ESAIM: M2AN 44 (2010) 455–484
DOI: 10.1051/m2an/2010009

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

SKIPPING TRANSITION CONDITIONS IN *A POSTERIORI* ERROR ESTIMATES FOR FINITE ELEMENT DISCRETIZATIONS OF PARABOLIC EQUATIONS

Stefano Berrone[1]

Abstract. In this paper we derive *a posteriori* error estimates for the heat equation. The time discretization strategy is based on a θ-method and the mesh used for each time-slab is independent of the mesh used for the previous time-slab. The novelty of this paper is an upper bound for the error caused by the coarsening of the mesh used for computing the solution in the previous time-slab. The technique applied for deriving this upper bound is independent of the problem and can be generalized to other time dependent problems.

Mathematics Subject Classification. 65N30, 65N15, 65N50, 65J15.

Received April 15, 2008. Revised July 28, 2009.
Published online February 4, 2010.

1. Introduction

A posteriori error estimates and adaptive algorithms are important keys in pursuing efficient and accurate discretizations of partial differential equations. Since the pioneering work of Babuška and Rheinboldt [1], these topics have become an important field for scientific computing [2,8,11,15,19,22]. Many works were devoted to elliptic problems and some important results were obtained in the parabolic case, too [3,10,17,20].

In particular, in [17] a residual based *a posteriori* error estimator and an adaptive strategy without coarsening were proposed. In [20] a robust residual based *a posteriori* error estimator was proposed and, in order to admit a moderate coarsening, a transition condition was introduced. The assembling of the linear system for each timestep and the computation of the error estimator are performed therein on a mesh that is a common refinement of the meshes used for the previous timestep and the current timestep. This introduces great difficulties in dealing with general adapted meshes. In [5] similar results were extended to the case of the heat equation with discontinuous, piecewise constant, coefficients.

In this work we present robust *a posteriori* error estimates for a discretization of the heat equation by conforming finite elements and a classical A-stable θ-scheme, avoiding the transition condition and the need of a common refinement of the meshes used for the previous timestep and the current timestep. This can lead to a great simplification in coding the method and make the implementation of an adaptive approach in an existing code much easier. In fact, assembling the linear system and computing the error estimators require information from a single mesh at each adaptive iteration and the transmission of the solution of the previous

Keywords and phrases. A posteriori error estimates, transition condition, parabolic problems.

[1] Dipartimento di Matematica, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy.
sberrone@calvino.polito.it

Article published by EDP Sciences

timestep on the mesh used for the current timestep is performed by a suitable projection operator for coarsened elements and by interpolation for refined elements.

We consider the proposed approach a remarkable simplification for a rigorous coding of a space-time adaptive algorithm, because we remove the difficulty of dealing with two meshes at every stage of the code.

In our work we explicitly consider the effect of changing the mesh from a timestep to the next one and, in particular, the effect of the coarsening, taking into account that the new degrees of freedom introduced by refinement whose values are obtained by interpolation do not introduce any transition error. Then we describe a bound of the discretization error arising from the coarsening and a bound for the total error in an arbitrary time interval. We prove an upper bound for the coarsening error by a coarsening error estimator composed by two terms involving a H^1-norm and a H^{-1}-norm, respectively. In Section 4, introducing a suitable projection of the old solution onto the new mesh, we also provide an upper bound for the H^{-1}-term leading to a fully computable upper bound for the coarsening error that, in our approach, is considered as a *data approximation error*. For this reason we do not need a lower bound for it.

The estimates proposed allow us to perform a control of the space-discretization and of the timestep-length used in each timestep. Further, they allow us to ensure the error of the full discretization to be bounded from above and from below in each timestep *via* global-in-time and global-in-space upper and lower bounds. The ratio between the upper and the lower bounds for the error is independent of any mesh-size, timestep-length and diffusivity parameter. No hypothesis on the changes of meshes used for two consecutive timesteps are made. The bounds involve a data-approximation-error in space and a data-approximation-error in time that can be used to adapt the mesh and the timestep-length in each timestep.

The main results of this paper are Theorems 3.5 and 3.9 that give an upper and a lower bound for the discretization error. In these theorems we give an upper bound of the effects in the discretization error caused by the coarsening applied between two subsequent timesteps.

The *a posteriori* error estimates proposed can easily be extended to the case of piecewise constant discontinuous diffusivity coefficients assuming a *quasi monotonicity* condition [4,9,16] as in [5].

Let us introduce some notation. The symbol $a \precsim b$ means that there exists a constant c independent of any mesh-size, timestep-length and problem parameter such that $a \leq c\, b$. The symbol $a \asymp b$ means that $a \precsim b$ and $b \precsim a$.

The paper is organized as follows: in Section 2 we define the considered problem and the discretization method. In Section 3 we define the error estimators and derive lower and upper bounds for the error considering the error introduced by the coarsening in transmission of the solution from one mesh to the next one. In Section 4 we describe a suitable projection operator that allows to bound from above the H^{-1}-norm present in the coarsening error estimator. In Section 5 we present some numerical test comparing the performances of a simple adaptive algorithm based on the proposed estimates and on the *common refinement approach* based on the two meshes. In the Appendix we report an example of construction of the dual basis functions introduced in Section 4 for the 2D linear finite elements when the mesh modification can introduce hanging nodes.

2. The heat equation

2.1. The continuous problem

Let Ω be a bounded Lipschitz continuous domain in $\mathbb{R}^d$ with boundary $\partial\Omega$ and let $(0,T)$ be the time interval of interest. For any $f \in \mathrm{L}^2((0,T);\mathrm{L}^2(\Omega))$ and $U^0 \in \mathrm{L}^2(\Omega)$, we want to find $U : \Omega \times (0,T) \to \mathbb{R}$ such that

$$\frac{\partial U}{\partial t} - \kappa \triangle U = f, \qquad \text{in } \Omega \times (0,T), \tag{2.1}$$

$$U(x,t) = 0, \qquad \text{on } \partial\Omega \times (0,T), \tag{2.2}$$

$$U(x,0) = U^0(x), \qquad \text{in } \Omega. \tag{2.3}$$

Let the domain Ω be a polygonal domain and $\kappa > 0$ the diffusivity parameter.

Setting

$$W = \left\{ w \in \mathrm{L}^2((0,T);\mathrm{H}^1_0(\Omega)) \ : \ \frac{\partial w}{\partial t} \in \mathrm{L}^2((0,T);\mathrm{H}^{-1}(\Omega)) \right\}$$

the variational continuous formulation of the above problem is:
Find $U \in W$ *such that* $U(.,0) = U^0$ *and*

$$\left\langle \frac{\partial U}{\partial t}, v \right\rangle + (\kappa \nabla U, \nabla v) = (f, v), \quad \forall v \in \mathrm{H}^1_0(\Omega), \text{ a.e. in } (0,T). \tag{2.4}$$

Here $\langle\,.\,,\,.\,\rangle$ stands for the duality pairing between $\mathrm{H}^{-1}(\Omega)$ and $\mathrm{H}^1_0(\Omega)$, and $(\,.\,,\,.\,)$ is the usual inner product in $\mathrm{L}^2(\Omega)$. If $U \in W$, then $U \in \mathcal{C}^0([0,T];\mathrm{L}^2(\Omega))$ and the initial condition $U(.,0) = U^0$ is meaningful in $\mathrm{L}^2(\Omega)$.

2.2. **The numerical discretization**

Let us consider a partition of $(0,T)$ into subintervals $\left(t^{n-1}, t^n\right)$ of length $\tau^n = t^n - t^{n-1}$, with $0 = t^0 < t^1 < \ldots < t^N = T$; set $I^n = \left[t^{n-1}, t^n\right]$. In each time-slab $\Omega \times I^n$, $n \geq 1$, we consider a regular family of partitions $\mathcal{T}_h^n$ of $\overline{\Omega}$ into elements $K \in \mathcal{T}_h^n$ which satisfy the usual conformity and minimal-angle conditions [6] and we denote by h_K^n the diameter of the element $K \in \mathcal{T}_h^n$.

Let $\mathrm{V}_h^n \subset \mathrm{V} = \mathrm{H}^1_0(\Omega)$ be a family of conforming finite element spaces based on the partitions $\mathcal{T}_h^n$:

$$\mathrm{V}_h^n = \left\{ v_h \in \mathrm{H}^1_0(\Omega) \cap \mathcal{C}^0(\overline{\Omega}) : \ v_{|_K} \in \mathbb{P}_i(K), \ \forall K \in \mathcal{T}_h^n \right\}.$$

We recall that $\mathbb{P}_i(K)$ is the space of polynomials of degree $i \geq 1$ on the element $K \in \mathcal{T}_h^n$. In case we allow different polynomial degrees of the finite element spaces in different timesteps, we assume a global in time upper bound for that.

Given an approximation $u_{h,\tau}^{n-1} \in \mathrm{V}_h^{n-1}$ of $U^{n-1} = U(.,t^{n-1})$ we define $P^n u_{h,\tau}^{n-1}$ its projection or interpolation onto V_h^n. For the moment we do not make any assumption on the operator P^n.

We introduce the continuous in space, piecewise affine in time approximation of the solution $U(x,t)$ on the time-slab $\Omega \times I^n$:

$$u^n(x,t) = \frac{t - t^{n-1}}{t^n - t^{n-1}}\, u_{h,\tau}^n(x) + \frac{t^n - t}{t^n - t^{n-1}}\, P^n u_{h,\tau}^{n-1}(x), \qquad t \in I^n, \qquad x \in \Omega, \tag{2.5}$$

being $u^n \in \mathrm{V}_h^n \times I^n$ and $u_{h,\tau}^n,\ P^n u_{h,\tau}^{n-1} \in \mathrm{V}_h^n$. Moreover, let us define

$$u(x,t) = \sum_{n=1}^{N} \left(\frac{t - t^{n-1}}{t^n - t^{n-1}}\, u_{h,\tau}^n(x) + \frac{t^n - t}{t^n - t^{n-1}}\, u_{h,\tau}^{n-1}(x) \right) \chi_{I^n}, \qquad x \in \Omega. \tag{2.6}$$

The approximation u is a continuous function in time on the whole time interval $(0,T)$ and in each time-slab is a continuous transition from $u_{h,\tau}^{n-1} \in \mathrm{V}_h^{n-1}$ to $u_{h,\tau}^n \in \mathrm{V}_h^n$. We remark that *for* $n = 1, \ldots, N$ and $\forall t \in I^n$

$$u = u^n - \frac{t^n - t}{t^n - t^{n-1}}\, \delta u_{h,\tau}^{n-1}, \tag{2.7}$$

where $\delta u_{h,\tau}^{n-1} = P^n u_{h,\tau}^{n-1} - u_{h,\tau}^{n-1}$.

Then, we introduce the discretization based on the classical θ-scheme for the time integration:
Find $u^n_{h,\tau} \in \mathrm{V}^n_h$ *such that*

$$\left(\frac{\partial u^n}{\partial t}, v_h\right) + \theta\left(\kappa \nabla u^n_{h,\tau}, \nabla v_h\right) + (1-\theta)\left(\kappa \nabla P^n u^{n-1}_{h,\tau}, \nabla v_h\right) = \theta\left(\Pi_{\mathcal{T}_h} f^n, v_h\right) + (1-\theta)\left(\Pi_{\mathcal{T}_h} f^{n-1}, v_h\right), \qquad \forall v_h \in \mathrm{V}^n_h,\ t \in I^n,\ n = 1, \ldots, N. \quad (2.8)$$

In the last scalar products of the previous equation we assume that $f \in \mathcal{C}^0([0,T]; \mathrm{L}^2(\Omega))$ and we set $f^r = f(., t^r)$ for $r = n$ or $n-1$. Moreover we introduce an arbitrary polynomial approximation $\Pi_{\mathcal{T}_h} f$ of the data f.

Remark 2.1. All the approximations of the solution U involved in the scheme (2.8) are defined on the mesh of the current timestep $\mathcal{T}^n_h$ and not on the two meshes $\mathcal{T}^n_h$ and $\mathcal{T}^{n-1}_h$ as in [20].

3. A RESIDUAL-BASED *A POSTERIORI* ERROR ESTIMATOR

In this section we derive a residual-based error estimator for the fully discretized model problem following the work in [5,17,20]. In particular, we shall derive global-in-space, local-in-time upper and lower bounds, as well as global-in-space, global-in-time upper and lower bounds. In our estimates some terms depend on the possible coarsening of the mesh from the previous timestep to the next one. Our target is to set up an estimator bounding the error introduced by the coarsening of the mesh and the transition of the solution from the mesh used in the $(n-1)$th timestep and the mesh used in the nth timestep.

First, we introduce some notation which will be used for the construction of the estimator, more detail can be found in [5].

3.1. Definitions and general results

For any $K \in \mathcal{T}^n_h$ we denote by $\mathcal{E}(K)$ the set of its faces (edges if $d = 2$); we denote by $\mathcal{E}^n_h = \cup_{K \in \mathcal{T}^n_h} \mathcal{E}(K)$ the set of all faces of the partition $\mathcal{T}^n_h$. Moreover, we split $\mathcal{E}^n_h$ into the form $\mathcal{E}^n_h = \mathcal{E}^n_{h,\Omega} \cup \mathcal{E}^n_{h,\partial\Omega}$ with $\mathcal{E}^n_{h,\Omega} = \{E \in \mathcal{E}^n_h : E \not\subset \partial\Omega\}$, $\mathcal{E}^n_{h,\partial\Omega} = \{E \in \mathcal{E}^n_h : E \subset \partial\Omega\}$. For any face $E \in \mathcal{E}^n_h$ we define: $\omega^n_E = \bigcup_{\{K' :\ E \in \mathcal{E}(K')\}} K'$, and to any face $E \in \mathcal{E}^n_{h,\Omega}$ we associate an orthogonal unit vector n_E and denote by $[.]_E$ the jump across E in the direction n_E. Let us denote by $\hat{K}$ the reference element and by $\hat{E}$ the reference face as shown in Figure 1 on the left for $d = 2$. Let λ_i, $i = 0, \ldots, d$ be the barycentric coordinates on the reference element, then the *reference element bubble function* is $\hat{b}_{\hat{K}} = (d+1)^{d+1} \lambda_0 \lambda_1 \ldots \lambda_d$, and the *reference face bubble function* is $\hat{b}_{\hat{E}} = 4\hat{x}_1(1-\hat{x}_1)$ for $d = 2$ and $\hat{b}_{\hat{E}} = 27\hat{x}_1\hat{x}_2(1-\hat{x}_1-\hat{x}_2)$ for $d = 3$. $F_K : \hat{K} \to K$ is the affine mapping from the reference element to the element $K \in \mathcal{T}^n_h$ [6,19]. Here we indicate with $b^n_K = \hat{b}_{\hat{K}} \circ F^{-1}_K$ the *element bubble function*, that does not depend on time on a fixed timestep.

Let us introduce the tools for defining a suitable set of orthogonal face cut of functions for triangles and tetrahedra. Given any $E \in \mathcal{E}^n_{h,\Omega}$, let $K^\sharp$ and $K^\flat$ the two elements of $\mathcal{T}^n_h$ such that $\omega^n_E = K^\sharp \cup K^\flat$, let us enumerate the vertices of $K^\sharp$ and $K^\flat$ in such a way that the vertices of E are numbered first. Let K be one of the elements $K^\sharp$ and $K^\flat$, assume that E has vertices $\overline{a}_0, \ldots, \overline{a}_{d-1}$ and denote by $\overline{a}_c$ be the barycentre of the element K; let us partition K into the elements K_0, $K_1, \ldots, K_d$ with K_d having E as a face (for the 2D case see Fig. 1, right). Let $F_{E,K} : \hat{K} \to K_d$ be the invertible affine mapping that maps the reference element $\hat{K}$ onto the element K_d. Then we define the *face bubble function* b^n_E by patching the two bubble functions: $b^n_{E,K^\sharp} = \hat{b}_{\hat{E}} \circ F^{-1}_{E,K^\sharp}$, $b^n_{E,K^\flat} = \hat{b}_{\hat{E}} \circ F^{-1}_{E,K^\flat}$, each one being not zero only on $K^\sharp_d$ and $K^\flat_d$, respectively. We need to define the set $\mathring{\omega}^n_E = K^\sharp_d \cup K^\flat_d$ that is the dashed area in Figure 2. For the boundary face E that belongs to the element K only, we naturally identify b^n_E with $b^n_{E,K} = \hat{b}_{\hat{E}} \circ F^{-1}_{E,K}$. With this definition of *face*

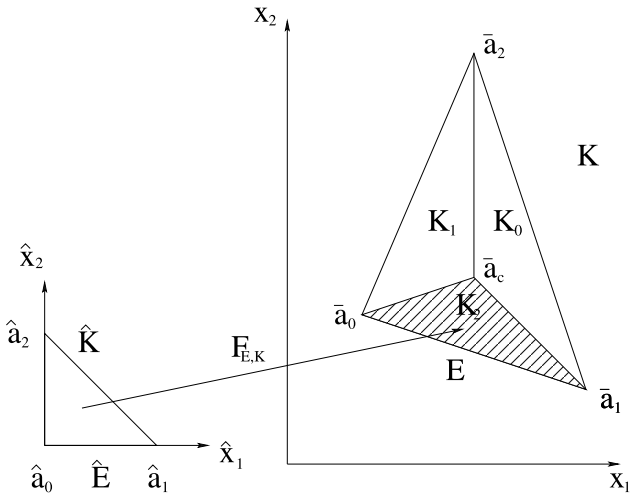

FIGURE 1. The mapping $F_{E,K} : \hat{K} \to K_2$.

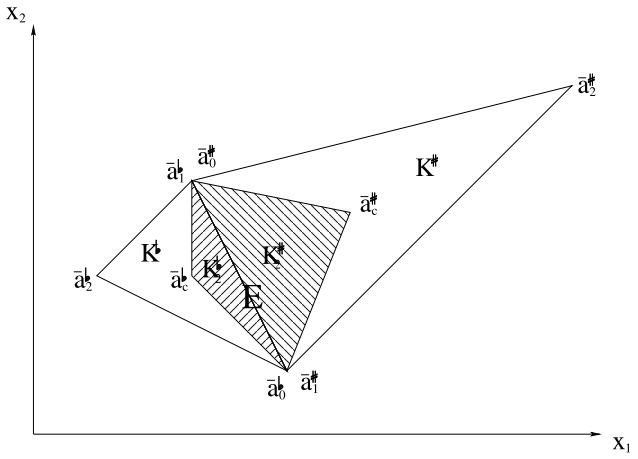

FIGURE 2. The support of the function b_E^n.

bubble functions we have a set of orthogonal functions. This property is also true for the set of *element bubble functions*.

Moreover, for the reference face $\hat{E}$ we define the extension operator $\hat{\mathcal{P}}_{\hat{E}} : \mathbb{P}_i(\hat{E}) \to \mathbb{P}_i(\hat{K})$ which extends a polynomial of degree i defined on the face $\hat{E}$ to a polynomial of the same degree defined on $\hat{K}$ with constant values along lines orthogonal to the face $\hat{E}$. Then, we define the extension operator $\mathcal{P}_E : \mathbb{P}_i(E) \to \mathbb{P}_i(\omega_E)$ which extends a polynomial of degree i defined on the face E to a continuous piecewise polynomial of the same degree defined on ω_E^n by patching the two operators: $\mathcal{P}_{E|_{K^\sharp}} = F_{K^\sharp} \circ \hat{\mathcal{P}}_{\hat{E}} \circ F^{-1}_{K^\sharp|_E}$ and $\mathcal{P}_{E|_{K^\flat}} = F_{K^\flat} \circ \hat{\mathcal{P}}_{\hat{E}} \circ F^{-1}_{K^\flat|_E}$.

Thanks to the regularity hypothesis, there exist constants only depending on the quality of the partition such that for each $n = 1, \ldots, N$ we have $h_K^n \asymp h_E^n$, $\forall E \in \mathcal{E}(K)$.

Moreover, let us recall the following properties of the *quasi-interpolation operator of Clément* $I_h : \mathrm{V} \to \mathrm{V}_h^n$, [7] in which the sets $\tilde{\omega}_K^n$ and $\tilde{\omega}_E^n$ are suitable patches of neighbouring adjacent elements to the considered element K or an edge/face E.

Lemma 3.1. *Let $K \in \mathcal{T}_h$ and $E \in \mathcal{E}_h$ be arbitrary. Then we have the following interpolation error estimates:*

$$\| v - I_h v \|_{0,K} \leq Cl_R\, h_K^n \| \nabla v \|_{0,\tilde{\omega}_K^n}, \quad \forall v \in \mathrm{H}^1(\tilde{\omega}_K^n), \tag{3.1}$$

$$\| v - I_h v \|_{0,E} \leq Cl_E \sqrt{h_E^n} \| \nabla v \|_{0,\tilde{\omega}_E^n}, \ \forall v \in \mathrm{H}^1(\tilde{\omega}_E^n), \tag{3.2}$$

the constants Cl_R and Cl_E depending only on the smallest angle in the partition.

Let us consider the space $\mathrm{H}_0^1(\Omega)$ equipped with the norm

$$\| v \|_{\kappa,1}^2 = \left\| \sqrt{\kappa}\, \nabla v \right\|_0^2 = \kappa \| \nabla v \|_0^2$$

and denote by

$$\| F \|_{\kappa,-1} = \sup_{v \in \mathrm{H}_0^1(\Omega)} \frac{\langle F, v \rangle}{\| v \|_{\kappa,1}} = \frac{\| F \|_{-1}}{\sqrt{\kappa}}$$

the norm of the dual space $\mathrm{H}^{-1}(\Omega)$.

Let us define the error of our approximation u^n in the interval I^n as

$$e^n = u^n - U$$

and the error of our continuous in time approximation u as

$$e = u - U.$$

We remark the following relation between e^n and e derived from (2.7):

$$e^n = e + \frac{t^n - t}{t^n - t^{n-1}}\, \delta u_{h,\tau}^{n-1}, \qquad \forall t \in I^n. \tag{3.3}$$

Definition 3.2. Let us define the residuals and inter-element jumps of our approximation u^n:

$$R_K^n(u^n) = \left. \frac{\partial u^n}{\partial t} - \theta\, \kappa \triangle u_{h,\tau}^n - (1-\theta)\, \kappa \triangle P^n u_{h,\tau}^{n-1} - \theta\, \Pi_{\mathcal{T}_h} f^n - (1-\theta)\, \Pi_{\mathcal{T}_h} f^{n-1} \right|_K,$$

$$R_\Omega^n(u^n) = \sum_{K \in \mathcal{T}_h^n} R_K^n(u^n), \qquad J_E^n(u^n) = \left[\!\left[\theta\, \kappa\, \frac{\partial\, u_{h,\tau}^n}{\partial n_E} + (1-\theta)\, \kappa\, \frac{\partial\, P^n u_{h,\tau}^{n-1}}{\partial n_E} \right]\!\right]_E.$$

Definition 3.3. Let us define the following local-in-space-and-time estimators:

$$\left(\eta_{R,K}^n\right)^2 = \tau^n \left((h_K^n)^2 \left\| \frac{1}{\sqrt{\kappa}}\, R_K^n(u^n) \right\|_{0,K}^2 + \frac{1}{2} \sum_{E \in \mathcal{E}(K) \cap \mathcal{E}_{h,\Omega}^n} h_E^n \left\| \frac{1}{\sqrt{\kappa}}\, J_E^n(u^n) \right\|_{0,E}^2 \right),$$

$$\left(\eta_{\tau,K}^n\right)^2 = \tau^n \left\| \sqrt{\kappa}\, \nabla \left(u_{h,\tau}^n - P^n u_{h,\tau}^{n-1} \right) \right\|_{0,K}^2.$$

Then, we define the following global-in-space-and-local-in-time estimators

$$\left(\eta_R^n\right)^2 = \sum_{K \in \mathcal{T}_h^n} \left(\eta_{R,K}^n\right)^2, \qquad \left(\eta_\tau^n\right)^2 = \sum_{K \in \mathcal{T}_h^n} \left(\eta_{\tau,K}^n\right)^2,$$

$$\left(\eta_{f,\theta,\tau^n}^n\right)^2 = \int_{t^{n-1}}^{t^n} \| \Pi_{\mathcal{T}_h} f - \theta\, \Pi_{\mathcal{T}_h} f^n - (1-\theta)\, \Pi_{\mathcal{T}_h} f^{n-1} \|_{\kappa,-1}^2 \mathrm{d}t,$$

$$\left(\eta_{f,\Pi_{\mathcal{T}_h}}^n\right)^2 = \int_{t^{n-1}}^{t^n} \| f - \Pi_{\mathcal{T}_h} f \|_{\kappa,-1}^2 \mathrm{d}t, \qquad \left(\eta_f^n\right)^2 = \left(\eta_{f,\theta,\tau^n}^n\right)^2 + \left(\eta_{f,\Pi_{\mathcal{T}_h}}^n\right)^2.$$

Finally, we define the following global-in-space-and-time estimators:

$$\eta_R^2 = \sum_{n=1}^{N} (\eta_R^n)^2, \qquad \eta_\tau^2 = \sum_{n=1}^{N} (\eta_\tau^n)^2, \qquad \eta_f^2 = \sum_{n=1}^{N} \left(\eta_f^n\right)^2.$$

In the sequel we will derive upper and lower bounds for the error e involving the following norms:

$$\| e \|_{\kappa,I^n} = \left\{ \int_{t^{n-1}}^{t^n} \left\| \frac{\partial e}{\partial t} \right\|_{\kappa,-1}^2 \mathrm{d}t + \int_{t^{n-1}}^{t^n} \| e \|_{\kappa,1}^2 \mathrm{d}t \right\}^{\frac{1}{2}},$$

$$\| e \|_{\kappa,(0,T)} = \left\{ \sum_{n=1}^{N} \left(\int_{t^{n-1}}^{t^n} \left\| \frac{\partial e}{\partial t} \right\|_{\kappa,-1}^2 \mathrm{d}t + \int_{t^{n-1}}^{t^n} \| e \|_{\kappa,1}^2 \mathrm{d}t \right) \right\}^{\frac{1}{2}}.$$

Remark 3.4. Following considerations of [17,20], and numerical experiments of [5], we can say that η_R^n is a space error estimator related to the quality of the partition $\mathcal{T}_h^n$. When the mesh is suitably adapted, the quantity η_τ^n gives information on the error due to time discretization. Moreover, $\eta_{f,\Pi_{\mathcal{T}_h}}^{[n]}$ gives information essentially on the space-data-approximation error, and $\eta_{f,\theta,\tau^n}^{[n]}$ on time-data-approximation error.

Let us define $t^{\theta,n} = \theta\, t^n + (1-\theta)\, t^{n-1}$, in the following we will use often the properties:

$$\begin{aligned} u^n - u_{h,\tau}^n &= \frac{t - t^n}{t^n - t^{n-1}} \left(u_{h,\tau}^n - P^n u_{h,\tau}^{n-1} \right), \\ u^n - P^n u_{h,\tau}^{n-1} &= \frac{t - t^{n-1}}{t^n - t^{n-1}} \left(u_{h,\tau}^n - P^n u_{h,\tau}^{n-1} \right), \\ \theta \left(u^n - u_{h,\tau}^n \right) + (1-\theta) \left(u^n - P^n u_{h,\tau}^{n-1} \right) &= \frac{t - t^{\theta,n}}{t^n - t^{n-1}} \left(u_{h,\tau}^n - P^n u_{h,\tau}^{n-1} \right). \end{aligned} \tag{3.4}$$

3.2. **Global-in-space, global-in-time upper bound**

The following theorem gives the final upper bound of the error:

Theorem 3.5. *Under the assumptions on the continuous problem (2.4) and on the discrete formulation (2.8), there exist constants $C_{t^{n-1}}^{\uparrow,t^n}$ for each $n = 1, \ldots, N$, and $C_0^{\uparrow,T} \max_{n=1,\ldots,N} C_{t^{n-1}}^{\uparrow,t^n}$ independent of any meshsize, timestep-length, problem-parameter, but depending on the quality of partitions $\mathcal{T}_h^n$ such that the inequality*

$$\begin{aligned} 9 \| e(.,t^m) \|_0^2 + \| e \|_{\kappa,[0,t^m]}^2 \leq 9 \left\| e(.,t^0) \right\|_0^2 + C_0^{\uparrow,T} \sum_{n=1}^{m} \Big[& (\eta_R^n)^2 + (\eta_\tau^n)^2 \\ & + \left(\eta_f^n\right)^2 + \frac{1}{\tau^n} \left\| \delta u_{h,\tau}^{n-1} \right\|_{\kappa,-1}^2 + \tau^n \left\| \delta u_{h,\tau}^{n-1} \right\|_{\kappa,1}^2 \Big], \quad \forall m = 1, \ldots, N \end{aligned} \tag{3.5}$$

holds true.

The last two terms in (3.5) represent a coarsening error estimator, in Section 4 we shall provide a suitable upper bound for the H^{-1} term. To prove this theorem we first start with the following proposition:

Proposition 3.6. *Under the assumptions on the continuous problem (2.4) and on the discrete formulation (2.8), for each $n = 1, \dots, N$, there exists a constant $\tilde{C}_{t^{n-1}}^{t^n}$ independent of any meshsize, timestep-length, problem-parameter and depending only on the quality of the partition $\mathcal{T}_h^n$ and on the parameter θ, such that*

$$\| e(., t^m) \|_0^2 + \int_{t^0}^{t^m} \| \sqrt{\kappa}\, \nabla e \|_0^2 \mathrm{d}t \le \| e(., t^0) \|_0^2 + \sum_{n=1}^{m} \tilde{C}_{t^{n-1}}^{t^n} \left((\eta_R^n)^2 + (\eta_\tau^n)^2 + (\eta_f^n)^2 + \tau^n \left\| \delta u_{h,\tau}^{n-1} \right\|_{\kappa,1}^2 + \frac{1}{\tau^n} \left\| \delta u_{h,\tau}^{n-1} \right\|_{\kappa,-1}^2 \right). \quad (3.6)$$

Proof. Let us define

$$E^{0,m} = \int_{t^0}^{t^m} \left[\left\langle \frac{\partial e}{\partial t}, e \right\rangle + (\kappa \nabla e, \nabla e) \right] \mathrm{d}t = \frac{1}{2} \| e(., t^m) \|_0^2 - \frac{1}{2} \| e(., t^0) \|_0^2 + \int_{t^0}^{t^m} \| \sqrt{\kappa}\, \nabla e \|_0^2 \mathrm{d}t\,.$$

Moreover we write

$$E^{0,m} = \sum_{n=1}^{m} E^{n-1,n},$$

where

$$\begin{aligned} E^{n-1,n} &= \int_{t^{n-1}}^{t^n} \left[\left(\frac{\partial u}{\partial t}, e \right) + (\kappa \nabla u, \nabla e) \right] \mathrm{d}t - \int_{t^{n-1}}^{t^n} \left[\left\langle \frac{\partial U}{\partial t}, e \right\rangle + (\kappa \nabla u, \nabla e) \right] \mathrm{d}t \\ &= \int_{t^{n-1}}^{t^n} \left[\left(\frac{\partial u^n}{\partial t}, e \right) + \left(\frac{\delta u_{h,\tau}^{n-1}}{\tau^n}, e \right) \right] \mathrm{d}t + \int_{t^{n-1}}^{t^n} \left[(\kappa \nabla u^n, \nabla e) - \frac{t^n - t}{t^n - t^{n-1}} \left(\kappa \nabla \delta u_{h,\tau}^{n-1}, e \right) - (f, e) \right] \mathrm{d}t \end{aligned}$$

and we subtract to $E^{n-1,n}$ the integral over I^n of the discrete formulation (2.8) with $I_h e_{h,\tau}$ as test function. We get

$$\begin{aligned} E^{n-1,n} = \int_{t^{n-1}}^{t^n} & \left[\left(\frac{\partial u^n}{\partial t}, e - I_h e_{h,\tau} \right) + \left(\theta \kappa \nabla u_{h,\tau}^n + (1-\theta) \kappa \nabla P^n u_{h,\tau}^{n-1}, \nabla (e - I_h e_{h,\tau}) \right) \right. \\ & \left. - \left(\theta\, \Pi_{\mathcal{T}_h} f^n + (1-\theta)\, \Pi_{\mathcal{T}_h} f^{n-1}, e - I_h e_{h,\tau} \right) \right] \mathrm{d}t \\ & + \int_{t^{n-1}}^{t^n} \theta \left(\kappa \nabla \left(u^n - u_{h,\tau}^n \right), \nabla e \right) \mathrm{d}t + \int_{t^{n-1}}^{t^n} (1-\theta) \left(\kappa \nabla \left(u^n - P^n u_{h,\tau}^{n-1} \right), \nabla e \right) \mathrm{d}t \\ & - \int_{t^{n-1}}^{t^n} (f - \Pi_{\mathcal{T}_h} f, e) \mathrm{d}t - \int_{t^{n-1}}^{t^n} \left(\Pi_{\mathcal{T}_h} f - \theta\, \Pi_{\mathcal{T}_h} f^n - (1-\theta)\, \Pi_{\mathcal{T}_h} f^{n-1}, e \right) \mathrm{d}t \\ & + \int_{t^{n-1}}^{t^n} \left(\frac{\delta u_{h,\tau}^{n-1}}{\tau^n}, e \right) \mathrm{d}t - \int_{t^{n-1}}^{t^n} \frac{t^n - t}{t^n - t^{n-1}} \left(\kappa \nabla \delta u_{h,\tau}^{n-1}, e \right) \mathrm{d}t\,. \end{aligned}$$

After integration by parts of the term $\left(\theta \kappa \nabla u_{h,\tau}^n + (1-\theta) \kappa \nabla P^n u_{h,\tau}^{n-1}, \nabla (e - I_h e_{h,\tau}) \right)$ we apply Cauchy-Schwarz's inequality, inequalities of Lemma 3.1 and Young's inequality with a suitable choice of constants, noting that

$\int_{t^{n-1}}^{t^n}\left(\frac{t^n-t}{t^n-t^{n-1}}\right)^2\mathrm{d}t=\frac{\tau^n}{3}$ and $\int_{t^{n-1}}^{t^n}\left(\frac{t-t^{\theta,n}}{t^n-t^{n-1}}\right)^2\mathrm{d}t=\left(\theta^2-\theta+\frac{1}{3}\right)\tau^n$, summing over the time-slabs $n=1,\ldots,m$, we conclude that there exist constants $\tilde{\tilde{C}}^{t^n}_{t^{n-1}}$ such that

$$\begin{aligned}\|\,e(.,t^m)\,\|_0^2+\int_{t^0}^{t^m}\|\,\sqrt{\kappa}\,\nabla\,e\,\|_0^2\mathrm{d}t\le\|\,e(.,t^0)\,\|_0^2\\+\sum_{n=1}^{m}\tilde{\tilde{C}}^{t^n}_{t^{n-1}}\left[\tau^n\left(\sum_{K\in\mathcal{T}_h^n}(h_K^n)^2\left\|\frac{1}{\sqrt{\kappa}}\,R_K^n(u^n)\right\|_{0,K}^2+\sum_{E\in\mathcal{E}_{h,\Omega}^n}h_E^n\left\|\frac{1}{\sqrt{\kappa}}\,J_E^n(u^n)\right\|_{0,K}^2\right.\right.\\\left.+\sum_{K\in\mathcal{T}_h^n}\left\|\sqrt{\kappa}\,\nabla\left(u_{h,\tau}^n-P^nu_{h,\tau}^{n-1}\right)\right\|_{0,K}^2+\left\|\,\delta u_{h,\tau}^{n-1}\,\right\|_{\kappa,1}^2\right)+\frac{1}{\tau^n}\left\|\,\delta u_{h,\tau}^{n-1}\,\right\|_{\kappa,-1}^2\\+\int_{t^{n-1}}^{t^n}\left\|\,\Pi_{\mathcal{T}_h}f-\theta\,\Pi_{\mathcal{T}_h}f^n-(1-\theta)\,\Pi_{\mathcal{T}_h}f^{n-1}\,\right\|_{\kappa,-1}^2\mathrm{d}t+\left.\int_{t^{n-1}}^{t^n}\|\,f-\Pi_{\mathcal{T}_h}f\,\|_{\kappa,-1}^2\mathrm{d}t\right]\end{aligned}$$

and we get (3.6). □

The result given by Proposition 3.6 is an upper bound of the error measured in the $\mathrm{L}^2(\Omega)$-norm at time t^m and in the $\mathrm{L}^2((t^0,t^m);\mathrm{H}_0^1(\Omega)\,)$-norm. We also need an upper bound in the $\mathrm{L}^2((t^0,t^m);\mathrm{H}^{-1}(\Omega)\,)$-norm for $\partial\,e\,/\partial t$.

Lemma 3.7. *Under the assumptions on the continuous problem (2.4) and on the discrete formulation (2.8) for each $n=1,\ldots,N$ and for each $t\in\left(t^{n-1},t^n\right)$, we have*

$$\begin{aligned}\left\|\frac{\partial\,e}{\partial t}\right\|_{\kappa,-1}\le Cl_R\left\{\sum_{K\in\mathcal{T}_h^n}(h_K^n)^2\left\|\frac{1}{\sqrt{\kappa}}\,R_K^n(u^n)\right\|_{0,K}^2\right\}^{\frac{1}{2}}+Cl_E\left\{\sum_{E\in\mathcal{E}_{h,\Omega}^n}h_E^n\left\|\frac{1}{\sqrt{\kappa}}\,J_E^n(u^n)\right\|_{0,E}^2\right\}^{\frac{1}{2}}\\+\|\,f-\Pi_{\mathcal{T}_h}f\,\|_{\kappa,-1}+\left\|\,\Pi_{\mathcal{T}_h}f-\theta\,\Pi_{\mathcal{T}_h}f^n-(1-\theta)\,\Pi_{\mathcal{T}_h}f^{n-1}\,\right\|_{\kappa,-1}+\frac{t-t^{\theta,n}}{t^n-t^{n-1}}\left\|\sqrt{\kappa}\,\nabla\left(u_{h,\tau}^n-P^nu_{h,\tau}^{n-1}\right)\right\|_0\\+\left\|\,\sqrt{\kappa}\,\nabla\,e\,\right\|_0+\frac{1}{\tau^n}\left\|\,\delta u_{h,\tau}^{n-1}\,\right\|_{\kappa,-1}+\frac{t^n-t}{t^n-t^{n-1}}\left\|\,\delta u_{h,\tau}^{n-1}\,\right\|_{\kappa,1}.\end{aligned}\tag{3.7}$$

Proof. Let us notice that

$$\begin{aligned}\left\langle\frac{\partial\,e}{\partial t},v\right\rangle+(\kappa\,\nabla\,e,\nabla v)&=\left(\frac{\partial\,u^n}{\partial t},v\right)+(\kappa\,\nabla\,u^n,\nabla v)-(f,v)\\&\quad+\left(\frac{\delta u_{h,\tau}^{n-1}}{\tau^n},v\right)-\frac{t^n-t}{t^n-t^{n-1}}\left(\kappa\,\nabla\,\delta u_{h,\tau}^{n-1},\nabla v\right)\\&=\left(\frac{\partial\,u^n}{\partial t},v\right)+\left(\kappa\,\nabla\left(\theta\,u_{h,\tau}^n+(1-\theta)\,P^nu_{h,\tau}^{n-1}\right),\nabla v\right)-\left(\theta\,\Pi_{\mathcal{T}_h}f^n+(1-\theta)\,\Pi_{\mathcal{T}_h}f^{n-1},v\right)\\&\quad+\frac{t-t^{\theta,n}}{t^n-t^{n-1}}\left(\kappa\,\nabla\left(u_{h,\tau}^n-P^nu_{h,\tau}^{n-1}\right),\nabla v\right)-\left(\Pi_{\mathcal{T}_h}f-\theta\,\Pi_{\mathcal{T}_h}f^n-(1-\theta)\,\Pi_{\mathcal{T}_h}f^{n-1},v\right)\\&\quad-(f-\Pi_{\mathcal{T}_h}f,v)+\left(\frac{\delta u_{h,\tau}^{n-1}}{\tau^n},v\right)-\frac{t^n-t}{t^n-t^{n-1}}\left(\kappa\,\nabla\,\delta u_{h,\tau}^{n-1},\nabla v\right).\end{aligned}$$

Moreover, using (2.8) with $v_h = I_h v$, $v \in \mathrm{H}^1_0(\Omega)$, applying Cauchy-Schwarz's and Hölder inequalities and Lemma 3.1 we get

$$\left\| \frac{\partial e}{\partial t} \right\|_{\kappa,-1} \leq \sup_{v\in \mathrm{H}^1_0(\Omega)} \frac{1}{\| v \|_{\kappa,1}} Cl_R \left\{ \sum_{K\in\mathcal{T}_h^n} (h_K^n)^2 \left\| \frac{1}{\sqrt{\kappa}} R_K^n(u^n) \right\|^2_{0,K} \right\}^{\frac{1}{2}} \left\{ \sum_{K\in\mathcal{T}_h^n} \| \sqrt{\kappa}\, \nabla v \|^2_{0,\omega_K^n} \right\}^{\frac{1}{2}}$$

$$+ \sup_{v\in \mathrm{H}^1_0(\Omega)} \frac{1}{\| v \|_{\kappa,1}} Cl_E \left\{ \sum_{K\in\mathcal{T}_h^n} h_E^n \left\| \frac{1}{\sqrt{\kappa}} J_E^n(u^n) \right\|^2_{0,E} \right\}^{\frac{1}{2}} \left\{ \sum_{K\in\mathcal{T}_h^n} \| \sqrt{\kappa}\, \nabla v \|^2_{0,\omega_E^n} \right\}^{\frac{1}{2}}$$

$$+ \sup_{v\in \mathrm{H}^1_0(\Omega)} \frac{1}{\| v \|_{\kappa,1}} \frac{t - t^{\theta,n}}{t^n - t^{n-1}} \left(\kappa \nabla \left(u^n_{h,\tau} - P^n u^{n-1}_{h,\tau} \right), \nabla v \right)$$

$$+ \sup_{v\in \mathrm{H}^1_0(\Omega)} \frac{1}{\| v \|_{\kappa,1}} \left(\Pi_{\mathcal{T}_h} f - \theta\, \Pi_{\mathcal{T}_h} f^n - (1-\theta)\, \Pi_{\mathcal{T}_h} f^{n-1}, v \right) + \sup_{v\in \mathrm{H}^1_0(\Omega)} \frac{1}{\| v \|_{\kappa,1}} \left(f - \Pi_{\mathcal{T}_h} f, v \right)$$

$$+ \sup_{v\in \mathrm{H}^1_0(\Omega)} \frac{1}{\| v \|_{\kappa,1}} \left(\kappa \nabla e^n, \nabla v \right) + \sup_{v\in \mathrm{H}^1_0(\Omega)} \frac{1}{\| v \|_{\kappa,1}} \left(\frac{\delta u^{n-1}_{h,\tau}}{\tau^n}, v \right) + \sup_{v\in \mathrm{H}^1_0(\Omega)} \frac{1}{\| v \|_{\kappa,1}} \frac{t^n - t}{t^n - t^{n-1}} \left(\kappa \nabla \delta u^{n-1}_{h,\tau}, \nabla v \right)$$

and then we easily get (3.7). □

Proposition 3.8. *Under the assumptions on the continuous problem (2.4) and on the discrete formulation (2.8), for each $n = 1, \ldots, N$, there exists a constant C independent of any meshsize, timestep-length, problem-parameter such that*

$$\int_{t^{n-1}}^{t^n} \left\| \tfrac{\partial e}{\partial t} \right\|^2_{\kappa,-1} \mathrm{d}t \leq C \left[(\eta_R^n)^2 + (\eta_f^n)^2 + \left(\theta^2 - \theta + \frac{1}{3} \right) (\eta_\tau^n)^2 \right.$$

$$\left. + \frac{1}{\tau^n} \left\| \delta u^{n-1}_{h,\tau} \right\|^2_{\kappa,-1} + \frac{\tau^n}{3} \left\| \delta u^{n-1}_{h,\tau} \right\|^2_{\kappa,1} \right] + 8 \int_{t^{n-1}}^{t^n} \| \sqrt{\kappa}\, \nabla e^n \|^2_0 \mathrm{d}t\,. \tag{3.8}$$

Proof. Applying Young inequality to (3.7) and integrating in time on the n-th time interval we get (3.8). □

Theorem 3.5 is proved by obtaining (3.5) from equations (3.6) and (3.8).

3.3. Lower bound for the error e

We prove that the terms $(\eta_R^n)^2$ and $(\eta_\tau^n)^2$ bound from below the error e between the solution of the discretized problem and the exact solution of the continuous variational formulation.

Theorem 3.9. *Under the assumptions on the continuous problem (2.4) and on the discrete formulation (2.8), there exist constants $C^{t^n}_{\downarrow,t^{n-1}}$, $n = 1, \ldots, N$, and $C^T_{\downarrow,0} = \max_{n=1,\ldots,N} C^{t^n}_{\downarrow,t^{n-1}}$ independent of any meshsize, timestep-length, problem-parameter, but depending on the parameter θ and on the quality of the partitions $\mathcal{T}_h^n$ such that the inequalities*

$$(\eta_R^n)^2 + (\eta_\tau^n)^2 \leq C^{t^n}_{\downarrow,t^{n-1}} \left[|\!|\!| e |\!|\!|^2_{\kappa,I^n} + (\eta_f^n)^2 + \frac{1}{\tau^n} \left\| \delta u^{n-1}_{h,\tau} \right\|^2_{\kappa,-1} + \tau^n \left\| \delta u^{n-1}_{h,\tau} \right\|^2_{\kappa,1} \right] \tag{3.9}$$

and for each $m = 1, \dots, N$

$$\sum_{n=1}^{m} (\eta_R^n)^2 + \sum_{n=1}^{m} (\eta_\tau^n)^2 \le C_{\downarrow,0}^T \left[\| e \|_{\kappa,[0,t^m]}^2 + \sum_{n=1}^{m} \left(\eta_f^n\right)^2 + \sum_{n=1}^{m} \frac{1}{\tau^n} \left\| \delta u_{h,\tau}^{n-1} \right\|_{\kappa,-1}^2 + \sum_{n=1}^{m} \tau^n \left\| \delta u_{h,\tau}^{n-1} \right\|_{\kappa,1}^2 \right] \tag{3.10}$$

hold true.

To prove this theorem we consider separately the contributions of the equation residual, the inter-element jumps and the time error estimator. Then, proceeding as in [5,20], we collect them to get Theorem 3.9.

3.3.1. *Equation residual*

Here we show how the residual of the equation can bound the error from below on the time interval I^n. For any element $K \in \mathcal{T}_h^n$, let us define the following function in Ω

$$\begin{aligned} w_{R,K}^n(x) &= \begin{cases} (h_K^n)^2 \, \frac{1}{\sqrt{\kappa}} \, R_K^n(u^n) \, b_K^n(x), & \text{if } x \in K, \\ 0, & \text{if } x \notin K, \end{cases} \\ w_{R,\Omega}^n &= \sum_{K \in \mathcal{T}_h^n} w_{R,K}^n \, . \end{aligned}$$

In what follows we will use the property supp $w_{R,K}^n \subseteq K$.

Lemma 3.10. *There exist constants* C_R *and* C_R^* *independent of any meshsize, timestep-length and problem-parameter such that*

$$(h_K^n)^2 \left\| \frac{1}{\sqrt{\kappa}} \, R_K^n(u^n) \right\|_{0,K}^2 \le C_R \left(\frac{1}{\sqrt{\kappa}} \, R_K^n(u^n), w_{R,K}^n \right)_K , \tag{3.11}$$

$$\| w_{R,K}^n \|_{0,K} \le (h_K^n)^2 \left\| \frac{1}{\sqrt{\kappa}} \, R_K^n(u^n) \right\|_{0,K} , \tag{3.12}$$

$$\| \nabla w_{R,K}^n \|_{0,K} \le C_R^* \frac{1}{h_K^n} \| w_{R,K}^n \|_{0,K} \, . \tag{3.13}$$

Proof. These results are obtained exploiting the properties of bubble functions and the finite dimensionality of the residual function [19]. The constants C_R and C_R^* depend on the maximal polynomial degree of the finite element spaces. □

Proposition 3.11. *Under the assumptions on the continuous problem* (2.4) *and on the discrete formulation* (2.8)*, on each time interval* $\left(t^{n-1}, t^n\right)$*,* $\forall \alpha \ge 0$*, we have*

$$\begin{aligned} \left\{ \sum_{K \in \mathcal{T}_h^n} (h_K^n)^2 \left\| \tfrac{1}{\sqrt{\kappa}} \, R_K^n(u^n) \right\|_{0,K}^2 \tau^n \right\}^{\frac{1}{2}} \le C_R C_R^* \Bigg[& 2 \frac{\alpha+1}{\sqrt{2\alpha+1}} \left\{ \| e \|_{\kappa,I^n}^2 + \left(\eta_f^n\right)^2 \right\}^{\frac{1}{2}} \\ & + \left| \theta - \frac{\alpha+1}{\alpha+2} \right| \eta_\tau^n + \frac{\sqrt{\tau^n}}{\alpha+2} \left\| \delta u_{h,\tau}^{n-1} \right\|_{\kappa,1} + \frac{1}{\sqrt{\tau^n}} \left\| \delta u_{h,\tau}^{n-1} \right\|_{\kappa,-1} \Bigg] . \end{aligned} \tag{3.14}$$

Proof. In the following we introduce an arbitrary function of time $b^n(t) \geq 0,\ \forall t \in I^n$. We start by subtracting to $\left(\frac{1}{\sqrt{\kappa}} R^n_\Omega(u^n), w^n_{R,\Omega}\right)$ the continuous variational formulation (2.4) tested against $w^n_{R,\Omega}/\sqrt{\kappa}$ and integrating on the time interval I^n the result times b^n:

$$\int_{t^{n-1}}^{t^n} \sum_{K\in\mathcal{T}_h^n} \left(\frac{1}{\sqrt{\kappa}} R^n_\Omega(u^n), w^n_{R,K}\right)_K b^n \mathrm{d}t = \int_{t^{n-1}}^{t^n} \left\langle \frac{1}{\sqrt{\kappa}} \frac{\partial e^n}{\partial t}, w^n_{R,\Omega}\right\rangle b^n \mathrm{d}t$$
$$- \int_{t^{n-1}}^{t^n} \left(\theta\sqrt{\kappa}\nabla\left(u^n - u^n_{h,\tau}\right) + (1-\theta)\sqrt{\kappa}\nabla\left(u^n - P^n u^{n-1}_{h,\tau}\right), \nabla w^n_{R,\Omega}\right) b^n \mathrm{d}t$$
$$+ \int_{t^{n-1}}^{t^n} \left(\sqrt{\kappa}\nabla e^n, \nabla w^n_{R,\Omega}\right) b^n \mathrm{d}t + \int_{t^{n-1}}^{t^n} \left(\frac{1}{\sqrt{\kappa}}\left(f - \Pi_{\mathcal{T}_h} f\right), w^n_{R,\Omega}\right) b^n \mathrm{d}t$$
$$+ \int_{t^{n-1}}^{t^n} \left(\frac{1}{\sqrt{\kappa}}\left(\Pi_{\mathcal{T}_h} f - \theta\,\Pi_{\mathcal{T}_h} f^n - (1-\theta)\,\Pi_{\mathcal{T}_h} f^{n-1}\right), w^n_{R,\Omega}\right) b^n \mathrm{d}t\,.$$

Now we use equation (3.3) to link the error e^n to the error e and $\delta u^{n-1}_{h,\tau}$.

Then, we apply Cauchy-Schwarz's and Hölder inequality, inequalities (3.11)–(3.13). Then we define the function b^n as in [20]

$$b^n = (\alpha+1)\left(\frac{t - t^{n-1}}{t^n - t^{n-1}}\right)^\alpha \qquad \alpha \geq 0 \tag{3.15}$$

and we proceed as in [5,20] to get (3.14). □

3.3.2. *Inter-element jumps*

Now we consider the faces $E \in \mathcal{E}^n_{h,\Omega}$ and we show how the jumps $J^n_E(u^n)$ can bound the error from below. Let us define

$$w^n_{J,E}(x) = \begin{cases} h^n_E\,\mathcal{P}_E\left(\frac{1}{\sqrt{\kappa}} J^n_E(u^n)\right) b^n_E(x), & \text{if } x \in \mathring{\omega}^n_E, \\ 0, & \text{if } x \notin \mathring{\omega}^n_E, \end{cases}$$
$$w^n_{J,\Omega} = \sum_{E\in\mathcal{E}^n_{h,\Omega}} w^n_{J,E}\,.$$

We remark that $w^n_{J,\Omega}$ vanishes on the faces/edges of the elements $K' \in \mathcal{T}_{h^n,\mathring{\omega}^n}$ inside the elements $K \in \mathcal{T}^n_h$.

In what follows we will use the orthogonality of the face bubble functions.

Lemma 3.12. *There exist constants C_E and C^*_E independent of any meshsize, timestep-length and problem-parameter such that*

$$h^n_E \left\| \frac{1}{\sqrt{\kappa}} J^n_E(u^n)\right\|^2_{0,E} \leq C_E \left(\frac{1}{\sqrt{\kappa}} J^n_E(u^n), w^n_{J,E}\right)_E, \tag{3.16}$$
$$\left\| w^n_{J,E}\right\|_{0,\mathring{\omega}^n_E} \leq \sqrt{h^n_E}\, h^n_E \left\| \frac{1}{\sqrt{\kappa}} J^n_E(u^n)\right\|_{0,E}, \tag{3.17}$$
$$\left\| \nabla w^n_{J,E}\right\|_{0,\mathring{\omega}^n_E} \leq C^*_E \frac{1}{h^n_E} \left\| w^n_{J,E}\right\|_{0,\mathring{\omega}^n_E}. \tag{3.18}$$

Proof. The previous results are derived exploiting the properties of bubble functions and inverse inequalities for the jump functions. The constants C_E and C^*_E depend on the maximal polynomial degree of the finite element spaces. □

Proposition 3.13. *Under the assumptions on the continuous problem (2.4) and on the discrete formulation (2.8), on each time interval I^n we have*

$$\left\{\sum_{E\in\mathcal{E}^n_{h,\Omega}} h^n_E \left\| \frac{1}{\sqrt{\kappa}}\, J^n_E(u^n)\right\|^2_{0,E} \tau^n\right\}^{\frac12} \leq C_E\,(C^*_E + C_R C^*_R)\left[2\frac{\alpha+1}{\sqrt{2\alpha+1}}\left\{|\!|\!| e |\!|\!|^2_{\kappa,I^n} + \left(\eta^n_f\right)^2\right\}^{\frac12} + \left|\theta - \frac{\alpha+1}{\alpha+2}\right| \eta^n_\tau + \frac{\sqrt{\tau^n}}{\alpha+2}\left\|\delta u^{n-1}_{h,\tau}\right\|_{\kappa,1} + \frac{1}{\sqrt{\tau^n}}\left\|\delta u^{n-1}_{h,\tau}\right\|_{\kappa,-1}\right]. \quad (3.19)$$

Proof. We start by integrating $\left(J^n_E(u^n)/\sqrt{\kappa}, w^n_{J,E}\right)$ against the arbitrary bubble function $b^n(t)$. We introduce the continuous variational formulation (2.4) tested against $w^n_{J,E}/\sqrt{\kappa}$.

$$\begin{aligned}
&\int_{t^{n-1}}^{t^n} \sum_{E\in\mathcal{E}^n_{h,\Omega}} \left(\frac{1}{\sqrt{\kappa}}\, J^n_E(u^n), w^n_{J,E}\right)_E b^n \mathrm{d}t \\
&\qquad = \int_{t^{n-1}}^{t^n} \sum_{E\in\mathcal{E}^n_{h,\Omega}} \sum_{K'\in\mathring{\omega}^n_E} \int_{K'} \nabla\cdot\left[\frac{1}{\sqrt{\kappa}}\left(\theta\kappa\nabla\, u^n_{h,\tau} + (1-\theta)\kappa\nabla\, P^n u^{n-1}_{h,\tau}\right) w^n_{J,\Omega}\right] \mathrm{d}\Omega\; b^n \mathrm{d}t \\
&= \int_{t^{n-1}}^{t^n} \left\langle \frac{1}{\sqrt{\kappa}}\,\frac{\partial\, e^n}{\partial t}, w^n_{J,\Omega}\right\rangle b^n \mathrm{d}t + \int_{t^{n-1}}^{t^n} \left(\frac{\kappa}{\sqrt{\kappa}}\nabla\, e^n, \nabla\, w^n_{J,\Omega}\right) b^n \mathrm{d}t + \int_{t^{n-1}}^{t^n} \left(\frac{1}{\sqrt{\kappa}}\left(f - \Pi_{\mathcal{T}_h} f\right), w^n_{J,\Omega}\right) b^n \mathrm{d}t \\
&+ \int_{t^{n-1}}^{t^n} \left(\frac{1}{\sqrt{\kappa}}\left(\Pi_{\mathcal{T}_h} f - \theta\, f^n - (1-\theta)\, f^{n-1}\right), w^n_{J,\Omega}\right) b^n \mathrm{d}t - \int_{t^{n-1}}^{t^n} \sum_{E\in\mathcal{E}^n_{h,\Omega}} \sum_{K'\in\mathring{\omega}^n_E} \left(\frac{1}{\sqrt{\kappa}}\, R^n_K(u^n), w^n_{J,\Omega}\right)_{K'} b^n \mathrm{d}t \\
&\qquad - \int_{t^{n-1}}^{t^n} \frac{t - t^{\theta,n}}{t^n - t^{n-1}} \left(\sqrt{\kappa}\,\nabla\left(u^{n,n}_{h,\tau} - u^{n,n-1}_{h,\tau}\right), \nabla\, w^n_{J,\Omega}\right) b^n \mathrm{d}t\,.
\end{aligned}$$

We apply Cauchy-Schwarz's inequality and (3.3). Then we apply Hölder's inequality, inequalities (3.17), (3.18), definition (3.15) of b^n, inequalities (3.16), (3.14) and regularity hypotheses to get (3.19). □

3.3.3. *Time discretization estimator*

Now we consider the elements $K \in \mathcal{T}^n_h$ and we show how the norm $\left\|\sqrt{\kappa}\nabla\left(u^n_{h,\tau} - P^n u^{n-1}_{h,\tau}\right)\right\|_0$ can bound the error from below. Let us define

$$w^n_{\tau,\Omega} \;=\; \frac{t - t^{\theta,n}}{t^n - t^{n-1}}\left(u^n_{h,\tau} - P^n u^{n-1}_{h,\tau}\right).$$

Proposition 3.14. *Under the assumptions on the continuous problem (2.4) and on the discrete formulation (2.8), on each time interval I^n the following inequality*

$$\eta^n_\tau \leq 2\sqrt{3} C l_E \left\{\tau^n \sum_{E\in\mathcal{E}^n_{h,\Omega}} h^n_E \left\|\frac{1}{\sqrt{\kappa}}\, J^n_E(u^n)\right\|^2_{0,E}\right\}^{\frac12} + 2\sqrt{3} C l_R \left\{\tau^n \sum_{K\in\mathcal{T}^n_h} {h^n_K}^2 \left\|\frac{1}{\sqrt{\kappa}}\, R^n_K(u^n)\right\|^2_{0,K}\right\}^{\frac12} + 4\sqrt{3}\left\{|\!|\!| e |\!|\!|^2_{\kappa,I^n} + \left(\eta^n_f\right)^2\right\}^{\frac12} + 2\sqrt{3}\left(\frac{1}{\sqrt{\tau^n}}\left\|\delta u^{n-1}_{h,\tau}\right\|_{\kappa,-1} + \frac{\sqrt{\tau^n}}{\sqrt{3}}\left\|\delta u^{n-1}_{h,\tau}\right\|_{\kappa,1}\right) \quad (3.20)$$

holds true.

Proof. In the proof we use the relation (3.4) and

$$\int_{t^{n-1}}^{t^n} \left\langle \nabla \cdot (\kappa \nabla u), w^n_{\tau,\Omega} \right\rangle \mathrm{d}t \quad = \quad - \int_{t^{n-1}}^{t^n} \left(\kappa \nabla u, \nabla w^n_{\tau,\Omega} \right) \mathrm{d}t \,.$$

Then we write

$$\left(\theta^2 - \theta + \frac{1}{3}\right) \tau^n \left\| \sqrt{\kappa} \nabla \left(u^n_{h,\tau} - P^n u^{n-1}_{h,\tau} \right) \right\|^2_0 = \int_{t^{n-1}}^{t^n} \frac{t - t^{\theta,n}}{t^n - t^{n-1}} \left(\kappa \nabla \left(u^n_{h,\tau} - P^n u^{n-1}_{h,\tau} \right), \nabla w^n_{\tau,\Omega} \right) \mathrm{d}t$$

$$= - \int_{t^{n-1}}^{t^n} \sum_{E \in \mathcal{E}^n_{h,\Omega}} \left(J^n_E(u^n), \gamma^{[n]}_E(w^n_{\tau,\Omega}) \right)_E \mathrm{d}t - \int_{t^{n-1}}^{t^n} \sum_{K \in \mathcal{T}^n_h} \left(R^n_K(u^n), w^n_{\tau,\Omega} \right)_K \mathrm{d}t - \int_{t^{n-1}}^{t^n} \left\langle \frac{\partial\, e^n}{\partial t}, w^n_{\tau,\Omega} \right\rangle \mathrm{d}t$$

$$+ \int_{t^{n-1}}^{t^n} \left(\kappa \nabla e^n, \nabla w^n_{\tau,\Omega} \right) \mathrm{d}t - \int_{t^{n-1}}^{t^n} \left(f - \Pi_{\mathcal{T}_h} f, w^n_{\tau,\Omega} \right) \mathrm{d}t - \int_{t^{n-1}}^{t^n} \left(\Pi_{\mathcal{T}_h} f - \theta\, \Pi_{\mathcal{T}_h} f^n - (1-\theta)\, \Pi_{\mathcal{T}_h} f^{n-1}, w^n_{\tau,\Omega} \right) \mathrm{d}t \,.$$

Now we observe that the discrete equation (2.8) can be written as

$$\sum_{K \in \mathcal{T}^n_h} \left(R^n_K(u^n), v_h \right)_K + \sum_{E \in \mathcal{E}^n_{h,\Omega}} \left(J^n_E(u^n), \gamma^{[n]}_E(v_h) \right)_E = 0,$$

that we use with the test function $v_h = I_h(w^n_{\tau,\Omega})$. Using (3.3), applying Clément quasi-interpolation inequalities of Lemma 3.1 and Hölder inequality we derive (3.20). □

3.3.4. *Final lower bounds*

Now we collect the results of Propositions 3.11, 3.13 and 3.14 to get a lower bound for the error in terms of η^n_τ. Then we use this result to derive a lower bound for the same quantity with respect to η^n_R.

Proposition 3.15. *Under the assumptions on the continuous problem (2.4) and on the discrete formulation (2.8), on each time interval I^n there exist constants $\tilde{c}^{t^n}_{t^{n-1}}$ and $\tilde{\tilde{c}}^{t^n}_{t^{n-1}}$ independent of any meshsize, timestep-length and problem-parameter, but depending on the quality of the mesh $\mathcal{T}^n_h$ such that*

$$\eta^n_\tau \quad \leq \quad \tilde{c}^{t^n}_{t^{n-1}} \left[\left\{ \| e \|^2_{\kappa, I^n} + \left(\eta^n_f \right)^2 \right\}^{\frac{1}{2}} + \frac{1}{\sqrt{\tau^n}} \left\| \delta u^{n-1}_{h,\tau} \right\|_{\kappa,-1} + \sqrt{\tau^n} \left\| \sqrt{\kappa} \nabla\, \delta u^{n-1}_{h,\tau} \right\|_0 \right], \tag{3.21}$$

$$\eta^n_R \quad \leq \quad \tilde{\tilde{c}}^{t^n}_{t^{n-1}} \left[\left\{ \| e \|^2_{\kappa, I^n} + \left(\eta^n_f \right)^2 \right\}^{\frac{1}{2}} + \frac{1}{\sqrt{\tau^n}} \left\| \delta u^{n-1}_{h,\tau} \right\|_{\kappa,-1} + \sqrt{\tau^n} \left\| \delta u^{n-1}_{h,\tau} \right\|_{\kappa,1} \right]. \tag{3.22}$$

Proof. The proof follows from (3.14), (3.19) and (3.20) proceeding as in [5,20]. □

We finally obtain the lower bound stated by Theorem 3.9 by Propositions 3.15.

4. Computable bound for the coarsening error

In Theorems 3.5 and 3.9 we have obtained a global in time upper bound and a global in time lower bound, respectively. The quantity

$$\sum_{n=1}^{m} \left(\frac{1}{\tau^n} \left\| \delta u^{n-1}_{h,\tau} \right\|^2_{\kappa,-1} + \tau^n \left\| \delta u^{n-1}_{h,\tau} \right\|^2_{\kappa,1} \right)$$

is a global upper bound of the error due to the coarsening and its control is necessary to guarantee a reliable control of global discretization error. In performing the computation in each time-slab we can apply a coarsening

of the mesh used in the previous time-slab and project the solution at the end of the previous time-slab on the new mesh for computing the solution at the end of the current time-slab.

To each coarsening applied from one time-slab to the next one is associated an error that we bound by

$$\frac{1}{\tau^n}\left\|\delta u_{h,\tau}^{n-1}\right\|_{\kappa,-1}^2+\tau^n\left\|\delta u_{h,\tau}^{n-1}\right\|_{\kappa,1}^2$$

and, if it is large, we can perform only a partial coarsening to keep it small, *i.e.*, we do not coarsen some element marked for coarsening. We will see how to choose those elements for which skipping the coarsening is recommended. An efficient computation of the two norms $\left\|\delta u_{h,\tau}^{n-1}\right\|_{\kappa,-1}$ and $\left\|\delta u_{h,\tau}^{n-1}\right\|_{\kappa,1}$ is a key ingredient for an efficient adaptive algorithm.

The H^1 norm can be easily computed, while the H^{-1} term is well defined, but, in general, not easily computable. To easily get a reliable bound for this term we need to reduce the generality in the definition of the approximation $P^n u_{h,\tau}^{n-1}$ kept up to here and define the approximation operator P^n as a quasi-interpolation operator as in [18].

We give details on the computation of $\left\|\delta u_{h,\tau}^{n-1}\right\|_{\kappa,-1}^2$ for the case of P_1 finite elements and constant coefficients following results given in [18,21]. These results can be generalized to $d=3$ and to higher order finite elements having dual spaces with the same properties as the ones used here; examples of constructions of these dual spaces can be found in [13,14]. To deal with discontinuous coefficients a modification of these operators could be applied following [16] in the definition of a local projector on a suitable Lipschitz subset of triangles neighboring a vertex.

Let us temporarily drop the superscript denoting the timestep and let $\mathcal{N}_h$ be the set of the internal nodes of the triangulation $\mathcal{T}_h$ and let $\{\phi_i\}_{i=1,\ldots,\dim\mathcal{N}_h}$ be the set of the standard nodal hat basis functions associated with a given triangulation $\mathcal{T}_h$. In [18] local dual basis functions have been used to define a projection operator satisfying stability and approximation properties. Let us denote by $\{\psi_i\}_{i=1,\ldots,\dim\mathcal{N}_h}$ these dual basis functions satisfying the following properties:

(1) $\operatorname{supp}\phi_i=\operatorname{supp}\psi_i,\quad i=1,\ldots,\dim\mathcal{N}_h$;
(2) $\psi_i\in\mathrm{L}^2(\Omega)$;
(3) $\displaystyle\sum_{i=1}^{\dim\mathcal{N}_h}\psi_i=1$;
(4) $\int_\Omega\psi_i\phi_j=\delta_{i,j}\int_\Omega\phi_i$.

For example, for the P_1 finite elements in 2D these dual basis functions can be defined as follows: on the reference element $\hat{K}$: $\hat{\psi}_i=3\hat{\phi}_i-\sum_{j=1,j\neq i}^3\hat{\phi}_j$, $i=1,\ldots,3$; on the domain Ω as $\psi_i|_K=\hat{\psi}_i\circ F_K^{-1}$, $\forall K\in\omega_i$ and $\psi_i=0$ elsewhere, where $\omega_i=\{K\in\mathcal{T}_h : x_i\in K\}$.

Let us define $\mathrm{V}_h=\operatorname{span}\{\phi_i : i=1,\ldots,\dim\mathcal{N}_h\}\subset\mathrm{H}_0^1(\Omega)$ the standard conforming finite element space and $\mathrm{W}_h=\operatorname{span}\{\psi_i : i=1,\ldots,\dim\mathcal{N}_h\}\subset\mathrm{L}^2(\Omega)\not\subset\mathrm{H}^1(\Omega)$ the dual space. Moreover, $\forall u\in\mathrm{L}^2(\Omega)$ let us define the following projection operators $P_{h,1}:\mathrm{L}^2(\Omega)\to\mathrm{V}_h$ and $Q_{h,1}:\mathrm{L}^2(\Omega)\to\mathrm{W}_h$ by

$$(P_{h,1}\,u,w_h)=(u,w_h),\quad\forall w_h\in\mathrm{W}_h,\qquad(Q_{h,1}\,u,v_h)=(u,v_h),\quad\forall v_h\in\mathrm{V}_h\,.$$

These operators are locally defined and satisfy the following properties whose proofs can be found in or follows from [18,21]:

- $P_{h,1}\,v_h=v_h,\ \forall v_h\in\mathrm{V}_h$ and $Q_{h,1}\,w_h=w_h,\ \forall w_h\in\mathrm{W}_h$;
- $(P_{h,1}\,v,w)=(P_{h,1}\,v,Q_{h,1}\,w)=(v,Q_{h,1}\,w),\ \forall v,w\in\mathrm{L}^2(\Omega)$,

$$(v-P_{h,1}\,v,w)=(v-P_{h,1}\,v,w-Q_{h,1}\,w)=(v,w-Q_{h,1}\,w),\quad\forall v,w\in\mathrm{L}^2(\Omega);$$

– L^2 stability: let us define $\omega_K = \bigcup_{\{K' \in \mathcal{T}_h:\ K' \cap K \neq \emptyset\}} K'$

$$\| P_{h,1}\, v \|^2_{0,T} = \| v \|^2_{0,\omega_K},\ \forall v \in \mathrm{L}^2(\omega_K), \qquad \| Q_{h,1}\, w \|^2_{0,T} = \| w \|^2_{0,\omega_K},\ \forall w \in \mathrm{L}^2(\omega_K);$$

– Optimal approximation properties:

$$\| u - P_{h,1}\, u \|_{0,K} \precsim h_K \,|\, u \,|_{1,\omega_K}, \quad \| u - Q_{h,1}\, u \|_{0,K} \precsim h_T \,|\, u \,|_{1,\omega_K}, \quad \forall u \in \mathrm{H}^1(\omega_K).$$

Let us define $\mathcal{C}_{\mathcal{T}_h^n}$ as the union of all the triangles $K \in \mathcal{T}_h^n$ contained in the support of dual basis functions ψ^n with a non-empty intersection with regions in which $\mathcal{T}_h^{n-1}$ differs from $\mathcal{T}_h^n$ due to a coarsening.

Let us reintroduce the superscript denoting the timestep, if we define P^n as $P^n_{h,1}$ and the value in the node i of the projected solution $P^n u^{n-1}_{h,\tau}$ as

$$\left(P^n u^{n-1}_{h,\tau}\right)_i = \frac{\displaystyle\sum_{K \in \mathcal{C}_{\mathcal{T}_h^n}} \int_K u^{n-1}_{h,\tau}\, \psi^n_i \mathrm{d}\Omega}{\displaystyle\sum_{K \in \mathcal{C}_{\mathcal{T}_h^n}} \int_K \phi^n_i \mathrm{d}\Omega}$$

we have

$$\begin{aligned}
\left\| \delta u^{n-1}_{h,\tau} \right\|_{\kappa,-1} &= \sup_{v \in \mathrm{H}^1_0(\Omega)} \frac{\left(P^n_{h,1}\, u^{n-1}_{h,\tau} - u^{n-1}_{h,\tau}, v\right)}{\| v \|_{1,\kappa}} = \sup_{v \in \mathrm{H}^1_0(\Omega)} \frac{\left(u^{n-1}_{h,\tau} - P^n_{h,1}\, u^{n-1}_{h,\tau}, v - Q^n_{h,1}\, v\right)}{\| v \|_{1,\kappa}} \\
&\leq \frac{\sum_{K \in \mathcal{C}_{\mathcal{T}_h^n}} \left\| u^{n-1}_{h,\tau} - P^n_{h,1}\, u^{n-1}_{h,\tau} \right\|_{0,K} \left\| v - Q^n_{h,1}\, v \right\|_{0,K}}{\| v \|_{1,\kappa}} \\
&\precsim \frac{\sum_{T \in \mathcal{C}_{\mathcal{T}_h^n}} (h^n_K)^2 \left| u^{n-1}_{h,\tau} \right|_{1,\omega_K} |\, v \,|_{1,\omega_K}}{\| v \|_{1,\kappa}} \precsim \left\{ \sum_{K \in \mathcal{C}_{\mathcal{T}_h^n}} \frac{(h^n_K)^4}{\kappa} \left\| u^{n-1}_{h,\tau} \right\|^2_{K,1} \right\}^{\frac{1}{2}}.
\end{aligned}$$

In the previous estimate we have exploited the locality property of the projector operator $P^n_{h,1}$. Exploiting again this property we can easily check which are the elements of $\mathcal{T}_h^{n-1}$ coarsened that cause a large coarsening error and we can chose to skip the coarsening of these elements if the coarsening error is too large.

5. Numerical tests

In this section we describe a simple adaptive algorithm and some numerical results obtained applying a coarsening error control and a suitable projection of the solution at the end of the previous timestep on the mesh used for the current timestep. The obtained results of our strategy will be compared with the results obtained by a similar code, based on the same marking strategy and similar *a posteriori estimates* and a *common refinement approach* [5,20], so that we do not have any coarsening error thanks to the use of the full solution of the previous timestep.

Our code is based on the library LibMesh [12] that allows the presence of hanging nodes, for this reason we need to derive suitable dual basis functions for all the possible constrained primal basis functions. Within the library [12] refinement is performed by splitting an element in four similar children elements; coarsening, when possible, consists of removing the children and set as active the parent element. Neighbouring elements can

differ at most for one level (*level one condition*). Every element marked for refinement is refined, whereas not all the elements marked for coarsening are really coarsened.

In the Appendix we give some details on the construction of primal and dual constrained basis functions.

5.1. **Adaptive algorithm**

Our adaptive approach is based on equidistribution and on the splitting between space and time error estimators described in Remark 3.4. In our coding of the proposed approach to coarsening, we always use a single mesh. We use the described projection for the (degrees of freedom) dofs contained in the set $\mathcal{C}_{\mathcal{T}_h^n}$, whereas we simply interpolate all the other dofs. If to perform a timestep we need to change the mesh several times, we always project/interpolate on the new (current) mesh the old solution obtained on the previous mesh. With this simplifying choice the norm of the coarsening error is bounded (by triangle inequality) by the projection error of each iteration.

For all our computations we start from an initial very coarse mesh and we apply an initial uniform refinement four times. The obtained mesh is the initial mesh for adaptivity of the first timestep that can be refined and coarsened (initially at most of four levels) by the adaptive algorithm.

5.1.1. *Marking strategy*

We choose a space tolerance TOL_Ω and a time tolerance TOL_{I^n} for each timestep.

Then, we initially mark for refining or coarsening each triangle $K \in \mathcal{T}_h^n$ according to the following rules:

– `if`

$$\frac{(1+\alpha)^2\,\mathrm{TOL}_\Omega^2}{N_K}\,|\!|\!|\,u\,|\!|\!|^2_{\kappa,I^n} < \left(\eta^n_{R,K}\right)^2 + \frac{\left(\eta^n_{f,\Pi_{\mathcal{T}_h}}\right)^2}{N_K}$$

mark for refinement;

– `else if`

$$\left(\eta^n_{R,K}\right)^2 < \frac{(1-\alpha)^2\,\mathrm{TOL}_\Omega^2}{N_K}\,|\!|\!|\,u\,|\!|\!|^2_{\kappa,I^n}$$

mark for coarsening.

After this marking stage, we set a variable `buff` $=$ `0` and we perform a loop over the active elements. For each element marked for coarsening we check if the parent element was already visited, if not, if all the children are active and none of the neighbouring elements of the parent is marked for refinement, we count its children not marked for coarsening. If only one of its children is not marked for coarsening and $\left(\eta^n_{R,K}\right)^2$ for that children is less than `buff`, we set `buff` $=$ `buff` $-\left(\eta^n_{R,K}\right)^2$ and we mark for coarsening that element. If we can not mark for coarsening this element or if the children not marked for coarsening are more than one, we unmark for coarsening all the children marked for coarsening and add their $\left(\eta^n_{R,K}\right)^2$ to `buff`.

For the code based on the coarsening error control, we choose a tolerance $\mathrm{TOL}_{\mathrm{coarse}}$ and after the stage previously described we check if an upper bound for the estimated coarsening error is less than the coarsening tolerance. Let us define the estimated coarsening error and its upper bound as

$$\left(\eta^n_{\delta,K}\right)^2 = \frac{\left(h^n_K\right)^4}{\tau^n\,\kappa}\left\|\,u^{n-1}_{h,\tau}\,\right\|^2_{K,1} + \tau^n\,\kappa\left\|\,\delta u^{n-1}_{h,\tau}\,\right\|^2_{K,1}, \qquad \left(\overline{\eta}^n_{\delta,K}\right)^2 = \left(\frac{\left(h^n_K\right)^4}{\tau^n\,\kappa} + \tau^n\,\kappa\right)\left\|\,u^{n-1}_{h,\tau}\,\right\|^2_{K,1}.$$

Every time we apply a coarsening, *i.e.*, we remove the four children K_c and make active their parent K we compute the norms $\left\|\,\delta u^{n-1}_{h,\tau}\,\right\|^2_{K,1}$ and $\sum_{K'\in\mathcal{C}^n_{\mathcal{T}_{h,K}}}\left\|\,u^{n-1}_{h,\tau}\,\right\|^2_{K',1}$, where $\mathcal{C}^n_{\mathcal{T}_{h,K}}$ is the set of the neighbouring elements contained in the support of a dual basis function ψ^n of the parent K. If the children were not already coarsened

we set

$$\left(\mathrm{err}_{\mathrm{c}K,1}^{n}\right)^2 = \left\| \delta u_{h,\tau}^{n-1} \right\|_{K,1}^2,$$
$$\left(\mathrm{err}_{\mathrm{c}K,-1}^{n}\right)^2 = \sum_{K' \in \mathcal{C}_{\mathcal{T}_{h,K}}^{n}} \left\| u_{h,\tau}^{n-1} \right\|_{K',1}^2,$$

else, if the children were obtained by coarsening in a previous adaptive iteration of the current timestep

$$\left(\mathrm{err}_{\mathrm{c}K,1}^{n}\right)^2 = \left\| \delta u_{h,\tau}^{n-1} \right\|_{K,1}^2 + \sum_{c=0}^{3} \left(\mathrm{err}_{\mathrm{c}K_c,1}^{n}\right)^2,$$
$$\left(\mathrm{err}_{\mathrm{c}K,-1}^{n}\right)^2 = \sum_{K' \in \mathcal{C}_{\mathcal{T}_{h,K}}^{n}} \left\| u_{h,\tau}^{n-1} \right\|_{K',1}^2 + \sum_{c=0}^{3} \left(\mathrm{err}_{\mathrm{c}K_c,1}^{n}\right)^2.$$

If the four children K_c of a parent K are marked for coarsening and

$$\sum_{c=0}^{3} \left(\tau^n \kappa \left(\mathrm{err}_{\mathrm{c}K_c,1}^{n}\right)^2 + \frac{\left(h_K^n\right)^4}{\tau^n \kappa} \left(\mathrm{err}_{\mathrm{c}K_c,-1}^{n}\right)^2 \right) + \left(\overline{\eta}_{\delta,K}^{n}\right)^2 \leq \frac{\mathrm{TOL}_{\mathrm{coarse}}^2}{N_K} \left\| u \right\|_{\kappa,I^n}^2 \tag{5.1}$$

we leave the markers unchanged else we unmark for coarsening the children. In (5.1), the sum in the left-hand-side is an upper bound of the coarsening error of the previous coarsening performed in the same timestep. The last term in the left-hand-side is an estimate of the coarsening error we are going to introduce when coarsening the children of K. If we really perform this coarsening, this upper bound of the coarsening error will be replaced by a sharper upper bound included in $\left(\mathrm{err}_{\mathrm{c}K,1}^{n}\right)^2$ and $\left(\mathrm{err}_{\mathrm{c}K,-1}^{n}\right)^2$. If we refine an element previously obtained by coarsening, we transfer to each of its children the coarsening error $\left(\mathrm{err}_{\mathrm{c}K,1}^{n}\right)^2/4$ and $\left(\mathrm{err}_{\mathrm{c}K,-1}^{n}\right)^2/4$.

We mark for enlarging or shortening the timestep-length according to the following rules:

- `if`
$$\left(1+\alpha\right)^2 \mathrm{TOL}_{I^n}^2 \left\| u \right\|_{\kappa,I^n}^2 < \left(\eta_\tau^n\right)^2$$
`then` $\tau^n := \dfrac{\tau^n}{\rho}$;
- `else if`
$$\left(\eta_\tau^n\right)^2 < \left(1-\alpha\right)^2 \mathrm{TOL}_{I^n}^2 \left\| u \right\|_{\kappa,I^n}^2$$
`then` $\tau^n := \dfrac{\tau^n}{\rho}$,

where

$$\rho = \min\left\{ \frac{\left(\eta_\tau^n\right)^2}{\mathrm{TOL}_{I^n}^2 \left\| u \right\|_{\kappa,I^n}^2}, 2 \right\}.$$

We repeat at most ten times the same timestep performing the mesh changes and the timestep-length corrections required by the previous marking strategy if any of the following condition occurs:

- $\left(1+\alpha\right)^2 \mathrm{TOL}_{\Omega}^2 \left\| u \right\|_{\kappa,I^n}^2 < \left(\eta_R^n\right)^2$;
- the marked elements for coarsening are more than $\%C \cdot N_K$, where $\%C$ is a given percentage of the active elements;
- the marked elements for refinement are more than $\%R \cdot N_K$, where $\%R$ is a given percentage of the active elements;
- $\left(\eta_R^n\right)^2 < \left(1-\alpha\right)^2 \mathrm{TOL}_{\Omega}^2 \left\| u \right\|_{\kappa,I^n}^2$;
- $\left(1+\alpha\right)^2 \mathrm{TOL}_{I^n}^2 \left\| u \right\|_{\kappa,I^n}^2 < \left(\eta_\tau^n\right)^2$.

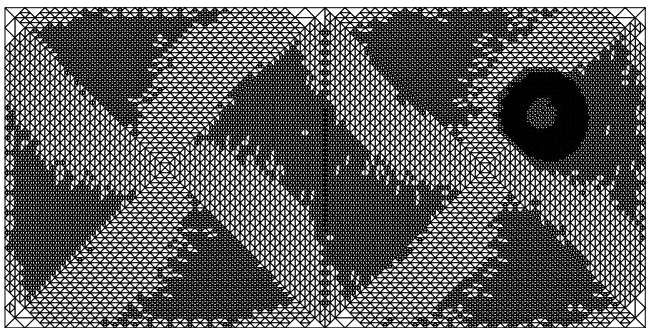

FIGURE 3. First test problem, $\alpha = 0.25$: adapted mesh at $t = 0.2$.

If none of the previous situations occurs, we keep the same mesh and computed solution and we possibly enlarge the timestep-length for the next timestep according to the previous marking strategy for enlarging the timestep-length. In all the simulations we have neglected the data approximation errors.

5.2. **First test problem**

The problem (2.1)–(2.3) is solved in the domain $\Omega = (-1, 1) \times (0, 1)$ and in the time interval $(0, T) = (0, 1)$ with $\kappa = 1$. This test problem was already considered in [5] with discontinuous κ. The initial condition $U^0(x, y)$ and the forcing function $f(x, y, t)$ are defined such that the solution is

$$u(x,y,t) = \begin{cases} u_1(x,y,t), & \text{if } x \geq 0,\ y \geq 0, \\ u_2(x,y,t), & \text{if } x < 0,\ y \geq 0, \end{cases}$$

with

$$u_1(x,y,t) = 500x^2(1-x)^2y^2\,(1-y)^2 \mathrm{e}^{-\mathrm{e}^{R_3-R_1t}\left(\left(x-\frac{1}{2}-\frac{1}{4}\cos(2\pi(1+\sin(2\pi R_1t)))\right)^2+\left(y-\frac{1}{2}-\frac{1}{4}\sin(2\pi(1+\sin(2\pi R_1t)))\right)^2\right)^4}$$

$$\times \frac{1-\mathrm{e}^{-R_4\left(\left(x-\frac{1}{2}\right)^2+\left(y-\frac{1}{2}\right)^2\right)x(1-x)y(1-y)}}{1+\ln(1+R_1t)} + \left(\left(-\frac{R_5}{R_1}-\sin(2\pi t)\right)x^2+\frac{R_5x}{R_1}+\sin(2\pi t)\right)y\,(1-y)$$

and

$$u_2(x,y,t) = \left(\left(\frac{R_5}{R_2}-\sin(2\pi t)\right)x^2+\frac{R_5x}{R_2}+\sin(2\pi t)\right)y\,(1-y)\,.$$

The solution and its derivatives on the edge $x = 0,\ 0 \leq y \leq 1$ are continuous. Moreover, the function u in $x > 0,\ y > 0$ includes a small circular Gaussian peak whose center moves on a circle of radius 0.25 and centre $(0.5, 0.5)$. The parameters R_i are: $R_1 = R_2 = 1$ (R_1 represent the coefficient κ in the square $x \geq 0$ and R_2 is κ in the square $x < 0$), $R_3 = 18$, $R_4 = 100$, $R_5 = 10$.

We have chosen $\mathrm{TOL}_\Omega = 0.35$, $\mathrm{TOL}_{I^n} = 0.20$, $\alpha = 0.25$ and $\mathrm{TOL}_{\mathrm{coarse}} = 0.5(\mathrm{TOL}_\Omega + \mathrm{TOL}_{I^n})$ with a maximum τ^n equal to 0.01, $\%C = 0.1$ and $\%R = 0.4$. In Figure 3 we report the mesh accepted by the adaptive algorithm with control of coarsening error at the time $t = 0.2$ and in Figure 4 the corresponding computed solution.

In Figure 5 we compare the number of active degrees of freedom of the meshes used for the simulation with the proposed coarsening error approach described in this paper and with the approach that uses the common refinement of the meshes $\mathcal{T}_h^n$ and $\mathcal{T}_h^{n-1}$ [5,20]. In Figure 6 we compare the timestep-lengths produced by the two

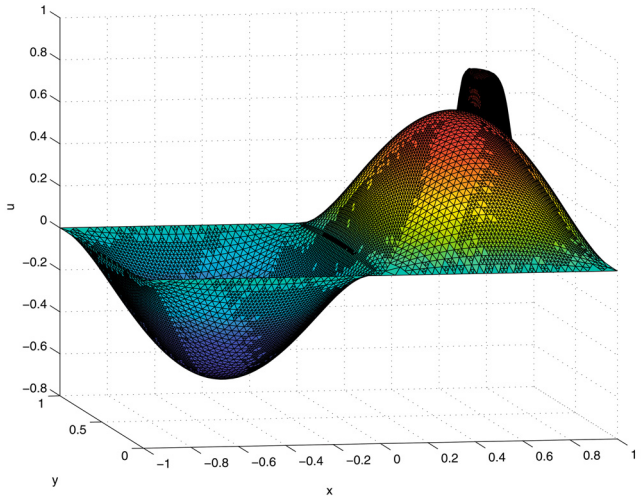

FIGURE 4. First test problem, $\alpha = 0.25$: computed solution at $t = 0.2$.

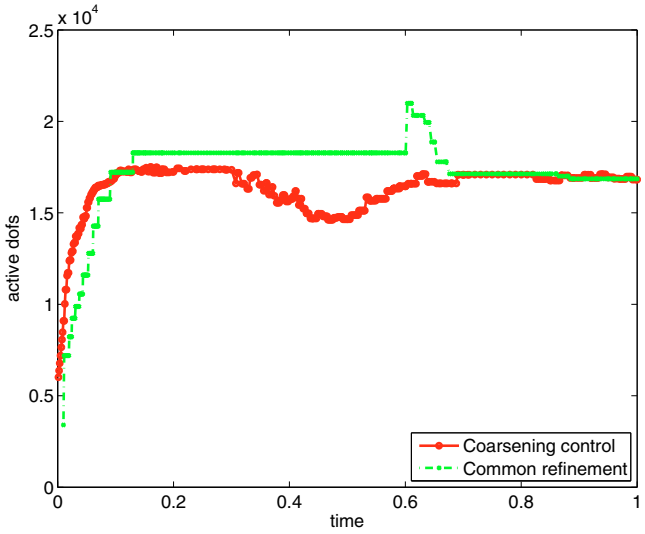

FIGURE 5. First test problem, $\alpha = 0.25$: comparison of required dofs.

adaptive algorithms with and without the coarsening error control, and in Figure 7 we compare the true error $\|\!|\!| e |\!|\!|_{\kappa, I^n}$ produced by the two algorithms. The coarsening control approach requires slightly more dofs up to $t \approx 0.1$, then uses less dofs up to $t \approx 0.7$. After this time the two approaches have similar requirements in terms of dofs. Concerning the timestep-length τ^n the coarsening control approach is slightly more efficient usually requiring comparable or larger τ^n.

In Figures 8–10 we report the same quantities corresponding to the same tolerances, but with $\alpha = 0.5$. In this case the coarsening control approach and the common refinement approach work in a slightly different way on the mesh and the τ^n, producing comparable values of the true error.

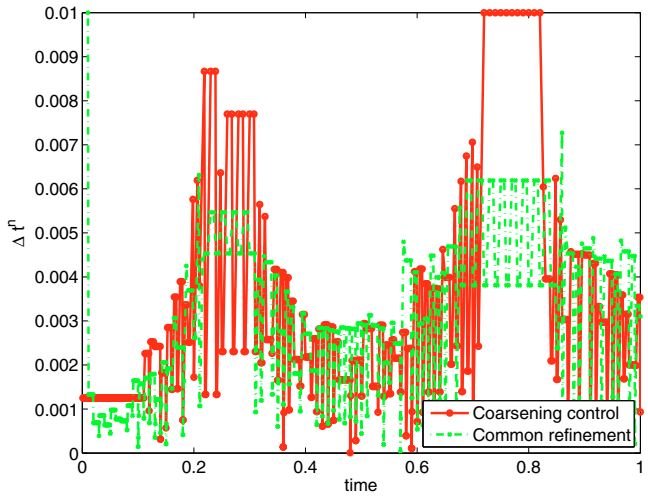

FIGURE 6. First test problem, $\alpha = 0.25$: comparison of required τ^n.

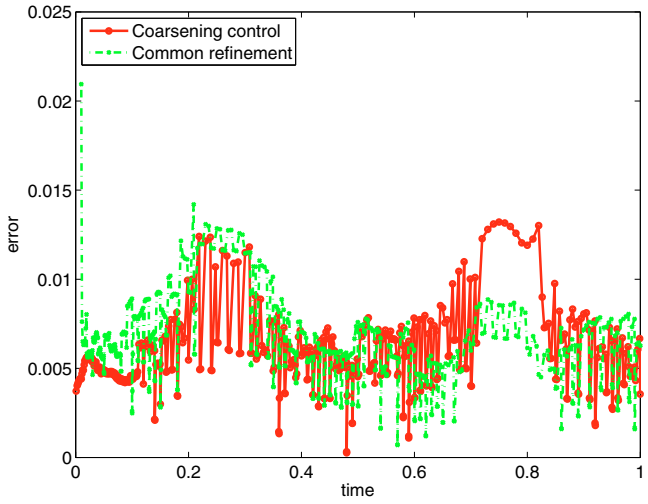

FIGURE 7. First test problem, $\alpha = 0.25$: comparison of true errors $\| e \|_{\kappa, I^n}$.

Figures 11 and 12 compare the effectivity indices

$$e.i. = \frac{\sqrt{(\eta_R^n)^2 + (\eta_\tau^n)^2}}{\| e \|_{\kappa, I^n}}$$

of the two approaches and for the two simulations performed. We can see that the effectivity indices are comparable and bounded between values 3 and 8. This property confirm the validity of the estimators and, thanks the previous results, of the whole adaptive algorithms.

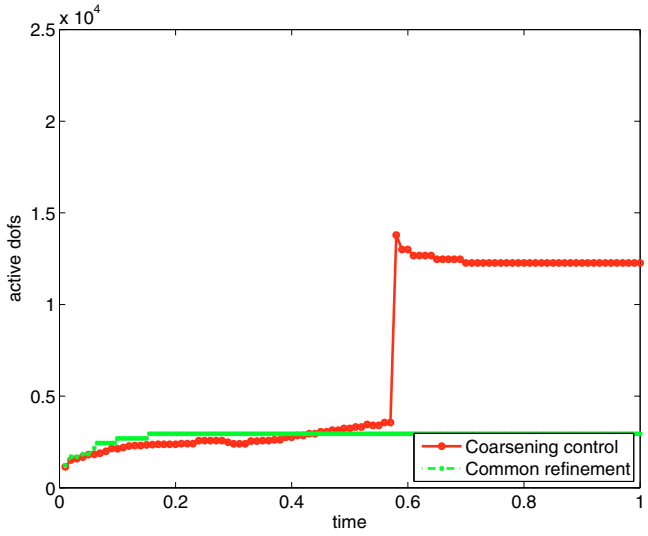

FIGURE 8. First test problem, $\alpha = 0.5$: comparison of required dofs.

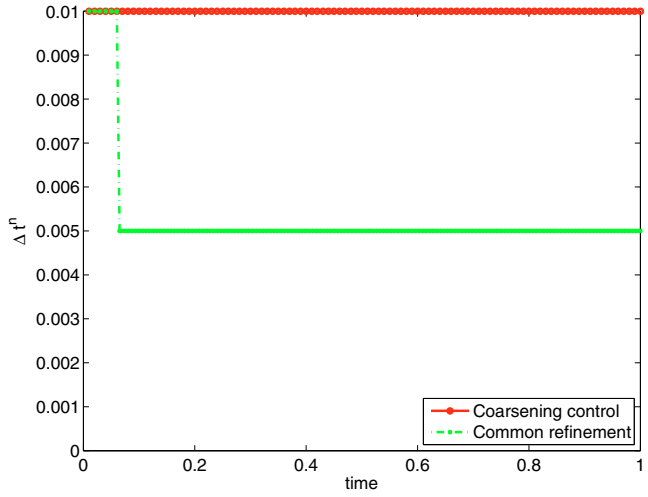

FIGURE 9. First test problem, $\alpha = 0.5$: comparison of required τ^n.

5.3. **Second test problem**

The problem (2.1)–(2.3) is solved in the domain $\Omega = (0,1) \times (0,1)$ and in the time interval $(0,T) = (0,1)$ with $\kappa = 1$. The initial condition $U^0(x,y)$ and the forcing function $f(x,y,t)$ are defined such that the solution is

$$
\begin{aligned}
u(x,y,t) \;=\; & R_5\, x\,(1-x)\,y\,(1-y) + 2000\, x^2\,(1-x)^2\,y^2\,(1-y)^2 \\
& \times \mathrm{e}^{-\mathrm{e}^{(R_3-t)}\left(\left(x-\frac{1}{2}-\frac{1}{4}\cos(2\pi t)\right)^2+\left(y-\frac{1}{2}-\frac{1}{4}\sin(2\pi t)\right)^2\right)^4} \left(1-\mathrm{e}^{-R_4\left(\left(x-\frac{1}{2}\right)^2+\left(y-\frac{1}{2}\right)^2\right)x(1-x)y(1-y)}\right) \\
& \times \mathrm{e}^{-10\,000\,{R_2}^2(t-0.1)^2(t-0.5)^2(t-0.9)^2}.
\end{aligned}
$$

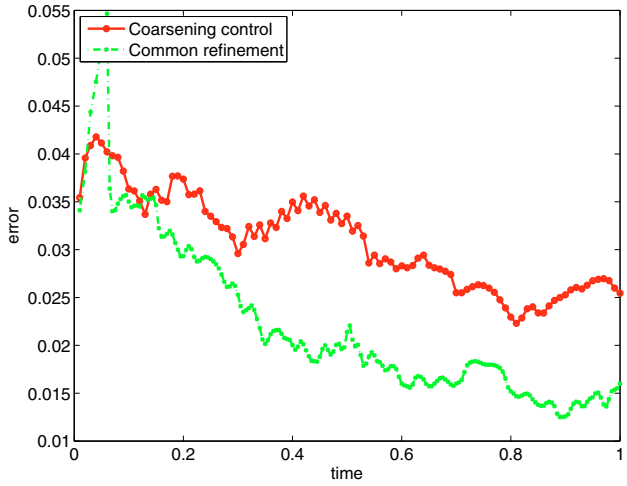

FIGURE 10. First test problem $\alpha = 0.5$: comparison of true errors $\| e \|_{\kappa, I^n}$.

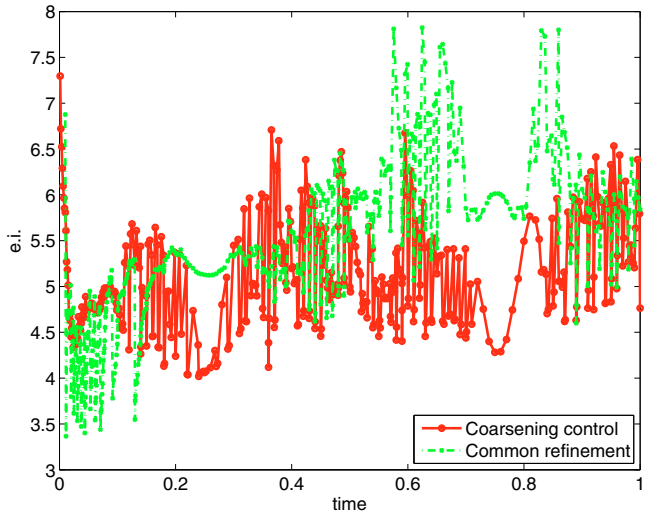

FIGURE 11. First test problem, $\alpha = 0.25$: comparison of effectivity indices.

The solution is the superposition of a global bubble whose maximum value is controlled by R_5 and a small circular Gaussian peak as the one of Section 5.2 running around the centre of the domain. This peak is usually dumped to zero except small time intervals around times $t = 0.1$, $t = 0.5$ and $t = 0.9$. This test case aims at investigating the performances of the two algorithms when the solution displays strongly different behaviours in time evolution.

The parameters R_i are: $R_1 = 100$, $R_2 = 2$ (this parameter controls the size of the time windows in which the Gaussian peak is not dumped), $R_3 = 18$, $R_4 = 100$, $R_5 = 2$.

First, we have chosen $\mathrm{TOL}_\Omega = 0.4$, $\mathrm{TOL}_{I^n} = 0.4$, $\alpha = 0.2$ and $\mathrm{TOL}_{\mathrm{coarse}} = 0.5(\mathrm{TOL}_\Omega + \mathrm{TOL}_{I^n})$ with a maximal τ^n equal to 0.01, $\%C = 0.5$ and $\%R = 0.4$.

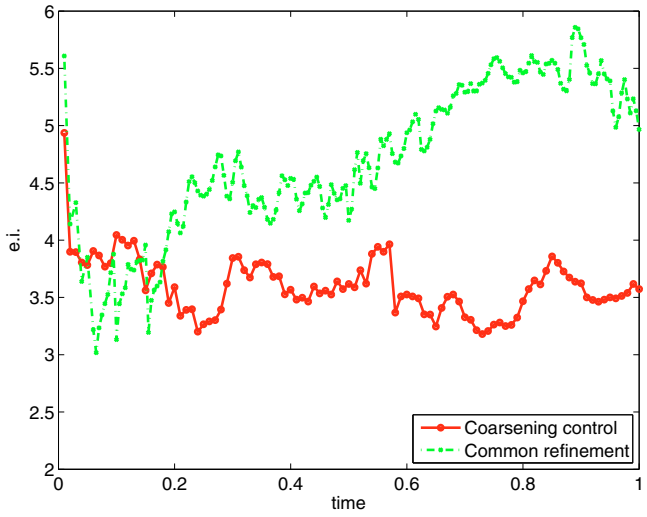

FIGURE 12. First test problem, $\alpha = 0.5$: comparison of effectivity indices.

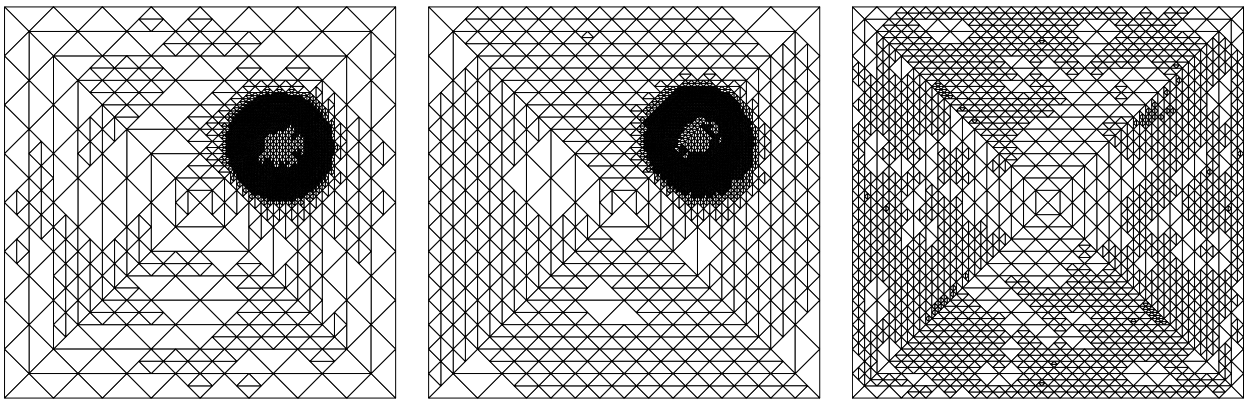

FIGURE 13. Second test problem: adapted meshes at $t = 0.1, 0.12, 0.17$.

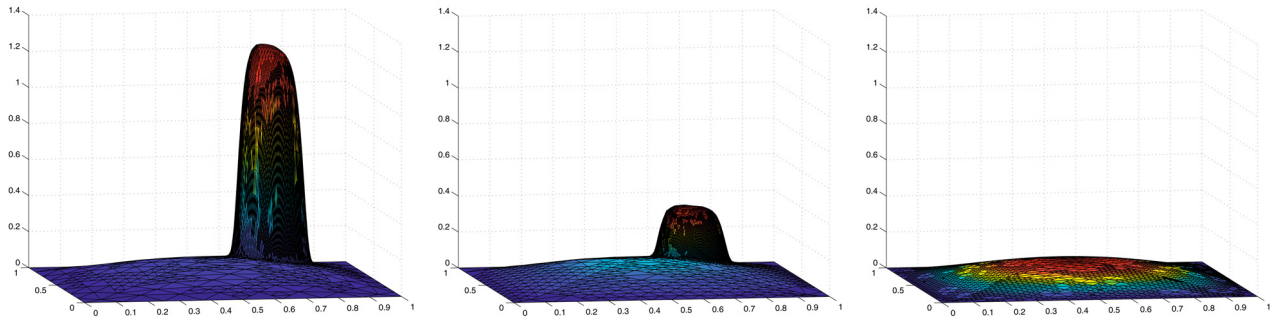

FIGURE 14. Second test problem: computed solutions at $t = 0.1, 0.12, 0.17$.

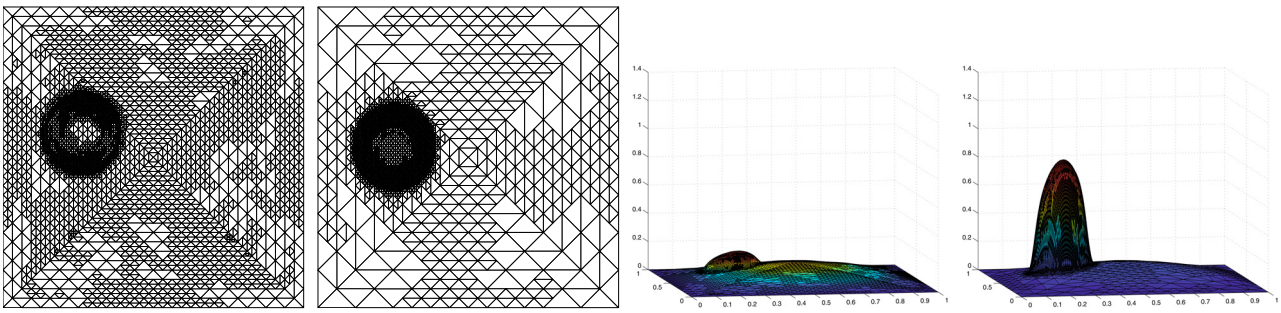

FIGURE 15. Second test problem: computed meshes and solutions at $t = 0.45, 0.48$.

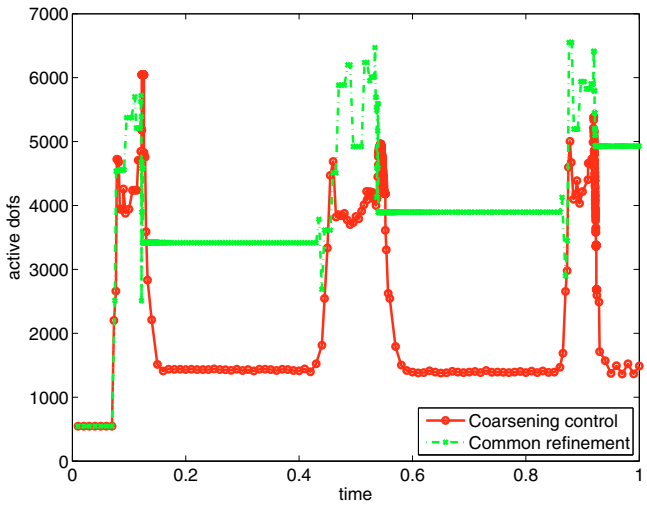

FIGURE 16. Second test problem: comparison of required dofs.

Figure 13 displays the meshes produced by the adaptive algorithm during the first abrupt dumping of the Gaussian peak. Figures 14 displays the solutions produced by the adaptive algorithm corresponding to the previous mesh. Figure 15 displays the meshes and the solutions produced by the adaptive algorithm during the second abrupt growing of the Gaussian peak.

Figure 16 displays the number of dofs requested by the two algorithms and Figures 17 and 18 the corresponding τ^n and true errors. Figure 19 report the two components of the bound of the coarsening error.

Figure 20 displays the dofs requested by the two compared algorithms, the corresponding τ^n and true errors when $\mathrm{TOL}_\Omega = 0.2$, $\mathrm{TOL}_{I^n} = 0.3$ and $\alpha = 0.2$.

6. Conclusions

In this paper we present *a posteriori error estimates* for the heat equation involving terms for the control of a coarsening error due to a projection/interpolation of the solution at the end of the previous timestep on the new adapted mesh produced to perform the current timestep. This approach aims at avoiding the introduction of a common refinement of these two meshes, that can be very expensive in term of coding and computation to be constructed and maintained at each adaptive iteration within the same timestep. In Section 5 we propose

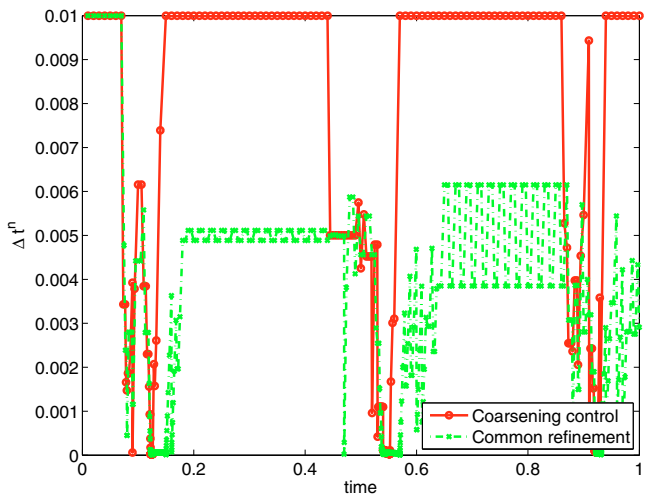

FIGURE 17. Second test problem: comparison of required τ^n.

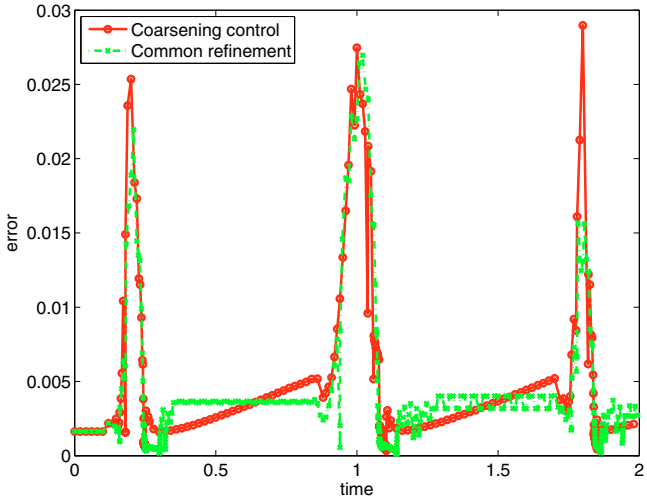

FIGURE 18. Second test problem: comparison of true errors $\| e \|_{\kappa, I^n}$.

a very simple adaptive algorithm based on this coarsening control that skips this common refinement mesh. Although this approach introduces an error at each adaptive iteration, trough a suitable control of it, the true error and the resources required by the adaptive algorithm are not larger than the error and the resources required by an adaptive algorithm that exactly deals with the solution at the end of the previous timestep *via* the common refinement mesh. Numerical computations have shown that we have a slight redistribution of the mesh and timestep-length requirements without any remarkable effect on the true error. Both the adaptive approaches perform very well on both the test problems considered, with different properties in term of mesh and timestep-length adaptation.

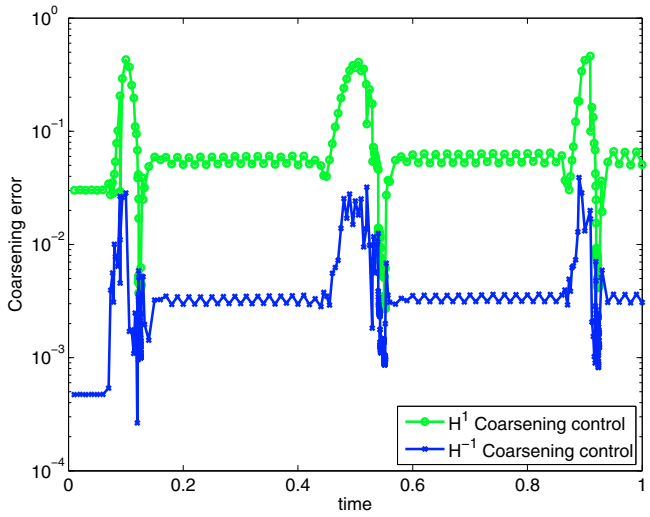

FIGURE 19. Second test problem: coarsening estimated errors.

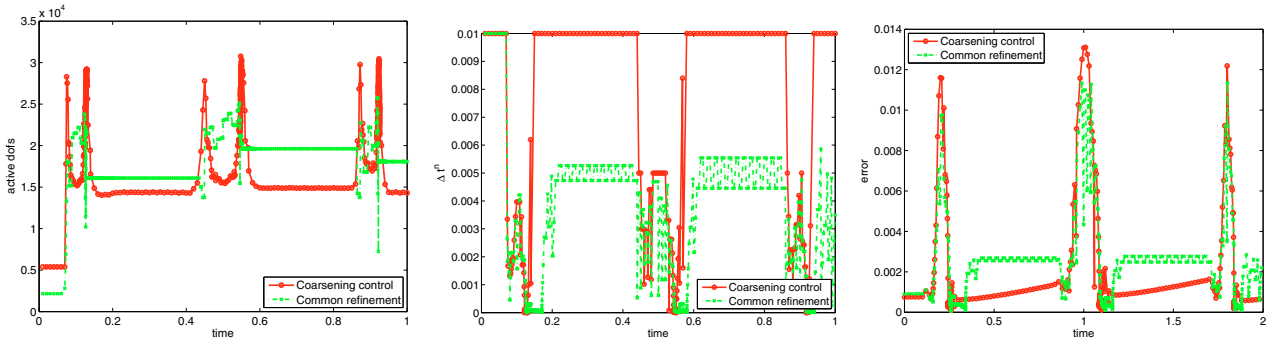

FIGURE 20. Second test problem, $\mathrm{TOL}_\Omega = 0.2$, $\mathrm{TOL}_{I^n} = 0.3$: comparison of active dofs, τ^n and true errors $\| e \|_{\kappa, I^n}$.

A. Appendix

In this section we focus on 2D finite elements and give some details on the construction of primal and dual constrained basis functions.

A.1. Constrained primal and dual basis functions

We consider linear finite elements, so the possible constrained degrees of freedom (dofs) are related to the vertices of the triangle that are midpoints for the parent element. It is convenient to write the basis functions of the constrained active children elements as suitable basis functions defined on the parent element. For example, let us consider the parent reference element in Figure 21: if the dof in the node n_3 is constrained, the solution on the edge between the nodes n_0-n_1 has to be linear and the active children K_0, K_1 and K_3 are constrained. We can simply deal with this constraint defining suitable Lagrange linear basis functions on the parent element associated to the nodes n_0, n_1, n_2, n_4, n_5. We remark that, thanks to the level one condition we can always deal with the constraints writing suitable basis functions on the parent element. If a two or more level mismatch

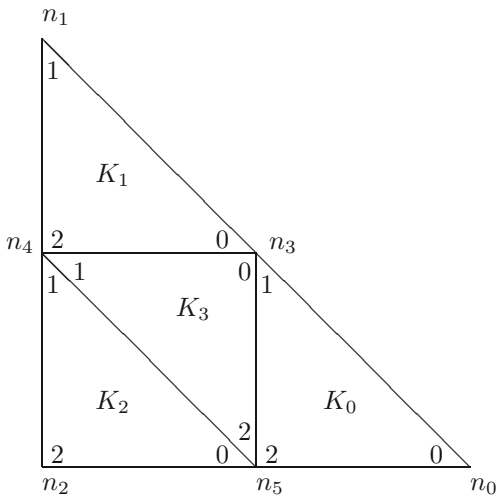

FIGURE 21. Parent reference linear element $\hat{K}$ with its children.

was allowed, we would need to go from our active constrained element to its ancestor of degree equal to the level of mismatch.

Let $\hat{\phi}_i$ be the primal basis function on the parent reference element $\hat{K}$ associated to the node n_i, and let $\hat{\psi}_i$ be the corresponding dual basis function. Furthermore, let $\hat{\varphi}_j$ be the unconstrained primal basis functions of the (active) children elements.

The constrained primal and dual basis functions are described as

$$\hat{\phi}_i = \sum_{c=0}^{3}\sum_{j=0}^{2} a_{\hat{\phi}_i}(c,j)\hat{\varphi}_j \circ F^{-1}_{\hat{K}_c,\hat{K}}, \qquad \hat{\psi}_i = \sum_{c=0}^{3}\sum_{j=0}^{2} a_{\hat{\psi}_i}(c,j)\hat{\varphi}_j \circ F^{-1}_{\hat{K}_c,\hat{K}}, \tag{A.1}$$

where the coefficients $a_{\hat{\phi}_i}(c,j)$ and $a_{\hat{\psi}_i}(c,j)$, $c = 0,\dots,3$, $j = 0,\dots,2$ are given in the sequel and differ according to the number of constrained dofs.

A.1.1. *One constrained dof on the parent element* $\hat{K}$

Let n_3 be the constrained dof. The coefficients in (A.1) are given in the following matrices:

$$a_{\hat{\phi}_0} = \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{bmatrix} \quad a_{\hat{\phi}_1} = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{bmatrix} \quad a_{\hat{\phi}_4} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad a_{\hat{\phi}_5} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$a_{\hat{\psi}_0} = \begin{bmatrix} \frac{103}{26} & \frac{25}{26} & \frac{-41}{26} \\ \frac{25}{26} & \frac{-53}{26} & \frac{11}{26} \\ 0 & 0 & 0 \\ \frac{25}{26} & \frac{11}{26} & \frac{-41}{26} \end{bmatrix} \quad a_{\hat{\psi}_1} = \begin{bmatrix} \frac{-53}{26} & \frac{25}{26} & \frac{11}{26} \\ \frac{25}{26} & \frac{103}{26} & \frac{-41}{26} \\ 0 & 0 & 0 \\ \frac{25}{26} & \frac{-41}{26} & \frac{11}{26} \end{bmatrix} \quad a_{\hat{\psi}_4} = \begin{bmatrix} \frac{22}{65} & \frac{-6}{13} & \frac{-34}{65} \\ \frac{-6}{13} & \frac{-82}{65} & \frac{174}{65} \\ 0 & 0 & 0 \\ \frac{-6}{13} & \frac{174}{65} & \frac{-34}{65} \end{bmatrix} \quad a_{\hat{\psi}_5} = \begin{bmatrix} \frac{-82}{65} & \frac{-6}{13} & \frac{174}{65} \\ \frac{-6}{13} & \frac{22}{65} & \frac{-34}{65} \\ 0 & 0 & 0 \\ \frac{-6}{13} & \frac{-34}{65} & \frac{174}{65} \end{bmatrix}.$$

In this case we choose to define the constrained basis functions on the parent reference element with support restricted to the children $K_0 \cup K_1 \cup K_3$ and to leave unchanged the basis functions on the element K_2. With this

choice, if the element K_2 is refined we can use the same constrained reference basis functions on the element K_2 that become the parent for the elements obtained by its refinement.

A.1.2. *Two constrained dofs on the parent element* $\hat{K}$

Let n_4 and n_5 be the constrained dofs. The coefficients in (A.1) are given in the following matrices:

$$a_{\hat{\phi}_0} = \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix} \quad a_{\hat{\phi}_1} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix} \quad a_{\hat{\phi}_2} = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \quad a_{\hat{\phi}_3} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

$$a_{\hat{\psi}_0} = \begin{bmatrix} \frac{85}{24} & \frac{-25}{24} & \frac{25}{24} \\ \frac{-25}{24} & \frac{25}{24} & \frac{-5}{24} \\ \frac{25}{24} & \frac{-5}{24} & \frac{-35}{24} \\ \frac{-25}{24} & \frac{-5}{24} & \frac{25}{24} \end{bmatrix} \quad a_{\hat{\psi}_1} = \begin{bmatrix} \frac{25}{24} & \frac{-25}{24} & \frac{-5}{24} \\ \frac{-25}{24} & \frac{85}{24} & \frac{25}{24} \\ \frac{-5}{24} & \frac{25}{24} & \frac{-35}{24} \\ \frac{-25}{24} & \frac{25}{24} & \frac{-5}{24} \end{bmatrix} \quad a_{\hat{\psi}_4} = \begin{bmatrix} \frac{-7}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{-7}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{11}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{bmatrix} \quad a_{\hat{\psi}_5} = \begin{bmatrix} \frac{-5}{4} & \frac{11}{4} & \frac{-1}{2} \\ \frac{11}{4} & \frac{-5}{4} & \frac{-1}{2} \\ \frac{-1}{2} & \frac{-1}{2} & \frac{1}{4} \\ \frac{11}{4} & \frac{-1}{2} & \frac{-1}{2} \end{bmatrix}.$$

In this case the support of the primal and dual constrained basis functions is the whole element $\hat{K}$.

A.1.3. *Tree constrained dofs on the parent element* $\hat{K}$

The constrained dofs are n_3, n_4 and n_5, the primal and dual basis functions coincide with the unconstrained basis functions on the parent element and can be defined by the following matrices:

$$a_{\hat{\phi}_0} = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} \quad a_{\hat{\phi}_1} = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \quad a_{\hat{\phi}_2} = \begin{bmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

$$a_{\hat{\psi}_0} = \begin{bmatrix} 3 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \\ 1 & -1 & 1 \end{bmatrix} \quad a_{\hat{\psi}_1} = \begin{bmatrix} -1 & 1 & -1 \\ 1 & 3 & 1 \\ -1 & 1 & -1 \\ 1 & 1 & -1 \end{bmatrix} \quad a_{\hat{\psi}_2} = \begin{bmatrix} -1 & -1 & 1 \\ -1 & -1 & 1 \\ 1 & 1 & 3 \\ -1 & 1 & 1 \end{bmatrix}.$$

Acknowledgements. The author would like to thank Prof. Barbara Wohlmuth for her fruitful suggestions and discussions about the computation of the coarsening error and the anonymous Referees for several helpful comments and suggestions which significantly improved the paper. This work was supported by Italian funds MIUR-PRIN-2006 "Nuove tecniche di accoppiamento di modelli e di metodi numerici nel trattamento di problemi differenziali ", coordinated by Professor Franco Brezzi and by Regione Piemonte *via* the project "AirToLyMi": "Modeling and simulating sustainable mobility strategies. A study of three real test cases: Turin, Lyon, Milan" (CIPE grant 2006).

REFERENCES

[1] I. Babuška and W.C. Rheinboldt, Error estimates for adaptive finite element method. *SIAM J. Numer. Anal.* **15** (1978) 736–754.

[2] R. Becker and R. Rannacher, An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numer. 10* (2001) 1–102.

[3] A. Bergam, C. Bernardi and Z. Mghazli, A posteriori analysis of the finite element discretization of some parabolic equations. *Math. Comp.* **74** (2004) 1117–1138.

[4] C. Bernardi and R. Verfürth, Adaptive finite element methods for elliptic equations with non-smooth coefficients. *Numer. Math.* **85** (2000) 579–608.
[5] S. Berrone, Robust *a posteriori* error estimates for finite element discretizations of the heat equation with discontinuous coefficients. *ESAIM: M2AM* **40** (2006) 991–1021.
[6] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*. North-Holland Publishing Company, Amsterdam (1978).
[7] P. Clément, Approximation by finite element functions using local regularization. *RAIRO Anal. Numér.* **9** (1975) 77–84.
[8] W. Dörfler, A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.* **33** (1996) 1106–1124.
[9] M. Dryja, M.V. Sarkis and O.B. Widlund, Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions. *Numer. Math.* **72** (1996) 313–348.
[10] K. Eriksson and C. Johnson, Adaptive finite element methods for parabolic problems. V. Long-time integration. *SIAM J. Numer. Anal.* **32** (1995) 1750–1763.
[11] K. Eriksson, D. Estep, P. Hansbo and C. Johnson, Introduction to adaptive methods for differential equations. *Acta Numer.* **4** (1995) 105–158.
[12] B.S. Kirk, J.W. Peterson, R. Stogner and S. Petersen, *LibMesh*. The University of Texas, Austin, CFDLab and Technische Universität Hamburg, Hamburg, http://libmesh.sourceforge.net.
[13] B.P. Lamichhane and B.I. Wohlmuth, Higher order dual Lagrange multiplier spaces for mortar finite element discretizations. *Calcolo* **39** (2002) 219–237.
[14] B.P. Lamichhane, R.P. Stevenson and B.I. Wohlmuth, Higher order mortar finite element methods in 3D with dual Lagrange multiplier bases. *Numer. Math.* **102** (2005) 93–121.
[15] P. Morin, R.H. Nochetto and K.G. Siebert, Convergence of adaptive finite element methods. *SIAM Rev.* **44** (2002) 631–658.
[16] M. Petzoldt, *A posteriori* error estimators for elliptic equations with discontinuous coefficients. *Adv. Comput. Math.* **16** (2002) 47–75.
[17] M. Picasso, Adaptive finite elements for a linear parabolic problem. *Comput. Methods Appl. Mech. Engrg.* **167** (1998) 223–237.
[18] L.R. Scott and S. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comput.* **54** (1990) 483–493.
[19] R. Verfürth, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. John Wiley & Sons, Chichester-New York (1996).
[20] R. Verfürth, *A posteriori* error estimates for finite element discretization of the heat equations. *Calcolo* **40** (2003) 195–212.
[21] B.I. Wohlmuth, A mortar finite element method using dual spaces for the Lagrange multiplier. *SIAM J. Numer. Anal.* **38** (2000) 989–1012.
[22] O.C. Zienkiewicz and J.Z. Zhu, A simple error estimator and adaptive procedure for practical engineering analysis. *Internat. J. Numer. Methods Engrg.* **24** (1987) 337–357.