AN ADAPTIVE FINITE ELEMENT ALGORITHM
WITH RELIABLE AND EFFICIENT ERROR CONTROL
FOR LINEAR PARABOLIC PROBLEMS

ZHIMING CHEN AND JIA FENG

ABSTRACT. An efficient and reliable a posteriori error estimate is derived for
linear parabolic equations which does not depend on any regularity assump-
tion on the underlying elliptic operator. An adaptive algorithm with variable
time-step sizes and space meshes is proposed and studied which, at each time
step, delays the mesh coarsening until the final iteration of the adaptive pro-
cedure, allowing only mesh and time-step size refinements before. It is proved
that at each time step the adaptive algorithm is able to reduce the error indi-
cators (and thus the error) below any given tolerance within a finite number
of iteration steps. The key ingredient in the analysis is a new coarsening strat-
egy. Numerical results are presented to show the competitive behavior of the
proposed adaptive algorithm.

1. INTRODUCTION

A posteriori error estimates are computable quantities in terms of the discrete
solution and data that measure the actual discrete errors without the knowledge of
exact solutions. They are essential in designing algorithms for mesh and time-step
size modifications which equidistribute the computational effort and optimize the
computation. Ever since the pioneering work of Babuška and Rheinboldt [4], the
adaptive finite element methods based on a posteriori error estimates have become
a central theme in scientific and engineering computations. There are considerable
efforts in the literature devoted to the development of efficient adaptive algorithms
for various linear and nonlinear parabolic partial differential equations. In par-
ticular, a posteriori error estimates are derived in Bieterman and Babuška [2], [3]
and Moore [14] for one-dimensional and in Eriksson and Johnson [12], [13], Picasso
[17], and Verfürth [20] for multidimensional linear and mildly nonlinear parabolic
problems. Efficient adaptive procedures based on a posteriori error estimates are
also developed in Chen and Dai [5], Chen, Nochetto and Schmidt [7], Nochetto,
Schmidt and Verdi [16] for solving nonlinear partial differential equations arising
from physical and industrial processes.

Let Ω be a polyhedron domain in $\mathbf{R}^d(d = 1, 2, 3)$, $\Gamma = \partial\Omega$, and $T > 0$. We consider the following linear parabolic equation:

$$(1.1) \qquad \begin{aligned} &\frac{\partial u}{\partial t} - \operatorname{div}(a(x)\nabla u) = f \quad \text{in } \Omega \times (0, T), \\ &u = 0 \quad \text{on } \Gamma \times (0, T), \quad u(x, 0) = u_0(x) \quad \text{in } \Omega, \end{aligned}$$

where $f \in L^2(0, T; L^2(\Omega))$, $u_0 \in L^2(\Omega)$, and $a(x)$ is a piecewise constant function. Let τ_n be the n-th time-step size and $t^n = \sum_{i=1}^n \tau_i$. Denote by $\mathcal{M}^n$ a uniformly regular partition of Ω into simplexes such that $a(x)$ is a constant on each element $K \in \mathcal{M}^n$. We use refinement/coarsening procedures based on the bisection algorithm (cf., e.g., Bänsch [1]), which lead to compatible consecutive meshes whose minimum angles are bounded uniformly away from zero. Two meshes are compatible if one is the local refinement by bisection of the other. Let V^n indicate the usual space of conforming linear finite elements over $\mathcal{M}^n$ and $V_0^n = V^n \cap H_0^1(\Omega)$. Let $U_h^0 = P_0 u_0$, where $P_0 : L^2(\Omega) \to V_0^0$ is the L^2 projection operator into the linear finite element space V_0^0 over the initial mesh $\mathcal{M}^0$. We consider the following simple fully discrete finite element scheme for solving (1.1). For $n = 1, 2, \ldots$, find $U_h^n \in V_0^n$ such that

$$(1.2) \qquad \left\langle \frac{U_h^n - U_h^{n-1}}{\tau_n}, v \right\rangle + \langle a\nabla U_h^n, \nabla v \rangle = \langle \bar{f}^n, v \rangle \quad \forall v \in V_0^n,$$

where $\langle \cdot, \cdot \rangle$ stands for the inner product on $L^2(\Omega)$, and $\bar{f}^n = \frac{1}{\tau_n} \int_{t^{n-1}}^{t^n} f(x, t)\, dt$.

The main tool in deriving a posteriori error estimates in [5], [7], [12], [13], [16] is the analysis of linear dual problems of the corresponding *error* equations. The derived a posteriori error estimates, however, depend on the H^2 regularity assumption on the underlying elliptic operator. Without using this regularity assumption, the energy method is used in [17] to derive an a posteriori error estimate for the total energy error of the approximate solution. A lower bound for the local error is also derived for the associated a posteriori error indicator in that paper. We remark that while the energy method applies to less regular solutions than the duality method, the latter yields a better rate in the case of regular solutions. Also the norms that the error is being measured in are different. The duality technique applies to lower order norms and yields better results for them when the underlying elliptic operators are regular.

The results in [17] are obtained under the conditions that (1) the space-time mesh $\{(t^{n-1}, t^n) \times K : K \in \mathcal{M}^n\}$ is regular in the sense of [9], and (2) the finite element spaces are nested $V^{n-1} \subset V^n$; that is, no mesh coarsening is allowed. In this paper, without these two conditions we derive an a posteriori error estimate for the total energy error based on a direct energy estimate argument which has been used in Chen, Nochetto and Schmidt [8] for the phase relaxation model, a system of one parabolic equation coupled with one variational inequality. This energy estimate argument is slightly different from that of [17]. We also prove a lower bound for the local error for our a posteriori error indicator without using the space-mesh regularity assumption in [17].

Let $U^n \in H_0^1(\Omega)$ be the solution of the following semidiscrete problem

$$(1.3) \qquad \left\langle \frac{U^n - U^{n-1}}{\tau_n}, \varphi \right\rangle + \langle a\nabla U^n, \nabla \varphi \rangle = \langle \bar{f}^n, \varphi \rangle \quad \forall \varphi \in H_0^1(\Omega).$$

We observe that by modifying the time-step size τ_n, we are essentially controlling the error between $u^n = u(x, t^n)$ and U^n, not between u^n and U_h^n. Moreover, if we let $U_*^n \in H_0^1(\Omega)$ be the solution of the auxiliary problem

$$(1.4) \qquad \left\langle \frac{U_*^n - U_h^{n-1}}{\tau_n}, \varphi \right\rangle + \langle a\nabla U_*^n, \nabla\varphi \rangle = \langle \bar{f}^n, \varphi \rangle \quad \forall \varphi \in H_0^1(\Omega),$$

then we also observe that for fixed time-step size τ_n, by adapting the mesh $\mathcal{M}^n$ we are essentially controlling the error between U_h^n and U_*^n, not between U_h^n and U^n (or the exact solution u). These two observations play an important role in the subsequent analysis to prove the lower bound for the space a posteriori error indicator and in the development of the refinement/coarsening strategy.

Based on the local error indicators, the usual adaptive algorithm solving the parabolic problem (1.1) at the n-th time step reads as

$$\text{Solve} \;\to\; \text{Estimate} \;\to\; \text{Refine/Coarsen}.$$

Here the refinement/coarsening procedure includes both the mesh and time-step size modifications. There are several possibilities proposed in the literature on the strategies of the adaptive control of meshes and time-step sizes. We refer to Schmidt and Siebert [18] for a nice review of various adaptive algorithms and their implementation details. In this paper, at each time step n, we propose the following algorithm to modify the time-step size τ_n and mesh $\mathcal{M}^n$ starting from the initial time-step size $\tau_{n,0} = \tau_{n-1}$ and initial mesh $\mathcal{M}^{n,0} = \mathcal{M}^{n-1}$:

1. Refine the time-step size $\tau_{n.0}$ and the mesh $\mathcal{M}^{n,0}$ so that for the solution $U_h^n \in V_0^n$ of (1.2) on the final mesh $\mathcal{M}^n$ with final time-step size τ_n, the associated space and time error indicators are less than the prescribed tolerances respectively;
2. Coarsen the mesh $\mathcal{M}^n$ so that for the solution U_H^n on the coarsened mesh $\mathcal{M}_H^n$, the associated coarsening error indicator is less than some prescribed tolerance;
3. Enlarge the initial time-step size $\tau_{n+1,0}$ for next time step if the current time error indicator is much less than the tolerance.

We will extend the convergence analysis of adaptive finite element methods developed for linear elliptic problems in Dörfler [11] and Morin, Nochetto and Siebert [15] to prove that the iteration in Step 1 of the above algorithm terminates in a finite number of steps. The main difficulty now is the treatment of the oscillation of the residual $R^n = \bar{f}^n - (U_h^n - U_h^{n-1})/\tau_n$ which changes at each refinement procedure. The oscillation of any function $\varphi \in L^2(\Omega)$ over the mesh $\mathcal{M}^n$ is defined as

$$(1.5) \qquad \text{osc}(\varphi, \mathcal{M}^n) = \left(\sum_{K \in \mathcal{M}^n} h_K^2 \| \varphi - P_K\varphi \|_{L^2(K)}^2 \right)^{1/2},$$

where $P_K\varphi = \frac{1}{|K|} \int_K \varphi dx$. For fixed time step size τ_n, let $\mathcal{M}^{n,k+1}$ be a refinement of the mesh $\mathcal{M}^{n,k}$, and $U_k^{n,k+1}, U_h^{n,k}$ be the corresponding solutions of (1.2) over the meshes $\mathcal{M}^{n,k+1}, \mathcal{M}^{n,k}$. Note that since $R^{n,k+1} \neq R^{n,k}$, the following relation, which is crucial in the argument in [15], is no longer valid:

$$\text{osc}(R^{n,k+1}, \mathcal{M}^{n,k+1})^2 \leq \zeta \; \text{osc}(R^{n,k}, \mathcal{M}^{n,k})^2, \quad \text{for some constant } \zeta < 1.$$

Instead, we show in Lemma 4.5 that

$$(1.6) \quad \operatorname{osc}(R^{n,k+1}, \mathcal{M}^{n,k+1})^2 \leq \zeta_1 \operatorname{osc}(R^{n,k}, \mathcal{M}^{n,k})^2 + \zeta_2 \| U_h^{n,k+1} - U_h^{n,k} \|_{\tau_n,\Omega}^2$$

for some constants $0 < \zeta_1 < 1$ and $\zeta_2 > 0$. Here $\| \cdot \|_{\tau_n,\Omega}$ is the weighted norm of $H^1(\Omega)$ with parameter $\tau_n > 0$,

$$(1.7) \quad \| \varphi \|_{\tau_n,\Omega} = \left(\frac{1}{\tau_n} \| \varphi \|_{L^2(\Omega)}^2 + \|\varphi\|_\Omega^2 \right)^{1/2} \quad \forall \varphi \in H^1(\Omega)$$

with $\|\varphi\|_\Omega = \langle a\nabla\varphi, \nabla\varphi \rangle^{1/2}$, the energy seminorm of φ. The additional term $\| U_h^{n,k+1} - U_h^{n,k} \|_{\tau_n,\Omega}^2$ in (1.6) can be controlled by extending the technique in [15].

The choice of the coarsening error indicator and coarsening strategy in Step 2 is a subtle issue. The error incurred due to the over-coarsening can only be reduced by re-refining the coarsened mesh. Thus over-coarsening leads to the unnecessary solution of discrete problems, that is usually expensive and undesirable. Our coarsening error indicator is based on a direct control of the error between the coarsened discrete solution and the continuous solution U_*^n of (1.4). More precisely, let $\mathcal{M}_H^n$ be a coarsening of $\mathcal{M}^n$, and U_H^n, U_h^n be the corresponding solutions of (1.2) with fixed time-step size τ_n. Since $V_0^{n,H} \subset V_0^n$, we deduce by Galerkin orthogonality (see Theorem 3.1 below) that

$$\| U_*^n - U_H^n \|_{\tau_n,\Omega}^2 \leq \| U_*^n - U_h^h \|_{\tau_n,\Omega}^2 + \| U_h^n - I_H^n U_h^n \|_{\tau_n,\Omega}^2,$$

where $I_H^n : C(\bar{\Omega}) \to V^{n,H}$ is the usual linear finite element interpolant over $\mathcal{M}_H^n$. Our coarsening error indicator is then given by

$$\eta_{\text{coarse}}^n := \frac{1}{\tau_n} \| U_h^n - I_H^n U_h^n \|_{L^2(\Omega)}^2 + \| U_h^n - I_H^n U_h^n \|_\Omega^2,$$

which does not depend on the coarsened solution U_H^n. We will show that this direct error control leads to only one coarsening step.

The rest of the paper is organized as follows: In §2 we derive an a posteriori error estimate between the solution U_h^n of the discrete problem (1.2) and the solution u of the continuous problem (1.1). A lower bound for the local error in terms of our space error indicator will be established which indicates that over-refinement will not occur for the refinement procedure based on this indicator. In §3 we introduce our adaptive algorithm and justify its various steps. In §4 we prove that at each time step our adaptive algorithms are able to reduce the error indicators below any given tolerance within a finite number of iterations. In §5 we discuss the coarsening algorithm and present one numerical example to show the competitive behavior of the proposed adaptive algorithm.

2. A POSTERIORI ERROR ANALYSIS

We begin with introducing some notation. For any open set $G \subset \mathbf{R}^d$, we denote by $H^1(G)$ the standard Sobolev space of functions in $L^2(G)$ whose first derivatives are also in $L^2(G)$. The coefficient $a(x)$ is assumed to be piecewise constant and positive. Thus the seminorm $\| \cdot \|_G$ defined by $\|\varphi\|_G^2 = \langle a\nabla\varphi, \nabla\varphi \rangle_G$ is equivalent to the $H_0^1(\Omega)$ norm when $G = \Omega$. Here $\langle \cdot, \cdot \rangle_G$ stands for the $L^2(G)$ inner product or the duality pair between $H^{-1}(G)$ and $H_0^1(G)$. Throughout we write $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_\Omega$.

Given $f \in L^2(0,T;L^2(\Omega))$ and $u_0 \in L^2(\Omega)$, the weak formulation of (1.1) reads as follows. Find $u \in L^2(0,T;H_0^1(\Omega)) \cap H^1(0,T;H^{-1}(\Omega))$ such that $u(\cdot,0) = u_0(\cdot)$,

and for a.e. $t \in (0, T)$ the relation

$$(2.1) \qquad \left\langle \frac{\partial u}{\partial t}, \varphi \right\rangle + \langle a\nabla u, \nabla\varphi \rangle = \langle f, \varphi \rangle \quad \forall \varphi \in H_0^1(\Omega)$$

holds.

In this section we consider the a posteriori error analysis for the full time-space discretization in §2.1. The lower bound of the local error in terms of the space error indicator is derived in §2.2.

2.1. The upper bound.

We introduce the implicit Euler scheme for solving (2.1). Let τ_n be the n-th time-step size and set

$$t^n := \sum_{i=1}^{n} \tau_i, \quad \varphi^n(\cdot) = \varphi(\cdot, t^n)$$

for any function φ continuous in $(t^{n-1}, t^n]$. Let N be the total number of steps; that is $t^N \geq T$. At each time step n, $n = 1, 2, \ldots, N$, we denote by $\mathcal{M}^n$ a uniformly regular partition of Ω into simplexes which is obtained from $\mathcal{M}^{n-1}$ by using refinement/coarsening procedures. Given an element $K \in \mathcal{M}^n$, h_K stands for its diameter. Denote by $\mathcal{B}^n$ the collection of interior interelement sides e of $\mathcal{M}^n$ in Ω; h_e stands for the size of $e \in \mathcal{B}^n$. Let V^n indicate the usual space of conforming linear finite elements over $\mathcal{M}^n$, and $V_0^n = V^n \cap H_0^1(\Omega)$. Let $U_h^0 = P_0 u_0$, where $P_0 : L^2(\Omega) \to V_0^0$ is the L^2 projection operator into the linear finite element space V_0^0 over the initial mesh $\mathcal{M}^0$. Then the fully discrete finite element approximation at the n-th time step reads as follows. Given $U_h^{n-1} \in V_0^{n-1}$, then $\mathcal{M}^{n-1}$ and τ_{n-1} are modified as described below to give rise to $\mathcal{M}^n$ and τ_n and, thereafter, $U_h^n \in V_0^n$ computed according to the following prescription:

$$(2.2) \qquad \left\langle \frac{U_h^n - U_h^{n-1}}{\tau_n}, v \right\rangle + \langle a\nabla U_h^n, \nabla v \rangle = \langle \bar{f}^n, v \rangle \quad \forall v \in V_0^n.$$

We define the interior residual

$$R^n := \bar{f}^n - \frac{U_h^n - U_h^{n-1}}{\tau_n},$$

along with the jump residual across $e \in \mathcal{B}^n$

$$J_e^n := [\![a\nabla U_h^n]\!]_e \cdot \nu_e = (a\nabla U_h^n|_{K_1} - a\nabla U_h^n|_{K_2}) \cdot \nu_e, \quad e = \partial K_1 \cap \partial K_2,$$

using the convention that the unit normal vector ν_e to e points from K_2 to K_1. We observe that integration by parts implies

$$(2.3) \qquad \langle a\nabla U_h^n, \nabla\varphi \rangle = -\sum_{e \in \mathcal{B}^n} \int_e J_e^n \varphi ds \quad \forall \varphi \in H_0^1(\Omega).$$

Theorem 2.1. *For any integer $1 \leq m \leq N$, there exists a constant $C > 0$ depending only on the minimum angle of meshes $\mathcal{M}^n, n = 1, 2, \ldots, m$, and the coefficient*

a(x) such that the a posteriori error estimate

$$(2.4) \qquad \frac{1}{2}\| u^m - U_h^m \|^2_{L^2(\Omega)} + \sum_{n=1}^{m} \int_{t^{n-1}}^{t^n} \| u - U_h^n \|^2_\Omega dt$$

$$\leq \quad \| u_0 - U_h^0 \|^2_{L^2(\Omega)} + \sum_{n=1}^{m} \tau_n \eta^n_{\text{time}} + C \sum_{n=1}^{m} \tau_n \eta^n_{\text{space}}$$

$$+ \quad 2\Big(\sum_{n=1}^{m} \int_{t^{n-1}}^{t^n} \| f - \bar{f}^n \|_{L^2(\Omega)} dt \Big)^2$$

holds, where the time error indicator η^n_{time} and space error indicator η^n_{space} are given by

$$\eta^n_{\text{time}} = \frac{1}{3}\|U_h^n - U_h^{n-1}\|^2_\Omega, \qquad \eta^n_{\text{space}} = \sum_{e \in \mathcal{B}^n} \eta^n_e$$

with the local error indicator η^n_e defined as

$$\eta^n_e = \frac{1}{2} \sum_{K \in \Omega_e} h_K^2 \| R^n \|^2_{L^2(K)} + h_e\| J_e^n \|^2_{L^2(e)}.$$

Here Ω_e is the collection of two elements sharing the common side $e \in \mathcal{B}^n$.

Proof. For the sake of completeness, we sketch the proof here. From (2.2) we know that, for any $\varphi \in H_0^1(\Omega)$ and $v \in V_0^n$,

$$\Big\langle \frac{U_h^n - U_h^{n-1}}{\tau_n}, \varphi \Big\rangle + \langle a\nabla U_h^n, \nabla\varphi \rangle = -\langle R^n, \varphi - v \rangle + \langle a\nabla U_h^n, \nabla(\varphi - v) \rangle + \langle \bar{f}^n, \varphi \rangle.$$

For any $t \in (t^{n-1}, t^n]$, we denote by

$$U_h(t) = l(t)U_h^n + (1 - l(t))U_h^{n-1}, \qquad l(t) = (t - t^{n-1})/\tau_n.$$

Then from (2.1) and (2.5) we have, for a.e. $t \in (t^{n-1}, t^n]$, and for any $\varphi \in H_0^1(\Omega), v \in V_0^n$,

$$\Big\langle \frac{\partial(u - U_h)}{\partial t}, \varphi \Big\rangle + \langle a\nabla(u - U_h^n), \nabla\varphi \rangle = \langle R^n, \varphi - v \rangle - \langle a\nabla U_h^n, \nabla(\varphi - v) \rangle + \langle f - \bar{f}^n, \varphi \rangle.$$

Now we resort to the Clément interpolant $\Pi^n : H_0^1(\Omega) \to V_0^n$, which satisfies the following local approximation properties, for any $\varphi \in H_0^1(\Omega)$:

$$(2.5) \qquad \| \varphi - \Pi^n\varphi \|_{L^2(K)} + h_K\| \nabla(\varphi - \Pi^n\varphi) \|_{L^2(K)} \leq C^* h_K\| \nabla\varphi \|_{L^2(N(K))},$$

$$(2.6) \qquad \| \varphi - \Pi^n\varphi \|_{L^2(e)} \leq C^* h_e^{1/2}\| \nabla\varphi \|_{L^2(N(e))},$$

where $N(A)$ is the union of all elements in $\mathcal{M}^n$ surrounding the sets $A = K \in \mathcal{M}^n$ or $A = e \in \mathcal{B}^n$ (see Clément [10]). The constant C^* depends only on the minimum angle of mesh $\mathcal{M}^n$. Based on this interpolant, by taking $\varphi = u - U_h \in H_0^1(\Omega)$, $v = \Pi^n(u - U_h) \in V_0^n$, and using (2.3) and the identity

$$\langle a\nabla(u - U_h^n), \nabla(u - U_h) \rangle = \frac{1}{2}\|u - U_h^n\|^2_\Omega + \frac{1}{2}\|u - U_h\|^2_\Omega - \frac{1}{2}\|U_h - U_h^n\|^2_\Omega,$$

we deduce that

$$(2.7) \qquad \frac{1}{2}\frac{d}{dt}\| u - U_h \|_{L^2(\Omega)}^2 + \frac{1}{2}\|u - U_h^n\|_\Omega^2 + \frac{1}{2}\|u - U_h\|_\Omega^2$$

$$= \quad \frac{1}{2}\|U_h - U_h^n\|_\Omega^2 + \langle R^n, (u - U_h) - \Pi^n(u - U_h)\rangle$$

$$+ \quad \sum_{e\in\mathcal{B}^n}\int_e J_e^n[(u - U_h) - \Pi^n(u - U_h)]ds + \langle f - \bar{f}^n, u - U_h\rangle.$$

For any $t^* \in (t^{m-1}, t^m]$, by integrating (2.7) in time from 0 to t^*, using (2.5)-(2.6), and exploiting the standard argument in finite element a posteriori analysis, we have

$$\frac{1}{2}\| (u - U)(t^*) \|_{L^2(\Omega)}^2 + \frac{1}{2}\sum_{n=1}^m \int_{t^{n-1}}^{t^n \wedge t^*} \left(\|u - U_h^n\|_\Omega^2 + \|u - U_h\|_\Omega^2\right)dt$$

$$\leq \quad \frac{1}{2}\| u_0 - U_h^0 \|_{L^2(\Omega)}^2 + \frac{1}{2}\sum_{n=1}^m \int_{t^{n-1}}^{t^n} \|U_h - U_h^n\|_\Omega^2 dt$$

$$+ \quad C\sum_{n=1}^m \int_{t^{n-1}}^{t^n} (\eta_{\text{space}}^n)^{1/2}\|u - U_h\|_\Omega dt + \frac{1}{4}\max_{0\leq t\leq t^*}\| u - U_h \|_{L^2(\Omega)}^2$$

$$+ \quad \left(\sum_{n=1}^m \int_{t^{n-1}}^{t^n} \| f - \bar{f}^n \|_{L^2(\Omega)}dt\right)^2,$$

where $t^n \wedge t^* = \min(t^n, t^*)$. This implies the desired estimate (2.4) by using the fact that

$$\int_{t^{n-1}}^{t^n} \|U_h - U_h^n\|_\Omega^2 dt = \int_{t^{n-1}}^{t^n} (1 - l(t))^2\|U_h^n - U_h^{n-1}\|_\Omega^2 dt = \frac{1}{3}\tau_n\|U_h^n - U_h^{n-1}\|_\Omega^2.$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

To conclude this section we give a few remarks about the a posteriori error estimates derived.

Remark 2.1. Let $P^n : L^2(\Omega) \to V_0^n$ be the L^2 projection operator. The coarsening errors involving $U_h^{n-1} - P^n U_h^{n-1}$ which appeared in previous studies [7], [12], [16] are not present in our a posteriori error estimates. They are hidden in the space error term $\| h_n R^n \|_{L^2(\Omega)}^2$ and the time error term $\|U_h^n - U_h^{n-1}\|_\Omega^2$. Here h_n is the piecewise constant function which is equal to h_K on each $K \in \mathcal{M}^n$.

Remark 2.2. The constant C in the a posteriori error estimate (2.4) is proportional to the maximum jump of the coefficients $\Lambda = \max_{x\in\bar{\Omega}} a(x)/\min_{x\in\bar{\Omega}} a(x)$. This dependence is rather undesirable for strongly discontinuous coefficients and indeed can be removed by modifying the associated space error indicator (see Chen and Dai [6] for a study on a posteriori error analysis and adaptivity for elliptic problems with strongly discontinuous coefficients). In order to avoid unnecessary and inessential complications of the argument, we will not elaborate on this issue in this paper.

Remark 2.3. In our a posteriori error estimate at the n-th time step, the time discretization error is controlled by $\|U_h^n - U_h^{n-1}\|_\Omega$ and $\int_{t^{n-1}}^{t^n} \| f - \bar{f}^n \|_{L^2(\Omega)}dt$, which can only be reduced by reducing the time-step sizes τ_n. On the other hand, the time-step size τ_n essentially controls the semidiscretization error: the error between the exact solution u and the solution U^n of (1.3). Thus $\|U_h^n - U_h^{n-1}\|_\Omega$ is not a

good error indicator for time discretization unless the space-discretization error is sufficiently resolved. In the adaptive method for time-dependent problems, we must do space mesh and time-step size adaptation simultaneously. Ignoring either one of them may not provide correct error control of approximation to the problem.

Remark 2.4. For given $U_h^{n-1} \in V_0^{n-1}$, let $U_*^n \in H_0^1(\Omega)$ be the solution of the continuous problem (1.4). Then the space error indicator η_{space}^n controls only the error between U_h^n and U_*^n, not between U_h^n and U^n (or the exact solution u). This observation plays an important role in the subsequent analysis.

2.2. **A lower bound (space error).** The objective of this subsection is to prove the following estimate for the local error which ensures over-refinement will not occur for the refinement strategy based on our space error indicator. For any $K \in \mathcal{M}^n$ and $\varphi \in L^2(\Omega)$, we define $P_K \varphi = \frac{1}{|K|} \int_K \varphi dx$, the average of φ over K. For any $n = 1, 2, \ldots$, we also need the notation

$$\text{(2.8)} \qquad \hat{C}_n = \max_{K \in \mathcal{M}^n} (h_K^2/\tau_n), \quad h_K = \mathrm{diam}(K).$$

Theorem 2.2. *There exist constants $C_2, C_3 > 0$ depending only on the minimum angle of $\mathcal{M}^n$ and the coefficient $a(x)$ such that for any $e \in \mathcal{B}^n$, the following estimate holds:*

$$\text{(2.9)} \qquad \eta_e^n \ \leq \ C_2 \hat{C}_n \sum_{K \in \Omega_e} \left(\frac{1}{\tau_n} \| U_*^n - U_h^n \|_{L^2(K)}^2 + \| U_*^n - U_h^n \|_K^2 \right)$$

$$+ C_3 \sum_{K \in \Omega_e} h_K^2 \| R^n - P_K R^n \|_{L^2(K)}^2.$$

Before we proceed to the proof, we remark that the dependence of C_2, C_3 on the coefficient in the case of strongly discontinuous coefficients can be traced by the techniques developed in [6] for elliptic problems. We also remark that our form of the lower bound (2.9) differs from the one in [17], where the space error indicator is bounded by the sum of the energy-error and the time-error indicator $\| U_h^n - U_h^{n-1} \|_\Omega$.

Proof. The proof makes use of the idea in Verfürth [19]. For any $K \in \mathcal{M}^n$, let $\psi_K = (d+1)^{d+1} \lambda_1 \cdots \lambda_{d+1}$ be the bubble function, where $\lambda_1, \ldots, \lambda_{d+1}$ are the barycentric coordinate functions. By the standard scaling argument, we have the following inf-sup relation that holds for some constant β depending only on the minimum angle of $K \in \mathcal{M}^n$:

$$\inf_{v_h \in P_1(K)} \sup_{\varphi_h \in P_1(K)} \frac{\displaystyle\int_K v_h \varphi_h \psi_K dx}{\| \varphi_h \|_{L^2(K)} \| v_h \|_{L^2(K)}} \geq \beta > 0.$$

Thus there exists a function $\varphi^n \in P_1(K)$ with $\| \varphi^n \|_{L^2(K)} = 1$ such that

$$\beta \| P_K R^n \|_{L^2(K)}$$

$$\leq \int_K (P_K R^n) \psi_K \varphi^n dx$$

$$= \int_K (P_K R^n - R^n) \psi_K \varphi^n dx + \int_K \left(\bar{f}^n - \frac{U_h^n - U_h^{n-1}}{\tau_n} \right) \psi_K \varphi^n dx$$

$$= \int_K (P_K R^n - R^n) \psi_K \varphi^n dx + \int_K \frac{U_*^n - U_h^n}{\tau_n} \psi_K \varphi^n dx + \langle a \nabla U_*^n, \nabla(\psi_K \varphi^n) \rangle_K,$$

where we have used (1.4) in the last identity. Since $U_h^n \in P_1(K)$ and $\psi_K = 0$ on ∂K, simple integration by parts implies that $\langle a\nabla U_h^n, \nabla(\psi_K\varphi^n)\rangle_K = 0$. Thus, we have

$$
\begin{aligned}
\| P_K R^n \|_{L^2(K)} &\leq C\| R^n - P_K R^n \|_{L^2(K)} + C\tau_n^{-1}\| U_*^n - U_h^n \|_{L^2(K)} \\
&\quad + C\|U_*^n - U_h^n\|_K \|\psi_K\varphi^n\|_K.
\end{aligned}
$$

Since $\|\psi_K\varphi^n\|_K \leq Ch_K^{-1}$ by inverse estimate, we conclude by the definition of $\hat{C}_n$ in (2.8) that

$$
\begin{aligned}
\| P_K R^n \|_{L^2(K)} &\leq C\| R^n - P_K R^n \|_{L^2(K)} \\
&\quad + C\hat{C}_n^{1/2}h_K^{-1}\Big(\frac{1}{\tau_n}\| U_*^n - U_h^n \|_{L^2(K)}^2 + \|U_*^n - U_h^n\|_K^2\Big)^{1/2}.
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
h_K^2\| R^n \|_{L^2(K)}^2 &\leq Ch_K^2\| R^n - P_K R^n \|_{L^2(K)}^2 \\
&\quad + C\hat{C}_n\Big(\frac{1}{\tau_n}\| U_*^n - U_h^n \|_{L^2(K)}^2 + \|U_*^n - U_h^n\|_K^2\Big).
\end{aligned}
$$

For any $e \in \mathcal{B}^n$, let $\psi_e = d^d\lambda_1\cdots\lambda_d$ be the bubble function, where $\lambda_1,\ldots,\lambda_d$ are the barycentric coordinate functions associated with the nodes of e. Denote by $\psi^n = J_e^n\psi_e \in H_0^1(\Omega)$. Then, since J_e^n is constant on $e \in \mathcal{B}^n$, we get, after integration by parts, that

$$
\begin{aligned}
\| J_e^n \|_{L^2(e)}^2 &\leq C\int_e J_e^n\psi^n dx = -C\sum_{K\in\Omega_e}\int_K a(x)\nabla U_h^n\nabla\psi^n dx \\
&= C\sum_{K\in\Omega_e}\int_K a(x)\nabla(U_*^n - U_h^n)\nabla\psi^n dx - C\sum_{K\in\Omega_e}\int_K R^n\psi^n dx,
\end{aligned}
$$

where we have used the definition of U_*^n in (1.4). Moreover, it is easy to see that

$$
\| \nabla\psi^n \|_{L^2(K)} \leq Ch_e^{-1/2}\| J_e^n \|_{L^2(e)}, \qquad \| \psi^n \|_{L^2(K)} \leq Ch_e^{1/2}\| J_e^n \|_{L^2(e)} \quad \forall K \in \Omega_e.
$$

Thus

$$
h_e\| J_e^n \|_{L^2(e)}^2 \leq C\sum_{K\in\Omega_e}(h_K^2\| R^n \|_{L^2(K)}^2 + \|U_*^n - U_h^n\|_K^2).
$$

This completes the proof. $\qquad\square$

Summing up (2.9) for all $e \in \mathcal{B}^n$, we obtain that the following important estimate

$$
\tag{2.10} \eta_{\text{space}}^n \leq 2C_2\hat{C}_n\| U_*^n - U_h^n \|_{\tau_n,\Omega}^2 + 2C_3 \ \text{osc}(R^n, \mathcal{M}^n)^2,
$$

where the weighted norm $\| U_*^n - U_h^n \|_{\tau_n,\Omega}$ is defined in (1.7), and the oscillation of the residual $\text{osc}(R^n, \mathcal{M}^n)$ is defined in (1.5).

3. Adaptive algorithm

We start by considering the algorithm for time-step size control. The adjustment of the time-step size is based on the error equidistribution strategy: the time discretization error should be equally distributed to each time interval

$(t^{n-1}, t^n), n = 1, \ldots, N$. Let $\mathtt{TOL_{time}}$ be the total tolerance allowed for the part of a posteriori error estimate in (2.4) related to the time discretization; that is,

$$(3.1) \qquad \sum_{n=1}^{N} \tau_n \eta_{\text{time}}^n + 2 \Big(\sum_{n=1}^{N} \int_{t^{n-1}}^{t^n} \| f - \bar{f}^n \|_{L^2(\Omega)} dt \Big)^2 \leq \mathtt{TOL_{time}}.$$

A natural way to achieve (3.1) is to adjust the time-step size τ_n such that the following relations are satisfied:

$$(3.2) \qquad \eta_{\text{time}}^n \leq \frac{\mathtt{TOL_{time}}}{2T}, \qquad \frac{1}{\tau_n} \int_{t^{n-1}}^{t^n} \| f - \bar{f}^n \|_{L^2(\Omega)} dt \leq \frac{1}{2T} \sqrt{\mathtt{TOL_{time}}}.$$

The following algorithm can be used to control the time-step size at each time step n (cf., e.g., [18]).

Algorithm 3.1 (Time-step size control). Prescribe tolerance $\mathtt{TOL_{time}}$, parameters $\delta_1 \in (0, 1), \delta_2 > 1$, and $\theta_{\text{time}} \in (0, 1)$.

> $\tau_n := \tau_{n-1}$
>   solve the time discretization problem and estimate the error
>   while (3.2) is not satisfied do
> $\qquad \tau_n := \delta_1 \tau_n$
> $\qquad$ solve the time discretization problem and estimate the error
>   end while
>   if the relations

$$(3.3) \qquad \eta_{\text{time}}^n \leq \theta_{\text{time}} \frac{\mathtt{TOL_{time}}}{2T}, \qquad \frac{1}{\tau_n} \int_{t^{n-1}}^{t^n} \| f - \bar{f}^n \|_{L^2(\Omega)} dt \leq \frac{1}{2T} \sqrt{\theta_{\text{time}} \mathtt{TOL_{time}}}$$

>   are satisfied then
> $\qquad \tau_n := \delta_2 \tau_n$
>   end if

A good choice of the parameters in the algorithm above for the backward Euler scheme in time is to take $\delta_1 = 0.5, \delta_2 = 2$, and $\theta_{\text{time}} = 0.5$. Algorithm 3.1 controls only the time-step size, which must be combined with some algorithm for the mesh adaptation in practical applications. The main difficulty now is the mesh coarsening, which increases the error. If the mesh adaptation procedure iterates several times, it may occur that elements which were marked for coarsening in the beginning must be refined at the end to reduce the error, which is certainly undesirable. One possibility, which is proposed in [18], for overcoming this difficulty is to delay the mesh coarsening until the final iteration of the adaptive procedure, allowing refinements before only. Let $\mathtt{TOL_{space}}$ be the tolerance allowed for the part of the a posteriori error estimate in (2.4) related to the space discretization. Then the usual stopping criterion for the mesh adaptation is to satisfy the following relation at each time step n:

$$(3.4) \qquad\qquad\qquad\qquad \eta_{\text{space}}^n \leq \frac{\mathtt{TOL_{space}}}{T}.$$

This stopping rule is appropriate for mesh refinements but not for mesh coarsening. In this paper we will use a new coarsening error indicator based on the following theorem.

Theorem 3.1. *Given $U_h^{n-1} \in V_0^{n-1}$ and $\tau_n > 0$, let $\mathcal{M}_H^n$ be a coarsening of the mesh $\mathcal{M}^n$. Let $U_H^n \in V_0^{n,H}$, $U_h^n \in V_0^n$ be the solutions of the discrete problem (1.2) over meshes $\mathcal{M}_H^n$ and $\mathcal{M}^n$, respectively. Then the following error estimate is valid*

$$\| U_*^n - U_H^n \|_{\tau_n,\Omega}^2 \le \| U_*^n - U_h^n \|_{\tau_n,\Omega}^2 + \| U_h^n - I_H^n U_h^n \|_{\tau_n,\Omega}^2,$$

where $I_H^n : C(\bar{\Omega}) \to V^{n,H}$ is the standard linear finite element interpolant, and the weighted norm $\| \cdot \|_{\tau_n,\Omega}$ is defined in (1.7).

Proof. By definition, $U_H^n \in V_0^{n,H}$ and $U_h^n \in V_0^n$ satisfy the relations

$$(3.5) \qquad \left\langle \frac{U_H^n - U_h^{n-1}}{\tau_n}, v \right\rangle + \langle a\nabla U_H^n, \nabla v \rangle = \langle \bar{f}^n, v \rangle \quad \forall v \in V_0^{n,H},$$

$$(3.6) \qquad \left\langle \frac{U_h^n - U_h^{n-1}}{\tau_n}, v \right\rangle + \langle a\nabla U_h^n, \nabla v \rangle = \langle \bar{f}^n, v \rangle \quad \forall v \in V_0^n,$$

Since $\mathcal{M}_H^n$ is a coarsening of $\mathcal{M}^n$, we have $V_0^{n,H} \subset V_0^n$. Thus $U_H^n - U_h^n \in V_0^n$. Now the equation (3.6) together with (1.4) implies the Galerkin orthogonal identity

$$\left\langle \frac{U_*^n - U_h^n}{\tau_n}, U_H^n - U_h^n \right\rangle + \langle a\nabla(U_*^n - U_h^n), \nabla(U_H^n - U_h^n) \rangle = 0.$$

Hence

$$(3.7) \qquad \| U_*^n - U_H^n \|_{\tau_n,\Omega}^2 = \| U_*^n - U_h^n \|_{\tau_n,\Omega}^2 + \| U_H^n - U_h^n \|_{\tau_n,\Omega}^2.$$

Next, by subtracting (3.5) from (3.6) and taking $v = U_H^n - I_H^n U_h^n \in V_0^{n,H}$, we obtain the Galerkin orthogonal relation

$$\left\langle \frac{U_H^n - U_h^n}{\tau_n}, U_H^n - I_H^n U_h^n \right\rangle + \langle a\nabla(U_H^n - U_h^n), \nabla(U_H^n - I_H^n U_h^n) \rangle = 0,$$

which implies

$$\begin{aligned} \| U_H^n - U_h^n \|_{\tau_n,\Omega}^2 &= \| U_h^n - I_H^n U_h^n \|_{\tau_n,\Omega}^2 - \| U_H^n - I_H^n U_h^n \|_{\tau_n,\Omega}^2 \\ &\le \| U_h^n - I_H^n U_h^n \|_{\tau_n,\Omega}^2. \end{aligned}$$

This completes the proof by using (3.7). $\qquad\square$

Theorem 3.1 suggests that we introduce the coarsening error indicator

$$(3.8) \qquad \eta_{\text{coarse}}^n = \frac{1}{\tau_n} \| I_H^n U_h^n - U_h^n \|_{L^2(\Omega)}^2 + \| I_H^n U_h^n - U_h^n \|_\Omega^2.$$

The nice feature of this indicator is that it does not depend on U_H^n, the solution of the coarsened problem. This property allows us to coarsen only once, without checking whether the coarsened solution U_H^n satisfies some stopping criterion such as (3.4). Combining the ideas above, we arrive at the following adaptive algorithm for one single time step, which modifies the procedure proposed in [18].

Algorithm 3.2 (Time and space adaptive algorithm)**.** Prescribe tolerances TOL_{time}, $\text{TOL}_{\text{space}}$ and $\text{TOL}_{\text{coarse}}$, parameters $\delta_1 \in (0,1), \delta_2 > 1$ and $\theta_{\text{time}} \in (0,1)$. Let U_h^{n-1} be computed from the previous time step at time t^{n-1} with the mesh $\mathcal{M}^{n-1}$ and the time-step size τ_{n-1}.

    1. $\mathcal{M}^n := \mathcal{M}^{n-1}, \tau_n := \tau_{n-1}, t^n := t^{n-1} + \tau_n$
       solve the discrete problem for U_h^n on $\mathcal{M}^n$ using data U_h^{n-1}
       compute error estimates on $\mathcal{M}^n$

2. while (3.2) is not satisfied do
     $\tau_n := \delta_1 \tau_n$, $t^n := t^{n-1} + \tau_n$
     solve the discrete problem for U_h^n on $\mathcal{M}^n$ using data U_h^{n-1}
     compute error estimates on $\mathcal{M}^n$
   end while
3. while $\eta_{\mathrm{space}}^n > \mathtt{TOL}_{\mathrm{space}}/T$ do
     refine mesh $\mathcal{M}^n$ producing a modified $\mathcal{M}^n$
     solve the discrete problem for U_h^n on $\mathcal{M}^n$ using data U_h^{n-1}
     compute error estimates on $\mathcal{M}^n$
     while (3.2) is not satisfied do
       $\tau_n := \delta_1 \tau_n$, $t^n := t^{n-1} + \tau_n$
       solve the discrete problem for U_h^n on $\mathcal{M}^n$ using data U_h^{n-1}
       compute error estimates on $\mathcal{M}^n$
     end while
   end while
4. coarsen $\mathcal{M}^n$ producing a modified mesh $\mathcal{M}^n$ according to $\eta_{\mathrm{coarse}}^n \leq \dfrac{\mathtt{TOL}_{\mathrm{coarse}}}{T}$
   solve the discrete problem for U_h^n on $\mathcal{M}^n$ using data U_h^{n-1}
5. if (3.3) is satisfied then
     $\tau_n := \delta_2 \tau_n$
   end if

The goal of the first three steps in the algorithm above is to reduce the time-step size and refine the mesh so that the time and space error indicators become smaller than the respective tolerances. We achieve this goal by first reducing the time-step size to have the time error estimate below the tolerance while keeping the mesh unchanged. In Step 5, when the time error indicator is much smaller than the tolerance, the step size is enlarged (coarsened) by a factor $\delta_2 > 1$. In this case, the actual time step is not recalculated; only the initial time-step size for the next time step is changed.

We have the following theorem which guarantees the reliability of the algorithm above in terms of error control.

Theorem 3.2. *For $n \geq 1$, assume Algorithm 3.2 terminates and generates the final mesh $\mathcal{M}_H^n$, time-step size τ_n, and the the corresponding discrete solution U_H^n. The mesh $\mathcal{M}_H^n$ is coarsened from the mesh $\mathcal{M}^n$ produced by the first three steps. Then for any integer $1 \leq m \leq N$, there exists a constant C depending only on the minimum angles of $\mathcal{M}^n, n = 1, 2, \ldots, m$, and the coefficient $a(x)$ such that the estimate*

$$(3.9) \quad \frac{1}{2} \| u^m - U_H^m \|_{L^2(\Omega)}^2 + \sum_{n=1}^m \int_{t^{n-1}}^{t^n} \| u - U_H^n \|_\Omega^2 \, dt$$

$$\leq \| u_0 - U_h^0 \|_{L^2(\Omega)}^2 + \frac{t^m}{T} \mathtt{TOL}_{\mathrm{time}} + C \frac{t^m}{T} \mathtt{TOL}_{\mathrm{space}} + C \hat{C}_H^m \frac{t^m}{T} \mathtt{TOL}_{\mathrm{coarse}}.$$

holds, where $\hat{C}_H^m = \max\{h_K^2 / \tau_n : K \in \mathcal{M}_H^n, n = 1, 2, \ldots, m\}$.

Proof. Let U_h^n be the solution of the discrete problem (2.2) over the mesh $\mathcal{M}^n$ and with the time-step size τ_n. Then upon the termination of Algorithm 3.2 we have that

$$\eta_{\mathrm{time}}^n \leq \frac{\mathtt{TOL}_{\mathrm{time}}}{2T}, \quad \frac{1}{\tau_n} \int_{t^{n-1}}^{t^n} \| f - \bar{f}^n \|_{L^2(\Omega)} \, dt \leq \frac{1}{2T} \sqrt{\mathtt{TOL}_{\mathrm{time}}}, \quad \eta_{\mathrm{space}}^n \leq \frac{\mathtt{TOL}_{\mathrm{space}}}{T}.$$

From (2.2) we know that for any $\varphi \in H_0^1(\Omega)$

$$(3.10) \qquad \left\langle \frac{U_H^n - U_h^{n-1}}{\tau_n}, \varphi \right\rangle + \langle a\nabla U_H^n, \nabla\varphi \rangle$$

$$= \left\langle \frac{U_H^n - U_h^n}{\tau_n}, \varphi \right\rangle + \langle a\nabla(U_H^n - U_h^n), \nabla\varphi \rangle$$

$$- \langle R^n, \varphi \rangle + \langle a\nabla U_h^n, \nabla\varphi \rangle + \langle \bar{f}^n, \varphi \rangle.$$

Since $\mathcal{M}_H^n$ is a coarsening of $\mathcal{M}^n$, by the Galerkin orthogonal relation, as in Theorem 3.1, we have

$$\left\langle \frac{U_H^n - U_h^n}{\tau_n}, v_H \right\rangle + \langle a\nabla(U_H^n - U_h^n), \nabla v_H \rangle = 0 \quad \forall v_H \in V_0^{n,H}.$$

On the other hand, since U_h^n is the discrete solution over mesh $\mathcal{M}^n$, we have

$$-\langle R^n, v \rangle + \langle a\nabla U_h^n, \nabla v \rangle = 0 \quad \forall v \in V_0^n.$$

Thus from (2.1) and (3.10) we deduce that for a.e. $t \in (t^{n-1}, t^n]$ and for any $\varphi \in H_0^1(\Omega)$, $v_H \in V_0^{n,H}$, $v \in V_0^n$,

$$\left\langle \frac{\partial(u - U_H)}{\partial t}, \varphi \right\rangle + \langle a\nabla(u - U_H^n), \nabla\varphi \rangle$$

$$= \langle R^n, \varphi - v \rangle - \langle a\nabla U_h^n, \nabla(\varphi - v) \rangle + \langle f - \bar{f}^n, \varphi \rangle$$

$$- \left\langle \frac{U_H^n - U_h^n}{\tau_n}, \varphi - v_H \right\rangle - \langle a\nabla(U_H^n - U_h^n), \nabla(\varphi - v_H) \rangle,$$

where for any $t \in (t^{n-1}, t^n]$, $U_H(t) = l(t)U_H^n + (1 - l(t))U_h^{n-1}$ with $l(t) = (t - t^{n-1})/\tau_n$. By taking $v_H = \Pi_H^n\varphi \in V_0^{n,H}$, the Clément interpolation of $\varphi \in H_0^1(\Omega)$ in $V_0^{n,H}$, we get, after using the estimate (2.5) for the Clément interpolant, that

$$\left| \left\langle \frac{U_H^n - U_h^n}{\tau_n}, \varphi - \Pi_H^n\varphi \right\rangle + \langle a\nabla(U_H^n - U_h^n), \nabla(\varphi - \Pi_H^n\varphi) \rangle \right|$$

$$\leq C\Big(\sum_{K \in \mathcal{M}_H^n} h_K^2 \tau_n^{-2} \| U_H^n - U_h^n \|_{L^2(K)}^2 + \| U_H^n - U_h^n \|_\Omega^2 \Big)^{1/2} \|\varphi\|_\Omega$$

$$\leq C(\hat{C}_H^m)^{1/2} \| U_H^n - U_h^n \|_{\tau_n, \Omega} \|\varphi\|_\Omega.$$

Again, since $\mathcal{M}_H^n$ is a coarsening of $\mathcal{M}^n$, from the proof of Theorem 3.1 and Step 4 in Algorithm 3.2 we know that

$$\| U_H^n - U_h^n \|_{\tau_n, \Omega} \leq \| I_H^n U_h^n - U_h^n \|_{\tau_n, \Omega} \leq (\eta_{\text{coarse}}^n)^{1/2} \leq \sqrt{\frac{\text{TOL}_{\text{coarse}}}{T}},$$

which yields

$$\left| \left\langle \frac{U_H^n - U_h^n}{\tau_n}, \varphi - \Pi_H^n\varphi \right\rangle + \langle a\nabla(U_H^n - U_h^n), \nabla(\varphi - \Pi_H^n\varphi) \rangle \right| \leq C\sqrt{\frac{\hat{C}_H^m \text{TOL}_{\text{coarse}}}{T}} \|\varphi\|_\Omega.$$

The rest of the proof is similar to that of Theorem 2.1. Here we omit the details. $\quad \square$

In practical computations, it is natural to choose the coarsening tolerance $\text{TOL}_{\text{coarse}}$ much smaller than the space tolerance $\text{TOL}_{\text{space}}$. However, the additional factor $\hat{C}_H^m$ in estimate (3.9) suggests that the ratio between the coarsening tolerance and the time tolerance should also be small; see Remark 4.7.

4. TERMINATION OF THE ADAPTIVE ALGORITHM

The goal of this section is to show that for any given tolerances $\mathtt{TOL_{time}}, \mathtt{TOL_{space}}$ and $\mathtt{TOL_{coarse}}$, Algorithm 3.2 will terminate in a finite number of steps. This is a non-standard result in finite element theory for parabolic problems as it does not require the underlying finite element mesh size and the time-step size tends to zero. Since the algorithm only coarsens the mesh once in Step 4, we need only consider the termination of the refinements in the first three steps.

Let $\tau_{n,k}$ and $\mathcal{M}^{n,k}$ be the time-step size and mesh at the k-th iteration produced by the first three steps in the algorithm. The associated finite element space and discrete solution are denoted by $V^{n,k}$ and $U_h^{n,k}$, respectively. Let $t^{n,k} = t^{n-1} + \tau_{n,k}$. Throughout this section we will use the notation that $\phi^{n,k} = \phi(x, t^{n,k})$ if ϕ is continuous in $(t^{n-1}, t^n]$ and $\bar{\phi}^{n,k} = \frac{1}{\tau_{n,k}} \int_{t^{n-1}}^{t^{n,k}} \phi(x,t)\, dt$ if $\phi \in L^2(t^{n-1}, t^{n,k}; L^2(\Omega))$.

By (1.2) we know that $U_h^{n,k} \in V_0^{n,k} = V^{n,k} \cap H_0^1(\Omega)$ satisfies

$$(4.1) \qquad \left\langle \frac{U_h^{n,k} - U_h^{n-1}}{\tau_{n,k}}, v \right\rangle + \langle a\nabla U_h^{n,k}, \nabla v \rangle = \langle \bar{f}^{n,k}, v \rangle \quad \forall v \in V_0^{n,k}.$$

From Algorithm 3.2 we know that $\mathcal{M}^{n,0} = \mathcal{M}^{n-1}$, $\tau_{n,0} = \tau_{n-1}$ and

$$\mathcal{M}^{n,k} \text{ is a refinement of } \mathcal{M}^{n,k-1}, \quad \tau_{n,k} = \delta_1 \tau_{n,k-1} \text{ or } \tau_{n,k} = \tau_{n,k-1} \quad \forall k \geq 1.$$

4.1. Termination of time-step size refinements.
Our first objective is to show that the iteration to find the time-step size in Algorithm 3.2 can be terminated in a finite number of steps. To begin with, we recall the following result.

Lemma 4.1. *Assume that $f \in L^2(0, T; L^2(\Omega))$ and $w_0 \in H_0^1(\Omega)$. Let w be the solution of the linear parabolic problem*

$$(4.2) \qquad \begin{aligned} \frac{\partial w}{\partial t} - \mathrm{div}(a(x)\nabla w) &= f \quad \text{in } \Omega \times (0, T), \\ w = 0 \quad \text{on } \Gamma \times (0, T), & \quad w(x, 0) = w_0(x) \quad \text{in } \Omega. \end{aligned}$$

Let $\bar{w}(x,t) = \frac{1}{t}\int_0^t w(x,s)ds$. Then we have

$$\frac{1}{t}\| w(t) - w_0 \|_{L^2(\Omega)}^2 + \|\bar{w}(t) - w_0\|_\Omega^2 \to 0 \quad \text{as } t \to 0.$$

Proof. Since $w_0 \in H_0^1(\Omega)$, it is well known that $\partial_t w \in L^2(0, T; L^2(\Omega))$ and $w \in L^\infty(0, T; H_0^1(\Omega))$. Thus

$$(4.3) \qquad \frac{1}{t}\| \bar{w}(t) - w_0 \|_{L^2(\Omega)}^2 + \frac{1}{t}\| w(t) - w_0 \|_{L^2(\Omega)}^2$$

$$\leq 2\int_0^t \| \partial_t w \|_{L^2(\Omega)}^2 ds \to 0 \quad \text{as } t \to 0.$$

Multiplying (4.2) by $\partial_t w$ and integrating over $\Omega \times (0, t)$, we get

$$\int_0^t \left\| \frac{\partial w}{\partial t} \right\|_{L^2(\Omega)}^2 dt + \frac{1}{2}\|w(t)\|_\Omega^2 - \frac{1}{2}\|w_0\|_\Omega^2 = \int_0^t \left\langle f, \frac{\partial w}{\partial t} \right\rangle ds.$$

By letting $t \to 0$ in the equality above, we know that $\|w(t)\|_\Omega \to \|w_0\|_\Omega$ and thus $\|\bar{w}(t)\|_\Omega \to \|w_0\|_\Omega$ as $t \to 0$. Now we integrate (4.2) in time from 0 to t to obtain

$$(4.4) \qquad \left\langle \frac{w(t) - w_0}{t}, \varphi \right\rangle + \langle a\nabla\bar{w}(t), \nabla\varphi \rangle = \langle \bar{f}(t), \varphi \rangle \quad \forall \varphi \in H_0^1(\Omega).$$

By (4.3), we have

$$\frac{1}{t}\langle w(t) - w_0, \bar{w}(t) - w_0 \rangle \to 0 \quad \text{and} \quad \langle \bar{f}(t), \bar{w}(t) - w_0 \rangle \to 0 \quad \text{as } t \to 0.$$

Thus, by taking $\varphi = \bar{w} - w_0$ in (4.4) and letting $t \to 0$, we get

$$\lim_{t \to 0} \left(\frac{1}{2} \|\bar{w}(t)\|_{\Omega}^2 - \frac{1}{2} \|w_0\|_{\Omega}^2 + \frac{1}{2} \|\bar{w}(t) - w_0\|_{\Omega}^2 \right) = 0,$$

which yields $\|\bar{w}(t) - w_0\|_{\Omega} \to 0$ since $\|\bar{w}(t)\|_{\Omega} \to \|w_0\|_{\Omega}$ as $t \to 0$. This completes the proof. $\square$

Lemma 4.2. *Assume $f \in L^{\infty}(0, T; L^2(\Omega))$ and $\partial_t f \in L^{\infty}(0, T; L^2(\Omega))$. Then there exists an integer $L \geq 1$ depending only on the discrete solution U_h^{n-1} at time t^{n-1}, the mesh $\mathcal{M}^{n,0} = \mathcal{M}^{n-1}$, the source function f, and the coefficient $a(x)$ such that $\tau_{n,k} = \tau_{n,L}$ for any $k \geq L$; that is, at each time step, the algorithm to determine the time-step size can be terminated in a finite number of steps.*

Proof. Since $\partial_t f \in L^{\infty}(0, T; L^2(\Omega))$, we have

$$\frac{1}{\tau_{n,k}} \int_{t^{n-1}}^{t^{n,k}} \| f - \bar{f}^{n,k} \|_{L^2(\Omega)} dt \leq \tau_{n,k} \sup_{0 \leq t \leq T} \left\| \frac{\partial f}{\partial t} \right\|_{L^2(\Omega)}.$$

Thus for sufficiently small $\tau_{n,k}$, e.g.,

$$\tau_{n,k} \leq \frac{1}{2T} \sqrt{\text{TOL}_{\text{time}}} \left(\sup_{0 \leq t \leq T} \left\| \frac{\partial f}{\partial t} \right\|_{L^2(\Omega)} \right)^{-1},$$

the second inequality in (3.2) can be fulfilled. It remains to check the first inequality in (3.2). Let ψ be the solution of the parabolic problem

$$\frac{\partial \psi}{\partial t} - \text{div}(a(x)\nabla\psi) = f \quad \text{in } \Omega \times (t^{n-1}, T),$$
$$\psi = 0 \quad \text{on } \Gamma \times (t^{n-1}, T), \quad \psi(x, t^{n-1}) = U_h^{n-1}(x) \quad \text{on } \Omega.$$

Integrating the equation above in time from t^{n-1} to $t^{n-1} + \tau_{n,k}$, we know that

$$(4.5) \quad \left\langle \frac{\psi^{n,k} - U_h^{n-1}}{\tau_{n,k}}, \varphi \right\rangle + \langle a\nabla\bar{\psi}^{n,k}, \nabla\varphi \rangle = \langle \bar{f}^{n,k}, \varphi \rangle \quad \forall \varphi \in H_0^1(\Omega).$$

From Lemma 4.1 we have that for $\epsilon = \sqrt{3\text{TOL}_{\text{time}}/2T}$, there exists a constant $\tau > 0$ such that as long as $\tau_{n,k} \leq \tau$, we have

$$(4.6) \quad \left(\frac{1}{\tau_{n,k}} \| \psi^{n,k} - U_h^{n-1} \|_{L^2(\Omega)}^2 + \|\bar{\psi}^{n,k} - U_h^{n-1}\|_{\Omega}^2 \right)^{1/2} \leq \frac{\epsilon}{3}.$$

Now for any given $\tau_{n,k} > 0$, let $\hat{U}_h^{n,k} \in V_0^{n,0}$ be the solution of the problem

$$(4.7) \quad \left\langle \frac{\hat{U}_h^{n,k} - U_h^{n-1}}{\tau_{n,k}}, v \right\rangle + \langle a\nabla\hat{U}_h^{n,k}, \nabla v \rangle = \langle \bar{f}^{n,k}, v \rangle \quad \forall v \in V_0^{n,0}.$$

Since $\mathcal{M}^{n,k}$ is a refinement of $\mathcal{M}^{n,0}$, we have $U_h^{n,k} - \hat{U}_h^{n,k} \in V_0^{n,k}$. Subtracting (4.5) from (4.1) and taking the test function as $U_h^{n,k} - \hat{U}_h^{n,k}$, we then have

$$\left\langle \frac{\psi^{n,k} - U_h^{n,k}}{\tau_{n,k}}, U_h^{n,k} - \hat{U}_h^{n,k} \right\rangle + \langle a\nabla(\bar{\psi}^{n,k} - U_h^{n,k}), \nabla(U_h^{n,k} - \hat{U}_h^{n,k}) \rangle = 0.$$

Using the identity $(a-c)^2 = (a-b)^2 - (b-c)^2 + 2(a-c)(b-c)$ for any $a, b, c \in \mathbf{R}$, we obtain

$$(4.8) \qquad \frac{1}{\tau_{n,k}}\|\,\psi^{n,k} - U_h^{n,k}\,\|_{L^2(\Omega)}^2 + \|\bar\psi^{n,k} - U_h^{n,k}\|_\Omega^2$$

$$= \quad \frac{1}{\tau_{n,k}}\|\,\psi^{n,k} - \hat U_h^{n,k}\,\|_{L^2(\Omega)}^2 + \|\bar\psi^{n,k} - \hat U_h^{n,k}\|_\Omega^2$$

$$- \quad \frac{1}{\tau_{n,k}}\|\,\hat U_h^{n,k} - U_h^{n,k}\,\|_{L^2(\Omega)}^2 - \|\hat U_h^{n,k} - U_h^{n,k}\|_\Omega^2$$

$$\leq \quad \frac{1}{\tau_{n,k}}\|\,\psi^{n,k} - \hat U_h^{n,k}\,\|_{L^2(\Omega)}^2 + \|\bar\psi^{n,k} - \hat U_h^{n,k}\|_\Omega^2.$$

But, by the triangle inequality and (4.6), we have, for $\tau_{n,k} \leq \tau$,

$$(4.9) \qquad \left(\frac{1}{\tau_{n,k}}\|\,\psi^{n,k} - \hat U_h^{n,k}\,\|_{L^2(\Omega)}^2 + \|\bar\psi^{n,k} - \hat U_h^{n,k}\|_\Omega^2\right)^{1/2}$$

$$\leq \quad \left(\frac{1}{\tau_{n,k}}\|\,\psi^{n,k} - U_h^{n-1}\,\|_{L^2(\Omega)}^2 + \|\bar\psi^{n,k} - U_h^{n-1}\|_\Omega^2\right)^{1/2}$$

$$+ \quad \left(\frac{1}{\tau_{n,k}}\|\,\hat U_h^{n,k} - U_h^{n-1}\,\|_{L^2(\Omega)}^2 + \|\hat U_h^{n,k} - U_h^{n-1}\|_\Omega^2\right)^{1/2}$$

$$\leq \quad \frac{\epsilon}{3} + \left(\frac{1}{\tau_{n,k}}\|\,\hat U_h^{n,k} - U_h^{n-1}\,\|_{L^2(\Omega)}^2 + \|\hat U_h^{n,k} - U_h^{n-1}\|_\Omega^2\right)^{1/2}.$$

Thus, combining (4.6), (4.8) and (4.9) and using the triangle inequality, we have, for $\tau_{n,k} \leq \tau$,

$$(4.10) \qquad \|U_h^{n,k} - U_h^{n-1}\|_\Omega$$

$$\leq \quad \|U_h^{n,k} - \bar\psi^{n,k}\|_\Omega + \|\bar\psi^{n,k} - U_h^{n-1}\|_\Omega$$

$$\leq \quad \frac{2}{3}\epsilon + \left(\frac{1}{\tau_{n,k}}\|\,\hat U_h^{n,k} - U_h^{n-1}\,\|_{L^2(\Omega)}^2 + \|\hat U_h^{n,k} - U_h^{n-1}\|_\Omega^2\right)^{1/2}.$$

It remains to estimate the error $\hat U_h^{n,k} - U_h^{n-1}$. From (4.7) we know that, for any $v \in V_0^{n,0}$,

$$\left\langle \frac{\hat U_h^{n,k} - U_h^{n-1}}{\tau_{n,k}}, v\right\rangle + \langle a\nabla(\hat U_h^{n,k} - U_h^{n-1}), \nabla v\rangle = \langle \bar f^{n,k}, v\rangle - \langle a\nabla U_h^{n-1}, \nabla v\rangle.$$

Since $U_h^{n-1} \in V_0^{n-1} = V_0^{n,0}$, we can take $v = \hat U_h^{n,k} - U_h^{n-1} \in V_0^{n,0}$ in the identity above and obtain

$$(4.11) \qquad \frac{1}{\tau_{n,k}}\|\,\hat U_h^{n,k} - U_h^{n-1}\,\|_{L^2(\Omega)}^2 + \|\hat U_h^{n,k} - U_h^{n-1}\|_\Omega^2$$

$$= \quad \langle \bar f^{n,k}, \hat U_h^{n,k} - U_h^{n-1}\rangle - \langle a\nabla U_h^{n-1}, \nabla(\hat U_h^{n,k} - U_h^{n-1})\rangle.$$

By the Cauchy-Schwarz inequality we have

$$\langle \bar f^{n,k}, \hat U_h^{n,k} - U_h^{n-1}\rangle \quad \leq \quad \|\,\bar f^{n,k}\,\|_{L^2(\Omega)}\|\,\hat U_h^{n,k} - U_h^{n-1}\,\|_{L^2(\Omega)}$$

$$\leq \quad \frac{1}{4\tau_{n,k}}\|\,\hat U_h^{n,k} - U_h^{n-1}\,\|_{L^2(\Omega)}^2 + \tau_{n,k}\|\,\bar f^{n,k}\,\|_{L^2(\Omega)}^2.$$

By the inverse estimate, we get

$$\langle a\nabla U_h^{n-1}, \nabla(\hat{U}_h^{n,k} - U_h^{n-1})\rangle$$

$$\leq \sum_{K\in\mathcal{M}^{n,0}} \bar{C}h_K^{-1}\| a\nabla U_h^{n-1}\|_{L^2(K)}\|\hat{U}_h^{n,k} - U_h^{n-1}\|_{L^2(K)}$$

$$\leq \frac{1}{4\tau_{n,k}}\|\hat{U}_h^{n,k} - U_h^{n-1}\|_{L^2(\Omega)}^2 + \tau_{n,k}\bar{C}^2\sum_{K\in\mathcal{M}^{n,0}} h_K^{-2}\| a\nabla U_h^{n-1}\|_{L^2(K)}^2,$$

where $\bar{C}$ is a constant depending on the minimum angle of $\mathcal{M}^{n,0}$. Substituting the estimates above into (4.11), we finally get

$$(4.12)\qquad \frac{1}{2\tau_{n,k}}\|\hat{U}_h^{n,k} - U_h^{n-1}\|_{L^2(\Omega)}^2 + \|\hat{U}_h^{n,k} - U_h^{n-1}\|_{\Omega}^2$$

$$\leq \tau_{n,k}\Big(\| f\|_{L^\infty(0,T;L^2(\Omega))}^2 + \bar{C}^2\sum_{K\in\mathcal{M}^{n,0}} h_K^{-2}\| a\nabla U_h^{n-1}\|_{L^2(K)}^2\Big) =: \tau_{n,k}\Delta_n.$$

The quantity Δ_n depends only on $\mathcal{M}^{n,0}$, U_h^{n-1}, $f(x,t)$, and $a(x)$, which are fixed through the adaptive procedure at the n-th time step. Thus, for sufficiently small $\tau_{n,k}$ (e.g., $\tau_{n,k} \leq \epsilon^2/(18\Delta_n)$), we get

$$(4.13)\quad \frac{1}{\tau_{n,k}}\|\hat{U}_h^{n,k} - U_h^{n-1}\|_{L^2(\Omega)}^2 + \|\hat{U}_h^{n,k} - U_h^{n-1}\|_{\Omega}^2 \leq 2\tau_{n,k}\Delta_n \leq \frac{1}{9}\epsilon^2.$$

Substituting (4.13) into (4.10), we finally get that if $\tau_{n,k} \leq \min(\tau, \epsilon^2/(18\Delta_n))$, then

$$\eta_{\text{time}}^{n,k} = \frac{1}{3}\|U_h^{n,k} - U_h^{n-1}\|_{\Omega}^2 \leq \frac{1}{3}\epsilon^2 = \frac{\text{TOL}_{\text{time}}}{2T}.$$

This completes the proof. $\square$

To conclude this subsection, we remark that the condition $\partial_t f \in L^\infty(0,T;L^2(\Omega))$ can be weakened. In fact, Lemma 4.2 holds if f is piecewise smooth in time: f has jumps at $0 < T_1 < T_2 < \cdots < T_{I-1} < T$ and $\partial_t f \in L^\infty(T_{i-1},T_i;L^2(\Omega))$. Here $T_0 = 0$ and $T_I = T$. In this case, we require additionally that $T_i, i = 1,2\ldots,I$ are the nodes of the time discretization; that is, $T_i = t^{n_i}$ for some n_i, where $t^n, n = 0,\ldots,N$, is the partition of the time interval $[0,T]$.

4.2. **Termination of space mesh refinements.** Now we turn to the finite step termination of mesh refinements. By Lemma 4.2 we can, without loss of generality, fix the time-step size $\tau_{n,k} = \tau_{n,L}$. In the following we will write $\tau_{n,k} = \tau_n$. Let $\hat{C}_{n,k} = \max\{h_K^2/\tau_n : K \in \mathcal{M}^{n,k}\}$. Note that $\hat{C}_{n,k+1} \leq \hat{C}_{n,k}$ if $\mathcal{M}^{n,k+1}$ is a refinement of $\mathcal{M}^{n,k}$. Thus there exists a constant $\tilde{C}_n$ which is fixed in the procedure for doing mesh refinements such that

$$(4.14)\qquad\qquad\qquad \hat{C}_{n,k} \leq \tilde{C}_n \quad \forall k \geq L.$$

To begin with, we first introduce the following refinement strategy by Morin, Nochetto and Siebert [15] for linear elliptic problems. This strategy is an improvement of the Guaranteed Convergence Strategy of Dörfler [11].

MNS Refinement Strategy. Given two parameters $0 < \theta, \hat{\theta} < 1$, mesh $\mathcal{M}^{n,k}$, and solution $U_h^{n,k} \in V_0^{n,k}$ over the mesh $\mathcal{M}^{n,k}$,

1. Select a subset $\hat{\mathcal{B}}^{n,k}$ of sides in $\mathcal{B}^{n,k}$ such that

$$\Big(\sum_{e \in \hat{\mathcal{B}}^{n,k}} \eta_e^{n,k} \Big)^{1/2} \geq \theta \Big(\sum_{e \in \mathcal{B}^{n,k}} \eta_e^{n,k} \Big)^{1/2}.$$

2. Let $\hat{\mathcal{M}}^{n,k}$ be the set of elements with one side in $\hat{\mathcal{B}}^{n,k}$. Enlarge $\hat{\mathcal{M}}^{n,k}$ such that

$$\mathrm{osc}(R^{n,k}, \hat{\mathcal{M}}^{n,k}) \geq \hat{\theta} \ \mathrm{osc}(R^{n,k}, \mathcal{M}^{n,k}).$$

3. Refine every element in $\hat{\mathcal{M}}^{n,k}$ in such a way that a node is created in the interior of the element.

Here $\mathcal{B}^{n,k}$ is the collection of all interior interelement sides, $\eta_e^{n,k}$ is the local space error indicator associated with $e \in \mathcal{B}^{n,k}$

$$\eta_e^{n,k} = \frac{1}{2} \sum_{K \in \Omega_e} h_K^2 \| R^{n,k} \|_{L^2(K)}^2 + h_e \| J_e^{n,k} \|_{L^2(e)}^2$$

with $R^{n,k} = \bar{f}^n - (U_h^{n,k} - U_h^{n-1})/\tau_n$, $J_e^{n,k} = [\![a \nabla U_h^{n,k}]\!]_e \cdot \nu_e$, and Ω_e is the collection of two elements sharing e as a common side. The oscillations $\mathrm{osc}(R^{n,k}, \hat{\mathcal{M}}^{n,k})$ and $\mathrm{osc}(R^{n,k}, \mathcal{M}^{n,k})$ are defined as in (1.5). For a discussion on the choices of the parameters $\theta, \hat{\theta}$ and their influence on the performance of adaptive methods, we refer to [6]. Usually, a smaller θ leads to better performance.

The starting point of the analysis is the following orthogonality result whose proof is similar to that of (3.7) in Theorem 3.1.

Lemma 4.3. *If $\mathcal{M}^{n,k+1}$ is a refinement of $\mathcal{M}^{n,k}$ such that $V^{n,k} \subset V^{n,k+1}$, then the relation*

$$\| U_*^n - U_h^{n,k+1} \|_{\tau_n,\Omega}^2 = \| U_*^n - U_h^{n,k} \|_{\tau_n,\Omega}^2 - \| U_h^{n,k} - U_h^{n,k+1} \|_{\tau_n,\Omega}^2$$

holds. Here $\| \cdot \|_{\tau_n,\Omega}$ is defined in (1.7).

For any $K \in \mathcal{M}^{n,k}$ and $\varphi \in L^2(\Omega)$, define $P_K \varphi = \frac{1}{|K|} \int_K \varphi dx$. The following lemma can be proved exactly as in [15, Lemma 4.2] by using a similar idea as in the proof of Theorem 2.2. Here we omit the details.

Lemma 4.4. *Let $\mathcal{M}^{n,k+1}$ be a refinement of $\mathcal{M}^{n,k}$ according to the MNS Refinement Strategy. Then there exist constants $C_4, C_5 > 0$, depending only on the minimum angle of $\mathcal{M}^{n,k}$ and on the coefficient $a(x)$ such that, for all $e \in \hat{\mathcal{B}}^{n,k}$, we have*

$$\eta_e^{n,k} \leq C_4 \hat{C}_{n,k} \sum_{K \in \Omega_e} \| U_h^{n,k+1} - U_h^{n,k} \|_{\tau_n,K}^2 + C_5 \sum_{K \in \Omega_e} h_K^2 \| R^{n,k} - P_K R^{n,k} \|_{L^2(K)}^2,$$

where $\hat{C}_{n,k}$ is defined at the beginning of this subsection.

Now since $U_h^{n,k} \in V_0^{n,k}$ is the solution of the discrete problem

$$\Big\langle \frac{U_h^{n,k} - U_h^{n-1}}{\tau_n}, v \Big\rangle + \langle a \nabla U_h^{n,k}, \nabla v \rangle = \langle \bar{f}^n, v \rangle \quad \forall v \in V_0^{n,k},$$

which is an approximation of equation (1.4), we deduce by the standard a posteriori error analysis that there exists a constant $C_1 > 0$ such that

$$\| U_*^n - U_h^{n,k} \|_{\tau_n,\Omega}^2 \leq C_1 \sum_{e \in \mathcal{B}^{n,k}} \eta_e^{n,k}.$$

From Step 1 in the MNS Refinement Strategy, we know that

$$\theta^2 \| U_*^n - U_h^{n,k} \|_{\tau_n,\Omega}^2 \le \theta^2 C_1 \sum_{e \in \mathcal{B}^{n,k}} \eta_e^{n,k} \quad \le \quad C_1 \sum_{e \in \hat{\mathcal{B}}^{n,k}} \eta_e^{n,k},$$

which yields, upon using Lemma 4.4 and (4.14),

$$(4.15) \qquad \theta^2 \| U_*^n - U_h^{n,k} \|_{\tau_n,\Omega}^2 \quad \le \quad 2C_1 C_4 \tilde{C}_n \| U_h^{n,k+1} - U_h^{n,k} \|_{\tau_n,\Omega}^2$$
$$+ \quad 2C_1 C_5 \ \mathrm{osc}(R^{n,k}, \mathcal{M}^{n,k})^2.$$

Let $\gamma \in (0,1)$ be the reduction factor of element size associated with one refinement step. Given $\hat{\theta} \in (0,1)$, let $\hat{\alpha} = (1 - (1-\gamma^2)\hat{\theta}^2)^{1/2} < 1$. Then by Step 2 in the MNS Refinement Strategy we obtain, as in [15, Lemma 3.8], that

$$(4.16) \qquad \mathrm{osc}(R^{n,k}, \mathcal{M}^{n,k+1}) \le \hat{\alpha} \ \mathrm{osc}(R^{n,k}, \mathcal{M}^{n,k}).$$

On the other hand, since

$$R^{n,k+1} = \bar{f}^n - \frac{U_h^{n,k+1} - U_h^{n-1}}{\tau_n} = R^{n,k} - \frac{U_h^{n,k+1} - U_h^{n,k}}{\tau_n},$$

we deduce from Young's inequality and (4.16) that

$$(4.17) \quad \mathrm{osc}(R^{n,k+1}, \mathcal{M}^{n,k+1})^2$$

$$\le \quad (1+\delta)\,\mathrm{osc}(R^{n,k}, \mathcal{M}^{n,k+1})^2 + (1+\delta^{-1})\,\mathrm{osc}\Big(\frac{U_h^{n,k+1} - U_h^{n,k}}{\tau_n}, \mathcal{M}^{n,k+1}\Big)^2$$

$$\le \quad \hat{\alpha}^2(1+\delta)\,\mathrm{osc}(R^{n,k}, \mathcal{M}^{n,k})^2 + (1+\delta^{-1})\,\mathrm{osc}\Big(\frac{U_h^{n,k+1} - U_h^{n,k}}{\tau_n}, \mathcal{M}^{n,k+1}\Big)^2,$$

where $\delta > 0$ is some constant such that $\hat{\alpha}^2(1+\delta) < 1$.

Lemma 4.5. *Set* $\zeta_1 = \hat{\alpha}^2(1+\delta) < 1$ *and* $\zeta_2 = (1+\delta^{-1})\tilde{C}_n$. *Let* $\alpha, \beta \in (0,1)$ *be the constants defined by*

$$\beta = 1 - \frac{2}{2 + (1-\zeta_1)(1+\delta^{-1})^{-1}C_4 C_5^{-1}}, \qquad \alpha = \Big(1 - \frac{\theta^2 \beta}{2C_1 C_4 \tilde{C}_n}\Big)^{1/2}.$$

Then if $\mathcal{M}^{n,k+1}$ *is a refinement of* $\mathcal{M}^{n,k}$ *according to the MNS Refinement Strategy, the following estimate is valid:*

$$\| U_*^n - U_h^{n,k+1} \|_{\tau_n,\Omega}^2 + (1-\beta)\zeta_2^{-1}\,\mathrm{osc}(R^{n,k+1}, \mathcal{M}^{n,k+1})^2$$

$$\le \quad \max(\alpha^2, \tfrac{1}{2}(1+\zeta_1))\Big(\| U_*^n - U_h^{n,k} \|_{\tau_n,\Omega}^2 + (1-\beta)\zeta_2^{-1}\,\mathrm{osc}(R^{n,k}, \mathcal{M}^{n,k})^2\Big).$$

Proof. First we note that

$$\mathrm{osc}\Big(\frac{U_h^{n,k+1} - U_h^{n,k}}{\tau_n}, \mathcal{M}^{n,k+1}\Big)^2$$

$$= \quad \sum_{K \in \mathcal{M}^{n,k+1}} h_K^2 \tau_n^{-2} \| (U_h^{n,k+1} - U_h^{n,k}) - P_K(U_h^{n,k+1} - U_h^{n,k}) \|_{L^2(K)}^2$$

$$\le \quad \hat{C}_{n,k+1} \sum_{K \in \mathcal{M}^{n,k+1}} \tau_n^{-1} \| U_h^{n,k+1} - U_h^{n,k} \|_{L^2(K)}^2$$

$$\le \quad \tilde{C}_n \tau_n^{-1} \| U_h^{n,k+1} - U_h^{n,k} \|_{L^2(\Omega)}^2$$

$$\le \quad \tilde{C}_n \| U_h^{n,k+1} - U_h^{n,k} \|_{\tau_n,\Omega}^2.$$

Thus, by (4.17) we get

$$(4.18) \quad \operatorname{osc}(R^{n,k+1}, \mathcal{M}^{n,k+1})^2 \leq \zeta_1 \operatorname{osc}(R^{n,k}, \mathcal{M}^{n,k})^2 + \zeta_2 \| U_h^{n,k+1} - U_h^{n,k} \|_{\tau_n, \Omega}^2.$$

From (4.15) we obtain

$$\| U_h^{n,k+1} - U_h^{n,k} \|_{\tau_n, \Omega}^2 \geq \frac{\theta^2}{2C_1 C_4 \tilde{C}_n} \| U_*^n - U_h^{n,k} \|_{\tau_n, \Omega}^2 - \frac{C_5}{C_4 \tilde{C}_n} \operatorname{osc}(R^{n,k}, \mathcal{M}^{n,k})^2,$$

which, together with Lemma 4.3, yields

$$\begin{aligned}
\| U_*^n - U_h^{n,k+1} \|_{\tau_n, \Omega}^2 \\
= \| U_*^n - U_h^{n,k} \|_{\tau_n, \Omega}^2 \\
\quad - \beta \| U_h^{n,k+1} - U_h^{n,k} \|_{\tau_n, \Omega}^2 - (1 - \beta) \| U_h^{n,k+1} - U_h^{n,k} \|_{\tau_n, \Omega}^2 \\
\leq \alpha^2 \| U_*^n - U_h^{n,k} \|_{\tau_n, \Omega}^2 \\
\quad + \beta \frac{C_5}{C_4 \tilde{C}_n} \operatorname{osc}(R^{n,k}, \mathcal{M}^{n,k})^2 - (1 - \beta) \| U_h^{n,k+1} - U_h^{n,k} \|_{\tau_n, \Omega}^2.
\end{aligned}$$

Multiplying (4.18) by $\zeta_2^{-1}(1 - \beta)$ and adding it to the estimate above, we get

$$\begin{aligned}
& \| U_*^n - U_h^{n,k+1} \|_{\tau_n, \Omega}^2 + (1 - \beta)\zeta_2^{-1} \operatorname{osc}(R^{n,k+1}, \mathcal{M}^{n,k+1})^2 \\
\leq & \; \alpha^2 \| U_*^n - U_h^{n,k} \|_{\tau_n, \Omega}^2 + (\beta C_5 C_4^{-1} \tilde{C}_n^{-1} + (1 - \beta)\zeta_1 \zeta_2^{-1}) \operatorname{osc}(R^{n,k}, \mathcal{M}^{n,k})^2 \\
\leq & \; \max(\alpha^2, \mu)\Big(\| U_*^n - U_h^{n,k} \|_{\tau_n, \Omega}^2 + (1 - \beta)\zeta_2^{-1} \operatorname{osc}(R^{n,k}, \mathcal{M}^{n,k})^2 \Big),
\end{aligned}$$

where

$$\mu = \frac{\beta C_5 C_4^{-1} \tilde{C}_n^{-1} + (1 - \beta)\zeta_1 \zeta_2^{-1}}{(1 - \beta)\zeta_2^{-1}} = \zeta_1 + \frac{\beta}{1 - \beta}(1 + \delta^{-1})C_5 C_4^{-1}.$$

This completes the proof by observing that $\frac{\beta}{1-\beta} = \frac{1}{2}(1 - \zeta_1)(1 + \delta^{-1})^{-1} C_4 C_5^{-1}$. □

Now by the lower bound (2.10) proved in §2.2 and using (4.14), we know that

$$\begin{aligned}
\eta_{\text{space}}^{n,k+1} & \leq & 2C_2 \hat{C}_{n,k+1} \| U_*^n - U_h^{n,k+1} \|_{\tau_n, \Omega}^2 + 2C_3 \operatorname{osc}(R^{n,k+1}, \mathcal{M}^{n,k+1})^2 \\
& \leq & 2C_2 \tilde{C}_n \| U_*^n - U_h^{n,k+1} \|_{\tau_n, \Omega}^2 + 2C_3 \operatorname{osc}(R^{n,k+1}, \mathcal{M}^{n,k+1})^2.
\end{aligned}$$

Since $\alpha^2 < 1$ and $\frac{1}{2}(1 + \zeta_1) < 1$, we conclude from Lemma 4.5 that

$$\lim_{k \to \infty} \| U_*^n - U_h^{n,k} \|_{\tau_n, \Omega} = 0, \qquad \lim_{k \to \infty} \operatorname{osc}(R^{n,k}, \mathcal{M}^{n,k}) = 0.$$

Thus $\eta_{\text{space}}^{n,k} \to 0$ as $k \to \infty$. This together with Lemma 4.2 implies the following theorem, which is the main result of this section.

Theorem 4.6. *Assume that $f \in L^\infty(0, T; L^2(\Omega))$ and $\partial_t f \in L^\infty(0, T; L^2(\Omega))$. Then, at each time step $n \geq 1$, if the MNS Refinement Strategy is used in doing mesh refinements, Algorithm 3.2 will stop in a finite number of steps for any given tolerances $\mathtt{TOL_{time}}, \mathtt{TOL_{space}}$ and $\mathtt{TOL_{coarse}}$.*

Remark 4.7. The following observation, which is made to us by the referee of the paper, points out an important consequence of the analysis in this section on the choice of the coarsening tolerance and the time tolerance. If the exact solution has a positive energy and is within a good tolerance from the approximate solution, then

$\| a\nabla U_h^{n-1} \|_{L^2(\Omega)}^2 \geq c_0$ for some positive constant c_0. It follows from the definition of Δ_n in (4.12) that

$$\Delta_n \geq \frac{c_0\bar{C}^2}{h_{n,0}^2}, \qquad \text{where } h_{n,0} = \max_{K \in \mathcal{M}^{n,0}} h_K.$$

This yields, from the proof of Lemma 4.2, that

$$\tau_n = \tau_{n,L} \leq \frac{\epsilon^2}{18\Delta_n} \leq \frac{\epsilon^2 h_{n,0}^2}{18c_0\bar{C}^2} \leq \frac{h_{n,0}^2}{c_0\bar{C}^2 T} \text{ TOL}_{\text{time}}.$$

On the other hand, according to the definition of $\hat{C}_{n,0}$ at the beginning of §4.2, we have

$$h_{n,0}^2 \leq \hat{C}_{n-1}\tau_{n-1}.$$

This implies that the constant $\hat{C}_H^m$ in Theorem 3.2 satisfies

$$(4.19) \qquad \hat{C}_H^m = \max_{n \leq m+1} \hat{C}_{n-1} \geq \frac{c_0\bar{C}^2 T}{\text{TOL}_{\text{time}}} \max_{n \leq m+1} \frac{\tau_n}{\tau_{n-1}} \geq \frac{c_0\bar{C}^2 T}{\text{TOL}_{\text{time}}},$$

whenever $\max_{n \leq m+1} (\tau_n/\tau_{n-1}) \geq 1$. Notice that if the global algorithm is ever to terminate, that is, to reach time T with the estimated error below the prescribed tolerances, then it is necessary that $\max (\tau_n/\tau_{n-1}) \geq 1$. A comparison of (4.19) with estimate (3.9) implies that the ratio between the coarsening tolerance and the time tolerance has to be small. In other words, a small time tolerance puts a severe restriction on the coarsening tolerance if the overall error has to be small.

We also remark that Theorem 4.6 states that Algorithm 3.2 can be terminated in a finite number of iteration steps at any time step t^n starting from any given solution U_h^{n-1} at t^{n-1}. Whether the adaptive algorithm can reach the final time T is an open question which is not addressed in this paper and requires further research.

5. NUMERICAL EXPERIMENT

The implementation of Algorithm 3.2 is based on a 2D adaptive finite element solver for elliptic problems developed by Shibin Dai. The MNS Refinement Strategy is implemented as in [15, Algorithm 5.1]. We use the method indicated in Figure 1 to refine a marked element which creates an interior node.

Now we turn to the algorithm for mesh coarsening in Step 4 in Algorithm 3.2. Let K be some element which is obtained by dividing some parent element $\hat{K}$ into two elements using the bisection algorithm. Denote by $\hat{e}$ the edge of $\hat{K}$ which is divided, and $\hat{K}'$ the neighboring macro-element sharing the common edge $\hat{e}$ with
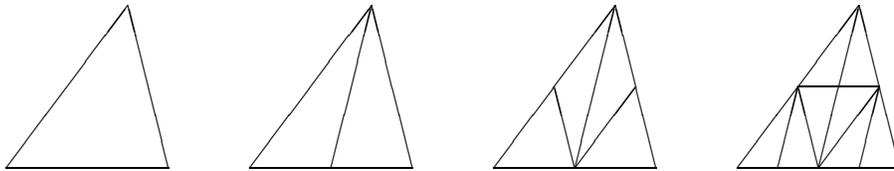

FIGURE 1. Refinement of a marked triangle which generates an interior node

$\hat{K}$. We define the patch of elements $\mathcal{P}(K)$ associated with K as the collection of all subelements included in $\hat{K}$ and $\hat{K}'$. We note that if K is marked for coarsening, then all elements in $\mathcal{P}(K)$ must also be marked for coarsening. For $n \geq 1$, we let $\mathcal{M}^n$ be the mesh obtained after doing the first three steps in Algorithm 3.2 which now is ready for possible coarsening, and U_h^n be the corresponding discrete solution. The elements in $\mathcal{M}^n$ which are obtained by refining some element (possibly many times) in $\mathcal{M}^{n,0} = \mathcal{M}^{n-1}$ are also not allowed for coarsening. All the elements in $\mathcal{M}^n$ which can be coarsened are called admissible coarsening elements. The collection of all admissible coarsening elements is denoted by $\hat{\mathcal{M}}^n$. For any $K \in \hat{\mathcal{M}}^n$, we denote by $\hat{K}$ its parent element such that K is obtained by refining $\hat{K}$ into two elements using the bisection algorithm. Then for each $K \in \hat{\mathcal{M}}^n$, we can compute the local coarsening error indicator as follows:

$$\eta_{\mathrm{coarse},K}^n = \frac{1}{\tau_n}\| U_h^n - \hat{U}_h^n \|_{L^2(K)}^2 + \|U_h^n - \hat{U}_h^n\|_K^2,$$

where $\hat{U}_h^n$ is the linear function on $\hat{K}$ whose values at the vertexes of $\hat{K}$ are the same as those of U_h^n. The guideline of the coarsening strategy is then to mark as many elements $K \in \hat{\mathcal{M}}^n$ as possible so that

$$\sum_{K\mathrm{marked}} \eta_{\mathrm{coarse},K}^n \leq \frac{\mathtt{TOL_{coarse}}}{T}$$

is satisfied. In a practical situation, it may happen that after all elements in $\hat{\mathcal{M}}^n$ marked for coarsening have been coarsened, the coarsening error is still below the tolerance; that is,

$$(5.1) \qquad \eta_{\mathrm{coarse}}^{n,1} = \frac{1}{\tau_n}\| U_h^n - I_H^{n,1}U_h^n \|_{L^2(\Omega)}^2 + \|U_h^n - I_H^{n,1}U_h^n\|_\Omega^2 < \frac{\mathtt{TOL_{coarse}}}{T},$$

where $\hat{\mathcal{M}}_H^{n,1}$ is the mesh produced by coarsening all the elements in $\mathcal{M}^n$ marked for coarsening, and $I_H^{n,1}$ is the associated linear finite element interpolant over $\mathcal{M}_H^{n,1}$. If $\mathcal{M}_H^n$ is a further coarsening of $\mathcal{M}_H^{n,1}$, then by the triangle inequality we have

$$\begin{aligned}
\| U_h^n - I_H^n U_h^n \|_{\tau_n,\Omega} &\leq \| U_h^n - I_H^{n,1}U_h^n \|_{\tau_n,\Omega} + \| I_H^{n,1}U_h^n - I_H^n U_h^n \|_{\tau_n,\Omega} \\
&= \| U_h^n - I_H^{n,1}U_h^n \|_{\tau_n,\Omega} + \| I_H^{n,1}U_h^n - I_H^n(I_H^{n,1}U_h^n) \|_{\tau_n,\Omega}.
\end{aligned}$$

Thus in the case of (5.1) we can set $\mathcal{M}^n = \mathcal{M}_H^{n,1}$, $U_h^n = I_H^{n,1}U_h^n$ and start the mesh coarsening iteration again. The algorithm to coarsen the mesh is terminated when either $\eta_{\mathrm{coarse}}^n \leq \mathtt{TOL_{coarse}}/T$ is violated or the marked coarsening set of $\mathcal{M}^n$ is empty. The following algorithm provides one possible implementation of the coarsening strategy.

Algorithm 5.1 (Mesh coarsening)**.** Prescribe the tolerance $\mathtt{TOL_{coarse}}$, the parameter $\nu \in (0,1)$, the mesh $\mathcal{M}^n$, and the solution U_h^n over the mesh $\mathcal{M}^n$.

  set $\sigma_1 = 0, \sigma_2 = 0, \gamma = 0, \rho = \mathtt{TOL_{coarse}}/4T$
  if $\sigma_2 < \rho$ do

      form the set $\hat{\mathcal{M}}^n$ of admissible coarsening elements
      compute $\eta_{\mathrm{coarse},K}^n$ for any $K \in \hat{\mathcal{M}}^n$
      $\eta_{\max} = \max\{\eta_{\mathrm{coarse},K}^n : K \in \hat{\mathcal{M}}^n\}$
      while $\gamma \leq 1$ and $\eta_{\max} > 0$ do
          $\gamma = \gamma + \nu$

for all K in $\hat{\mathcal{M}}^n$ do

    if K is not marked for coarsening

        if $\eta^n_{\mathrm{coarse},K} \leq \gamma\eta_{\max}$ and $K' \in \hat{\mathcal{M}}^n$ for any $K' \in \mathcal{P}(K)$

            $\sigma_2 = \sigma_2 + \sum_{K' \in \mathcal{P}(K)} \eta^n_{\mathrm{coarse},K'}$

            if $\sigma_2 \leq \rho$

                mark all elements in $\mathcal{P}(K)$ for coarsening

            else

                $\sigma_2 = \sigma_2 - \sum_{K' \in \mathcal{P}(K)} \eta^n_{\mathrm{coarse},K'}$

            end if

        end if

    end if

end for

end while

set $\sigma_1 = \sigma_1 + \sqrt{\sigma_2}$, $\sigma_2 = 0$

if $\eta_{\max} = 0$ do

    mark all elements in $\hat{\mathcal{M}}^n$ for coarsening

end if

set $\mathcal{M}^n$ as the mesh obtained by coarsening all elements marked for coarsening

set U^n_h as the linear interpolant of U^n_h on the coarsened mesh

if the number of elements before and after the coarsening is the same

    go to end

end if

$\rho = \max(0, \sqrt{\mathtt{TOL_{coarse}}/T} - \sigma_1)^2$

end if

if end

Now we report a numerical example computed by using Algorithm 3.2 with the MNS Refinement Strategy and Algorithm 5.1. We take the parameters $\delta_1 =$
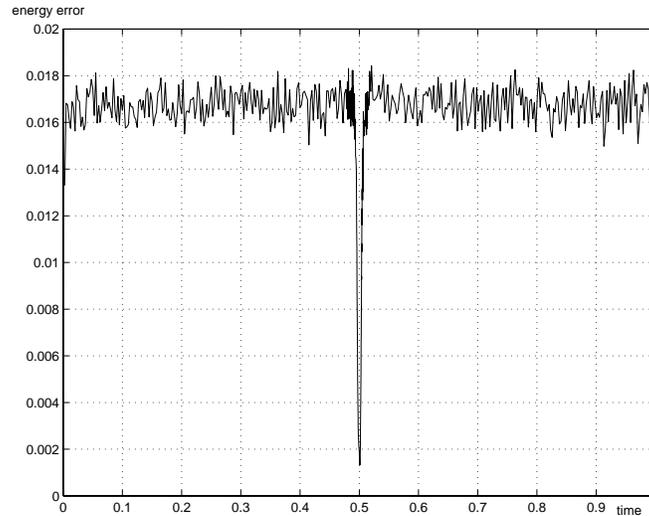

FIGURE 2. The energy error at each time step when $\mathtt{TOL_{space}} = 0.01$.

TABLE 5.1. The number of time steps N, the average number of nodes of the meshes $\mathrm{av}(\mathcal{M}^n)$, the total estimated error η, the total energy error $\mathcal{E}$, and the effectiveness index `eff` for different values of $\mathtt{TOL}_{\mathrm{space}}$.

$\mathtt{TOL}_{\mathrm{space}}$	N	$\mathrm{av}(\mathcal{M}^n)$	η	$\mathcal{E}$	`eff`
0.04	257	297	$2.095E-1$	$3.235E-2$	6.4740
0.01	512	786	$1.057E-1$	$1.671E-2$	6.3260
0.0025	1028	2736	$5.300E-2$	$8.470E-3$	6.2613
0.02	292	441	$1.580E-1$	$2.313E-2$	6.8326
0.005	587	1337	$7.971E-2$	$1.200E-2$	6.6421
0.00125	1175	4912	$3.985E-2$	$6.058E-3$	6.5784

$0.5, \delta_2 = 2, \theta_{\mathrm{time}} = 0.5$ in Algorithm 3.2, $\theta = 0.2, \hat{\theta} = 0.1$ in the MNS Refinement Strategy, and $\nu = 0.05$ in Algorithm 5.1. We set $\Omega = (-1,1) \times (-1,1)$ and $T = 1$. The exact solution of the continuous problem is prescribed as

$$u(x,t) = \beta(t) * e^{-[(x_1-t+0.5)^2+(x_2-t+0.5)^2]/0.04}$$

with

$$\beta(t) = 0.1 * (1 - e^{-10000*(t-0.5)^2}).$$

We observe that $\beta(t)$ drops exponentially around $t = 0.5$. The tolerances used are

(5.2) $\mathtt{TOL}_{\mathrm{time}} = \mathtt{TOL}_{\mathrm{space}}, \quad \mathtt{TOL}_{\mathrm{coarse}} = 0.03 * \mathtt{TOL}_{\mathrm{space}}.$

The effectiveness index `eff` of the a posteriori error estimate is defined as $\mathtt{eff} = \eta/\mathcal{E}$, where $\mathcal{E} = (\sum_{n=1}^{N} \tau_n \|u^n - U_h^n\|_\Omega^2)^{1/2}$ is the total energy error and η is the estimated error

$$\eta = \left\{ \sum_{n=1}^{N} \tau_n (\eta_{\mathrm{space}}^n + \eta_{\mathrm{time}}^n + \eta_{\mathrm{coarse}}^n) + 2 \left(\sum_{n=1}^{m} \int_{t^{n-1}}^{t^n} \| f - \bar{f}^n \|_{L^2(\Omega)} \right)^2 \right\}^{1/2}.$$
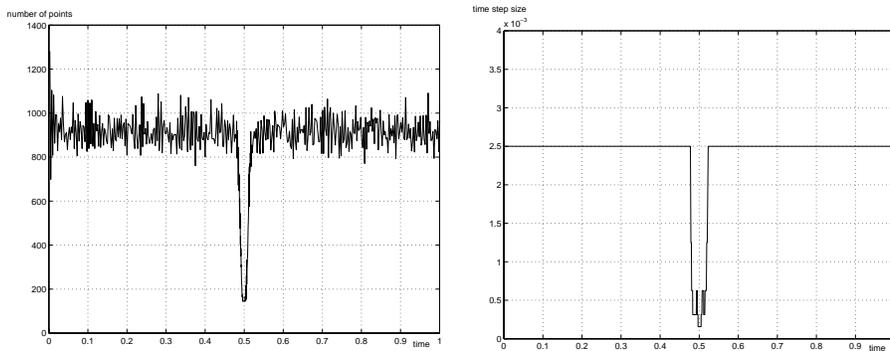

FIGURE 3. The number of nodes of $\mathcal{M}^n$ and the time-step size τ_n at each time step n when $\mathtt{TOL}_{\mathrm{space}} = 0.01$.
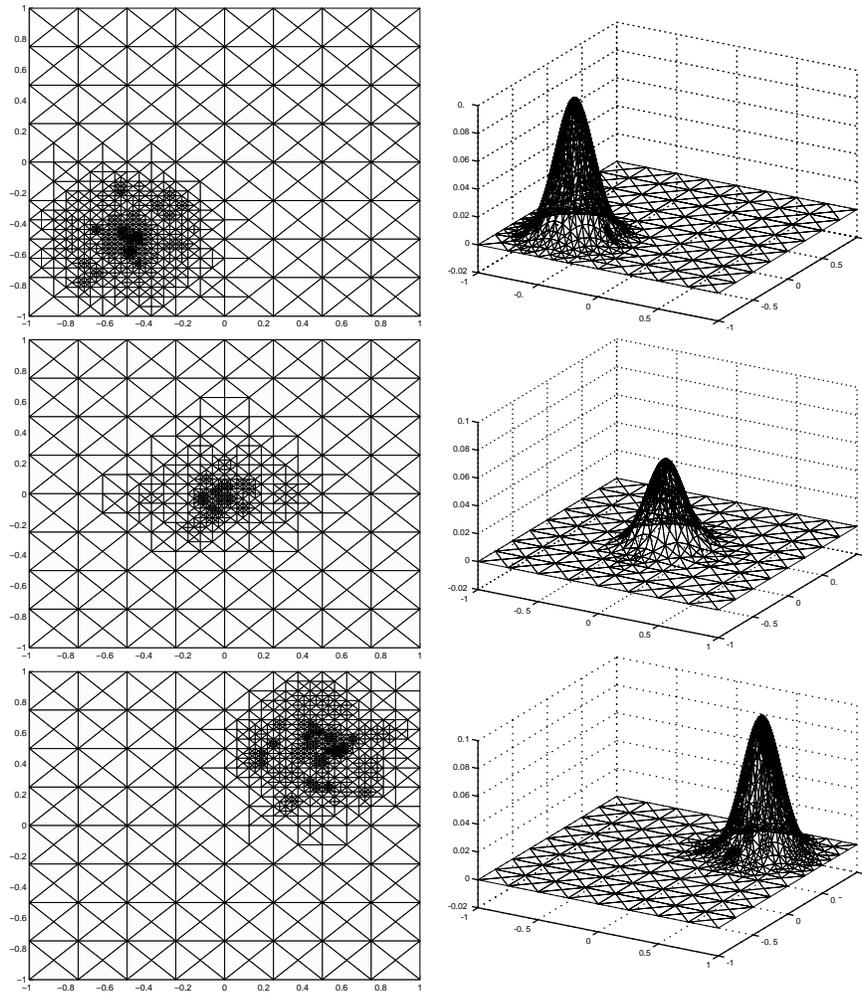
FIGURE 4. Surface plots of discrete solutions (right) and the corresponding meshes (left) at time $t = 0.01, 0.49, 0.998125$ (from top to bottom). The numbers of nodes of the meshes are $808, 459, 825$ for $t = 0.01, 0.49, 0.998125$, respectively.

In Table 5.1 we report the number of time steps N, the average number of nodes of the meshes $\mathrm{av}(\mathcal{M}^n)$, the estimated error η, the total energy error $\mathcal{E}$, and the effectiveness index `eff`, when running the algorithm for different values of $\mathtt{TOL_{space}}$. Recall from Theorem 3.2 (neglecting the initial error) that $\mathcal{E}^2 \leq C\mathtt{TOL_{space}}$ due to Theorem 3.2 and (5.2). Thus by reducing the tolerance $\mathtt{TOL_{space}}$ by one-fourth we are actually requiring the reduction of the total energy error $\mathcal{E}$ by half, which is indeed observed in Table 5.1. Moreover, we observe that when the error $\mathcal{E}$ is reduced by half, the number of time steps is increased by roughly twice, and the average number of nodes of the meshes is increased less than fourth. Most interestingly, the effectiveness indexes `eff` remain roughly constant in the computations.

Now we show more computational results when $\mathtt{TOL_{space}} = 0.01$. The energy error $\|u^n - U_h^n\|_\Omega$ at each time step n is depicted in Figure 2. The number of nodes and the time-step size at each time step n are shown in Figure 3. We note that the time-step size drops in a small time interval around $t = 0.5$ and is constant away from this interval, which is not surprising as $\beta(t)$ changes exponentially from 0.1 to 0 and then from 0 to 0.1 around $t = 0.5$, and away from $t = 0.5$, $\beta \approx 0.1$ and the peak of the exact solution moves at constant speed and the shape of the solution remains unchanged. This also explains why the energy errors and the numbers of nodes of the meshes are smaller around $t = 0.5$ as the exact solution u is smaller around $t = 0.5$. The surface plots of the computed solution and the associated adaptive mesh at time $t = 0.01, 0.49, 0.998125$ are displayed in Figure 4 which clearly indicates the ability of Algorithm 3.2 to capture the singularity of the solutions by locally refining and coarsening the meshes.

Acknowledgment

References

1. E. Bänsch, *Local mesh refinement in 2 and 3 dimensions,* IMPACT Comput. Sci. Engrg. 3 (1991), 181-191. MR **92h:**65150
2. M. Bieterman and I. Babuška, *The finite element method for parabolic equations: (I) a posteriori estimation,* Numer. Math. 40 (1982), 339-371. MR **85d:**65052a
3. M. Bieterman and I. Babuška, *The finite element method for parabolic equations: (II) a posteriori estimation and adaptive approach,* Numer. Math. 40 (1982), 373-406. MR **85d:**65052b
4. I. Babuška and C. Rheinboldt, *Error estimates for adaptive finite element computations,* SIAM J. Numer. Anal. 15 (1978), 736-754. MR **58:**3400
5. Z. Chen and S. Dai, *Adaptive Galerkin methods with error control for a dynamical Ginzburg-Landau model in superconductivity,* SIAM J. Numer. Anal. 38 (2001), 1961-1985. MR **2002g:**65114
6. Z. Chen and S. Dai, *On the efficiency of adaptive finite element methods for elliptic problems with discontinuous coefficients,* SIAM J. Sci. Comput. 24 (2002), 443-462.
7. Z. Chen, R.H. Nochetto and A. Schmidt, *A characteristic Galerkin method with adaptive error control for continuous casting problem,* Comput. Methods Appl. Mech. Engrg. 189 (2000), 249-276. MR **2001f:**80003
8. Z. Chen, R.H. Nochetto and A. Schmidt, *Error control and adaptivity for a phase relaxation model,* Math. Model. Numer. Anal. 34 (2000), 775-797. MR **2001i:**65117
9. P.G. Ciarlet, *The Finite Element Method for Elliptic Problems,* North-Holland, Amsterdam, 1978. MR **58:**25001
10. Ph. Clément, *Approximation by finite element functions using local regularization,* RAIRO Anal. Numer. 9 (1975), 77-84. MR **53:**4569
11. W. Dörfler, *A convergent adaptive algorithm for Possion's equations,* SIAM J. Numer. Anal. 33 (1996), 1106-1124. MR **97e:**65139
12. K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems I: A linear model problem,* SIAM J. Numer. Anal. 28 (1991), 43-77. MR **91m:**65274
13. K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems IV: Nonlinear problems,* SIAM J. Numer. Anal. 32 (1995), pp. 1729-1749. MR **96i:**65081
14. P.K. Moore, *A posteriori error estimation with finite element semi- and fully discrete methods for nonlinear parabolic equations in one space dimension,* SIAM J. Numer. Anal. 31 (1994), 149-169. MR **94m:**65145
15. P. Morin, R.H. Nochetto and K.G. Siebert, *Data oscillation and convergence of adaptive FEM,* SIAM J. Numer. Anal. 38 (2000), 466-488. MR **2001g:**65157
16. R.H. Nochetto, A. Schmidt and C. Verdi, *A posteriori error estimation and adaptivity for degenerate parabolic problems,* Math. Comp. 69 (2000), 1-24. MR **2000i:**65136

17. M. Picasso, *Adaptive finite elements for a linear parabolic problem,* Comput. Methods Appl. Mech. Engrg. 167 (1998), 223-237. MR **2000b:**65188
18. A. Schmidt and K.G. Siebert, *ALBERT: An adaptive hierarchical finite element toolbox,* IAM, University of Freiburg, 2000. `http://www.mathematik.uni-freiburg.de/IAM/Research/projectsdz/albert.`
19. R. Verfürth, *A Review of A Posteriori Error Estimation and Adaptive Mesh Refinement Techniques*, Teubner (1996).
20. R. Verfürth, *A posteriori error estimates for nonlinear problems: $L^{p,r}(0,T;W^{1,\rho}(\Omega)$ Finite element discretization of parabolic equations,* Numerical Methods in Partial Differential Equations 14 (1998), 487-518. MR **99g:**65099

LSEC, Institute of Computational Mathematics, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing 100080, Peoples Republic of China
E-mail address: `zmchen@lsec.cc.ac.cn`

Institute of Computational Mathematics, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, P.O. Box 2719, Beijing 100080, Peoples Republic of China
E-mail address: `fjia@lsec.cc.ac.cn`