-d/2} \end{aligned} \quad (4.36)$$

and

$$\overline{I}_j = d_j^2 \|\Delta_h P_h \tilde{\delta}_{x_0}\|_{1,\Omega'_j} + d_j \|\Delta_h P_h \tilde{\delta}_{x_0}\|_{\Omega'_j} \leq Ch^2 d_j^{-3-d/2}, \quad (4.37)$$

$$\begin{aligned} \overline{X}_j &= d_j^3 \|I_h \partial_{tt} \Gamma - \partial_{tt} \Gamma\|_{1,Q'_j} + d_j^2 \|I_h \partial_{tt} \Gamma - \partial_{tt} \Gamma\|_{Q'_j} \\ &\quad + d_j \|I_h \partial_t \Gamma - \partial_t \Gamma\|_{1,Q'_j} + \|I_h \partial_t \Gamma - \partial_t \Gamma\|_{Q'_j} \\ &\leq (d_j^3 h^\alpha + d_j^2 h^{1+\alpha}) \|\partial_{tt} \Gamma\|_{L^2 H^{1+\alpha}(Q'_j)} + (d_j h^\alpha + h^{1+\alpha}) \|\partial_t \Gamma\|_{L^2 H^{1+\alpha}(Q'_j)} \\ &\leq Ch^\alpha d_j^{-\alpha-1-d/2}. \end{aligned} \quad (4.38)$$

By substituting (4.33)-(4.38) into the expression of \mathcal{K} in (4.32), we have

$$\begin{aligned} \mathcal{K} &= \sum_j \left(\frac{h}{d_j} \right)^{-\frac{1}{2}} d_j^{1+\frac{d}{2}} (d_j^{-1} \|F\|_{1,Q_j} + \|\partial_t F\|_{Q_j} + d_j \|\partial_t F\|_{1,Q_j} + d_j^2 \|\partial_{tt} F\|_{Q_j}) \\ &\leq C \sum_j \left(\frac{h}{d_j} \right)^{-\frac{1}{2}} d_j^{1+\frac{d}{2}} \epsilon^{-3} (h^2 d_j^{-3-d/2} + h^\alpha d_j^{-\alpha-1-d/2} + d_j^{-2} \|F\|_{Q'_j}) \\ &\quad + \sum_j \left[C + \epsilon_* \left(\frac{h}{d_j} \right)^{-\frac{1}{2}} \right] d_j^{1+\frac{d}{2}} (d_j^{-1} \|F\|_{1,Q'_j} + \|\partial_t F\|_{Q'_j} + d_j \|\partial_t F\|_{1,Q'_j} + d_j^2 \|\partial_{tt} F\|_{Q'_j}) \\ &\leq C + C\epsilon^{-3} \sum_j \left(\frac{h}{d_j} \right)^{-\frac{1}{2}} d_j^{-1+\frac{d}{2}} \|F\|_{Q'_j} \\ &\quad + \sum_j \left[C + \epsilon_* \left(\frac{h}{d_j} \right)^{-\frac{1}{2}} \right] d_j^{1+\frac{d}{2}} (d_j^{-1} \|F\|_{1,Q'_j} + \|\partial_t F\|_{Q'_j} + d_j \|\partial_t F\|_{1,Q'_j} + d_j^2 \|\partial_{tt} F\|_{Q'_j}). \end{aligned} \quad (4.39)$$

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:

$$\begin{aligned}
\mathcal{K} &\leq C + C\epsilon^{-3} \left(\frac{h}{d_*}\right)^{-\frac{1}{2}} d_*^{-1+\frac{d}{2}} \|F\|_{Q_*} + C\epsilon^{-3} \sum_j \left(\frac{h}{d_j}\right)^{-\frac{1}{2}} d_j^{-1+\frac{d}{2}} \|F\|_{Q_j} \\
&\quad + \sum_j \left[C + \epsilon_* \left(\frac{h}{d_j}\right)^{-\frac{1}{2}} \right] d_*^{1+\frac{d}{2}} (d_*^{-1} \|F\|_{1,Q_*} + \|\partial_t F\|_{Q_*} + d_* \|\partial_t F\|_{1,Q_*} + d_*^2 \|\partial_{tt} F\|_{Q_*}) \\
&\quad + \sum_j \left[C + \epsilon_* \left(\frac{h}{d_j}\right)^{-\frac{1}{2}} \right] d_j^{1+\frac{d}{2}} (d_j^{-1} \|F\|_{1,Q_j} + \|\partial_t F\|_{Q_j} + d_j \|\partial_t F\|_{1,Q_j} + d_j^2 \|\partial_{tt} F\|_{Q_j}) \\
&\leq C + C\epsilon^{-3} C_*^{\frac{d-1}{2}} + \sum_j C\epsilon^{-3} \left(\frac{h}{d_j}\right)^{-\frac{1}{2}} d_j^{-1+\frac{d}{2}} \|F\|_{Q_j} + \left[C + \epsilon_* \left(\frac{h}{d_*}\right)^{-\frac{1}{2}} \right] C_*^{3+\frac{d}{2}} \\
&\quad + (\epsilon_* + CC_*^{-1/2}) \mathcal{K}. \tag{4.40}
\end{aligned}$$

where we have used $d_j \geq C_* h$ and (4.25)-(4.29) 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 \mathcal{K} in (4.32) to bound the terms involving Q_j . Since $\epsilon_* = \epsilon + \epsilon^{-1}/C_*$, we can make ϵ_* sufficiently small by first choosing ϵ small enough and then choosing C_* large enough (ϵ can be fixed now and C_* will be determined later). Then the last term on the right-hand side of (4.40) can be absorbed by the left-hand side. Therefore, we obtain

$$\mathcal{K} \leq C + CC_*^{3+\frac{d}{2}} + \sum_j \left(\frac{h}{d_j}\right)^{-\frac{1}{2}} d_j^{-1+\frac{d}{2}} \|F\|_{Q_j}. \tag{4.41}$$

It remains to estimate $\|F\|_{Q_j}$. This was estimated in [21, inequality (5.32)] as

$$\|F\|_{Q_j} \leq Ch^{1+\alpha-d/2} d_j^{-\alpha} + C \sum_{*,i} (h^{1+\alpha} \|F_t\|_{Q_i} + h^\alpha \|F\|_{1,Q_i}) d_i^{1-\alpha} \left(\frac{\min(d_i, d_j)}{\max(d_i, d_j)}\right)^\alpha. \tag{4.42}$$

Note that

$$\sum_j \left(\frac{h}{d_j}\right)^{-\frac{1}{2}} d_j^{-1+\frac{d}{2}} \left(\frac{\min(d_i, d_j)}{\max(d_i, d_j)}\right)^\alpha \leq \begin{cases} C \left(\frac{h}{d_i}\right)^{-\frac{1}{2}} & \text{if } d = 2 \text{ and } r \geq 1, \\ C \left(\frac{h}{d_i}\right)^{-\frac{1}{2}} d_i^{\frac{1}{2}} & \text{if } d = 3 \text{ and } r \geq 2, \end{cases} \tag{4.43}$$

where the last inequality uses the assumption that in the case $d = 3$ the domain is convex so that $\alpha > 1$ can be chosen; see (P3) in Section 3.2. By substituting (4.42)-(4.43) into (4.41) we obtain

$$\begin{aligned}
\mathcal{K} &\leq C + CC_*^{3+d/2} + \sum_j C d_j^{-1+\frac{d}{2}} \|F\|_{Q_j} \\
&\leq C + CC_*^{3+d/2} + C \sum_j \left(\frac{h}{d_j}\right)^{\alpha+\frac{1}{2}-\frac{d}{2}} \quad \text{here we substitute (4.42)} \\
&\quad + C \sum_j \left(\frac{h}{d_j}\right)^{-\frac{1}{2}} d_j^{-1+\frac{d}{2}} \sum_{*,i} (h^{1+\alpha} \|F_t\|_{Q_i} + h^\alpha \|F\|_{1,Q_i}) d_i^{1-\alpha} \left(\frac{\min(d_i, d_j)}{\max(d_i, d_j)}\right)^\alpha.
\end{aligned}$$

In the case ‘ $d = 2$ (in this case $\alpha > \frac{1}{2}$)’ or ‘ $d = 3, r \geq 2$ and Ω is convex (in this case $\alpha > 1$)’, we have

$$\begin{aligned}
\mathcal{H} &\leq C + CC_*^{3+d/2} + C \quad (\text{here we exchange the order of summation}) \\
&\quad + C \sum_{*,i} (h^{1+\alpha} \|F_t\|_{Q_i} + h^\alpha \|F\|_{1,Q_i}) d_i^{1-\alpha} \sum_j \left(\frac{h}{d_j}\right)^{-\frac{1}{2}} d_j^{-1+\frac{d}{2}} \left(\frac{\min(d_i, d_j)}{\max(d_i, d_j)}\right)^\alpha \\
&\leq C + CC_*^{3+\frac{d}{2}} + C \sum_{*,i} (h^{1+\alpha} \|F_t\|_{Q_i} + h^\alpha \|F\|_{1,Q_i}) d_i^{\frac{d}{2}-\alpha} \left(\frac{h}{d_i}\right)^{-\frac{1}{2}} \quad ((4.43) \text{ is used}) \\
&= C + CC_*^{3+\frac{d}{2}} + C \sum_{*,i} d_i^{1+\frac{d}{2}} \left(\|F_t\|_{Q_i} \left(\frac{h}{d_i}\right)^{\frac{1}{2}+\alpha} + d_i^{-1} \|F\|_{1,Q_i} \left(\frac{h}{d_i}\right)^{\alpha-\frac{1}{2}} \right) \\
&\leq C + CC_*^{3+\frac{d}{2}} + C d_*^{1+\frac{d}{2}} \left(\|F_t\|_{Q_*} + d_j^{-1} \|F\|_{1,Q_*} \right) \\
&\quad + C \sum_i \left(\frac{h}{d_i}\right)^{-\frac{1}{2}} d_i^{1+\frac{d}{2}} \left(\|F_t\|_{Q_i} + d_j^{-1} \|F\|_{1,Q_i} \right) \left(\frac{h}{d_i}\right)^\alpha \\
&\leq C + CC_*^{3+d/2} + \frac{C\mathcal{H}}{C_*^\alpha}.
\end{aligned}$$

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

$$\mathcal{H} \leq C. \quad (4.44)$$

Substituting the last inequality into (4.31) yields

$$h^{-1} \|\nabla F\|_{L^1((0,1) \times D_h)} \leq C. \quad (4.45)$$

Part II. For $t \in [1, \infty)$, we consider the error equation

$$\begin{cases}
(\partial_t(\Gamma_h(t) - \Gamma(t)), v_h) + (\nabla(\Gamma_h(t) - \Gamma(t)), \nabla v_h) = 0, \forall v_h \in \dot{S}_h, \\
\Gamma_h(0) - \Gamma(0) = \delta_{h,x_0} - \tilde{\delta}_{x_0}.
\end{cases} \quad (4.46)$$

Let $\chi(t)$ be a smooth cut-off function such that

$$\chi(t) = \begin{cases} 1 & \text{for } t \geq \frac{1}{2}, \\ 0 & \text{for } t \leq \frac{1}{4}. \end{cases}$$

Then the function $\eta(t) = \chi(t)(\Gamma_h(t) - \Gamma(t))$ satisfies the following equation:

$$\begin{cases}
(\partial_t \eta(t), v_h) + (\nabla \eta(t), \nabla v_h) = (\partial_t \chi(\Gamma_h(t) - \Gamma(t)), v_h), \forall v_h \in \dot{S}_h, \\
\eta(0) = 0,
\end{cases} \quad (4.47)$$

which can also be written as

$$\begin{cases}
\partial_t R_h \eta - \Delta_h R_h \eta = -\partial_t [\chi(P_h - R_h)\Gamma] + \partial_t \chi(t)(\Gamma_h(t) - P_h \Gamma(t)), \\
R_h \eta(0) = 0,
\end{cases} \quad (4.48)$$

where $-\Delta_h : \dot{S}_h \rightarrow \dot{S}_h$ is the discrete Laplacian defined by

$$(-\Delta_h v_h, \chi) = (\nabla v_h, \nabla \chi), \quad \forall v_h \in \dot{S}_h. \quad (4.49)$$

The solution to (4.48) can be furthermore expressed in terms of the semigroup $e^{t\Delta_h}$, i.e.,

$$\begin{aligned} R_h\eta(t) &= - \int_0^t e^{(t-s)\Delta_h} \partial_s [\chi(s)(P_h - R_h)\Gamma(s)] ds \\ &\quad + \int_0^t e^{(t-s)\Delta_h} \partial_s \chi(s)(\Gamma_h(s) - P_h\Gamma(s)) ds. \end{aligned} \quad (4.50)$$

By applying the operator $(-\Delta_h)^{\frac{1}{2}}$ to the both sides of the above equality, we obtain

$$\begin{aligned} (-\Delta_h)^{\frac{1}{2}} R_h\eta(t) &= - \int_0^t (-\Delta_h)^{\frac{1}{2}} e^{(t-s)\Delta_h} \partial_s [\chi(s)(P_h - R_h)\Gamma(s)] ds \\ &\quad + \int_0^t (-\Delta_h)^{\frac{3}{2}} e^{(t-s)\Delta_h} (-\Delta_h)^{-1} \partial_s \chi(s)(\Gamma_h(s) - P_h\Gamma(s)) ds, \end{aligned} \quad (4.51)$$

and therefore

$$\begin{aligned} \|\nabla R_h\eta(t)\|_{L^2} &= \|(-\Delta_h)^{\frac{1}{2}} R_h\eta(t)\|_{L^2} \\ &\leq \left\| \int_0^t e^{(t-s)\Delta_h} (-\Delta_h)^{\frac{1}{2}} \partial_s [\chi(s)(P_h - R_h)\Gamma(s)] ds \right\|_{L^2} \\ &\quad + \left\| \int_0^t (-\Delta_h)^{\frac{3}{2}} e^{(t-s)\Delta_h} (-\Delta_h)^{-1} \partial_s \chi(s)(\Gamma_h(s) - P_h\Gamma(s)) ds \right\|_{L^2} \\ &\leq C \int_{\frac{1}{4}}^t e^{-\lambda_0(t-s)} \|\nabla(P_h - R_h)\partial_s [\chi(s)\Gamma(s)]\|_{L^2} ds \\ &\quad + C \int_{\frac{1}{4}}^{\frac{1}{2}} (t-s)^{-\frac{3}{2}} \|(-\Delta_h)^{-1}(\Gamma_h(s) - P_h\Gamma(s))\|_{L^2} ds, \end{aligned} \quad (4.52)$$

where $\lambda_0 > 0$ is the smallest eigenvalue of $-\Delta$ and we have used the following exponential decay estimate for the discrete semigroup in the last inequality:

$$\|e^{t\Delta_h} u_h\|_{L^2} \leq C e^{-\lambda_0 t} \|u_h\|_{L^2}, \quad (4.53)$$

for any $t \geq 0$ and $u_h \in \mathring{S}_h(\Omega)$. In addition, we have used the following result in (4.52):

$$\|(-\Delta_h)^{3/2} e^{(t-s)\Delta_h}\|_{L^2(\Omega) \rightarrow L^2(\Omega)} \leq C(t-s)^{-3/2},$$

which follows from the smoothing estimates for the analytic semigroup $e^{(t-s)\Delta_h}$, which means that (see [21, Theorem 2.1 and Remark 2.1])

$$\|\partial_t^l e^{t\Delta_h} \chi\|_{L^2(\Omega)} \leq C t^{-l} \|\chi\|_{L^2(\Omega)} \quad \text{for } \forall \chi \in \mathring{S}_h \text{ and } l = 0, 1, 2, \dots$$

Since $\partial_t^l e^{t\Delta_h} \chi = \Delta_h^l e^{t\Delta_h} \chi$ for $\chi \in \mathring{S}_h$, it follows that

$$\|(-\Delta_h)^l e^{t\Delta_h} \chi\|_{L^2(\Omega)} \leq C t^{-l} \|\chi\|_{L^2(\Omega)} \quad \text{for } \forall \chi \in \mathring{S}_h \text{ and } l = 1, 2,$$

which imply that

$$\begin{aligned} \|(-\Delta_h)^{3/2} e^{(t-s)\Delta_h} \chi\|_{L^2(\Omega)}^2 &= \left((-\Delta_h)^{3/2} e^{(t-s)\Delta_h} \chi, (-\Delta_h)^{3/2} e^{(t-s)\Delta_h} \chi \right) \\ &= \left((-\Delta_h)^2 e^{(t-s)\Delta_h} \chi, (-\Delta_h) e^{(t-s)\Delta_h} \chi \right) \\ &\leq \|\Delta_h^2 e^{(t-s)\Delta_h} \chi\|_{L^2(\Omega)} \|\Delta_h e^{(t-s)\Delta_h} \chi\|_{L^2(\Omega)} \\ &\leq \frac{C}{(t-s)^3} \|\chi\|_{L^2(\Omega)}^2. \end{aligned}$$

This proves the desired result which we use in (4.52).

By integrating the above inequality for $t \in [1, +\infty)$ and using Fubini's theorem and (3.6)-(3.7), we obtain

$$\begin{aligned}
& \|\nabla R_h \eta\|_{L^1(1, \infty; L^2)} \\
& \leq C \|\nabla(P_h - R_h) \partial_t(\chi \Gamma)\|_{L^1(\frac{1}{4}, +\infty; L^2)} + C \|(-\Delta_h)^{-1}(\Gamma_h - P_h \Gamma)\|_{L^1(0, \frac{1}{2}; L^2)} \\
& \leq Ch^\alpha \|\partial_t(\chi \Gamma)\|_{L^1(\frac{1}{4}, +\infty; H^{1+\alpha})} + C \|\Gamma_h - \Gamma\|_{L^1(\frac{1}{4}, \frac{1}{2}; L^1)} \\
& \leq Ch^{\frac{1}{2}}, \tag{4.54}
\end{aligned}$$

where we have used Lemma 4.3 in deriving the second and the last inequality.

From (4.54) and Poincaré inequality,

$$\|\Gamma_h - R_h \Gamma\|_{L^1(1, +\infty; H^1(\Omega))} \leq Ch^{\frac{1}{2}}.$$

By using the triangle inequality, Hölder's inequality on D_h , (3.7) and Lemma 4.3 (item 1), we obtain

$$\begin{aligned}
& \|\Gamma_h - \Gamma\|_{L^1(1, \infty; W^{1,1}(D_h))} \\
& \leq Ch^{\frac{1}{2}} \|\Gamma_h - R_h \Gamma\|_{L^1(1, \infty; H^1(D_h))} + Ch^{\frac{1}{2}} \|R_h \Gamma - \Gamma\|_{L^1(1, \infty; H^1(D_h))} \\
& \leq Ch + Ch^{\frac{1}{2} + \alpha} \|\Gamma\|_{L^1(1, \infty; H^{1+\alpha}(\Omega))} \\
& \leq Ch. \tag{4.55}
\end{aligned}$$

Estimates (4.45)-(4.55) imply (4.21). This completes the proof of Theorem 2.1. \square

5 Weak maximum principle of fully discrete FEM

To prove Theorem 2.2, it suffices to show

$$\|u_h^N\|_{L^\infty(\Omega)} \leq C \max_{1 \leq n \leq k-1} \|u_h^n\|_{L^\infty(\Omega)} + C \max_{k \leq n \leq N} \|g_h^n\|_{L^\infty(\partial\Omega)} \quad \forall N \geq k, \tag{5.1}$$

with a constant C independent of τ , h and N . To this end, we decompose the solution into three parts, i.e.,

$$u_h^N = u_{h,1}^N + u_{h,2}^N + u_{h,3}^N, \tag{5.2}$$

where $u_{h,1}^n$, $u_{h,2}^n$ and $u_{h,3}^n$ are finite element solutions of the following problems:

$$\begin{cases} \left(\frac{1}{\tau} \sum_{j=0}^k \delta_j u_{h,1}^{n-j}, v_h \right) + (\nabla u_{h,1}^n, \nabla v_h) = 0 & \forall v_h \in \mathring{S}_h, \quad n = k, \dots, N, \\ u_{h,1}^n = g_h^n (\delta_{n,N-1} + \delta_{n,N}) & \text{on } \partial\Omega, \quad n = k, \dots, N, \\ u_{h,1}^j = 0 & \text{in } \Omega, \quad j = 0, \dots, k-1, \end{cases} \tag{5.3}$$

$$\begin{cases} \left(\frac{1}{\tau} \sum_{j=0}^k \delta_j u_{h,2}^{n-j}, v_h \right) + (\nabla u_{h,2}^n, \nabla v_h) = 0 & \forall v_h \in \mathring{S}_h, \quad n = k, \dots, N, \\ u_{h,2}^n = g_h^n (1 - \delta_{n,N-1} - \delta_{n,N}) & \text{on } \partial\Omega, \quad n = k, \dots, N, \\ u_{h,2}^j = 0 & \text{in } \Omega, \quad j = 0, \dots, k-1, \end{cases} \tag{5.4}$$

$$\begin{cases} \left(\frac{1}{\tau} \sum_{j=0}^k \delta_j u_{h,3}^{n-j}, v_h \right) + (\nabla u_{h,3}^n, \nabla v_h) = 0 & \forall v_h \in \mathring{S}_h, \quad n = k, \dots, N, \\ u_{h,3}^n = 0 & \text{on } \partial\Omega, \quad n = k, \dots, N, \\ u_{h,3}^j = u_h^j & \text{in } \Omega, \quad j = 0, \dots, k-1, \end{cases} \tag{5.5}$$

with $\delta_{n,N}$ denoting the Kronecker symbol. In the following three subsections, we prove the following three estimates:

$$\|u_{h,1}^N\|_{L^\infty(\Omega)} \begin{cases} \leq C \max_{n=N} \|g_h^n\|_{L^\infty(\partial\Omega)}, & N = k, \\ \leq C \max_{n=N-1,N} \|g_h^n\|_{L^\infty(\partial\Omega)}, & N \geq k+1, \end{cases} \quad (5.6)$$

$$\|u_{h,2}^N\|_{L^\infty(\Omega)} \begin{cases} = 0, & N = k, k+1, \\ \leq C \max_{k \leq n \leq N-2} \|g_h^n\|_{L^\infty(\partial\Omega)}, & N \geq k+2, \end{cases} \quad (5.7)$$

$$\|u_{h,3}^N\|_{L^\infty(\Omega)} \leq C \max_{1 \leq n \leq k-1} \|u_h^n\|_{L^\infty(\Omega)}. \quad (5.8)$$

Then, substituting these estimates into (5.2), we obtain the desired result in Theorem 2.2.

5.1 Proof of (5.6)

We define the sector $\Sigma_\theta = \{z \in \mathbb{C} : |\arg(z)| \leq \theta\}$. The following result will be used in both this and next subsections.

Lemma 5.1 *For $\theta \in (-\frac{\pi}{2}, \frac{\pi}{2})$ and $z \in \Sigma_\theta$, the elliptic finite element problem*

$$\begin{cases} (z\phi_h, v_h) + (\nabla\phi_h, \nabla v_h) = 0 & \forall v_h \in \dot{S}_h, \\ \phi_h = f_h & \text{on } \partial\Omega, \end{cases}$$

is well defined for any $f_h \in S_h(\partial\Omega)$ and satisfies the following estimate:

$$\|\phi_h\|_{L^\infty(\Omega)} \leq C \|f_h\|_{L^\infty(\partial\Omega)},$$

where the constant C is independent of $z \in \Sigma_\theta$.

Proof Let $w_h \in S_h$ be the finite element solution of the elliptic problem

$$\begin{cases} (\nabla w_h, \nabla v_h) = 0 & \forall v_h \in \dot{S}_h, \\ w_h = f_h & \text{on } \partial\Omega. \end{cases} \quad (5.9)$$

Then $\phi_h - w_h \in \dot{S}_h$ is the finite element solution of the zero Dirichlet boundary value problem

$$(z(\phi_h - w_h), v_h) + (\nabla(\phi_h - w_h), \nabla v_h) = -(zw_h, v_h) \quad \forall v_h \in \dot{S}_h,$$

which can equivalently be written as

$$(z - \Delta_h)(\phi_h - w_h) = -zP_h w_h,$$

where P_h is the L^2 projection operator from $L^2(\Omega)$ onto its closed subspace \dot{S}_h . (Since $w_h \neq 0$ on $\partial\Omega$, it follows that $P_h w_h \neq w_h$.)

For the elliptic problem (5.9), it is known that the following weak maximum principle holds (cf. [26] and [19] for 2D and 3D cases, respectively):

$$\|w_h\|_{L^\infty(\Omega)} \leq C \|f_h\|_{L^\infty(\partial\Omega)}. \quad (5.10)$$

Hence, we have

$$\|\phi_h - w_h\|_{L^\infty(\Omega)} = \|z(z - \Delta_h)^{-1}P_h w_h\|_{L^\infty(\Omega)} \leq C \|w_h\|_{L^\infty(\Omega)}, \quad (5.11)$$

where the last inequality is due to the L^∞ stability of P_h (cf. [29, Lemma 6.1]) and the resolvent estimate (cf. [22, Theorem 3.1])

$$\|z(z - \Delta_h)^{-1}\|_{L^\infty(\Omega) \rightarrow L^\infty(\Omega)} \leq C \quad \forall z \in \mathbb{C}, \quad z \in \Sigma_\theta.$$

By using the triangle inequality and (5.10)–(5.11), we obtain (5.6). \square

For $N = k$, according to the initial condition, $u_{h,1}^n$ vanishes for $1 \leq n \leq k-1$, and $u_{h,1}^N$ is the solution of the discrete elliptic equation

$$\begin{cases} \left(\frac{u_{h,1}^N}{\tau}, v_h \right) + (\nabla u_{h,1}^N, \nabla v_h) = 0 & \forall v_h \in \mathring{S}_h, \\ u_{h,1}^N = g_h^N & \text{on } \partial\Omega, \end{cases} \quad (5.12)$$

where we have used the fact that $\delta_0 = 1$ for all BDF- k methods with $k = 1, \dots, 6$. Then Lemma 5.1 implies that the solution of (5.12) satisfies (5.6).

If $N \geq k+1$, the solution $u_{h,1}^n$ of (5.3) is actually zero for $1 \leq n \leq N-2$. Therefore $u_{h,1}^{N-1}$ satisfies the elliptic equation

$$\begin{cases} \left(\frac{u_{h,1}^{N-1}}{\tau}, v_h \right) + (\nabla u_{h,1}^{N-1}, \nabla v_h) = 0 & \forall v_h \in \mathring{S}_h, \\ u_{h,1}^{N-1} = g_h^{N-1} & \text{on } \partial\Omega, \end{cases} \quad (5.13)$$

and similarly $u_{h,1}^N$ is determined by

$$\begin{cases} \left(\frac{u_{h,1}^N}{\tau}, v_h \right) + (\nabla u_{h,1}^N, \nabla v_h) = - \left(\frac{\delta_1 u_{h,1}^{N-1}}{\tau}, v_h \right) & \forall v_h \in \mathring{S}_h, \\ u_{h,1}^N = g_h^N & \text{on } \partial\Omega. \end{cases} \quad (5.14)$$

Since $\frac{1}{\tau} \in \Sigma_\theta$, Lemma 5.1 implies that the solution of (5.13) satisfies

$$\|u_{h,1}^{N-1}\|_{L^\infty(\Omega)} \leq C \|g_h^{N-1}\|_{L^\infty(\partial\Omega)}. \quad (5.15)$$

At the time level N , we decompose $u_{h,1}^N = u_{h,1}^{N,\text{homo}} + u_{h,1}^{N,\text{inhomo}}$ into homogeneous and inhomogeneous parts. By applying Lemma 5.1 to homogeneous part of (5.14) to get

$$\|u_{h,1}^{N,\text{homo}}\|_{L^\infty(\Omega)} \leq C \|g_h^N\|_{L^\infty(\partial\Omega)},$$

and for the inhomogeneous part we have,

$$\begin{aligned} \|u_{h,1}^{N,\text{inhomo}}\|_{L^\infty(\Omega)} &= \|(\tau^{-1} - \Delta_h)^{-1} P_h \tau^{-1} \delta_1 u_{h,1}^{N-1}\|_{L^\infty(\Omega)} \leq C \|u_{h,1}^{N-1}\|_{L^\infty(\Omega)} \\ &\leq C \|g_h^{N-1}\|_{L^\infty(\partial\Omega)}, \end{aligned}$$

where we have used the resolvent estimate and the L^∞ stability of P_h in the first inequality, and have used (5.15) in the second inequality. Then by the triangle inequality, we conclude (5.6) as follows

$$\begin{aligned} \|u_{h,1}^N\|_{L^\infty(\Omega)} &\leq \|u_{h,1}^{N,\text{homo}}\|_{L^\infty(\Omega)} + \|u_{h,1}^{N,\text{inhomo}}\|_{L^\infty(\Omega)} \\ &\leq C (\|g_h^{N-1}\|_{L^\infty(\partial\Omega)} + \|g_h^N\|_{L^\infty(\partial\Omega)}). \end{aligned}$$

5.2 Proof of (5.7)

According to (5.4), $u_{h,2}^n, k \leq n \leq N$ is trivial for $N \leq k+1$. So in this subsection we will assume $N \geq k+2$ in order to get at least one meaningful time-stepping.

We extend g_h^n to be zero for $0 \leq n \leq k-1$ and $n \geq N+1$, and redefine $g_h^{N-1} = 0$ and $g_h^N = 0$. Moreover we extend $u_{h,2}^n, n \geq N+1$, to be the solution of the following equation at the time level $n \geq N+1$:

$$\begin{cases} \left(\frac{1}{\tau} \sum_{j=0}^k \delta_j u_{h,2}^{n-j}, v_h \right) + (\nabla u_{h,2}^n, \nabla v_h) = 0 & \forall v_h \in \mathring{S}_h, \quad n \geq k, \\ u_{h,2}^n = g_h^n & \text{on } \partial\Omega, \quad n \geq k, \\ u_{h,2}^j = 0 & \text{in } \Omega, \quad j = 0, \dots, k-1. \end{cases} \quad (5.16)$$

We denote by

$$\tilde{u}_{h,2}(\zeta) = \sum_{n=k}^{\infty} u_{h,2}^n \zeta^n \quad \text{and} \quad \tilde{g}_h(\zeta) = \sum_{n=k}^{\infty} g_h^n \zeta^n$$

the generating functions of the two sequences $(u_{h,2}^n)_{n=k}^{\infty}$ and $(g_h^n)_{n=k}^{\infty}$, respectively. By definition, $g_h^n = 0$ for $n \geq N$. From the standard energy decay estimate we know the boundedness $\sup_{n \geq 0} \|u_{h,2}^n\|_{L^2(\Omega)} \leq C \sup_{n \geq N} \|u_{h,2}^n\|_{L^2(\Omega)} \leq C_{h,N}$, and moreover via the inverse inequality and the Hölder's inequality, it holds that $\sup_{n \geq 0} \|u_{h,2}^n\|_{L^p(\Omega)} \leq C_{h,N}$ for all $p \in [1, \infty]$. Therefore the L^p -valued generating functions $\tilde{g}_h(\zeta)$ and $\tilde{u}_{h,2}(\zeta)$ are analytic in $\zeta \in \mathbb{C}$ and in the open unit disk $\zeta \in \mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ respectively for any $h > 0$ and $p \in [1, \infty]$.

For any $\zeta \in \mathbb{D}$, we observe the following relation of generating functions

$$\begin{aligned} \delta(\zeta) \tilde{u}_{h,2}(\zeta) &= \left(\sum_{j=0}^k \delta_j \zeta^j \right) \left(\sum_{j=k}^{+\infty} u_{h,2}^j \zeta^j \right) \\ &= \sum_{n=k}^{+\infty} \zeta^n \sum_{j=0, n-j \geq k}^k \delta_j u_{h,2}^{n-j} \\ &= \sum_{n=k}^{+\infty} \zeta^n \sum_{j=0}^k \delta_j u_{h,2}^{n-j} - \sum_{n=k}^{2k-1} \zeta^n \sum_{j=0, n-j < k}^k \delta_j u_{h,2}^{n-j}. \end{aligned} \quad (5.17)$$

Multiplying (2.2) by ζ^n , summing up the results for $n = k, k+1, \dots$ and using the identity (5.17) and the initial condition $u_{h,2}^n = 0$ for $n = 0, \dots, k-1$, we obtain the following discrete elliptic equation of the generating function $\tilde{u}_{h,2}(\zeta)$:

$$\begin{cases} \left(\delta_\tau(\zeta) \tilde{u}_{h,2}(\zeta), v_h \right) + (\nabla \tilde{u}_{h,2}(\zeta), \nabla v_h) = \left(-\frac{1}{\tau} \sum_{n=k}^{2k-1} \zeta^n \sum_{j=0, n-j < k}^k \delta_j u_{h,2}^{n-j}, v_h \right) = 0 \\ \tilde{u}_{h,2}(\zeta) = \tilde{g}_h(\zeta) \quad \text{on } \partial\Omega, \end{cases} \quad \forall v_h \in \mathring{S}_h \quad (5.18)$$

whose solution map is denoted by $\tilde{M}_h(\zeta) : S_h(\partial\Omega) \rightarrow S_h(\Omega)$, $\tilde{g}_h(\zeta) \mapsto \tilde{u}_{h,2}(\zeta)$.

We define the truncated sectors $\Sigma_\theta^\tau = \{z \in \Sigma_\theta : \text{Im}(z) \leq \pi/\tau\}$ and $\Sigma_{\theta,\sigma}^\tau = \{z \in \Sigma_\theta : |z| \geq \sigma^{-1} \text{ and } \text{Im}(z) \leq \pi/\tau\}$ for any $\sigma > \tau/\pi$. The following result of the generating function of BDF- k method is shown in [12, Lemma 3.1].

Lemma 5.2 *For any ϵ and $k = 1, \dots, 6$, there exists $\theta_\epsilon \in (\pi/2, \pi)$ such that for any $\theta \in (\pi/2, \theta_\epsilon)$, there exist positive constants C, C_1 , and C_2 (independent of τ) such that for all $z \in \Sigma_\theta^\tau$ we have the inclusion $\delta_\tau(e^{-\tau z}) \subseteq \Sigma_{\pi-\theta_k+\epsilon}$ (Recall that θ_k is the angle of A -stability region.) and the following estimates hold*

$$\begin{aligned} C_1 |z| &\leq |\delta_\tau(e^{-\tau z})| \leq C_2 |z|, \\ \delta_\tau(e^{-z\tau}) &\in \Sigma_{\pi-\theta_k+\epsilon}, \\ |\delta_\tau(e^{-z\tau}) - z| &\leq C\tau^k |z|^{k+1}. \end{aligned}$$

For BDF-1 (i.e. backward Euler) time-stepping method, the results in Lemma 5.2 can be furthermore improved.

Lemma 5.3 *Let $\theta \in (\frac{\pi}{2}, \operatorname{arccot}(-\frac{2}{\pi}))$. For $z \in \Sigma_\theta$, we have $\frac{1-e^{-\tau z}}{\tau} \in \Sigma_\theta$. Moreover there exist positive constants C, C_1 , and C_2 such that for $z \in \Sigma_\theta^\tau$*

$$C_1|z| \leq \left| \frac{1-e^{-\tau z}}{\tau} \right| \leq C_2|z|,$$

$$\left| \frac{1-e^{-\tau z}}{\tau} - z \right| \leq C\tau|z|^2.$$

Additionally, for $\sigma > \tau/\pi$, the following inequalities hold:

$$\left| 1 - \frac{e^{\tau z} - 1}{z\tau} \right| \leq C|z|\tau \quad \text{and} \quad \left| \frac{e^{\tau z} - 1}{z\tau} \right| \leq C \quad \forall z \in \Gamma_{\theta,\sigma}^\tau, \quad (5.19)$$

$$\left| \frac{e^{\tau z} - 1}{z\tau} \right| \leq C \quad \text{and} \quad |z| \leq C \left| \frac{1-e^{-\tau z}}{\tau} \right| \quad \forall z \in \Gamma_{\theta,\sigma} \setminus \Gamma_{\theta,\sigma}^\tau. \quad (5.20)$$

The first part of this lemma results from [10, Lemma 3.4] and [9, Appendix C]. For the second part, inequalities in (5.19) can be verified by using Taylor's expansion on a bounded domain. Inequalities in (5.20) can be proved by using the fact that $|z|\tau \geq \pi/|\sin(\theta)|$ for $z \in \Gamma_{\theta,\sigma} \setminus \Gamma_{\theta,\sigma}^\tau$.

From now on, we fix any $0 < \epsilon < \min_{1 \leq k \leq 6} \theta_k$ and an angle θ such that

$$\frac{\pi}{2} < \theta < \min\{\theta_\epsilon, \operatorname{arccot}(-\frac{2}{\pi})\} \quad (5.21)$$

where θ_ϵ is determined in Lemma 5.2. From Lemma 5.2, we know $\delta_\tau(e^{-z\tau}) \in \Sigma_{\pi-\theta_k+\epsilon} \subseteq \rho(\Delta_h)$ for $z \in \Sigma_\theta^\tau$, so the operator $\tilde{M}(e^{-\tau z})$ is well-defined and analytic for $z \in \Sigma_\theta^\tau$, satisfying the following estimate according to Lemma 5.1:

$$\|\tilde{M}_h(e^{-\tau z})\|_{L^\infty(\partial\Omega) \rightarrow L^\infty(\Omega)} \leq C \quad \forall z \in \Sigma_\theta^\tau. \quad (5.22)$$

Since $\tilde{u}_{h,2}(\cdot)$ is analytic in the open unit disk \mathbb{D} , by Cauchy's integral formula, we derive for any $0 < \rho < 1$ that

$$\begin{aligned} u_{h,2}^N &= \frac{1}{2\pi i} \int_{|\zeta|=\rho} \tilde{u}_{h,2}(\zeta) \zeta^{-N-1} d\zeta \\ &= \frac{\tau}{2\pi i} \int_{-\frac{\ln \rho}{\tau} - i\frac{\pi}{\tau}}^{-\frac{\ln \rho}{\tau} + i\frac{\pi}{\tau}} \tilde{u}_{h,2}(e^{-\tau z}) e^{tNz} dz \quad (\text{change of variable } \zeta = e^{-\tau z}) \\ &= \frac{\tau}{2\pi i} \int_{\Gamma_{\theta,\sigma}^\tau} \tilde{u}_{h,2}(e^{-\tau z}) e^{tNz} dz \quad (\text{deform the contour}) \\ &= \frac{\tau}{2\pi i} \int_{\Gamma_{\theta,\sigma}^\tau} \tilde{M}_h(e^{-\tau z}) \tilde{g}_h(e^{-\tau z}) e^{tNz} dz \quad (\text{by the definition of } M_h(\zeta)), \end{aligned} \quad (5.23)$$

where we have deformed the contour of integration from

$$\{z \in \mathbb{C} : \operatorname{Re}(z) = -\ln \rho/\tau, |\operatorname{Im}(z)| \leq \pi/\tau\}$$

to

$$\begin{aligned} \Gamma_{\theta,\sigma}^\tau &= \Gamma_{\theta,\sigma}^{\tau,1} \cup \Gamma_{\theta,\sigma}^{\tau,2} \quad \text{with } \Gamma_{\theta,\sigma}^{\tau,1} = \{z \in \mathbb{C} : |\arg(z)| = \theta, |z| \geq s^{-1} \text{ and } |\operatorname{Im}(z)| \leq \pi/\tau\}, \\ &\quad \text{and } \Gamma_{\theta,\sigma}^{\tau,2} = \{z \in \mathbb{C} : |z| = s^{-1} \text{ and } |\arg(z)| \leq \theta\}, \end{aligned} \quad (5.24)$$

where $\sigma > \tau/\pi$ can be arbitrary. The deformation of the contour is valid due to the analyticity of the integrand and the periodicity in $\operatorname{Im}(z)$ with period $2\pi/\tau$.

Let $u_{h,2} \in C([0, T] \times \Omega) \cap C_{\text{pw}}^\infty([0, T]; S_h(\Omega))$ be the globally continuous piecewise smooth solution of the semi-discrete FEM

$$\begin{cases} (\partial_t u_{h,2}(t), v_h) + (\nabla u_{h,2}(t), \nabla v_h) = 0 & \forall v_h \in \dot{S}_h, \quad \forall t \in (0, T] \\ u_{h,2}(t) = g_h(t) & \text{on } \partial\Omega, \quad \forall t \in [0, T], \\ u_{h,2}(0) = 0 & \text{in } \Omega, \end{cases} \quad (5.25)$$

where $g_h(t) = g_h^n$ is the $S_h(\partial\Omega)$ -valued globally continuous piecewise linear function whose nodal values satisfy $g_h(t_n) = g_h^n$ for $n \geq 0$. In particular, $g_h(t) = 0$ for $t \in [t_{N-1}, t_N]$, which implies that

$$u_{h,2}(t) = e^{(t-t_{N-1})\Delta_h} u_{h,2}(t_{N-1}) \quad \forall t \in [t_{N-1}, t_N].$$

Hence, we can furthermore decompose $u_{h,2}(t)$ into

$$u_{h,2}(t) = u_{h,21}(t) + u_{h,22}(t),$$

with

$$u_{h,21}(t) = \begin{cases} u_{h,2}(t) & \text{for } t \in [0, t_{N-1}], \\ 0 & \text{for } t \in (t_{N-1}, t_N], \end{cases}$$

$$u_{h,22}(t) = \begin{cases} 0 & \text{for } t \in [0, t_{N-1}], \\ e^{(t-t_{N-1})\Delta_h} u_{h,2}(t_{N-1}) & \text{for } t \in (t_{N-1}, t_N], \end{cases}$$

and denote by

$$\hat{v}(z) = \int_0^{+\infty} e^{-tz} v(t) dt$$

the Laplace transform of a function v in time. Then

$$\hat{u}_{h,2}(z) = \hat{u}_{h,21}(z) + \hat{u}_{h,22}(z) = \hat{u}_{h,21}(z) + (z - \Delta_h)^{-1} e^{-zt_{N-1}} u_{h,2}(t_{N-1}).$$

By considering the Laplace transform of (5.25) in time, we get

$$\begin{cases} (z\hat{u}_{h,2}(z), v_h) + (\nabla \hat{u}_{h,2}(z), \nabla v_h) = 0 & \forall v_h \in \dot{S}_h, \\ \hat{u}_{h,2}(z) = \hat{g}_h(z) & \text{on } \partial\Omega. \end{cases} \quad (5.26)$$

We also have the following formula for the Laplace transform of the piecewise linear function $g_h(t)$.

Lemma 5.4

$$\hat{g}_h(z) = \frac{e^{-\tau z} + e^{\tau z} - 2}{z^2 \tau} \tilde{g}_h(e^{-\tau z}). \quad (5.27)$$

Proof We use the change of variables $\zeta = e^{-\tau z}$. Straightforward calculations together with the initial condition $g_h^0 = 0$ lead to

$$\begin{aligned} \hat{g}_h(z) &= \int_0^{+\infty} g_h(t) e^{-zt} dt \\ &= \sum_{j=0}^{+\infty} \int_{t_j}^{t_{j+1}} \left(g_h^j + \frac{g_h^{j+1} - g_h^j}{\tau} (t - t_j) \right) e^{-zt} dt \\ &= \sum_{j=0}^{+\infty} g_h^j \frac{\zeta^{j+1} - \zeta^j}{-z} + \sum_{j=0}^{+\infty} \frac{g_h^{j+1} - g_h^j}{\tau} \left(-\frac{\tau \zeta^{j+1}}{z} - \frac{\zeta^{j+1} - \zeta^j}{z^2} \right) \\ &= \sum_{j=0}^{+\infty} \frac{1 - \zeta}{z} g_h^j \zeta^j - \sum_{j=0}^{+\infty} \frac{1 - \zeta}{z} g_h^j \zeta^j - \sum_{j=0}^{+\infty} \frac{(\zeta^{j+1} - \zeta^j)(g_h^{j+1} - g_h^j)}{z^2 \tau} \end{aligned}$$

$$\begin{aligned}
&= - \sum_{j=0}^{+\infty} \frac{(1-\zeta)(1-\zeta^{-1})}{z^{2\tau}} g_h^j \zeta^j \\
&= \frac{e^{-\tau z} + e^{\tau z} - 2}{z^{2\tau}} \tilde{g}_h(e^{-\tau z}).
\end{aligned}$$

□

Analogous to (5.19) and (5.20), the estimates below follow from Taylor's expansion.

$$\left| 1 - \frac{e^{-\tau z} + e^{\tau z} - 2}{z^{2\tau^2}} \right| \leq C|z|^{2\tau^2} \quad \text{and} \quad \left| \frac{e^{-\tau z} + e^{\tau z} - 2}{z^{2\tau^2}} \right| \leq C \quad \forall z \in \Gamma_{\theta, \sigma}^\tau, \quad (5.28)$$

$$\left| \frac{e^{-\tau z} + e^{\tau z} - 2}{z^{2\tau^2}} \right| \leq C|e^{-\tau z}| \quad \forall z \in \Gamma_{\theta, \sigma} \setminus \Gamma_{\theta, \sigma}^\tau. \quad (5.29)$$

Then using the above lemma, we can rewrite (5.26) as

$$\begin{cases} \left(\frac{1 - e^{-\tau z}}{\tau} \hat{u}_{h,2}(z), v_h \right) + (\nabla \hat{u}_{h,2}(z), \nabla v_h) = \left(\left(\frac{1 - e^{-\tau z}}{\tau} - z \right) \hat{u}_{h,2}(z), v_h \right) \\ \hspace{20em} \forall v_h \in \dot{S}_h, \\ \hat{u}_{h,2}(z) = \frac{e^{-\tau z} + e^{\tau z} - 2}{z^{2\tau}} \tilde{g}_h(e^{-\tau z}) \quad \text{on } \partial\Omega. \end{cases} \quad (5.30)$$

We denote by $\tilde{L}_h(e^{-\tau z}) : S_h(\partial\Omega) \rightarrow S_h(\Omega)$ the solution map of the homogeneous part of equation (5.30). By the property of θ (equation (5.21)) and Lemma 5.3, we have

$$\|\tilde{L}_h(e^{-\tau z})\|_{L^\infty(\partial\Omega) \rightarrow L^\infty(\Omega)} \leq C \quad \forall z \in \Sigma_\theta. \quad (5.31)$$

The estimate of the operator difference $\tilde{M}_h(e^{-\tau z}) - \tilde{L}_h(e^{-\tau z})$ is given in the following lemma.

Lemma 5.5 *Given any $z \in \Sigma_\theta^\tau$, it holds that*

$$\|\tilde{M}_h(e^{-\tau z}) - \tilde{L}_h(e^{-\tau z})\|_{L^\infty(\partial\Omega) \rightarrow L^\infty(\Omega)} \leq C\tau|z|.$$

Proof Again, we use the change of variables $\zeta = e^{-\tau z}$. For any fixed $z \in \Sigma_\theta^\tau$ and any given boundary data $f_h \in S_h(\partial\Omega)$, if we denote $v_h(\zeta) := \tilde{M}_h(\zeta)f_h$ and $w_h(\zeta) := \tilde{L}_h(\zeta)f_h$, by the definitions of \tilde{M}_h and \tilde{L}_h , we know

$$\begin{cases} \left(\delta_\tau(e^{-\tau z})v_h, \phi_h \right) + (\nabla v_h, \nabla \phi_h) = 0 \quad \forall \phi_h \in \dot{S}_h, \\ v_h = f_h \quad \text{on } \partial\Omega, \end{cases}$$

and

$$\begin{cases} \left(\frac{1 - e^{-\tau z}}{\tau} w_h, \phi_h \right) + (\nabla w_h, \nabla \phi_h) = 0 \quad \forall \phi_h \in \dot{S}_h, \\ w_h = f_h \quad \text{on } \partial\Omega. \end{cases}$$

Upon subtraction, the following equation holds for all $\phi_h \in \dot{S}_h(\Omega)$

$$\left(\frac{1 - e^{-\tau z}}{\tau} (v_h - w_h), \phi_h \right) + (\nabla (v_h - w_h), \nabla \phi_h) = - \left(\left(\delta_\tau(e^{-\tau z}) - \frac{1 - e^{-\tau z}}{\tau} \right) v_h, \phi_h \right),$$

and by inverting the elliptic operator

$$v_h - w_h = - \left(\frac{1 - e^{-\tau z}}{\tau} - \Delta_h \right)^{-1} P_h \left(\left(\delta_\tau(e^{-\tau z}) - \frac{1 - e^{-\tau z}}{\tau} \right) v_h \right).$$

Consequently, from the resolvent estimate, L^∞ stability of P_h , Lemma 5.2–5.3 and the mapping property of $\tilde{M}_h(\zeta)$, we get

$$\begin{aligned} \|(\tilde{M}_h(\zeta) - \tilde{L}_h(\zeta))f_h\|_{L^\infty(\Omega)} &= \|v_h - w_h\|_{L^\infty(\Omega)} \\ &\leq C|z|^{-1}\tau|z|^2\|\tilde{M}_h(\zeta)f_h\|_{L^\infty(\Omega)} \\ &\leq C\tau|z|\|f_h\|_{L^\infty(\partial\Omega)}. \end{aligned}$$

Therefore, it follows

$$\|\tilde{M}_h(e^{-\tau z}) - \tilde{L}_h(e^{-\tau z})\|_{L^\infty(\partial\Omega) \rightarrow L^\infty(\Omega)} \leq C\tau|z|.$$

□

The solution of (5.30) can be represented as

$$\begin{aligned} &\hat{u}_{h,2}(z) \\ &= \tilde{L}_h(e^{-\tau z})\frac{e^{-\tau z} + e^{\tau z} - 2}{z^2\tau}\tilde{g}_h(e^{-\tau z}) + \left(\frac{1 - e^{-\tau z}}{\tau} - \Delta_h\right)^{-1}P_h\left(\frac{1 - e^{-\tau z}}{\tau} - z\right)\hat{u}_{h,2}(z) \\ &= \tilde{L}_h(e^{-\tau z})\frac{e^{-\tau z} + e^{\tau z} - 2}{z^2\tau}\tilde{g}_h(e^{-\tau z}) + \left(\frac{1 - e^{-\tau z}}{\tau} - \Delta_h\right)^{-1}P_h\left(\frac{1 - e^{-\tau z}}{\tau} - z\right)\hat{u}_{h,21}(z) \\ &\quad + \left(\frac{1 - e^{-\tau z}}{\tau} - \Delta_h\right)^{-1}\left(\frac{1 - e^{-\tau z}}{\tau} - z\right)(z - \Delta_h)^{-1}e^{-z t_{N-1}}u_{h,2}(t_{N-1}). \end{aligned} \quad (5.32)$$

By using the inverse Laplace transform and the expression in (5.32), we have

$$\begin{aligned} u_{h,2}(t_N) &= \frac{1}{2\pi i} \int_{\Gamma_{\theta,\sigma}} \hat{u}_{h,2}(z)e^{t_N z} dz \\ &= \frac{\tau}{2\pi i} \int_{\Gamma_{\theta,\sigma}^\tau} \tilde{L}_h(e^{-\tau z})\frac{e^{-\tau z} + e^{\tau z} - 2}{z^2\tau^2}\tilde{g}_h(e^{-\tau z})e^{t_N z} dz \\ &\quad + \frac{\tau}{2\pi i} \int_{\Gamma_{\theta,\sigma} \setminus \Gamma_{\theta,\sigma}^\tau} \tilde{L}_h(e^{-\tau z})\frac{e^{-\tau z} + e^{\tau z} - 2}{z^2\tau^2}\tilde{g}_h(e^{-\tau z})e^{t_N z} dz \\ &\quad + \frac{1}{2\pi i} \int_{\Gamma_{\theta,\sigma}} \left(\frac{1 - e^{-\tau z}}{\tau} - \Delta_h\right)^{-1}P_h\left(\frac{1 - e^{-\tau z}}{\tau} - z\right)\hat{u}_{h,21}(z)e^{t_N z} dz \\ &\quad + \frac{1}{2\pi i} \int_{\Gamma_{\theta,\sigma}} \left(\frac{1 - e^{-\tau z}}{\tau} - \Delta_h\right)^{-1}\left(\frac{1 - e^{-\tau z}}{\tau} - z\right)(z - \Delta_h)^{-1}u_{h,2}(t_{N-1})e^{\tau z} dz, \end{aligned} \quad (5.33)$$

where $\sigma > \tau/\pi$ can be arbitrary and different in each of the integral above.

Subtracting (5.33) from (5.23) and using the boundary condition $g_h^n = 0$ for $n \leq k-1$ and $n \geq N-1$, we obtain

$$\begin{aligned} &u_{h,2}^N - u_{h,2}(t_N) \\ &= \frac{\tau}{2\pi i} \int_{\Gamma_{\theta,\sigma}^\tau} \left(1 - \frac{e^{-\tau z} + e^{\tau z} - 2}{z^2\tau^2}\right)\tilde{M}_h(e^{-\tau z})\tilde{g}_h(e^{-\tau z})e^{t_N z} dz \\ &\quad + \frac{\tau}{2\pi i} \int_{\Gamma_{\theta,\sigma}^\tau} \frac{e^{-\tau z} + e^{\tau z} - 2}{z^2\tau^2} \left(\tilde{M}_h(e^{-\tau z}) - \tilde{L}_h(e^{-\tau z})\right)\tilde{g}_h(e^{-\tau z})e^{t_N z} dz \\ &\quad - \frac{\tau}{2\pi i} \int_{\Gamma_{\theta,\sigma} \setminus \Gamma_{\theta,\sigma}^\tau} \frac{e^{-\tau z} + e^{\tau z} - 2}{z^2\tau^2} \tilde{L}_h(e^{-\tau z})\tilde{g}_h(e^{-\tau z})e^{t_N z} dz \\ &\quad - \frac{1}{2\pi i} \int_{\Gamma_{\theta,\sigma}} \left(\frac{1 - e^{-\tau z}}{\tau} - \Delta_h\right)^{-1}\left(\frac{1 - e^{-\tau z}}{\tau} - z\right)\hat{u}_{h,21}(z)e^{t_N z} dz \end{aligned}$$

$$\begin{aligned}
& -\frac{1}{2\pi i} \int_{\Gamma_{\theta, \sigma}} \left(\frac{1-e^{-\tau z}}{\tau} - \Delta_h \right)^{-1} \left(\frac{1-e^{-\tau z}}{\tau} - z \right) (z - \Delta_h)^{-1} u_{h,2}(t_{N-1}) e^{\tau z} dz \\
& = \sum_{j=k}^{N-2} F_h^{N-j} g_h^j + \sum_{j=k}^{N-2} G_h^{n-j} g_h^j - \sum_{j=k}^{N-2} L_h^{n-j} g_h^j \\
& \quad - \int_{t_{k-1}}^{t_{N-1}} K_h(t_N - s) P_h u_{h,2}(s) ds - Q_h u_{h,2}(t_{N-1}), \tag{5.34}
\end{aligned}$$

where

$$F_h^n = \frac{\tau}{2\pi i} \int_{\Gamma_{\theta, \sigma}^\tau} \left(1 - \frac{e^{-\tau z} + e^{\tau z} - 2}{z^2 \tau^2} \right) \tilde{M}_h(e^{-\tau z}) e^{t_n z} dz, \tag{5.35}$$

$$G_h^n = \frac{\tau}{2\pi i} \int_{\Gamma_{\theta, \sigma}^\tau} \frac{e^{-\tau z} + e^{\tau z} - 2}{z^2 \tau^2} \left(\tilde{M}_h(e^{-\tau z}) - \tilde{L}_h(e^{-\tau z}) \right) e^{t_n z} dz, \tag{5.36}$$

$$L_h^n = \frac{\tau}{2\pi i} \int_{\Gamma_{\theta, \sigma} \setminus \Gamma_{\theta, \sigma}^\tau} \frac{e^{-\tau z} + e^{\tau z} - 2}{z^2 \tau^2} \tilde{L}_h(e^{-\tau z}) e^{t_n z} dz, \tag{5.37}$$

$$K_h(s) = \frac{1}{2\pi i} \int_{\Gamma_{\theta, \sigma}} \left(\frac{1-e^{-\tau z}}{\tau} - \Delta_h \right)^{-1} \left(\frac{1-e^{-\tau z}}{\tau} - z \right) e^{sz} dz, \tag{5.38}$$

$$Q_h = \frac{1}{2\pi i} \int_{\Gamma_{\theta, \sigma}} \left(\frac{1-e^{-\tau z}}{\tau} - \Delta_h \right)^{-1} \left(\frac{1-e^{-\tau z}}{\tau} - z \right) (z - \Delta_h)^{-1} e^{\tau z} dz. \tag{5.39}$$

First, by choosing $\sigma = t_n$ in (5.35) and using (5.22) and (5.28), we have

$$\begin{aligned}
\|F_h^n\|_{L^\infty(\partial\Omega) \rightarrow L^\infty(\Omega)} & \leq C\tau \int_{\Gamma_{\theta, \sigma}^\tau} |z|^2 \tau^2 \|\tilde{M}_h(e^{-\tau z})\|_{L^\infty(\partial\Omega) \rightarrow L^\infty(\Omega)} e^{t_n \operatorname{Re}(z)} |dz| \\
& \leq C\tau^3 \int_{\Gamma_{\theta, t_n}^{\tau, 1}} |z|^2 e^{t_n \operatorname{Re}(z)} |dz| + C\tau^3 \int_{\Gamma_{\theta, t_n}^{\tau, 2}} |z|^2 e^{t_n \operatorname{Re}(z)} |dz| \\
& \leq C\tau^3 \int_{t_n^{-1}}^{+\infty} r^2 e^{-t_n r |\cos \theta|} dr + C\tau^3 \int_{-\theta}^{\theta} t_n^{-3} e^{-|\cos \theta|} d\theta \\
& \leq \frac{C\tau^3}{t_n^3} = \frac{C}{n^3}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\left\| \sum_{j=k}^{N-2} F_h^{N-j} g_h^j \right\|_{L^\infty(\Omega)} & \leq C \sum_{j=k}^{N-2} \frac{C}{(N-j)^3} \max_{k \leq n \leq N-2} \|g_h^n\|_{L^\infty(\partial\Omega)} \\
& \leq C \max_{k \leq n \leq N-2} \|g_h^n\|_{L^\infty(\partial\Omega)}. \tag{5.40}
\end{aligned}$$

Analogously, from (5.28), Lemma 5.5 and choosing $\sigma = t_n$, we get

$$\begin{aligned}
\|G_h^n\|_{L^\infty(\partial\Omega) \rightarrow L^\infty(\Omega)} & \leq C\tau \int_{\Gamma_{\theta, \sigma}^\tau} \|\tilde{M}_h(e^{-\tau z}) - \tilde{L}_h(e^{-\tau z})\|_{L^\infty(\partial\Omega) \rightarrow L^\infty(\Omega)} e^{t_n \operatorname{Re}(z)} |dz| \\
& \leq C\tau^2 \int_{\Gamma_{\theta, t_n}^{\tau, 1}} |z| e^{t_n \operatorname{Re}(z)} |dz| + C\tau^2 \int_{\Gamma_{\theta, t_n}^{\tau, 2}} |z| e^{t_n \operatorname{Re}(z)} |dz| \\
& \leq C\tau^2 \int_{t_n^{-1}}^{+\infty} r e^{-t_n r |\cos \theta|} dr + C\tau^2 \int_{-\theta}^{\theta} t_n^{-2} e^{-|\cos \theta|} d\theta \\
& \leq \frac{C\tau^2}{t_n^2} = \frac{C}{n^2}.
\end{aligned}$$

Therefore, it follows

$$\begin{aligned} \left\| \sum_{j=k}^{N-2} G_h^{N-j} g_h^j \right\|_{L^\infty(\Omega)} &\leq C \sum_{j=k}^{N-2} \frac{C}{(N-j)^2} \max_{k \leq n \leq N-2} \|g_h^n\|_{L^\infty(\partial\Omega)} \\ &\leq C \max_{k \leq n \leq N-2} \|g_h^n\|_{L^\infty(\partial\Omega)}. \end{aligned} \quad (5.41)$$

Then, by choosing $\sigma = t_{n-1}$ in (5.37) and using (5.31) and the estimates in (5.29), we have

$$\begin{aligned} \|L_h^n\|_{L^\infty(\partial\Omega) \rightarrow L^\infty(\Omega)} &\leq C\tau \int_{\Gamma_{\theta,\sigma} \setminus \Gamma_{\theta,\sigma}^\tau} e^{-\tau \operatorname{Re}(z)} \|\tilde{L}_h(e^{-\tau z})\|_{L^\infty(\partial\Omega) \rightarrow L^\infty(\Omega)} e^{t_n \operatorname{Re}(z)} |dz| \\ &\leq C\tau \int_{\Gamma_{\theta,\sigma} \setminus \Gamma_{\theta,\sigma}^\tau} e^{-t_{n-1}|z| \cos(\theta)} |dz| \\ &\leq C\tau \int_{\frac{\pi}{\tau \sin(\theta)}}^{+\infty} e^{-t_{n-1}r |\cos(\theta)|} dr \\ &\leq C \int_{\frac{\pi}{\sin(\theta)}}^{+\infty} e^{-\frac{t_{n-1} |\cos(\theta)|}{\tau} \tilde{r}} d\tilde{r} \quad (\text{change of variables}) \\ &\leq C \frac{1}{(n-1) |\cos \theta|} e^{-(n-1)\pi/|\tan(\theta)|} \quad \text{for } n \geq 2, \end{aligned}$$

and then

$$\begin{aligned} \left\| \sum_{j=k}^{N-2} L_h^{N-j} g_h^j \right\|_{L^\infty(\Omega)} &\leq C \sum_{j=k}^{N-2} (N-1-j)^{-1} e^{-(N-1-j)\pi/|\tan(\theta)|} \max_{k \leq n \leq N-2} \|g_h^n\|_{L^\infty(\partial\Omega)} \\ &\leq C \max_{k \leq n \leq N-2} \|g_h^n\|_{L^\infty(\partial\Omega)}. \end{aligned} \quad (5.42)$$

Since $\tau|z| \geq \pi/\sin(\theta)$ for $z \in \Gamma_{\theta,\sigma} \setminus \Gamma_{\theta,\sigma}^\tau$, by using the second inequality of (5.20) it is easy to verify that

$$\left| \frac{1 - e^{-\tau z}}{\tau} - z \right| \leq C \left| \frac{1 - e^{-\tau z}}{\tau} \right| \leq C\tau |z| \left| \frac{1 - e^{-\tau z}}{\tau} \right| \quad \forall z \in \Gamma_{\theta,\sigma} \setminus \Gamma_{\theta,\sigma}^\tau.$$

With this inequality and taking $\sigma = s$, we have

$$\begin{aligned} \|K_h(s)\|_{L^\infty(\Omega) \rightarrow L^\infty(\Omega)} &\leq C \int_{\Gamma_{\theta,s}^\tau} \left| \frac{1 - e^{-\tau z}}{\tau} \right|^{-1} \tau |z|^2 e^{s \operatorname{Re}(z)} |dz| \\ &\quad + C \int_{\Gamma_{\theta,s} \setminus \Gamma_{\theta,s}^\tau} \left| \frac{1 - e^{-\tau z}}{\tau} \right|^{-1} \tau |z| \left| \frac{1 - e^{-\tau z}}{\tau} \right| e^{s \operatorname{Re}(z)} |dz| \\ &\leq C\tau \int_{\Gamma_{\theta,s}} |z| e^{s \operatorname{Re}(z)} |dz| \\ &\leq C\tau \int_{s^{-1}}^{+\infty} r e^{-sr |\cos(\theta)|} dr + C\tau \int_{-\theta}^{\theta} s^{-2} e^{-|\cos \theta|} d\varphi \\ &\leq C\tau s^{-2} \quad \text{for } s \geq \tau. \end{aligned}$$

Hence,

$$\begin{aligned} \left\| \int_{t_{k-1}}^{t_{N-1}} K_h(t_N - s) P_h u_{h,2}(s) ds \right\|_{L^\infty(\Omega)} &\leq \int_{t_{k-1}}^{t_{N-1}} C\tau (t_N - s)^{-2} \|u_{h,2}(s)\|_{L^\infty(\Omega)} ds \\ &\leq C \|u_{h,2}\|_{L^\infty(t_{k-1}, t_{N-1}; L^\infty(\Omega))} \\ &\leq C \|g_h\|_{L^\infty(t_{k-1}, t_{N-1}; L^\infty(\partial\Omega))} \end{aligned}$$

$$= C \max_{k \leq n \leq N-2} \|g_h^n\|_{L^\infty(\partial\Omega)}, \quad (5.43)$$

where we have used Theorem 2.1 in the last inequality.

Similarly, by choosing $\sigma = \tau$ in the expression of Q_h , we have

$$\begin{aligned} \|Q_h\|_{L^\infty(\Omega) \rightarrow L^\infty(\Omega)} &\leq C \int_{\Gamma_{\theta,\sigma}^\tau} \left| \frac{1 - e^{-\tau z}}{\tau} \right|^{-1} \tau |z|^2 |z|^{-1} e^{\tau \operatorname{Re}(z)} |dz| \\ &\quad + C \int_{\Gamma_{\theta,\sigma} \setminus \Gamma_{\theta,\sigma}^\tau} \left| \frac{1 - e^{-\tau z}}{\tau} \right|^{-1} \tau |z| \left| \frac{1 - e^{-\tau z}}{\tau} \right| |z|^{-1} e^{\tau \operatorname{Re}(z)} |dz| \\ &\leq C \tau \int_{\Gamma_{\theta,\tau}} e^{\tau \operatorname{Re}(z)} |dz| \\ &\leq C \tau \int_{\tau^{-1}}^{+\infty} e^{-\tau r |\cos(\theta)|} dr + C \tau \int_{-\theta}^{\theta} \tau^{-1} e^{-|\cos \theta|} d\varphi \\ &\leq C. \end{aligned}$$

Hence,

$$\begin{aligned} \|Q_h u_{h,2}(t_{N-1})\|_{L^\infty(\Omega)} &\leq C \|u_{h,2}(t_{N-1})\|_{L^\infty(\Omega)} \leq C \|g_h\|_{L^\infty(0,t_{N-1};L^\infty(\partial\Omega))} \\ &= C \max_{k \leq n \leq N-2} \|g_h^n\|_{L^\infty(\partial\Omega)}, \quad (5.44) \end{aligned}$$

where we have used Theorem 2.1 again in the last inequality.

Finally, substituting (5.40)–(5.44) into (5.34), we obtain

$$\|u_{h,2}^N - u_{h,2}(t_N)\|_{L^\infty(\Omega)} \leq C \max_{k \leq n \leq N-2} \|g_h^n\|_{L^\infty(\partial\Omega)}. \quad (5.45)$$

Then, using the triangle inequality and Theorem 2.1, we obtain (5.7).

5.3 Proof of (5.8)

For $k \leq N \leq 2k - 1$, at the time t_k , we solve for $u_h^k \in \mathring{S}_h$ satisfying

$$\left(\frac{u_h^k}{\tau}, v_h \right) + (\nabla u_h^n, \nabla v_h) = - \left(\frac{1}{\tau} \sum_{j=1}^k \delta_j u_h^{n-j}, v_h \right) \forall v_h \in \mathring{S}_h,$$

which can be equivalently written as

$$u_h^k = - \left(\frac{1}{\tau} - \Delta_h \right)^{-1} P_h \frac{1}{\tau} \sum_{j=1}^k \delta_j u_h^{k-j}.$$

Then,

$$\begin{aligned} \|u_h^k\|_{L^\infty(\Omega)} &\leq \left\| \left(\frac{1}{\tau} - \Delta_h \right)^{-1} \right\|_{L^\infty(\Omega) \rightarrow L^\infty(\Omega)} \left\| P_h \frac{1}{\tau} \sum_{j=1}^k \delta_j u_h^{k-j} \right\|_{L^\infty(\Omega)} \\ &\leq C \sup_{0 \leq j \leq k-1} \|u_h^j\|_{L^\infty(\Omega)}. \end{aligned}$$

Analogously, at the time t_{k+1} , we have

$$u_h^{k+1} = - \left(\frac{1}{\tau} - \Delta_h \right)^{-1} P_h \frac{1}{\tau} \sum_{j=1}^k \delta_j u_h^{k+1-j},$$

with

$$\|u_h^{k+1}\|_{L^\infty(\Omega)} \leq C \sup_{1 \leq j \leq k} \|u_h^j\|_{L^\infty(\Omega)} \leq C \sup_{0 \leq j \leq k-1} \|u_h^j\|_{L^\infty(\Omega)}.$$

Recursively, we conclude

$$\|u_h^N\|_{L^\infty(\Omega)} \leq C \sup_{0 \leq j \leq k-1} \|u_h^j\|_{L^\infty(\Omega)},$$

for any $k \leq N \leq 2k-1$.

If $N \geq 2k$, multiplying (5.5) by ζ^n , summing up the results for $n = k, k+1, \dots$, and using the identity (5.17), we obtain the following elliptic equation for the generating function $\tilde{u}_{h,3}(\zeta) \in \dot{S}_h(\Omega)$ satisfying zero Dirichlet boundary condition:

$$\left(\delta_\tau(\zeta) \tilde{u}_{h,3}(\zeta), v_h \right) + (\nabla \tilde{u}_{h,3}(\zeta), \nabla v_h) = \left(-\frac{1}{\tau} \sum_{n=k}^{2k-1} \zeta^n \sum_{j=0, n-j < k}^k \delta_j u_{h,3}^{n-j}, v_h \right),$$

whose solution can be represented by

$$\tilde{u}_{h,3}(\zeta) = -\left(\delta_\tau(\zeta) - \Delta_h \right)^{-1} P_h \frac{1}{\tau} \sum_{n=k}^{2k-1} \zeta^n \sum_{j=0, n-j < k}^k \delta_j u_{h,3}^{n-j}.$$

$\tilde{u}_{h,3}(\zeta)$ is analytic in $\mathbb{D} = \{\zeta \in \mathbb{C} : |\zeta| < 1\}$ due to the standard energy estimate. By Cauchy's integral formula and the deformation of contour (cf. (5.23)), for any $0 < \rho < 1$ we have

$$\begin{aligned} u_{h,3}^N &= -\frac{1}{2\pi i} \int_{|\zeta|=\rho} \left(\delta_\tau(\zeta) - \Delta_h \right)^{-1} P_h \frac{1}{\tau} \sum_{n=k}^{2k-1} \zeta^n \sum_{j=0, n-j < k}^k \delta_j u_{h,3}^{n-j} \zeta^{-N-1} d\zeta \\ &= -\frac{1}{2\pi i} \int_{\Sigma_{\theta, \sigma}^\tau} \left(\delta_\tau(e^{-\tau z}) - \Delta_h \right)^{-1} P_h \sum_{n=k}^{2k-1} \sum_{j=0, n-j < k}^k \delta_j u_{h,3}^{n-j} e^{t_{N-n} z} dz. \end{aligned}$$

By choosing $\sigma = t_{N-n}$ ($N-n \geq 2k - (2k-1) = 1$) for each term of the summation in the representation formula above, it follows

$$\begin{aligned} &\|u_{h,3}^N\|_{L^\infty(\Omega)} \\ &\leq C \int_{\Gamma_{\theta, \sigma}^\tau} \|(\delta_\tau(\zeta) - \Delta_h)^{-1}\|_{L^\infty(\Omega) \rightarrow L^\infty(\Omega)} \sum_{n=k}^{2k-1} \sum_{j=0, n-j < k}^k \|P_h \delta_j u_{h,3}^{n-j}\|_{L^\infty(\Omega)} e^{t_{N-n} \operatorname{Re}(z)} dz \\ &\leq C \sup_{0 \leq j \leq k-1} \|u_h^j\|_{L^\infty(\Omega)} \sum_{n=k}^{2k-1} \left(\int_{\Gamma_{\theta, t_{N-n}}^{\tau, 1}} + \int_{\Gamma_{\theta, t_{N-n}}^{\tau, 2}} \right) |z|^{-1} e^{t_{N-n} \operatorname{Re}(z)} dz \\ &\leq C \sup_{0 \leq j \leq k-1} \|u_h^j\|_{L^\infty(\Omega)} \sum_{n=k}^{2k-1} \left(\int_{t_{N-n}^{-1}}^{+\infty} r^{-1} e^{-t_{N-n} r |\cos \theta|} dr + \int_{-\theta}^{\theta} e^{-|\cos \theta|} d\theta \right) \\ &\leq C \sup_{0 \leq j \leq k-1} \|u_h^j\|_{L^\infty(\Omega)}, \end{aligned}$$

where in the second inequality we have applied the resolvent estimate (Lemma 5.2) and the L^∞ stability of P_h .

This proves (5.8). The proof of Theorem 2.2 is complete.

6 Conclusions and further discussions

We have proved the weak maximum principle of semi-discrete FEM and fully discrete FEM with k -step BDF time-stepping method for $k = 1, \dots, 6$, using the techniques of generating polynomials and the discrete inverse Laplace transform (Cauchy integral formula), which requires using a uniform time step size. Therefore, the analysis in this article cannot be readily extended to k -step BDF methods or other single-step methods (such as the Runge–Kutta methods or discontinuous Galerkin (dG) methods)

with variable time step sizes τ_n , $n = 1, \dots, N$. For the dG time-stepping method with possibly variable step sizes satisfying the following conditions:

1. There are constants $c, \beta > 0$ independent of the maximal step size τ such that $\tau_{\min} \geq c\tau^\beta$.
2. There is a constant $\kappa > 0$ independent on τ such that $\kappa^{-1} \leq \frac{\tau_n}{\tau_{n+1}} \leq \kappa$ for $n = 1, \dots, N-1$.
3. It holds $\tau \leq \frac{1}{4}T$,

an application of the discrete semigroup estimates in [20] can yield the following result:

$$\|u_{kh}\|_{L^\infty(I; L^\infty(\Omega))} \leq C \ln \frac{T}{\tau} \|g_h\|_{L^\infty(I; L^\infty(\partial\Omega))},$$

with an additional logarithmic factor depending on the maximal step size τ . We present a proof for dG(0) in Appendix. The general dG(k) can be treated similarly by using the approach from [20]. The proof of weak maximum principle of dG with variable step sizes without the logarithmic factor is interesting and nontrivial.

Acknowledgment

We thank the anonymous referee for the valuable comments. The work of G. Bai and B. Li is supported in part by the Research Grants Council of the Hong Kong Special Administrative Region, China (GRF project no. PolyU15300519) and an internal grant at The Hong Kong Polytechnic University (Project ID: P0045404).

Appendix: Weak maximum principle of dG with variable step sizes

The space-time finite element space of degree 0 in time and degree $r \geq 1$ in space is defined as

$$X_{\tau,h}^{0,r} = \{v_{\tau h} : v_{\tau h}|_{I_n} \in \mathcal{P}_0(\hat{S}_h), n = 1, 2, \dots, N, r \geq 1\}. \quad (\text{A.1})$$

The dG(0) method for the homogeneous parabolic equation

$$\begin{cases} \partial_t v - \Delta v = 0 & \text{in } (0, T] \times \Omega, \\ v = 0 & \text{on } [0, T] \times \partial\Omega, \\ v|_{t=0} = v_0 & \text{in } \Omega. \end{cases} \quad (\text{A.2})$$

can be written as follows: For the given $v_0 \in L^p(\Omega)$, find $v_{\tau h} \in X_{\tau,h}^{0,r}$ such that

$$\begin{aligned} v_{\tau h,1} - \tau_1 \Delta_h v_{\tau h,1} &= v_0, \\ v_{\tau h,m} - \tau_m \Delta_h v_{\tau h,m} &= v_{\tau h,m-1}, \quad m = 2, 3, \dots, M. \end{aligned} \quad (\text{A.3})$$

The following result was shown in [20].

Proposition A (Discrete semigroup estimates in $L^p(\Omega)$) *Let $v_{\tau h} \in X_{\tau,h}^{0,r}$ be the solution of (A.3) with $v_0 \in L^p(\Omega)$, $1 \leq p \leq \infty$. Then there exists a constant C independent of τ such that*

$$\|v_{\tau h,m}\|_{L^p(\Omega)} + (t_m - t_l) \|\Delta_h v_{\tau h,m}\|_{L^p(\Omega)} \leq C \|v_{\tau h,l}\|_{L^p(\Omega)}, \quad \forall m > l \geq 1.$$

Now we consider $u_{\tau h} \in X_{\tau, h}^{0, r}$ to be the dG(0) solution to the parabolic equation with nonzero right hand side f , i.e., $u_{\tau h}$ satisfies,

$$\begin{aligned} u_{kh,1} - \tau_1 \Delta_h u_{kh,1} &= \tau_1 P_h f_1, \\ u_{kh,m} - \tau_m \Delta_h u_{kh,m} &= u_{kh,m-1} + \tau_m P_h f_m, \quad m = 2, 3, \dots, M, \end{aligned} \quad (\text{A.4})$$

where

$$f_m(\cdot) = \frac{1}{\tau_m} \int_{I_m} f(t, \cdot) dt$$

and $P_h : L^2(\Omega) \rightarrow \mathring{S}_h$ is the L^2 projection. Since f_m is the L^2 projection of f onto the piecewise constant functions on each subinterval I_m , we have

$$\max_{1 \leq m \leq M} \|f_m\|_{L^p(\Omega)} \leq C \|f\|_{L^\infty(I; L^p(\Omega))}, \quad 1 \leq p \leq \infty. \quad (\text{A.5})$$

Using (A.4), we can write the dG(0) solution as

$$u_{kh,m} = \sum_{l=1}^m \tau_l \left(\prod_{j=1}^{m-l+1} r(-\tau_{m-j+1} \Delta_h) \right) P_h f_l, \quad m = 1, 2, \dots, M, \quad (\text{A.6})$$

where $r(z) = (1+z)^{-1}$.

Consider the parabolic equation

$$\begin{cases} \partial_t u - \Delta u = 0 & \text{in } (0, T] \times \Omega, \\ u = g & \text{on } [0, T] \times \partial\Omega, \\ u|_{t=0} = 0 & \text{in } \Omega. \end{cases} \quad (\text{A.7})$$

With the help of the previous result we establish:

Theorem A (Weak maximum principle of dG(0)) *Let $u_0 = 0$. Then, there exists a constant C independent of k and h such that for u_{kh} be the fully discrete dG(0) solution to (A.7), we have*

$$\|u_{kh}\|_{L^\infty(I; L^\infty(\Omega))} \leq C \ln \frac{T}{\tau} \|g_h\|_{L^\infty(I; L^\infty(\partial\Omega))},$$

where g_h is some finite element approximation of g .

Proof First we define Let G_h be the discrete harmonic extension of g_h , i.e. $G_h(t)$ satisfies (pointwise in t)

$$\begin{cases} (\nabla G_h, \nabla \chi)_\Omega = 0 & \forall \chi \in \mathring{S}_h, \\ G_h = g_h & \text{on } \partial\Omega. \end{cases} \quad (\text{A.8})$$

Then the difference $v_h = u_h - G_h \in \mathring{S}_h$ (zero on $\partial\Omega$) satisfies the semi-discrete problem,

$$\begin{cases} (\partial_t v_h, \chi)_\Omega + (\nabla v_h, \nabla \chi)_\Omega = (\partial_t G_h, \chi)_\Omega & \text{in } (0, T] \times \Omega, \\ v_h|_{t=0} = 0 & \text{in } \Omega, \end{cases} \quad (\text{A.9})$$

for any $\chi \in \mathring{S}_h$, and as a result the fully discrete difference $v_{kh} = u_{kh} - P_h G_{kh} \in X_{k, h}^{0, r}$, where P_h is the L^2 -orthogonal projection onto \mathring{S}_h , using (A.4) satisfies

$$v_{kh,m} = \sum_{l=1}^m \tau_l \left(\prod_{j=1}^{m-l+1} r(-\tau_{m-j+1} \Delta_h) \right) \frac{P_h G_{kh,l} - P_h G_{kh,l-1}}{\tau_l}, \quad m = 1, 2, \dots, M,$$

where $r(z) = (1+z)^{-1}$ and with the convention that $G_{kh,0} = 0$. Using the discrete integration by parts, we have

$$\begin{aligned}
& \sum_{l=1}^m \left(\prod_{j=1}^{m-l+1} r(-\tau_{m-j+1} \Delta_h) \right) (P_h G_{kh,l} - P_h G_{kh,l-1}) \\
&= r(-\tau_m \Delta_h) P_h G_{kh,m} - \prod_{j=1}^m r(-\tau_{m-j+1} \Delta_h) P_h G_{kh,0} \\
&+ \sum_{l=2}^{m-1} (\text{Id} - r(-\tau_l \Delta_h)) \left(\prod_{j=1}^{m-l+1} r(-\tau_{m-j+1} \Delta_h) \right) P_h G_{kh,l} \\
&= r(-\tau_m \Delta_h) P_h G_{kh,m} + \sum_{l=2}^{m-1} \tau_l \Delta_h r(-\tau_l \Delta_h) \left(\prod_{j=1}^{m-l+1} r(-\tau_{m-j+1} \Delta_h) \right) P_h G_{kh,l},
\end{aligned}$$

where we used that $\text{Id} - r(-\tau_l \Delta_h) = \tau_l \Delta_h r(-\tau_l \Delta_h)$ and $P_h G_{kh,0} = 0$. Using the properties of the rational function r and the elliptic theory, we have

$$\|r(-\tau_m \Delta_h) G_{kh,m+1}\|_{L^\infty(\Omega)} \leq C \|g\|_{L^\infty(I; L^\infty(\partial\Omega))}.$$

Hence for $m = 1, 2, \dots, M$,

$$\|v_{kh,m}\|_{L^\infty(\Omega)} \leq C \|g\|_{L^\infty(I; L^\infty(\partial\Omega))} + \sum_{l=1}^m \tau_l \left\| \left(\Delta_h \prod_{j=1}^{m-l+1} r(-\tau_{m-j+1} \Delta_h) \right) P_h G_{kh,l} \right\|_{L^\infty(\Omega)}.$$

From Proposition A, since each term in the sum on the right-hand side can be thought of as a homogeneous solution with initial condition $P_h G_{kh,l}$ at $t = t_{l-1}$, we have

$$\left\| \left(\Delta_h \prod_{j=1}^{m-l+1} r(-\tau_{m-j+1} \Delta_h) \right) P_h G_{kh,l} \right\|_{L^\infty(\Omega)} \leq \frac{C}{t_m - t_{l-1}} \|G_{kh,l}\|_{L^\infty(\Omega)}.$$

Thus, we obtain

$$\|v_{kh,m}\|_{L^\infty(\Omega)} \leq C \|g\|_{L^\infty(I; L^\infty(\partial\Omega))} + C \sum_{l=1}^m \frac{\tau_l}{t_m - t_{l-1}} \|G_{kh,l}\|_{L^\infty(\Omega)}, \quad m = 1, 2, \dots, M. \quad (\text{A.10})$$

From the above estimate and using (A.5),

$$\begin{aligned}
\|v_{kh}\|_{L^\infty(I; L^\infty(\Omega))} &= \max_{1 \leq m \leq M} \|v_{kh,m}\|_{L^\infty(\Omega)} \\
&\leq C \|g\|_{L^\infty(I; L^\infty(\partial\Omega))} + C \max_{1 \leq m \leq M} \sum_{l=1}^m \frac{\tau_l}{t_m - t_{l-1}} \|G_{kh,l}\|_{L^\infty(\Omega)} \\
&\leq C \|g\|_{L^\infty(I; L^\infty(\partial\Omega))} + C \max_{1 \leq l \leq M} \|G_{kh,l}\|_{L^\infty(\Omega)} \max_{1 \leq m \leq M} \sum_{l=1}^m \frac{\tau_l}{t_m - t_{l-1}} \\
&\leq C \ln \frac{T}{\tau} \|g\|_{L^\infty(I; L^\infty(\partial\Omega))},
\end{aligned}$$

where in the last step we used that

$$\max_{1 \leq l \leq M} \|G_{kh,l}\|_{L^\infty(\Omega)} \leq C \|g\|_{L^\infty(I; L^\infty(\partial\Omega))}$$

and

$$\sum_{l=1}^m \frac{\tau_l}{t_m - t_{l-1}} \leq 1 + \int_0^{t_{m-1}} \frac{dt}{t_m - t} = 1 + \ln \frac{t_m}{\tau_m} \leq C \ln \frac{T}{\tau}, \quad (\text{A.11})$$

by using the assumption $\tau_{\min} \geq C\tau^\beta$ and $\tau \leq \frac{T}{4}$. Finally, using the triangle inequality, it follows that

$$\|u_{kh}\|_{L^\infty(I;L^\infty(\Omega))} \leq \|v_{kh}\|_{L^\infty(I;L^\infty(\Omega))} + \|P_h G_{kh}\|_{L^\infty(I;L^\infty(\Omega))} \leq C \ln \frac{T}{\tau} \|g\|_{L^\infty(I;L^\infty(\partial\Omega))}.$$

Thus we obtain the result. \square

References

1. W. Arendt, C. J. K. Batty, M. Hieber, and F. Neubrander. *Vector-valued Laplace Transforms and Cauchy Problems*. Monographs in Mathematics. Springer, Basel, Switzerland, 2 edition, Apr. 2011.
2. J. Brandts, S. Korotov, M. Křížek, and J. Olc. On nonobtuse simplicial partitions. *SIAM Rev.*, 51(2):317–335, 2009.
3. P. Chatzipantelidis, Z. Horváth, and V. Thomée. On preservation of positivity in some finite element methods for the heat equation. *Comput. Methods Appl. Math.*, 15(4):417–437, 2015.
4. M. Dauge. Neumann and mixed problems on curvilinear polyhedra. *Integral Equations and Operator Theory*, 15:227–261, 1992.
5. E. B. Davies. *Heat Kernels and Spectral Theory*. Number 92. Cambridge university press, 1989.
6. E. B. Davies. Non-Gaussian aspects of heat kernel behaviour. *J. London Math. Soc.*, 55(1):105–125, 1997.
7. H. Fujii. Some remarks on finite element analysis of time-dependent field problems. In *Theory and Practice in Finite Element Structural Analysis*, pages 91–106. University of Tokyo Press, Tokyo, 1973.
8. M. Geissert. Discrete maximal lp regularity for finite element operators. *SIAM J. Numer. Anal.*, 44(2):677–698, 2006.
9. M. Gunzburger, B. Li, and J. Wang. Convergence of finite element solutions of stochastic partial integro-differential equations driven by white noise. *Numer. Math.*, 141(4):1043–1077, 2019.
10. M. Gunzburger, B. Li, and J. Wang. Sharp convergence rates of time discretization for stochastic time-fractional pdes subject to additive space-time white noise. *Math. Comp.*, 88(318):1715–1741, 2019.
11. W. Höhn and H.-D. Mittelmann. Some remarks on the discrete maximum-principle for finite elements of higher order. *Computing*, 27(2):145–154, 1981.
12. B. Jin and Z. Zhou. *Numerical Treatment and Analysis of Time-Fractional Evolution Equations*, volume 214. Springer Nature, 2023.
13. T. Kemmochi and N. Saito. Discrete maximal regularity and the finite element method for parabolic equations. *Numer. Math.*, 138:905–937, 2018.
14. S. Korotov and M. Křížek. Acute type refinements of tetrahedral partitions of polyhedral domains. *SIAM J. Numer. Anal.*, 39(2):724–733, 2001.
15. S. Korotov, M. Křížek, and P. Neittaanmäki. Weakened acute type condition for tetrahedral triangulations and the discrete maximum principle. *Math. Comp.*, 70(233):107–119, 2001.
16. B. Kovács, B. Li, and C. Lubich. A-stable time discretizations preserve maximal parabolic regularity. *SIAM J. Numer. Anal.*, 54(6):3600–3624, 2016.
17. P. C. Kunstmann and L. Weis. Maximal L^p -regularity for parabolic equations, Fourier multiplier theorems and-functional calculus. In *Functional analytic methods for evolution equations*, pages 65–311. Springer, 2004.
18. D. Leykekhman and B. Li. Maximum-norm stability of the finite element Ritz projection under mixed boundary conditions. *Calcolo*, 54(2):541–565, 2017.
19. D. Leykekhman and B. Li. Weak discrete maximum principle of finite element methods in convex polyhedra. *Math. Comp.*, 90(327):1–18, 2021.
20. D. Leykekhman and B. Vexler. Discrete maximal parabolic regularity for galerkin finite element methods. *Numer. Math.*, 135(3):923–952, 2017.
21. B. Li. Analyticity, maximal regularity and maximum-norm stability of semi-discrete finite element solutions of parabolic equations in nonconvex polyhedra. *Math. Comp.*, 88(315):1–44, 2019.
22. B. Li. Maximal regularity of multistep fully discrete finite element methods for parabolic equations. *IMA J. Numer. Anal.*, 42(2):1700–1734, 2022.
23. B. Li. Maximum-norm stability of the finite element method for the Neumann problem in nonconvex polygons with locally refined mesh. *Math. Comp.*, 91(336):1533–1585, 2022.

24. G. M. Lieberman. *Second Order Parabolic Differential Equations*. World Scientific, 1996.
25. C. Lubich. Convolution quadrature and discretized operational calculus. I. *Numer. Math.*, 52(2):129–145, 1988.
26. A. H. Schatz. A weak discrete maximum principle and stability of the finite element method in L_∞ on plane polygonal domains. I. *Math. Comp.*, 34(149):77–91, 1980.
27. 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(11-12):1349–1385, 1998.
28. A. H. Schatz and L. B. Wahlbin. Interior maximum-norm estimates for finite element methods. II. *Math. Comp.*, 64(211):907–928, 1995.
29. V. Thomée. *Galerkin Finite Element Methods for Parabolic Problems*, volume 25. Springer Science & Business Media, 2007.
30. 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.
31. V. Thomée and L. B. Wahlbin. On the existence of maximum principles in parabolic finite element equations. *Math. Comp.*, 77(261):11–19, 2008.
32. J. Wang and R. Zhang. Maximum principles for P1-conforming finite element approximations of quasi-linear second order elliptic equations. *SIAM J. Numer. Anal.*, 50(2):626–642, 2012.
33. G. Wanner and E. Hairer. *Solving Ordinary Differential Equations II*, volume 375. Springer Berlin Heidelberg New York, 1996.
34. J. Xu and L. Zikatanov. A monotone finite element scheme for convection-diffusion equations. *Math. Comp.*, 68(228):1429–1446, 1999.