A mass-lumping finite element method for radially symmetric solution of a multidimensional semilinear heat equation with blow-up

Toru Nakanishi† and Norikazu Saito

Graduate School of Mathematical Sciences, The University of Tokyo, Tokyo, Japan

ABSTRACT
This study presents a non-standard mass-lumping finite element method for computing the radially symmetric solution of a semilinear heat equation in $N$ dimensional ball ($N \geq 2$). We provide two schemes, (ML–1) and (ML–2), and derive their error estimates through the discrete maximum principle. In the weighted $L^2$ norm, the convergence of (ML–1) was at the optimal order but that of (ML–2) was only at sub-optimal order. Nevertheless, scheme (ML–2) reproduces a blow-up of the solution of the original equation. In fact, in scheme (ML–2), we could accurately approximate the blow-up time. Our theoretical results were validated in numerical experiments.

ARTICLE HISTORY
Received 11 December 2020
Revised 9 November 2021
Accepted 3 December 2021

KEYWORDS
Finite element method; blow-up; radially symmetric solution

2010 AMS SUBJECT CLASSIFICATIONS
65M60; 35K58

1. Introduction

This paper applies the finite element method (FEM) to a semilinear parabolic equation with a singular convection term:

\begin{align}
    u_t &= u_{xx} + \frac{N-1}{x} u_x + f(u), & x \in I &= (0,1), \ t > 0, \\
    u_x(0,t) &= u(1,t) = 0, & t > 0, \\
    u(x,0) &= u^0(x), & x \in I.
\end{align}

Here, $u = u(x,t)$, $x \in \bar{I} = [0,1]$, $t \geq 0$ denotes the function to be found, $f$ is a given locally Lipschitz continuous function, and $u^0$ is a given continuous function. Throughout this paper, we assume that

\begin{equation}
    N \text{ is an integer } \geq 2.
\end{equation}

To compute the blow-up solution of (1), we apply Nakagawa’s time-increment control strategy (see [30] and Section 6 of the present paper), a powerful technique for approximating blow-up times. As recalled below, the standard finite element approximation is unuseful for achieving this purpose. We thus propose a non-standard mass-lumping finite element approximation, prove its convergence, and apply it to a blow-up analysis.

We first clarify the motivation of this study. In many engineering problems, the spatial dimension of a mathematical model is at most three. Solving partial differential equations (PDEs) in more than three spatial dimensions is usually motivated by mathematical interests. Mathematicians understand
that solving problems in a general setting can reveal the hidden natures of PDEs. One successful result is the discovery of Fujita’s blow-up exponent for the semilinear heat equation of \( U = U(x, t) \) given as

\[
U_t = \Delta U + f(U) \quad (x \in \mathbb{R}^N, \ t > 0),
\]

where \( N \) and \( f(U) \) are defined above. Assuming \( f(U) = U|U|^\alpha \) with \( \alpha > 0 \), Fujita showed that any positive solution blows up in finite time if \( 1 + \alpha < 1 + 2/N \), but a solution remains smooth at any time if the initial value is small and \( 1 + \alpha > 1 + 2/N \). The quantity \( p_c = 1 + 2/N \) is known as Fujita’s critical exponent, and Equation (3) is called Fujita’s equation. Since Fujita’s work, a huge number of studies have been devoted to critical phenomena in nonlinear PDEs of several kinds (see [14,24,29,35] for details). The knowledge gained by these studies has been applied to problems with spatial dimensions of three or fewer. However, many problems related to stochastic analysis are formulated as higher-dimensional PDEs. These problems have attracted much interest but are beyond the scope of the present study.

Non-stationary problems in four-dimensional space are difficult to solve by numerical methods, even on modern computers. Consequently, numerical analyses of the blow-up solutions of nonlinear PDEs have been restricted to two-dimensional space (see for example [1,3–5,9–11,17,20–23,25,31,33]). Although Nakagawa’s time-increment control strategy is applied to various nonlinear PDEs including the nonlinear heat, wave and Schrödinger equations, these equations are considered only in the one-space dimension; see [6,8,9,13,23,36,37]. We know two notable exceptions; the one is [31] where the FEM to a semilinear heat equation in a two-dimensional polygonal domain was considered, and the other is [7] where the finite difference method to the radially symmetric solution of the semilinear heat equation in \( N \) dimensional ball was studied.

Following [7], the present paper investigates radially symmetric solutions to Equation (3). Assuming radial symmetry of the solution and the given data, the \( N \)-dimensional equation reduces to a one-dimensional equation. More specifically, considering (3) in \( N \)-dimensional unit ball \( B = \{x \in \mathbb{R}^N \mid |x| < 1\} \) with the homogeneous Dirichlet boundary condition on the boundary and assuming \( U \) is expressed as \( u(x) = U(x) \) for \( x \in B \) and \( x = |x|\mathbb{R}^N \), we came to consider the problem (1).

After completing the present work, we learned that Cho and Okamoto [12] extended the work in [7]. The time dimension was discretized by the semi-implicit Euler method in [7], but Cho and Okamoto [12] explored the explicit scheme, then proved optimal-order convergence with Nakagawa’s strategy. Because their schemes use special approximations around the origin to maintain some analytical properties of the solution, they should be performed on a uniform spatial mesh. Conversely, when seeking the blow up solution, non-uniform partitions of the space variable are useful for examining highly concentrated solutions at the origin. For this purpose, we developed the FEM scheme.

FEM analyses of the linear case, in which \( f(u) = 0 \) in Equation (1a) is replaced by a given function \( f(x, t) \), are not new. Eriksson and Thomée [16] and Thomée [39] studied the convergence property of the elliptic equation and proposed two schemes: the symmetric scheme, in which the optimal-order error is estimated in the weighted \( L^2 \) norm, and the nonsymmetric scheme, in which the \( L^\infty \) error is estimated. However, their finite element schemes are not easily adaptable to the semilinear heat equation (1a), as reported in our earlier study [32]. Our earlier results are briefly summarized as follows:

- If \( f \) is globally Lipschitz continuous, the solution of the symmetric scheme converges to the solution of (1) in the weighted \( L^2 \) norm in space and in the \( L^\infty \) norm in time. Moreover, the convergence is at the optimal order (see Theorem 4.1 in [32]).
- If \( f \) is locally Lipschitz continuous and \( N \leq 3 \), the same conclusion holds (see Theorem 4.3 in [32]). However, if \( N \geq 4 \), the convergence properties are not guaranteed. For this reason, interest in radially symmetric problems has diminished.
If \( f(u) = u|u|^\alpha \) with \( \alpha \geq 1 \) and the time partition is uniform, the solution of the non-symmetric scheme converges to the solution of (1) in the \( L^\infty(0, T; L^\infty(I)) \) norm. Optimal-order convergence holds up to the logarithmic factor (see Theorem 4.6 in [32]). Nakagawa’s time-increment control strategy is difficult to apply in such cases.

As the non-symmetric scheme seems to be incompatible with Nakagawa’s time-increment control strategy, we pose the following question: Can the restriction \( N \leq 3 \) be removed from the symmetric scheme? In fact, this restriction is imposed by the inverse inequality Lemma 4.8 in [32] and the necessity of finding the boundedness of the finite element solution (see the proof of Theorem 4.3 in [32]). To surmount this difficulty, the \( L^\infty \) estimates for the FEM can be directly derived using the discrete maximum principle (DMP). As the DMP is based largely on the nonnegativity of the finite element solution, the time derivative term should be approximated by the mass-lumping approximation. Unfortunately, we tried but failed to prove the convergence property of the finite element solution by this approximation (see (9)). Therefore, we propose a non-standard mass-lumping approximation (10) in this paper. It should be noticed that a similar mass-lumping approximation to ours was proposed in [34], where parabolic problems with degenerate coefficients in two-dimensional spatial domain was studied. Using our mass-lumping approximation, we prove the DMP and the convergence property of the finite element solution, and perform the blow up analysis for any \( N \geq 2 \).

Our typical results are summarized further. Here, our schemes are denoted as (ML–1) and (ML–2).

- The solution of (ML–1) is non-negative if \( f \) and \( u^0 \) satisfy some conditions (Theorem 2.1). Furthermore, if the time increment satisfies condition (13), the solution of (ML–2) is also non-negative (Theorem 2.1). Theorem 2.2 gives a useful sufficient condition of (13).
- The solution of (ML–1) converges to the solution of (1) in the weighted \( L^2 \) norm in space and in the \( L^\infty \) norm in time. Moreover, the convergence is at the optimal order (Theorem 2.3). The proof is based on a sub-optimal estimate in the \( L^\infty(0, T; L^\infty(I)) \) norm (Theorem 2.4).
- If condition (13) is satisfied, then the solution of (ML–2) converges to the solution of (1) in the \( L^\infty(0, T; L^\infty(I)) \) norm (Theorem 2.5). Unfortunately, the order of the convergence is sub-optimal.
- The solution of (ML–2) reproduces the blow up property of the solution of (1) (Theorems 5.1 and 5.2).

This paper comprises six sections and an Appendix. Section 2 presents our finite element schemes and the convergence theorems (Theorems 2.3–2.5). After describing our preliminary results in Section 3, we prove our convergence theorems in Section 4. Section 5 reports the results of our blow-up analysis, and Section 6 validates our theoretical results with numerical examples. Appendix presents the proofs of some auxiliary results on the eigenvalue problems.

### 2. The schemes and their convergence results

Throughout this paper, \( f \) is assumed as a locally Lipschitz continuous function of \( \mathbb{R} \to \mathbb{R} \).

For some arbitrary \( \chi \in \dot{H}^1 = \{ v \in H^1(I) \mid v(1) = 0 \} \), we multiply both sides of (1a) by \( x^{N-1} \chi \) and integrate by parts over \( I \). We thus obtain

\[
\int_I x^{N-1} u_t \chi \, dx + \int_I x^{N-1} u_x \chi_x \, dx = \int_I x^{N-1} f(u) \chi \, dx. \tag{4}
\]

Therefore, a weak formulation of (1a) is stated as follows. For \( t > 0 \), find \( u(\cdot, t) \in \dot{H}^1 \) such that

\[
(u_t, \chi) + A(u, \chi) = (f(u), \chi) \quad (\forall \chi \in \dot{H}^1), \tag{5}
\]

where

\[
(w, v) = \int_I x^{N-1} w v \, dx, \quad A(w, v) = \int_I x^{N-1} w_x v_x \, dx. \tag{6}
\]
We now introduce the FEM. For a positive integer $m$, we introduce node points

$$0 = x_0 < x_1 < \cdots < x_{j-1} < x_j < \cdots < x_{m-1} < x_m = 1,$$

and set $I_j = (x_{j-1}, x_j)$ and $h_j = x_j - x_{j-1}$, where $j = 1, \ldots, m$. The granularity parameter is defined as $h = \max_{1 \leq j \leq m} h_j$. Let $P_k(f)$ be the set of all polynomials in an interval $f$ of degree $\leq k$. We use the quasi-uniformity condition on the partition $\{x_i\}_{i=0}^m$ of $\mathcal{T} = [0,1];$

$$h \leq \beta \min_{1 \leq j \leq m} h_j,$$  (7)

where $\beta$ is a positive constant independent of $h$. The $P_1$ finite element space is defined as

$$S_h = \{ v \in H^1(I) \mid v \in P_1(I_j) (j = 1, \ldots, m), v(1) = 0 \}. $$  (8)

The standard basis function $\phi_j \in S_h, j = 0, 1, \ldots, m - 1$ is defined as

$$\phi_j(x_i) = \delta_{ij},$$

where $\delta_{ij}$ denotes Kronecker’s delta. We note that $S_h \subset \mathring{H}^1$ and that any function of $\mathring{H}^1$ is identified with a continuous function. The Lagrange interpolation operator $\Pi_h w = \sum_{j=0}^{m-1} w(x_j) \phi_j$ for $w \in \mathring{H}^1$.

The mass-lumping approximation of the weighted $L^2$ norm $(\cdot, \cdot)$ can be naturally defined as

$$(w, v) \approx w(x_0)v(x_0) \int_0^{x_1/2} x^{N-1} \, dx + \sum_{i=1}^{m-1} w(x_i)v(x_i) \int_{x_{i-1/2}}^{x_{i+1/2}} x^{N-1} \, dx,$$  (9)

where $x_{i-1/2} = (x_i + x_{i-1})/2$. As mentioned in the Introduction, this standard formulation is useless for our purpose. Instead, we define

$$\langle w, v \rangle = \sum_{i=0}^{m-1} w(x_i)v(x_i) (1, \phi_i) \quad (w, v \in \mathring{H}^1).$$  (10)

This definition leads to the following result, which can be verified by direct calculation.

**Lemma 2.1:** We have $\langle 1, w \rangle = (1, \Pi_h w)$ for any $w \in \mathring{H}^1$.

The associated norms with $(\cdot, \cdot)$ and $\langle \cdot, \cdot \rangle$ are, respectively, given by

$$\|v\| = (v, v)^{1/2} \quad \text{and} \quad |||v||| = \langle v, v \rangle^{1/2}.$$

These norms are equivalent in $S_h$ (see Lemma 3.4 below).

The time discretization is non-uniformly partitioned as

$$t_0 = 0, \quad t_n = \sum_{j=0}^{n-1} \tau_j \quad (n \geq 1),$$

where $\tau_j > 0$ denotes the time increment. Furthermore, we set

$$\tau = \sup_{j \geq 0} \tau_j.$$

In general, we write $\partial_{\tau_n} u_h^{n+1} = (u_h^{n+1} - u_h^n) / \tau_n$. 
The finite element schemes are then stated as follows. 

\[(ML-1)\] Find \(u_h^{n+1} \in S_h, n = 0, 1, \ldots\), such that

\[
\{\partial_t u_h^{n+1}, \chi\} + A(u_h^{n+1}, \chi) = (f(u_h^n), \chi) \quad (\chi \in S_h),
\]

where \(u_h^0 \in S_h\) is assumed to be given.

\[(ML-2)\] Find \(u_h^{n+1} \in S_h, n = 0, 1, \ldots\), such that

\[
\{\partial_t u_h^{n+1}, \chi\} + A(u_h^n, \chi) = (f(u_h^n), \chi) \quad (\chi \in S_h).
\]

Below, we will show the optimal-order error estimate in the weighted \(L^2\) norm for the solution of (ML–1). On the other hand, we are able to show only a sub-optimal-order error estimate in the \(L^\infty\) norm for the solution of (ML–2). Nevertheless, we consider (ML–2) because it is suitable for the blow-up analysis (see Section 5).

We also summarize the well-posedness of our schemes. The proof is omitted because it is identical to Theorems 3.1 and 3.2 in [32].

**Theorem 2.1:** Suppose that \(n \geq 0\) and \(u_h^n \in S_h\) are given.

(i) Schemes (ML–1) and (ML–2) admit unique solutions \(u_h^{n+1} \in S_h\).

(ii) In addition to the basic assumption on \(f\), assume that \(f\) is a non-decreasing function with \(f(0) \geq 0\). If \(u_h^n \geq 0\), then the solution \(u_h^{n+1}\) of (ML–1) satisfies \(u_h^{n+1} \geq 0\).

(iii) Under the assumptions of (ii) above, further assume that

\[
\tau \leq \min_{0 \leq i \leq m-1} \frac{(1, \phi_i)}{A(\phi_i, \phi_i)}.
\]

Then the solution \(u_h^{n+1}\) of (ML–2) satisfies \(u_h^{n+1} \geq 0\).

The following theorem provides a useful sufficient condition under which (13) holds.

**Theorem 2.2:** Inequality (13) holds if

\[
\tau \leq \frac{\beta^2}{N+1} h^2.
\]

**Proof:** It suffices to prove

\[
\frac{(1, \phi_i)}{A(\phi_i, \phi_i)} \geq \frac{1}{N+1} \min_{1 \leq i \leq m} h_i^2 \quad (0 \leq i \leq m-1).
\]

In fact, this, together with (7), implies

\[
\min_{0 \leq i \leq m-1} \frac{(1, \phi_i)}{A(\phi_i, \phi_i)} \geq \frac{\beta^2}{N+1} h^2.
\]

Therefore, (14) follows from (13). First, a direct calculation gives

\[
\frac{(1, \phi_0)}{A(\phi_0, \phi_0)} = \frac{1}{N+1} h_1^2.
\]

Next we consider the case of \(i \neq 0\). We write

\[
(1, \phi_i) = \int_{x_{i-1}}^{x_{i+1}} x^{N-1} \phi_i(x) \, dx = A_1 + A_2,
\]
\[ A(\phi_i, \phi_i) = \int_{x_{i-1}}^{x_{i+1}} x^{N-1} \phi_i'(x) \phi_i(x) \, dx = B_1 + B_2, \]

where

\[ A_1 = \int_{x_{i-1}}^{x_i} x^{N-1} \frac{1}{h_i} (x - x_{i-1}) \, dx, \quad A_2 = \int_{x_i}^{x_{i+1}} x^{N-1} \frac{1}{h_{i+1}} (x_{i+1} - x) \, dx, \]

\[ B_1 = \int_{x_{i-1}}^{x_i} x^{N-1} \frac{1}{h_i^2} \, dx, \quad B_2 = \int_{x_i}^{x_{i+1}} x^{N-1} \frac{1}{h_{i+1}^2} \, dx. \]

Then we have

\[ A_1 = \frac{1}{N+1} h_i^N + \frac{N-1}{N} x_{i-1} h_i^{N-1} + \frac{N-2}{2} x_{i-1}^2 h_i^{N-2} \]

\[ + \sum_{j=3}^{N} \frac{N-j}{j} \frac{(N-1)(N-2) \cdots (N-j+2)}{(j-1)!} x_{i-1}^j h_i^{N-j}, \]

\[ A_2 = \frac{1}{N(N+1)} h_{i+1}^N + \frac{1}{N} x_i h_{i+1}^{N-1} + \frac{1}{2} x_i^2 h_{i+1}^{N-2} \]

\[ + \sum_{j=3}^{N} \frac{(N-1)(N-2) \cdots (N-j+2)}{j!} x_i^j h_{i+1}^{N-j}, \]

\[ B_1 = \frac{1}{N} \sum_{j=0}^{N-1} \binom{N}{j} x_{i-1}^j h_i^{N-j-2}, \]

\[ B_2 = \frac{1}{N} \sum_{j=0}^{N-1} \binom{N}{j} x_i^j h_{i+1}^{N-j-2}. \]

By comparing the coefficients of \( x_{i-1}^j h_i^{N-j} \) and \( x_i^j h_{i+1}^{N-j} \) (0 \( \leq j \leq N - 1 \)), we have

\[ \frac{A_1}{B_1} \geq \min_{3 \leq j \leq N-1} \left( \frac{N}{N+1}, \frac{N-1}{N}, \frac{N-2}{N-1}, \frac{N-j}{N-j+1} \right) h_i^2 = \frac{1}{2} h_i^2, \]

\[ \frac{A_2}{B_2} \geq \min_{3 \leq j \leq N-1} \left( \frac{1}{N+1}, \frac{1}{N}, \frac{1}{N-1}, \frac{1}{N-j+1} \right) h_{i+1}^2 = \frac{1}{N+1} h_{i+1}^2. \]

Therefore,

\[ \frac{(1, \phi_i)}{A(\phi_i, \phi_i)} \geq \frac{1}{N+1} \min(h_i^2, h_{i+1}^2). \]

Therefore, (15) is proved. 

We now proceed to the convergence analysis. Our results for (ML–1) and (ML–2) assume a smooth solution \( u \) of (1): given \( T > 0 \) and setting \( Q_T = [0, 1] \times [0, T] \), we assume that \( u \) is sufficiently smooth such that

\[ \kappa(u) = \sum_{k=0}^{2} ||\partial_k^k u||_{L^\infty(Q_T)} + \sum_{l=1}^{2} ||\partial_l^l u||_{L^\infty(Q_T)} + \sum_{k=1}^{2} ||\partial_k \partial_k^k u||_{L^\infty(Q_T)} < \infty. \]  

(16)

Here, we have used the conventional \( ||v||_{L^\infty(\omega)} = \max_{\omega} |v| \) for a continuous function \( v \) defined in a bounded set \( \omega \) in \( \mathbb{R}^p, p \geq 1. \)
Moreover, the approximate initial value $u_h^0$ is chosen as
\[ \| u_h^0 - u^0 \|_{L^\infty(I)} \leq C_0 h^2 \] (17)
for a positive constant $C_0$.

We now express positive constants $C = C(\gamma_1, \gamma_2, \ldots)$ depending only on the parameters $\gamma_1, \gamma_2, \ldots$. Particularly, $C$ is independent of $h$ and $\tau$. The following Theorem 2.3, 2.4 and 2.5 are all valid for a locally Lipschitz continuous $f$. The monotonicity assumption on $f$ (see Theorem 2.1 (ii)) will be used only in Section 5.

**Theorem 2.3 (Optimal $L^2$ error estimate for (ML-1))**: Assume that, for $T > 0$, the solution $u$ of (1) is sufficiently smooth that (16) holds. Moreover, assume that (7) and (17) are satisfied. Then, for sufficiently small $h$ and $\tau$, we have
\[ \sup_{0 \leq t_n \leq T} \| u_h^n - u(\cdot, t_n) \| \leq C(h^2 + \tau), \] (18)
where $C = C(T, f, \kappa(u), C_0, N, \beta)$ and $u_h^n$ is the solution of (ML-1).

The following result, which is worth a separate mention, gives only a sub-optimal error estimate but is useful for proving Theorem 2.3.

**Theorem 2.4 (Sub-optimal $L^\infty$ error estimate for (ML-1))**: Under the assumptions of Theorem 2.3 and for sufficiently small $h$ and $\tau$, we have
\[ \sup_{0 \leq t_n \leq T} \| u_h^n - u(\cdot, t_n) \|_{L^\infty(I)} \leq C(h + \tau), \] (19)
where $C = C(T, f, \kappa(u), C_0, N, \beta)$ and $u_h^n$ is the solution of (ML-1).

**Theorem 2.5 (Sub-optimal $L^\infty$ error estimate for (ML-2))**: Also under the assumptions of Theorem 2.3, assume that (13) is satisfied. Then, for sufficiently small $h$ and $\tau$, we have
\[ \sup_{0 \leq t_n \leq T} \| u_h^n - u(\cdot, t_n) \|_{L^\infty(I)} \leq C(h + \tau), \] (20)
where $C = C(T, f, \kappa(u), C_0, N, \beta)$ and $u_h^n$ is the solution of (ML-2).

**Remark 2.1**: Other schemes based on the mass-lumping $\langle \cdot, \cdot \rangle$ are possible. For example, the schemes
\[ \langle \partial_t u_h^{n+1}, \chi \rangle + A(u_h^{n+1}, \chi) = \langle f(u_h^n), \chi \rangle \quad (\chi \in S_h) \] (21)
and
\[ \langle \partial_t u_h^{n+1}, \chi \rangle + A(u_h^n, \chi) = \langle f(u_h^n), \chi \rangle \quad (\chi \in S_h) \] (22)
have very similar properties to those of (ML-1) and (ML-2). We omit the details because the modifications are easily performed.

### 3. Preliminaries

This section gives some preliminary results of the theorem proofs. The quasi-uniformity condition (7) is always assumed.
For some \( w \in \dot{H}^1 \), the projection operator \( P_\Delta \) of \( \dot{H}^1 \to S_h \) associated with \( A(\cdot,\cdot) \) is defined as

\[
P_\Delta w \in S_h, \quad A(P_\Delta w - w, \chi) = 0 \quad (\chi \in S_h).
\]  

(23)

The following error estimates are proved in [16,26].

**Lemma 3.1:** Letting \( w \in C^2(\overline{I}) \cap \dot{H}^1 \), and \( h \) be sufficient small, we obtain

\[
\| P_\Delta w - w \| \leq C h^2 \| w_{xx} \|,
\]

(24a)

\[
\| P_\Delta w - w \|_{L^\infty(I)} \leq C \left( \log \frac{1}{h} \right) h^2 \| w_{xx} \|_{L^\infty(I)},
\]

(24b)

where \( C = (\beta, N) > 0 \).

**Lemma 3.2 (Inverse estimate):** There exists a constant \( C = C(\beta, N) > 0 \) such that

\[
\| w_x \| \leq C h^{-1} \| w \| \quad (w \in S_h).
\]

**Proof:** The proof is identical to that of the standard inverse estimate. \( \blacksquare \)

**Lemma 3.3:** Let \( w \in C(I) \) be a piecewise quadratic function in \( I \), that is, \( w|_{I_j} \in P_2(I_j) \) \((j = 1, \ldots, m)\). Then, we have

\[
\int_I x^{N-1} |\Pi_h w - w| \, dx \leq C h^2 \| x^{N-1} w_{xx} \|_{L^1(I)},
\]

(25)

where \( C = C(\beta, N) > 0 \).

**Proof:** To prove (25), it suffices to replace \( I = I_j \) with \( j = 1, \ldots, m \). First, let \( j \geq 2 \). By Taylor’s theorem, we can write

\[
|\Pi_h w(x) - w(x)| \leq Ch_j \int_{I_j} |w_{xx}(\xi)| \, d\xi \quad (x \in I_j).
\]

(26)

Referring to (7), we see that

\[
\frac{x}{x_j-1} = 1 + \frac{x - x_j-1}{x_j-1} \leq 1 + \frac{h_j}{h_{j-1}} \leq 1 + \beta \quad (x \in \overline{I}_j).
\]

(27)

Combining (26) and (27), we deduce that

\[
x^{N-1} |\Pi_h w(x) - w(x)| \leq C(1 + \beta)^{N-1} h_j \int_{I_j} x^{N-1} |w_{xx}(\xi)| \, d\xi \quad (x \in I_j).
\]

Integrating both sides, we obtain (25) for \( I = I_j \) and \( j \geq 2 \). We now proceed to the case \( m = 1 \). Setting \( w(x) = ax^2 + bx + c \) for \( x \in I_1 \), where \( a, b \) and \( c \) are constants, we express \( \Pi_h w(x) - w(x) = ax(h_1 - x) \) for \( x \in I_1 \). Therefore, we directly obtain

\[
\int_{I_1} x^{N-1} |\Pi_h w(x) - w(x)| \, dx = \frac{|a|}{(N+1)(N+2)} h_1^{N+2},
\]

\[
h_1^2 \| x^{N-1} w_{xx} \|_{L^1(I_1)} = \frac{2}{N} |a| h_1^{N+2},
\]

which implies (25) for \( I = I_1 \). \( \blacksquare \)
Lemma 3.4: There exist constants $C = C(\beta, N)$ and $C' = C'(\beta, N)$ such that

$$C' \| w \| \leq \| w \| \leq C \| w \| \quad (w \in S_h).$$

Proof: Let $w \in S_h$. To prove the first inequality, we note that $w^2$ is a piecewise quadratic function and $(w^2)_{xx} = 2(w_x)^2$. By Lemmas 3.3 and 3.2, we get

$$| \langle w, w \rangle - \langle w, w \rangle | \leq Ch^2 \| x^{N-1} (w^2)_{xx} \|_{L^1(I)} \leq Ch^2 \| w_x \|^2 \leq C \| w \|^2,$$

which implies the first inequality.

To prove the second inequality, we estimate $\| w \|$ as

$$\| w \|^2 \leq \sum_{j=1}^{m} \left[ w(x_{j-1})^2 \int_{x_{j-1}}^{x_j} x^{N-1} \, dx + w(x_j)^2 \int_{x_j}^{x_{j+1}} x^{N-1} \, dx \right].$$

We also express $\| w \|$ as

$$\| w \|^2 = w(x_0)^2 \int_{x_0}^{x_1} x^{N-1} \phi_0(x) \, dx + \sum_{j=1}^{m-1} w(x_j)^2 \int_{x_{j-1}}^{x_{j+1}} x^{N-1} \phi_j(x) \, dx.$$

Therefore, it suffices to show that

$$\int_{x_0}^{x_1} x^{N-1} \, dx \leq C_1 \int_{x_0}^{x_1} x^{N-1} \phi_0 \, dx,$$

$$\int_{x_{j-1}}^{x_{j+1}} x^{N-1} \, dx \leq C_2 \int_{x_{j-1}}^{x_{j+1}} x^{N-1} \phi_j \, dx \quad (j = 1, \ldots, m - 1),$$

where $C_1 = C_1(N) > 0$ and $C_2 = C_2(N) > 0$.

Equation (28b) is directly verified using (27). Equation (28a) is obtained by the change-of-variables technique, setting $\xi = x/h_1$. 

We here introduce two auxiliary problems. Given $n \geq 0$, $g^n_h \in S_h$ and $u^n_h \in S_h$, we seek $u^{n+1}_h \in S_h$ such that

$$\{ \partial_{x^n} u^{n+1}_h, \chi \} + A(u^{n+1}_h, \chi) = \{ g^n_h, \chi \} \quad (\chi \in S_h),$$

and

$$\{ \partial_{x^n} u^{n+1}_h, \chi \} + A(u^n_h, \chi) = \{ g^n_h, \chi \} \quad (\chi \in S_h).$$

Lemma 3.5: Suppose that $n \geq 0$ and $u^n_h, g^n_h \in S_h$ are given. Then, problem (29) admits a unique solution $u^{n+1}_h \in S_h$ and it satisfies

$$\| u^{n+1}_h \|_{L^\infty(I)} \leq \| u^n_h \|_{L^\infty(I)} + \tau_n \| g^n_h \|_{L^\infty(I)}.$$

Problem (30) also admits a unique solution $u^{n+1}_h \in S_h$ that satisfies (31) under condition (13).
Proof: The unique existence of the solution of (29) can be verified by a standard approach (see Theorems 3.1, 3.2 in [32]). Substituting $\chi = \phi_i$, $i = 0, \ldots, m - 1$, in (29), we have

$$\frac{\tau_n a_{i,j} - 1}{m_i} u^{n+1}_{i-1} + \left(1 + \frac{\tau_n a_{i,i}}{m_i}\right) u^{n+1}_i + \frac{\tau_n a_{i,i} + 1}{m_i} u^{n+1}_{i+1} = u^n_i + \tau_n g^n_i,$$

where $u^n_i = u^n_h(x_i), g^n_i = g^n_h(x_i), a_{i,j} = A(\phi_j, \phi_i)$ and $m_i = (\chi), (\nu, \phi_i)$. Therein, we should understand that $a_{0,-1} = 0, m_0 = 1$ and $u^n_{-1} = 1$. Moreover, substituting $\chi = \phi_i$ in (30), we get

$$u^{n+1}_i = -\frac{\tau_n a_{i,i} - 1}{m_i} u^{n+1}_{i-1} + \left(1 - \frac{\tau_n a_{i,i}}{m_i}\right) u^n_i - \frac{\tau_n a_{i,i} + 1}{m_i} u^{n+1}_{i+1} + \tau_n g^n_i.$$

From these expressions, (31) is deduced by a standard argument.

4. Proofs of theorems 2.3, 2.4, and 2.5

Proof of Theorem 2.3 using Theorem 2.4: This proof is divided into the following two steps:

Step 1. We prove Theorem 2.3 under an additional assumption: $f$ is a globally Lipschitz function. That is, we assume

$$M = \sup_{s,s' \in \mathbb{R}, s \neq s'} \frac{|f(s) - f(s')|}{|s - s'|} < \infty.$$ \hspace{1cm} (32)

Using $P_A u$, we divide the error into the form

$$u^n_h - u(\cdot, t_n) = \frac{(u^n_h - P_A u(\cdot, t_n)) + (P_A u(\cdot, t_n) - u(\cdot, t_n))}{w^n} + w^n.$$ \hspace{1cm} (33)

From (24a), we know that

$$\|w^n\| \leq C h^2 \|u_{xx}(t_n)\| \leq C h^2 \|u_{xx}\|_{L^\infty(Q_T)}$$ \hspace{1cm} (34)

and that $\partial_{\tau_n} P_A v = P_A \partial_{\tau_n} v$ for $v \in C(\overline{T})$.

We now estimate $v^n_h$. Using the weak form (5) at $t = t_{n+1}$, scheme (ML–1), and the property of $P_A$, we obtain

$$\langle \partial_{\tau_n} v^{n+1}_h, \chi \rangle + A(v^{n+1}_h, \chi) = (I + II + III + IV + V)(\chi) \quad (\chi \in S_h),$$ \hspace{1cm} (35)

where

$$I(\chi) = (f(u^n_h), \chi) - (f(u(\cdot, t_n)), \chi),$$

$$II(\chi) = (u_h(t_{n+1}), \chi) - (\partial_{\tau_n} u(\cdot, t_{n+1}), \chi),$$

$$III(\chi) = (f(u(\cdot, t_n)), \chi) - (f(u(\cdot, t_{n+1})), \chi),$$

$$IV(\chi) = (\partial_{\tau_n} u(\cdot, t_{n+1}), \chi) - (P_A \partial_{\tau_n} u(\cdot, t_{n+1}), \chi) = (\partial_{\tau_n} w^{n+1}, \chi),$$

$$V(\chi) = (\partial_{\tau_n} P_A u(\cdot, t_{n+1}), \chi) - (\partial_{\tau_n} P_A u(\cdot, t_{n+1}), \chi).$$

The estimations of I–IV are straightforward. That is, we have

$$|I(\chi)| \leq M(\|w^n\| + \|v^n_h\|) \cdot \|\chi\|,$$

$$|II(\chi)| \leq \tau_n \|u_{tt}\|_{L^\infty(Q_T)} \|\chi\|,$$

$$|III(\chi)| \leq \tau_n M \|u_t\|_{L^\infty(Q_T)} \|\chi\|,$$

$$|IV(\chi)| \leq \tau_n \|u_t\|_{L^\infty(Q_T)} \|\chi\|.$$
To estimate $V$, we use Lemmas 2.1 and 3.3. Lemma 3.3 is applicable because $\partial_{\tau_n} P_A u(\cdot, t_{n+1}) \chi$ is a piecewise quadratic function. That is,

$$|IV(\chi)| \leq C h^2 \|v_{txx}\|_{L^\infty(Q_T)} \|\chi\|.$$ 

Substituting $\chi = v^{n+1}_h$ in (35) gives

$$\frac{1}{2\tau_n} \left( ||v^{n+1}_h||^2 - ||v^n_h||^2 \right) + \left( ||v^{n+1}_h||_x \right)^2 \leq C ||v^n_h|| \cdot ||v^{n+1}_h||_x + C(h^2 + \tau_n^2) \kappa(u) ||(v^{n+1}_h)_x||.$$ 

Herein, we have used Lemma 3.4 and the Poincaré inequality (Lemma 18.1 in [39]). By Young’s inequality, we then deduce that

$$\frac{1}{\tau_n} \left( ||v^{n+1}_h||^2 - ||v^n_h||^2 \right) \leq C ||v^n_h||^2 + C(h^2 + \tau_n^2) \kappa(u)^2.$$ 

Therefore,

$$||v^n_h||^2 \leq e^{CT} ||v^0_h||^2 + C(e^{CT} - 1)(h^2 + \tau_n^2) \kappa(u)^2,$$

which completes the proof.

Step 2. We now turn to the original assumption on $f$; let us assume that $f$ is a locally Lipschitz continuous function. Let $r = 1 + \|u\|_{L^\infty(Q_T)}$. Consider (1a) and (ML–1) with replacement $f(s)$ in:

$$\tilde{f}(s) = \begin{cases} f(r) & (s \geq r) \\ f(s) & (|s| \leq r) \\ f(-r) & (s \leq -r). \end{cases}$$

The function $\tilde{f}$ is a globally Lipschitz function satisfying

$$M = \sup_{s, s' \in \mathbb{R}, s \neq s'} \frac{\tilde{f}(s) - \tilde{f}(s')}{|s - s'|} = \sup_{|s|, |s'| \leq r, s \neq s'} \frac{|f(s) - f(s')|}{|s - s'|}.$$ 

Let $\tilde{u}$ and $\tilde{u}^n_h$ be the solutions of (1a) and (ML–1) with $\tilde{f}$, respectively. Applying Step 1 and Theorem 2.4 to $\tilde{u}$ and $\tilde{u}^n_h$, we obtain

$$\sup_{0 \leq t_n \leq T} \|\tilde{u}^n_h - \tilde{u}(\cdot, t_n)\| \leq C(h^2 + \tau), \quad (36a)$$

$$\sup_{0 \leq t_n \leq T} \|\tilde{u}^n_h - \tilde{u}(\cdot, t_n)\|_{L^\infty(I)} \leq C \left( h + h^2 \log \frac{1}{h} + \tau \right). \quad (36b)$$
By the definition of \( r \) and the uniqueness of the solution of (1a), we know that \( u = \tilde{u} \) in \( Q_T \). For sufficiently small \( h \) and \( \tau \), we have
\[
C \left( h + h^2 \log \frac{1}{h} + \tau \right) \leq 1.
\]
Consequently, \( \| \tilde{u}_h^n \|_{L^\infty(I)} \leq r \) for \( 0 \leq t_n \leq T \) and, by the uniqueness of the solution of (ML–1), we have \( u_h^n = \tilde{u}_h^n \). Therefore, (36a) implies the desired result.

We now proceed to the proof of Theorem 2.4.

**Proof of Theorem 2.4:** The notation is that of the previous proof. It suffices to prove Theorem 2.4 under assumption (32), which is generalizable to an arbitrary \( f \) as demonstrated in the previous proof. By (24b), we have \( \| v_h^0 \|_{L^\infty(I)} \leq C h^2 \log(1/h) \kappa(u) \) and \( \| w^n \|_{L^\infty(I)} \leq C h^2 \log(1/h) \kappa(u) \) for \( 0 \leq t_n \leq T \). Therefore, it remains to estimate \( v_h^n \) when \( 0 < t_n < T \). Setting
\[
G^n_h = \sum_{i=0}^{m-1} G^n_i \phi_i, \quad G^n_i = \frac{(I + II + III + IV + V)(\phi_i)}{(1, \phi_i)},
\]
we rewrite (35) as
\[
\{ \partial_{\tau_n} v_{h}^{n+1}, \chi \} + A(v_{h}^{n+1}, \chi) = \{ G^n_h, \chi \} \quad (\chi \in S_h).
\]
Showing that
\[
\| G^n_h \|_{L^\infty(I)} \leq M \| v^n_h \|_{L^\infty(I)} + C(h + \tau) \kappa(u),
\]
we can apply Lemma 3.5 to obtain
\[
\| v_{h}^{n+1} \|_{L^\infty(I)} \leq (1 + M \tau_n) \| v^n_h \|_{L^\infty(I)} + \tau_n \cdot C(h + \tau) \kappa(u),
\]
and consequently
\[
\| v^n_h \|_{L^\infty(I)} \leq e^{M \tau_n} \| v^n_0 \|_{L^\infty(I)} + \frac{e^{M \tau_n} - 1}{M} C(h + \tau) \kappa(u).
\]
Thereby, we deduce the desired estimate.

Below we prove the truth of (37). Recall that we assumed global Lipschitz continuity (32) on \( f, I(\phi_i) - IV(\phi_i) \) are straightforwardly estimated as follows:
\[
\| I(\phi_i) \| \leq M [\| v^n_0 \|_{L^\infty(I)} + C h^2 \log(1/h) \kappa(u) ] \cdot (1, \phi_i),
\]
\[
\| II(\phi_i) \| \leq \tau_n \kappa(u)(1, \phi_i),
\]
\[
\| III(\phi_i) \| \leq M \tau_n \kappa(u)(1, \phi_i),
\]
\[
\| IV(\phi_i) \| \leq C h^2 \log(1/h) \kappa(u)(1, \phi_i).
\]
To estimate \( V(\phi_i) \), we write
\[
V(\phi_i) = V_1(\phi_i) + V_2(\phi_i) + V_3(\phi_i),
\]
where
\[
V_1(\phi_i) = (\partial_{\tau_n} P_A u(\cdot, t_{n+1}), \phi_i) - (\partial_{\tau_n} u(\cdot, t_{n+1}), \phi_i),
\]
\[
V_2(\phi_i) = (\partial_{\tau_n} u(\cdot, t_{n+1}), \phi_i) - (\partial_{\tau_n} u(\cdot, t_{n+1}), \phi_i),
\]
\[
V_3(\phi_i) = (\partial_{\tau_n} u(\cdot, t_{n+1}), \phi_i) - (\partial_{\tau_n} u(\cdot, t_{n+1}), \phi_i),
\]
\[ V_3(\phi_i) = \{ \partial_{\tau_n} u(\cdot, t_{n+1}), \phi_i \} - \{ \partial_{\tau_n} P_A u(\cdot, t_{n+1}), \phi_i \}. \]

The above terms are, respectively, estimated as,

\[ |V_1(\phi_i)| \leq \| P_A (\partial_{\tau_n} u(\cdot, t_{n+1})) - \partial_{\tau_n} u(\cdot, t_{n+1}) \|_\infty (1, \phi_i) \]
\[ \leq Ch^2 \log(1/h) \kappa(u)(1, \phi_i); \]
\[ |V_2(\phi_i)| \leq \int_I x^{N-1} \left| \partial_{\tau_n} u(x, t_{n+1}) - \partial_{\tau_n} u(x_i, t_{n+1}) \right| \phi_i(x) \, dx \]
\[ \leq Ch \kappa(u)(1, \phi_i); \]
\[ |V_3(\phi_i)| \leq |\partial_{\tau_n} u(x_i, t_{n+1}) - P_A \partial_{\tau_n} u(x_i, t_{n+1})| (1, \phi_i) \]
\[ \leq Ch^2 \log(1/h) \kappa(u)(1, \phi_i). \]

We thereby deduce that

\[ \| G^n_h \|_{L^\infty(I)} \leq M \| v_h \|_{L^\infty(I)} + C \left( h + h^2 \log(1/h) + \tau \right) \kappa(u), \]

which implies (37). This step completes the proof. \(\blacksquare\)

**Proof of Theorem 2.5:** The proof is identical to that of Theorem 2.4. \(\blacksquare\)

### 5. Blow-up analysis

#### 5.1. Results

This section considers the special nonlinearity

\[ f(s) = s|s|^\alpha, \quad \alpha > 0. \]

As we are interested in non-negative solutions, we assume that

\[ u^0 \geq 0, \neq 0, \quad u^0_h \geq 0, \neq 0. \tag{38} \]

Therefore, the solution \( u \) of (1) is non-negative and the solution \( u^h_n \) of (ML–2) is also non-negative under condition (13). Generally, the solution of (1) blows up when the initial data \( u_0 \) are sufficiently large, and the blow up is controlled by the energy functional associated with (1). Herein, we study whether or not the numerical solution behaves similarly by initially defining some properties of the solution \( u \) of (1). In particular, we see that (ML–2) is suitable for this purpose. In fact, we were unable to prove Theorems 5.1 and 5.2 below using (ML–1); see also Remark 5.3.

The energy functionals associated with (1) are defined as

\[ K(v) = \frac{1}{2} \| v_x \|^2 - \frac{1}{\alpha + 2} \int_I x^{N-1} |v(x)|^{\alpha + 2} \, dx, \]
\[ I(v) = \int_I x^{N-1} v(x) \psi(x) \, dx, \]

where \( \psi \in \dot{H}^1 \) denotes the eigenfunction associated with the first eigenvalue \( \mu > 0 \) of the eigenvalue problem

\[ A(\psi, \chi) = \mu (\psi, \chi) \quad (\chi \in \dot{H}^1). \tag{39} \]

Without loss of generality, we assume that \( \psi \geq 0 \) in \( I \) and \( \int_I x^{N-1} \psi(x) \, dx = 1. \)
The following Propositions 5.1, 5.2 and 5.3 are often applied to the semilinear heat equation in a bounded domain; see [40] for the proof of Propositions 5.1 and 5.2, and [27] for that of Proposition 5.3. They are easily extended to the radially symmetric case.

**Proposition 5.1:** $K(u(t))$ is a non-increasing function of $t$, where $u$ is the solution of (1).

**Proposition 5.2:** Suppose that $u^0 \geq 0, \neq 0$ and $u$ is the solution of (1). Then, the following statements are equivalent:

(i) There exists $T_\infty > 0$ such that $u$ blows up at $t = T_\infty$ in the sense that $\lim_{t \to T_\infty} \|u(\cdot, t)\| = \infty$.

(ii) There exists $t_0 \geq 0$ such that $K(u(\cdot, t_0)) < 0$.

**Proposition 5.3:** Suppose that $u^0 \geq 0, \neq 0$ and that $u$ is the solution of (1). Then, the following statements are equivalent:

(i) There exists $T_\infty > 0$ such that $u$ blows up at $t = T_\infty$ in the sense that $\lim_{t \to T_\infty} I(u(\cdot, t)) = \infty$.

(ii) There exists $t_0 \geq 0$ such that $I(u(\cdot, t_0)) > \mu^{1/\alpha}$.

**Remark 5.1:** In Propositions 5.2 and 5.3, the blow up time $T_\infty$ is estimated, respectively, as

$$T_\infty \leq t_0 + \frac{\alpha + 2}{\alpha^2} N^{-\frac{\alpha}{2}} \|u(\cdot, t_0)\|^{-\alpha},$$

and

$$T_\infty \leq t_0 + \int_{I(u(\cdot, t_0))} ds - \mu s + s^{1+\alpha}.$$

We now proceed to the discrete energy functionals. To this end, we employ the finite element version of the eigenvalue problem:

$$A(\hat{\psi}_h, \chi_h) = \hat{\mu}_h \langle \hat{\psi}_h, \chi_h \rangle \quad (\chi_h \in S_h). \quad (40)$$

Let $\hat{\psi}_h \in S_h$ be the eigenfunction associated with the smallest eigenvalue $\hat{\mu}_h > 0$ of (40). For the eigenvalue problem (40), we state the following result, postponing the proof to the Appendix.

**Proposition 5.4:** If the partition $\{x_j\}_{j=0}^m$ is quasi-uniform, that is, satisfies (7), we have the following:

(i) $\hat{\mu}_h \to \mu$ as $h \to 0$.

(ii) The first eigenfunction $\hat{\psi}_h$ of (40) does not change sign.

(iii) $\|(\hat{\psi}_h - \psi)_x\| \to 0$ as $h \to 0$.

Therefore, without loss of generality, we can assume that $\hat{\psi}_h \geq 0$ and $\int_I x^{N-1} \hat{\psi}_h(x) \, dx = 1$. For $v \in S_h$, we set

$$K_h(v) = \frac{1}{2} \|v_x\|^2 - \frac{1}{\alpha + 2} \sum_{i=0}^{m} |v(x_i)|^{\alpha + 2}(1, \phi_i),$$

$$I_h(v) = \langle v, \hat{\psi}_h \rangle = \int_I x^{N-1} \Pi_h(v \hat{\psi}_h)(x) \, dx.$$
We introduce the approximate blow-up time $\hat{T}_\infty(h)$ by setting
\[
\hat{T}_\infty(h) = \lim_{n \to \infty} n = \lim_{n \to \infty} \sum_{j=0}^{n-1} \tau_j. \tag{41}
\]

We are now in a position to mention the main theorems in this section:

**Theorem 5.1:** Let (38) be satisfied. Suppose that the solution $u$ of (1) blows up at finite time $T_\infty$ in the sense that
\[
\|u(\cdot, t)\|_{L^\infty(I)} \to \infty \quad \text{and} \quad \|u(\cdot, t)\| \to \infty \quad (t \to T_\infty - 0). \tag{42}
\]
Assume that for any $T < T_\infty$, $u$ is sufficiently smooth that (16) holds. Assuming also that (7) is satisfied, we set
\[
\tau = \delta \frac{N^2}{N + 1} \tag{43}
\]
for some $\delta \in (0, 1]$. The time increment $\tau_n$ is iteratively defined as
\[
\tau_n = \tau_n(h) = \tau \min \left\{ 1, \frac{1}{\|u_h^n\|^\sigma} \right\}, \tag{44}
\]
where we have used the solution $u_h^n$ of (ML–2). Moreover, assume that (13) is satisfied and that
\[
\forall T < T_\infty, \quad \lim_{h \to 0} \sup_{0 \leq t_n \leq T} |K(u(\cdot, t_n)) - K_h(u_h^n)| = 0. \tag{45}
\]
We then have
\[
\lim_{h \to 0} \hat{T}_\infty(h) = T_\infty. \tag{46}
\]

**Theorem 5.2:** Let (38) be satisfied. Suppose that the solution $u$ of (1) blows up at finite time $T_\infty$ in the sense that
\[
I(u(\cdot, t)) \to \infty \quad \text{and} \quad \|u(\cdot, t)\|_{L^\infty(I)} \to \infty \quad (t \to T_\infty - 0). \tag{47}
\]
Assume that, for any $T < T_\infty$, $u$ is sufficiently smooth that (16) holds. Assuming also that (7) is satisfied, we set $\tau$ by (43) with some $\delta \in (0, 1]$. The time increment $\tau_n$ is iteratively defined as
\[
\tau_n = \tau_n(h) = \tau \min \left\{ 1, \frac{1}{I_h(u_h^n)^\sigma} \right\}, \tag{48}
\]
where we have used the solution $u_h^n$ of (ML–2) with (17). We then obtain (46).

**Remark 5.2:** The above theorems differ in that Theorem 5.1 requires the convergence property (45) of the discrete energy functional $K_h(u_h^n)$, whereas no convergence property of $I_h$ is necessary in Theorem 5.2.

**Remark 5.3:** Unfortunately, we could not prove Theorems 5.1 and 5.2 using the solution of (ML–1). In particular, the proof of the difference inequalities (49) and (52) failed in scheme (ML–1).
5.2. Proof of theorem 5.1

To prove Theorem 5.1, we follow Nakagawa’s blow-up analysis [30]. For this purpose, we must derive the difference inequality (49) and the boundedness (50) of \( \hat{T}_\infty \) (see Lemmas 5.2 and 5.3). The original proof in [30] immediately follows from these results; see also [6], [7] and [13]. Therefore, we concentrate our efforts on proving Lemmas 5.2 and 5.3.

Throughout this section, we take the same assumptions as Theorem 5.1; in particular, the time-increment control (44). Note that condition (13) is satisfied by the definition of \( \tau_n \). Consequently, the solution \( u \) of (1a) and the solution \( u^n_h \) of (ML–2) are non-negative.

**Lemma 5.1:** \( K_h(u^n_h) \) is a non-increasing sequence of \( n \).

**Proof:** Fixing some \( n \geq 0 \), we write \( w = u^{n+1}_h, u = u^n_h, w_j = w(x_j), \) and \( u_j = u(x_j) \). To show that \( K_h(w) - K_h(u) \leq 0 \), we perform the following division:

\[
K_h(w) - K_h(u) = X + Y,
\]

where

\[
X = \frac{1}{2} \| w_x \|^2 - \frac{1}{2} \| u_x \|^2,
\]

\[
Y = -\frac{1}{\alpha + 2} \sum_{j=0}^{m-1} u^\alpha_{j+2}(1, \phi_j) + \frac{1}{\alpha + 2} \sum_{j=0}^{m-1} u^\alpha_{j+2}(1, \phi_j).
\]

\( X \) is expressed as

\[
X = A(u, w - u) + \frac{1}{2} A(w - u, w - u).
\]

By the mean value theorem, there exists \( \theta_j \in [0, 1] \) such that

\[
w_j^\alpha+2 - u_j^\alpha+2 = (\alpha + 2) \tilde{u}_j^\alpha+1(w_j - u_j),
\]

where \( \tilde{u}_j = u_j + \theta_j(w_j - u_j) \). Therefore,

\[
Y = -\sum_{j=0}^{m-1} \tilde{u}_j^\alpha+1(1, \phi_j)(w_j - u_j)
\]

\[
= -\sum_{j=0}^{m-1} [\tilde{u}_j^\alpha+1 - u_j^\alpha+1](1, \phi_j)(w_j - u_j) - \sum_{j=0}^{m-1} u_j^\alpha+1(1, \phi_j)(w_j - u_j).
\]

Setting

\[
Y_1 = -\sum_{j=0}^{m-1} [\tilde{u}_j^\alpha+1 - u_j^\alpha+1](1, \phi_j)(w_j - u_j),
\]

\[
Y_2 = -\sum_{j=0}^{m-1} u_j^\alpha+1(1, \phi_j)(w_j - u_j),
\]

\[
Y = Y_1 + Y_2.
\]
we calculate
\[ A(u, w - u) + Y_2 = -\left( \frac{w - u}{\tau_n}, w - u \right) = -\frac{1}{\tau_n} ||w - u||^2 \]
and
\[ Y_1 = - \sum_{j=0}^{m-1} (\alpha + 1)\theta_j (1, \phi_j) \hat{u}^\alpha_j (w_j - u_j)^2 \leq 0, \]
where \( \hat{u}_j = u + \hat{\theta}_j (\tilde{u}_j - u_j) \) with some \( \hat{\theta}_j \in [0, 1] \).

Meanwhile, for \( v_h \in S_h(I) \), we write
\[ A(v_h, v_h) \leq 2 \sum_{j=1}^{m} \int_{I_j} \chi_{N-1} \cdot \frac{1}{h_j^2} \, dx \cdot (v_j^2 + v_{j-1}^2) = 2 \sum_{j=0}^{m-1} a_{j} v_j^2. \]

Using (13), we have
\[ A(v_h, v_h) \leq 2 \sum_{j=0}^{m-1} (1, \phi_j) \tau_n v_j^2 = \frac{2}{\tau_n} \langle v_h, v_h \rangle = \frac{2}{\tau_n} ||v_h||^2 \quad (v_h \in S_h). \]

We thereby deduce that
\[ X + Y = -\frac{1}{\tau_n} ||w - u||^2 + \frac{1}{2} A(w - u, w - u) + Y_1 \leq 0, \]
which implies that \( K_h(u^n_h) \) is non-increasing in \( n \).

**Lemma 5.2:** If there exists a non-negative integer \( n' \) such that \( K_h(u^n_h) \leq 0 \) for all \( n \geq n' \), we have
\[ \frac{1}{2} \tau_n ||u^{n+1}_h||^2 \geq \frac{\alpha}{\alpha + 2} N^{\frac{\alpha}{2}} ||u^n_h||^{\alpha + 2} \quad (n \geq n'). \quad (49) \]

**Proof:** Substituting \( \chi_h = u^n_h \) in (ML–2), we obtain
\[ \left\langle \frac{u^{n+1}_h - u^n_h}{\tau_n}, u^n_h \right\rangle + A(u^n_h, u^n_h) = \langle u^n_h(u^n_h)^\alpha, u^n_h \rangle. \]

We note that
\[ \langle u^{n+1}_h - u^n_h, u^n_h \rangle \leq \left( u^{n+1}_h - u^n_h, \frac{1}{2} (u^{n+1}_h + u^n_h) \right) = \frac{1}{2} (||u^{n+1}_h||^2 - ||u^n_h||^2). \]

By the decreasing property of \( K_h(u^n_h) \), we have
\[ ||(u^n_h)_x||^2 \leq \frac{2}{\alpha + 2} \langle (u^n_h)^{\alpha + 2}, 1 \rangle. \quad (n \geq n'). \]

Combining these results, we get
\[ \frac{1}{2} \tau_n (||u^{n+1}_h||^2 - ||u^n_h||^2) \geq \frac{\alpha}{\alpha + 2} \langle (u^n_h)^{\alpha + 2}, 1 \rangle. \]

Using Hölder’s inequality, we calculate
\[ ||u^n_h||^2 \leq (1/N)^{\frac{\alpha}{\alpha + 2}} \langle (u^n_h)^{\alpha + 2}, 1 \rangle^{\frac{2}{\alpha + 2}}. \]

We thereby deduce (49).
Lemma 5.3: If $K_h(u_h^{n_0}) \leq 0$ and $\|u_h^{n_0}\| \geq 1$ for some integer $n_0 \geq 0$, then we have

$$\hat{T}_\infty(h) \leq t_{n_0} + \left\{ \frac{\alpha + 2}{\alpha^2} \cdot N^{-\frac{\alpha}{2}} + \tau \left( 1 + \frac{2}{\alpha} \right) \right\} \|u_h^{n_0}\|^{-\alpha}. \quad (50)$$

Proof: From Lemma 5.2,

$$\|u_h^{n+1}\|^2 \geq (1 + 2\tau C\|u_h^n\|^\alpha)\|u_h^n\|^2 = (1 + 2\tau C)\|u_h^n\|^2,$$

where $C = \frac{\alpha}{\alpha + 2} N^{\frac{\alpha}{2}}$. Therefore,

$$\lim_{n \to \infty} \|u_h^n\| = \infty$$

and, for $n \geq n_0$,

$$t_n = t_{n_0} + \sum_{m=n_0}^{n-1} \tau_m = t_{n_0} + \sum_{m=n_0}^{n-1} \frac{\tau}{\|u_h^m\|^\alpha}. \quad (51)$$

The remainder is identical to that of Corollary 2.1 in [13], so the details are omitted here. ■

5.3. Proof of Theorem 5.2

To prove Theorem 5.2, we apply abstract theory by (Propositions 4.2 and 4.3) in [37]. In this section, we take the same assumptions as Theorem 5.2, in particular, the time-increment control (48).

Lemma 5.4: We have $T_\infty \leq \lim \inf_{h \to 0} \hat{T}_\infty(h)$.

Proof: The proof is shown by contradiction. Setting $S_\infty = \lim \inf_{h \to 0} \hat{T}_\infty(h)$, we assume that $S_\infty < T_\infty$. Then, there exists $h_0 > 0$ such that $\inf_{h \leq h_0} \hat{T}_\infty(h') < M$ for all $h \leq h_0$, where $M = \frac{T_\infty + S_\infty}{2} < T_\infty$. That is, for some fixed $h \leq h_0$, we have $t_n \leq \hat{T}_\infty(h) < M$ and $I_h(u_h^n) \to \infty$ as $n \to \infty$. This implies that for some $0 \leq j(n) \leq m - 1$, we have $u_h^n(x_{j(n)}) \to \infty$, and consequently $\|u_h^n\|_{L^\infty(I)} \to \infty$. However, from Theorem 2.5, we observe that

$$\lim_{h \to 0} \sup_{0 \leq t_n \leq M} \|u_h^n - u(\cdot, t_n)\|_{L^\infty(I)} = 0.$$

If this expression is true, then $T_\infty$ cannot be the blow-up time of the solution $u$ of (1a). This contradiction completes the proof. ■

Lemma 5.5: For any $T < T_\infty$, we have

$$\lim_{h \to 0} \sup_{0 \leq t_n \leq T} |I_h(u_h^n) - I(u(\cdot, t_n))| = 0.$$

Proof: We derive separate estimations for $|I_h(u_h^n) - \bar{I}_h(u_h^n)|$ and $|\bar{I}_h(u_h^n) - I(u(\cdot, t_n))|$, where $\bar{I}_h(v)$ denotes the auxiliary functional

$$\bar{I}_h(v) = \int_I x^{N-1} v(x) \hat{\psi}_h(x) \, dx.$$

From Theorem 2.5, Lemma 3.3 and Proposition 5.4 (iii), we first derive

$$|I_h(u_h^n) - \bar{I}_h(u_h^n)| \leq C h^2 \|x^{N-1}(u_h^n \hat{\psi}_h)_{xx}\|_{L^1(I)}$$
\[ \leq Ch^2 \|u^n_h\| \cdot \|\hat{\psi}'_h\| \]
\[ \leq Ch\|u^n_h\|_{L^\infty(I)} \|\hat{\psi}'_h\| \]
\[ \leq Ch(\|u(\cdot, t_n)\|_{L^\infty(I)} + 1)(\|\psi'\| + 1), \]

where we have used the elemental inequality \[ \|\psi_x\| \leq Ch^{-1}\|\psi\|_{L^\infty(I)} \] for \( \psi \in S_h \). This implies that \[ |I_h(u^n_h) - \tilde{I}_h(u^n_h)| \to 0 \text{ as } h \to 0. \]

On the other hand, as \( h \to 0 \), we have
\[
|\tilde{I}_h(u^n_h) - I(u(\cdot, t_n))| = \left| \int_I x^{N-1} (u^n_h - u(\cdot, t_n)) \hat{\psi}_h \, dx \right| + \left| \int_I x^{N-1} u(\cdot, t_n)(\hat{\psi}_h - \psi) \, dx \right|
\leq \|u^n_h - u(\cdot, t_n)\| \cdot \|\hat{\psi}_h\| + \|u(\cdot, t_n)\| \cdot \|\hat{\psi}_h - \psi\| \to 0.
\]

This expression concludes the proof. \[ \square \]

The following is a readily obtainable consequence of Lemma 5.5.

**Lemma 5.6:** For any \( s_0 \), there exists a nonnegative integer \( n_0 = n_0(h) \) such that \( I_h(u^n_0) > s_0 \).

The following lemma is elementary and was originally stated as Lemma 3.1 in [13].

**Lemma 5.7:** There exists \( s_0 > 1 \) satisfying
\[
\frac{1}{2}f(s) + (1 + \mu)s \leq f(s) \quad (s \geq s_0),
\]
where \( f(s) = s^{1+\alpha} \).

We can now prove Theorem 5.2.

**Proof of Theorem 5.2:** It remains to verify that
\[ T_\infty \geq \limsup_{h \to 0} \hat{T}_\infty(h). \] (51)

To this end, we apply abstract theory (Propositions 4.2 and 4.3) in [37]. Adopting the notation of [37], we, respectively, set \( X, X_h, G, H, J \) and \( J_h \) in [37] as
\[
X = \hat{H}, \quad \|v\|_X = \|v\|_{L^\infty(I)}, \quad X_h = S_h,
\]
\[
G(s) = \frac{1}{2}f(s) = \frac{1}{2}s^{1+\alpha}, \quad H(s) = s^\alpha \quad (s \geq 0),
\]
\[
J(t, v) = \int_I x^{N-1} v(x) \cdot \psi(x) \, dx, \quad (t, v) \in (0, \infty) \times X,
\]
\[
J_h(t, v_h) = I_h(v_h) = \sum_{j=0}^{m-1} v_j \hat{\psi}_h(x_j)(1, \phi_j), \quad (t, v_h) \in (0, \infty) \times X_h, \quad v_j = v_h(x_j).
\]

Here, \( G \) and \( H \) are functions of class \((G)\) and class \((H)\), respectively.

To avoid unnecessary complexity, we assume that \( n_0 = 0 \) (see Lemma 5.6). Our problem setting matches problem settings (I)–(VIII) in § 4 of [37].
It is readily verified that conditions (H1), (H3), (H4) and (H5) in § 4 of [37] hold. We need only check that condition (H2)
\[ \partial \tau_n I_h(u_h^{n+1}) \geq \frac{1}{2} f(I_h(u_h^n)) \quad (n \geq 0) \] (52)
also holds.

Substituting \( \chi_h = \hat{\psi}_h \in S_h \) in (ML–2) and using the relation (40), we have
\[ \partial \tau_n I_h(u_h^{n+1}) + \hat{\mu}_h I_h(u_h^{n+1}) = \left\{ f(u_h^n), \hat{\psi}_h \right\}. \]
From Proposition 5.4 (i), we know that \( \hat{\mu}_h < \mu + 1 \). Moreover, because \( \left\{ 1, \hat{\psi}_h \right\} = 1 \), we can apply Jensen’s inequality to get
\[ \partial \tau_n I_h(u_h^{1}) \geq -\left( \mu + 1 \right) \cdot I_h(u_h^0) + f(I_h(u_h^0)). \]
By Lemma 5.7, \(-\left( \mu + 1 \right) s + f(s) \geq \frac{1}{2} f(s)\) for \( s \geq s_0 \). Because \( I_h(u_h^0) > s_0 \) by Lemma 5.6, we deduce that
\[ \partial \tau_n I_h(u_h^{1}) \geq \frac{1}{2} f(I_h(u_h^0)). \]
We thus obtain \( I_h(v_1^1) > s_0 \). By this process, we finally obtain
\[ \partial \tau_n I_h(u_h^{n+1}) \geq \frac{1}{2} f(I_h(u_h^n)), \quad I_h(u_h^n) > s_0 \quad (n \geq 0). \]

Remark 5.4: Theorem 5.2 remains true after replacing (48) by
\[ \tau_n = \tau_n(h) = \tau \min \left\{ 1, \frac{1}{\| u_h^n \|^q_{L_\infty(I)}} \right\}. \]
However, this definition increases the computational cost over that based on (48). Therefore, in the following numerical evaluation, we adopt (48).

6. Numerical examples
This section validates our theoretical results with numerical examples.
We first examine the error estimates of the solutions on a uniform spatial mesh \( x_j = jh \) \( (j = 0, \ldots, m) \) with \( h = 1/m \), regarding the numerical solution with \( h' = 1/480 \) as the exact solution. The following quantities were compared:
\[ E_1(h) = \| u_h^n - u_h^0 \|_{L^1(I)}, \]
\[ E_2(h) = \| u_h^n - u_h^0 \| = \left\| X^{N-1} (u_h^n - u_h^0) \right\|_{L^2(I)}, \]
\[ E_\infty(h) = \| u_h^n - u_h^0 \|_{L_\infty(I)}. \]
Figure 1 shows the results for \( N = 3, \alpha = 4 \) and \( u(0,x) = \cos \frac{\pi}{2} x \). The time increment was uniform \( \tau_n = \tau = \lambda h^2, n = 0, 1, \ldots, \lambda = 1/2 \) and the iterations were continued until \( t \leq T = 0.005 \). Hereinafter, we set \( u_h^0 = \Pi_h u^0 \). In scheme (ML-1), the numerical convergence rate was \( h^2 + \tau \) in the \( \| \cdot \| \) norm (see Theorem 2.3), but was slightly deteriorated in the \( L_\infty \) norm.
In the case $N = 4$, which is not supported in the convergence property of the standard symmetric FEM [32], we chose $\alpha = 3$ and $u(0, x) = 3 \cos \frac{\pi}{2} x$. The errors were computed up to $T^* = 0.0033$ on the non-uniform meshes with $x_i = \sin \frac{(i-1)\pi}{2m}$ and $\tau_n$ with $\lambda = 0.11$. As shown in Figure 2, the $\| \cdot \|$ norm showed second-order convergence in both (ML-1) and standard FEM, but the $\| \cdot \|_{L^\infty(I)}$ norm showed first-order convergence in the standard FEM.

Secondly, we confirmed the non-increasing property of the energy functional $K_h(u_h^t)$ in scheme (ML-2) with $N = 5$, $\alpha = \frac{4}{3}$, and $u(0, x) = \cos \frac{\pi}{2} x$, $13 \cos \frac{\pi}{2} x$. The time increment $\tau_n$ was determined through Theorem 5.6 with $\beta = 1$ and $\delta = 1$. Simulations were performed on a uniform spatial mesh $x_j = jh$ with $h = 1/m$ and $m = 50$. The results (see Figure 3) support Lemma 5.9. As shown in Figure 4, the energy functional $I_h(u_h^t)$ increased exponentially after $t = 0.04$ when the initial data
was large, but vanishes when the initial data was small. For $I_h(u^n_h)$, we used the time increment in Theorem 5.7 with $\delta = 1$.

Finally, we calculated the numerical blow-up time in scheme (ML–2). Here, we set $h = 1/m$ ($m = 16, 32, 64$). The time increments were defined as

(K) $\tau_n = \frac{1}{N + 1} \min \left\{ 1, \frac{1}{\|u^n_h\|_\alpha} \right\}$ (see Theorem 5.1),

(I) $\tau_n = \frac{1}{N + 1} \min \left\{ 1, \frac{1}{I_h(u^n_h)^\alpha} \right\}$ (see Theorem 5.2).

We compared the numerical blow-up times by some numerical methods because the exact blow-up time of (1) is unknown. For a comparison evaluation, we executed the FDM of Chen [7] and the FDM of Cho and Okamoto [12]. Specifically, set $\tau_n = \frac{1}{2} h^2 \cdot \min \{ 1, \frac{1}{\|u^n_h\|_2} \}$ in Chen’s FDM and $\tau_n = \frac{1}{3N} h^2 \cdot \min \{ 1, \frac{1}{\|u^n_h\|_2} \}$, $\sigma = \frac{3}{2N}$ in Cho and Okamoto’s FDM.

We then introduced the truncated numerical blow-up time $\hat{T}_M(h)$:

$$\hat{T}_M(h) = \min \left\{ t_n | \|u^n_h\|_\infty > M = 10^8 \right\} .$$

Evaluations were performed with three parameter sets:
Figure 5. Truncated numerical blow-up times $\hat{T}_M(h)$ in the four schemes with three parameter settings. (a) Case 1, (b) Case 2 and (c) Case 3.

Case 1 $N = 5$, $\alpha = 0.39$, and $u(0, x) = 8000 \cos \frac{\pi}{2} x$;
Case 2 $N = 4$, $\alpha = 0.49$, and $u(0, x) = 800(1 - x^2)$;
Case 3 $N = 3$, $\alpha = 0.66$, and $u(0, x) = 1000(e^{-x^2} - e^{-1})$.

Note that if $1 > \frac{N\alpha}{2}$ holds true and $u^0 \geq 0$ is decreasing in $x$, then $I(u(t)), \|u(t)\|$ and $\|u(t)\|_{L^\infty(I)}$ blow up simultaneously; see [19]. We chose these settings so that the assumptions in Theorems 5.1 and 5.2 hold true.

Figure 5 compares the truncated numerical blow-up times $\hat{T}_M(h)$ as functions of $h$ in the four schemes. The solution of Chen’s FDM blew up later than the others schemes, whereas that of (ML–2) with $I_h(u_h^0)$ blew up sooner than the other schemes.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by JST CREST (Core Research for Evolutional Science and Technology) [grant number JPMJCR15D1], Japan, and JSPS KAKENHI [grant number 15H03635], Japan. In addition, the first author was supported by the Program for Leading Graduate Schools, MEXT, Japan.

References

[1] L.M. Abia, J.C. López-Marcos, and J. Martínez, The Euler method in the numerical integration of reaction-diffusion problems with blow-up, Appl. Numer. Math. 38(3) (2001), pp. 287–313.
[2] U. Banerjee and J.E. Osborn, Estimation of the effect of numerical integration in finite element eigenvalue approximation, Numer. Math. 56(8) (1990), pp. 735–762.
[3] C. Brändle, P. Groisman, and I.D. Rossi, Fully discrete adaptive methods for a blow-up problem, Math. Models Methods Appl. Sci. 14(10) (2004), pp. 1425–1450.
[4] C. Brändle, F. Quirós, and I.D. Rossi, An adaptive numerical method to handle blow-up in a parabolic system, Numer. Math. 102(1) (2005), pp. 39–59.
[5] A. Cangiani, E.H. Georgoulis, I. Kyza, and S. Metcalfe, Adaptivity and blow-up detection for nonlinear evolution problems, SIAM J. Sci. Comput. 38(6) (2016), pp. A3833–A3856.
Appendix. Approximate eigenvalue problems

This Appendix establishes the proof of Proposition 5.4. Recall that $\mu$ and $\mu_h$ are the smallest eigenvalues of (39) and (40), respectively. Functions $\psi$ and $\psi_h$ are the eigenfunctions associated with $\mu$ and $\mu_h$, respectively.

We introduce the following linear operators $T : H^1 \to H^1$, $T_h : S_h \to S_h$ and $\hat{T}_h : S_h \to S_h$ by

$$A(Tv, \chi) = \langle v, \chi \rangle \quad (\chi \in H^1, v \in H^1),$$

$$A(T_hvh, \chi_h) = \langle vh, \chi_h \rangle \quad (\chi_h \in S_h, vh \in S_h),$$

$$A(\hat{T}_hvh, \chi_h) = \langle vh, \chi_h \rangle \quad (\chi_h \in S_h, vh \in S_h).$$

We also write $\nu' = \nu_x$ for some function $\nu = \nu(x)$.

Lemma A.1 (Lemma 2.2 in [38], Theorems 13 and 14 in [18]): For any $fh \in S_h \subset \dot{H}^1$,

$$\| (T_h)^\nu \| \leq C\| fh \|.$$

For a linear operator $B : X \subset \dot{H} \to \dot{H}$, we define

$$\| B \|_{1,X} = \sup_{v \in X, v \neq 0} \frac{\| (Bv)' \|}{\| v \|}.$$

Lemma A.2 (Lemma 3.3 in [2]): $\| T - \hat{T}_h \|_{1,S_h} \to 0$ as $h \to 0$.

Remark A.1: Lemma A.2 does not exactly agree with Lemma 3.3 in [2], but its proof is identical to that of Lemma 3.3 in [2].

Let $I$ be the identity operator and $\rho(B)$ be the resolvent set of a linear operator $B$. The resolvent operator $R_z(\hat{T}_h)$ for $z \in \mathbb{C}$ is defined as

$$R_z(\hat{T}_h) = (zI - \hat{T}_h)^{-1} : S_h \to S_h,$$

where $\hat{T}_h : S_h \to S_h$ and $z \in \rho(\hat{T}_h)$.

Lemma A.3 ([15, Lemma 1]): For any closed set $F \subset \rho(T)$, there exists $h_0 > 0$ such that for any $h \leq h_0$, $R_z(\hat{T}_h)$ exists and

$$\| R_z(\hat{T}_h) \|_{1,S_h} \leq C \quad (\forall z \in F),$$

where $C$ is independent of $h$.

We now define spectral projections of $T$ and $\hat{T}_h$. Let $\Gamma \subset \mathbb{C}$ be a circle centered at $\frac{1}{h}$ enclosing no other points of $\sigma(T)$ which stands for the spectral set of $T$. Let $E = E(\frac{1}{h}) : H^1 \to H^1$ and $\hat{E}_h = \hat{E}_h(\frac{1}{h}) : S_h \to S_h$ be the spectral projection operators associated with $T$ and $\hat{T}_h$ and the parts of the corresponding spectrum enclosed by $\Gamma$, respectively:

$$E = \frac{1}{2\pi i} \int_\Gamma R_z(T) \, dz, \quad \hat{E}_h = \frac{1}{2\pi i} \int_\Gamma R_z(\hat{T}_h) \, dz.$$

Remark A.2: By Lemma A.3, when $h$ is sufficiently small, $\Gamma \subset \rho(\hat{T}_h)$ holds and $\| R_z(\hat{T}_z) \|_{1,S_h}$ is bounded for all $z \in \Gamma$. Thus the integral of $\hat{E}_h$ exists.

To examine the convergence property of $\hat{E}_h$, we use the following lemma.
Lemma A.4 (Lemma 2 in [15]): \( \| E - \hat{E}_h \|_{1,S_h} \to 0 \) as \( h \to 0 \).

We use the following symbols.

- \( \text{dist}(w,A) = \inf_{y \in A} \| w - y \| \quad (w \in H^1, A \subset H^1) \)
- \( \delta(\hat{E}_h(S_h), E(\hat{H}^1)) = \sup_{v_h \in \hat{E}_h(S_h), \| v_h \| = 1} \text{dist}(v_h, E(\hat{H}^1)) \)
- \( \delta(E(\hat{H}^1), \hat{E}_h(S_h)) = \sup_{v \in E(\hat{H}^1), \| v \| = 1} \text{dist}(v, \hat{E}_h(S_h)) \)
- \( \delta(H(\hat{H}^1), \hat{E}_h(S_h)) = \max[\delta(\hat{E}_h(S_h), E(\hat{H}^1)), \delta(E(\hat{H}^1), \hat{E}_h(S_h))] \)

The result follows from the property of the spectral projection operator.

Corollary A.1: \( \delta(\hat{E}_h(S_h), E(\hat{H}^1)) \to 0 \) as \( h \to 0 \).

Lemma A.5 (Theorem 2 in [15]):

\[
\lim_{h \to 0} \inf_{\chi_h \in S_h} \| (u - \chi_h)' \| = 0 \quad (u \in H^1).
\]

Corollary A.2 (Theorem 3 in [15]): \( \delta(E(\hat{H}^1), \hat{E}_h(S_h)) \to 0 \) as \( h \to 0 \).

Lemma A.6 (Corollary 2.6 in [28]): If \( \delta(E(\hat{H}^1), \hat{E}_h(S_h)) < 1 \), then \( \text{dim } E(\hat{H}^1) = \text{dim } \hat{E}_h(S_h) \).

For sufficiently small \( h \), we observe that \( \delta(E(\hat{H}^1), \hat{E}_h(S_h)) < 1 \), that is, \( \text{dim } E(\hat{H}^1) = \text{dim } \hat{E}_h(S_h) \).

We have \( \dim E(\hat{H}^1) = 1 \), because \( E(\hat{H}^1) \) is the eigenspace of the smallest eigenvalue of (39). Then, the unique eigenvalue of \( \hat{T}_h \) (denoted by \( \frac{1}{v_h} \)) is located inside \( \Gamma \). Then, there exists \( \hat{v}_h(\neq 0) \in S_h \) such that

\[
A(\hat{v}_h, \chi_h) = \hat{v}_h \left\langle \hat{v}_h, \chi_h \right\rangle, \quad \chi_h \in S_h.
\]

Corollary A.3: \( \hat{v}_h \to \mu \) as \( h \to 0 \).

Proof: For some arbitrary \( \epsilon > 0 \), we set \( \Gamma_\epsilon = B_1(\epsilon) = \{ z \in \mathbb{C} \mid |z - \frac{1}{\mu}| = \epsilon \} \). As stated above, there exists \( h_\epsilon > 0 \) such that the eigenvalue \( \frac{1}{v_h} \) of \( \hat{T}_h \) is inside \( \Gamma_\epsilon \) for all \( h < h_\epsilon \). Therefore,

\[
\left| \frac{1}{v_h} - \frac{1}{\mu} \right| \leq \epsilon.
\]

Because \( \mu \) is positive, the proof is complete.

Remark A.3: Similarly, we find that a unique eigenvalue \( \frac{1}{v_h} \) of \( T_h \) exists inside \( \Gamma \) and that \( v_h \to \mu \) as \( h \to 0 \).

Let \( \mu_h > 0 \) be the smallest eigenvalue of

\[
A(\psi_h, \chi_h) = \mu_h(\psi_h, \chi_h), \quad \chi_h \in S_h,
\]

where \( \psi_h \in S_h \).

Lemma A.7: For sufficiently small \( h > 0 \), we have \( v_h = \mu_h \). In particular, \( \mu_h \to \mu \) as \( h \to 0 \).

Proof: We know that \( \text{dim } E(H^1) = 1 \). By variational characterization, we obtain

\[
\mu = \inf_{v \in H^1, v \neq 0} \frac{\| v' \|^2}{\| v \|^2} \leq \inf_{v_h \in S_h, v_h \neq 0} \frac{\| v_h' \|^2}{\| v_h \|^2}.
\]

Here \( \mu_h \) is the smallest eigenvalue of (A1), that is,

\[
\mu_h = \inf_{v_h \in S_h, v_h \neq 0} \frac{\| v_h' \|^2}{\| v_h \|^2}.
\]

All eigenvalues of (A1) are greater than or equal to \( \mu \). As the eigenvalue of \( T_h \) enclosed by \( \Gamma \) is unique, we obtain \( v_h = \mu_h \) for sufficiently small \( h > 0 \), and \( \mu_h \to \mu \) as \( h \to 0 \).

}\]
We now state the following proof.

Proof of Proposition 5.4 (i): By variational characterization, we get

\[
\hat{\mu}_h = \inf_{v_h \in S_h, v_h \neq 0} \frac{\|v_h^\prime\|^2}{\|v_h\|^2} = \left( \sup_{v_h \in S_h, v_h \neq 0} \frac{A(T_h v_h, v_h)}{\|v_h\|^2} \right)^{-1},
\]

\[
\mu_h = \inf_{v_h \in S_h, v_h \neq 0} \frac{\|v_h^\prime\|^2}{\|v_h\|^2} = \left( \sup_{v_h \in S_h, v_h \neq 0} \frac{A(T_h v_h, v_h)}{\|v_h\|^2} \right)^{-1}.
\]

Then,

\[
\sup_{v_h \in S_h, v_h \neq 0} \frac{A(\hat{T}_h v_h, v_h)}{\|v_h\|^2} = \sup_{v_h \in S_h, v_h \neq 0} \left( \frac{A(T_h v_h, v_h)}{\|v_h\|^2} + \frac{A(\hat{T}_h - T_h) v_h, v_h)}{\|v_h\|^2} \right) \leq \sup_{v_h \in S_h, v_h \neq 0} \frac{A(T_h v_h, v_h)}{\|v_h\|^2} + \|\hat{T}_h - T_h\|_{1,S_h}.
\]

Similarly,

\[
\sup_{v_h \in S_h, v_h \neq 0} \frac{A(T_h v_h, v_h)}{\|v_h\|^2} \leq \sup_{v_h \in S_h, v_h \neq 0} \frac{A(\hat{T}_h v_h, v_h)}{\|v_h\|^2} + \|\hat{T}_h - T_h\|_{1,S_h}.
\]

Applying Lemma A.2, we obtain \(\hat{\mu}_h \rightarrow \mu\) as \(h \rightarrow 0\).

Remark A.4: As the eigenvalue of \(\hat{T}_h\) enclosed by \(\Gamma\) is unique, we conclude that \(\hat{v}_h = \hat{\mu}_h\) for sufficiently small \(h > 0\).

Proof of Proposition 5.4 (ii): We write (40) in matrix form:

\[
\mathcal{A} \psi = \hat{\mu}_h \mathcal{M} \psi,
\]

where \(\mathcal{M} = \text{diag}(\mu_i)_{0 \leq i \leq m-1}\) and \(\mathcal{A} = (a_{ij})_{0 \leq i,j \leq m-1}\) are defined as \(m_i = (1, \phi_i)\) and \(a_{ij} = A(\phi_j, \phi_i)\), respectively. Moreover, \(\psi = (\psi_i)_{0 \leq i \leq m-1} \in \mathbb{R}^m\), \(\psi_i = \hat{\psi}_h(x_i)\). We thus express (40) as

\[
\mathcal{M}^{-1} \mathcal{A} \psi = \hat{\mu}_h \psi.
\]

Because \(\mathcal{M}\) is a diagonal matrix, the diagonal components \(a_{ij}/\mu_{ii}\) of \(\mathcal{M}^{-1} \mathcal{A}\) are all positive and the non-diagonal components are all non-positive.

Writing

\[
\left( -\mathcal{M}^{-1} \mathcal{A} + \max_{0 \leq i \leq m-1} \frac{a_{ii}}{m_{ii}} I \right) \psi = \left( -\hat{\mu}_h + \max_{0 \leq i \leq m-1} \frac{a_{ii}}{m_{ii}} \right) \psi,
\]

where \(I \in \mathbb{R}^{m \times m}\) is the identity matrix, then we can see that all components of the matrix on the left-hand side are non-negative. We now consider the following eigenvalue problem:

\[
\left( -\mathcal{M}^{-1} \mathcal{A} + \max_{0 \leq i \leq m-1} \frac{a_{ii}}{m_{ii}} I \right) \tilde{x} = \hat{\mu}_h \tilde{x}.
\]

By the Perron–Frobenius theorem, we can take a positive eigenvector for the largest eigenvalue of (A3). \(-\hat{\mu}_h + \max_{0 \leq i \leq m-1} \frac{a_{ii}}{m_{ii}}\) is the largest eigenvalue of (A3), because \(\hat{\mu}_h\) is the smallest eigenvalue of (40).

Consequently, the sign of the first eigenfunction of (40) is unchanged.

Proof of Proposition 5.4 (iii): We assume that

\[
\hat{\psi}_h \geq 0 \quad \text{and} \quad \int_I x^{N-1} \hat{\psi}_h(x) \, dx = 1.
\]

Setting \(\phi = \psi / \|\psi\|\) and \(\hat{\phi}_h = \hat{\psi}_h / \|\hat{\psi}_h\|\), applying Corollary A.1 and Proposition 5.4 (i), and setting \(v_h = \hat{\phi}_h\), we obtain

\[
\text{dist}(\hat{\phi}_h, E(\hat{H}^1)) \rightarrow 0 \quad \text{as} \quad h \rightarrow 0.
\]

Because \(\dim E(\hat{H}^1) = 1\) and \(E(\hat{H}^1)\) is a closed subspace in \(\hat{H}^1\), we find that

\[
E(\hat{H}^1) = \{ z \phi \in \hat{H}^1 \mid z \in \mathbb{C} \}
\]

and

\[
\text{dist}(\hat{\phi}_h, E(\hat{H}^1)) = \|\hat{\phi}_h^\prime - c_h \phi\|,
\]

where \(c_h \in \mathbb{C}\).
Therefore, $|c_h| \to 1$ as $h \to 0$. Using $\hat{\phi}_h$, $\phi \geq 0$ and $\|\hat{\phi}_h - c_h\phi\| \leq \|\hat{\phi}_h' - c_h\phi'\|$, we find that $c_h \to 1$ as $h \to 0$. That is, as $h \to 0$,

$$\|\hat{\phi}_h' - \phi'\| \leq \|\hat{\phi}_h' - c_h\phi'\| + \|c_h\phi' - \phi'\|
= \|\hat{\phi}_h' - c_h\phi'\| + |c_h - 1| \cdot \|\phi'\| \to 0.$$ 

On the other hand,

$$\|(\psi - \hat{\psi}_h)'\| \leq \frac{1}{\int_I x^{N-1} \phi(x) \, dx} \|\phi' - \hat{\phi}_h'\| + \left| \frac{1}{\int_I x^{N-1} \phi(x) \, dx} - \frac{1}{\int_I x^{N-1} \hat{\phi}_h(x) \, dx} \right|.$$ 

This, together with

$$\left| \int_I x^{N-1} \phi(x) \, dx - \int_I x^{N-1} \hat{\phi}_h(x) \, dx \right|$$

$$\leq \left( \frac{1}{N} \right)^{1/2} \cdot \|\phi - \hat{\phi}_h\| \leq \left( \frac{1}{N} \right)^{1/2} \cdot \|\phi' - \hat{\phi}_h'\| \to 0 \quad \text{as } h \to 0$$

implies that $\|(\psi - \hat{\psi}_h)'\| \to 0$ as $h \to 0$, which completes the proof. ■